

# *B***→***Xs* **study with the hadronic tagging method at** *Belle*

### **Hanjin Kim**

Institute of Physics and Applied Physics Yonsei Univ. [hanjin@yhep.yonsei.ac.kr](mailto:hanjin@yhep.yonsei.ac.kr?subject=)



# INDEX

## 1. Introduction

## 2. Simulation

2.(1) Signal generation

2.(2) Signal & Bkg. simulation, and pre-selection

### 3. Event Selection

- 3.(1) Off-timing veto
- 3.(2)  $P(\pi^0)$ & $P(\eta)$  Cut
- 3.(3) Continuum Suppression
- 3.(4) ECL variables, MVA
- 3.(5) cos $\theta_e$  cut
- 3.(6) Selection efficiency
- 4. Expected Yield
- 5. Summary & Plan
- 6. Back Up

# 1. INTRODUCTION



degraded since the recent NNLO correction was done. Measurements of the branching fraction and a comparison and a comparison **with the SM prediction at NNLO. A. Limosani, BN 1035**

### **A good probe to search for the New Physics**

**[1] T. Hurth, E. Lunghi, and W. Porod, Nucl. Phys. B704, 56 (2005), hep-ph/0312260. [2] M. Misiak et al., Phys. Rev. Lett. 98, 022002 (2007).**

# 1. INTRODUCTION

## This study will measure  $E<sub>\gamma</sub>$  spectrum, isospin and CP asymmetry

*The hadronic tagging method makes use of "EKP fullrecon"* 

This will let us analyze in the B-rest frame with the informations of fully-reconstructed tag side *B*

- Isospin asymmetry can be directly measured Provided the charge info. of tagged B, we can directly measure isospin asymmetry, *AI*.

- Improvement in the systematic uncertainty Resolutions of signal is significantly improved! ( As Yook said.. )

### - Improvement in the statistical uncertainty

 With about 4 times more data available, we can expected to be able to reduce the statistical uncertainty compared to the study by Babar using the same method.



### Related studies and their results from *Belle*

 $A_{CP} = 0.002 \pm 0.050_{stat} \pm 0.030_{syst}$ *S. Nishida(2004), PRL 93, 0318038*

*P. Koppenburg(2004), PRL 061803*

*A. Limosani(2009), PRL 103, 241801* **Inclusive results** 

# 2. SIMULATION



(1)24 mil. of events was generated using Evtgen and simulated using Gsim

 $(2)B \rightarrow X_s \gamma$  MC samples go through the random selection

defined by the Kagan-Neubert model<sup>[1]</sup>

Heavy Quark Parameters are used for the K-N dist. parameters.

**HFAG values**  $m_b = 4.574$  GeV,  $\mu_{\pi}^2 = 0.459$  GeV<sup>2</sup>

 $(3)B \rightarrow K^* \gamma$  MC samples are generated according to the ratio  $Xs: K^* = 88: 12$ , and simulated.



Table 1. The number of signal MC events

# 2. SIMULATION *(2) Signal & Bkg. MC simulation, and Pre-selection*



**SYW2014, Jan. 14th, 2014**

Table 3. The number of event before/after the fullrecon & pre-selection of sig/Generic/Continuum.

Continuum  $\vert$  5 streams  $\vert$  266364

**Figure 2. sig/bkg. simulation & pre-cut results**

10

 $-0.4 - 0.2$ 

# 3. EVENT SELECTION

## *Selection Criteria*

### π<sup>0</sup>&η and addbg(911, off-timing) particles are dominant.



### Selection Criteria



✻ In each step, tag-correction is applied.

### Analysis Regions

Region  $1:1.3 < E<sub>Y</sub>B < 1.8$  GeV

Region  $11 : 1.8 \le E_Y^B \le 2.0$  GeV (Optimization regions)

Region  $111:2.0 \text{ GeV} < E_1^B < 2.8 \text{ GeV}$ 

# 3. EVENT SELECTION

## *(1) Off-timing veto*

911(addbg) particles in MC correspond to the off-timing events, overlay beam backgrounds(QED bkg) in DATA. All the MC candidates with idhep=911 ( isthep= -91 ) are vetoed. All the off-timing on Triger Count is vetoed in the case of DATA.

## *(2) P(***π***0)<0.3 & P(***η***)<0.3 Cut*

Probability of the candidates to be a daughter of  $\pi^0$  or  $\eta$  was calculated, according to the distribution of the mass combined by the candidate gamma paired with another gamma.

( Only signal-side pairs were considered since there was not a significant advantage in using the others )



Fig 4. The related distributions to *P(*π*0)<0.3 & P(*η*)<0.3 Cut*

# 3. EVENT SELECTION *(3) |costhrust| < 0.8 Cut*

To suppress the continuum events, thrust angle cut was employed.

$$
\text{Thrust \; Axis} \; = \; \frac{\sum \stackrel{\longrightarrow \; longitude}{p_i}}{\stackrel{\longrightarrow}{p_i}}
$$

Angle btw. two thrust axis of sig-side and tag-side was calculated.

$$
\cos\theta_{\text{thrust}} = \overrightarrow{A}_{sig} \cdot \overrightarrow{A}_{tag}
$$

### *(4)* ECL variables, MVA



Gamma shower

Hadronic showe



So we decided to make use of **ECL shower parameters** to veto hadronic shower events.

### ECL shower parameters

- Shower width : The RMS width of shower shape
- Shower mass: The combined mass of showers
- E9/E25 : (E Deposited in 3X3 ECL cluster ) / (E Deposited in 5X5 ECL cluster)

Fig 6. The different character of gamma showers and hadronic showers. (from K. Bernlöhr, Imaging very high energy gamma-ray telescopes)

**9**



# 3. EVENT SELECTION

## *(4)* ECL variables, MVA



Fig 6. ECL shower parameters in Optimization region, SIG/BB ( 1.8~2.0 GeV)

These variables are correlated ( correlation matrix in back-up ), we tried Multi-Variate Analysis (MVA) using TMVA, the root built-in tool for MVA.

**Training Set:** 6 mil. of Signal / 1 stream of BB&qq Tested method : Rectangular Cuts, Fisher(Linear) Discriminant, Multi-dimensional Likelihood(PDERS)

PDERS shows a better ROC in general.



# 3. EVENT SELECTION

## *(4)* ECL variables, MVA

After overtraining test and FoM test for optimization region, we decided the cut at PDERS > 0.2. ( Details are shown in the BACK-UP)



**11**

## $(5) \cos\theta_e \leq 0.8$  cut



Angle btw. candidate gamma & the closest electron was tested to veto the events oriented by electron's emission.  $e \rightarrow e \gamma$  events have a peak around  $cos \theta_e = 1$ 

# 3. EVENT SELECTION

## *(6) Selection Efficiency*

### Region 1,  $1.3 < E\gamma^B < 1.8$  GeV



### Region 3,  $2.0 < E\gamma^B < 2.8$  GeV



### Region 2,  $1.8 < E\gamma^B < 2.0$  GeV





Table4. Signal Efficiency of selection criteria

# 4. EXPECTED YIELD



**Fig 12. Mbc Fit result in the bin 1.9-2.0 GeV**

# 4. EXPECTED YIELD



Table 5. Results of signal ftting and subtracted yield for each bin btw  $1.6 < Er^B < 2.6$  GeV



Fig 13. bkg-subtracted yield of each bin btw  $1.6 < E<sub>y</sub>B < 2.6 GeV$ 

# 4. EXPECTED YIELD



# 5. SUMMARY & PLANS

- 1. Selection Criteria is decided.
- 2. Expected Yields is calculated. (will be updated with including RareB & Ulnu) We could expect a better result than that of Babar.
- 3.  $\pi^0$ &n (dominant) background study should be done.

# 6. BACK UP

## ECL variables

OK, apparently, it seems hard to generally apply MVA(distances,  $cos\theta_e$ ) cut to my samples of interest. (But I want to use this MVA somehow!) So I checked the linear correlations btw ECL variables again, but this time, the whole MC samples were considered. (I only used samples in  $1.8 \sim 2.0$  GeV region in the last time)



**Wow, it seems hopeful!**

## ECL variables - MVA results



**Fisher(linear discriminant) & PDERS(Multi-dimensional Likelihood) are generally better than just Rectangular Cuts.** 

**And MVA can be a good choice for the optimization of rectangular cuts only.**

Fisher =  $-0.602^*$ (width)  $-0.104^*$ (e9/e25) - 0.534 $*$ (mass) +12.38

## ECL variables - MVA results

### FoM Comparison with Rectangular Cut  **of random test set (# of sig , # of bkg) = (1000, 1000)**





## ECL variables - MVA results

## Overtraining Check - Training Samples vs Testing Samples

**(TMVA guide) Proper training and validation requires three statistically independent data sets: one for the parameter optimi- sation, another one for the overtraining detection, and the last one for the performance validation. In TMVA, the last two samples have been merged to increase statistics. The (usually insignificant) bias introduced by this on the evaluation results does not affect the analysis as far as classification cut efficiencies or the regression resolution are** 



**TMVA overtraining check for classifier: PDERS**

TMVA overtraining check for classifier: Fisher



Kolmogorov-Smirnov test : A sort of comparison like  $\chi^2$  btw H<sub>0</sub>(trained PDF) & H<sub>1</sub>(testing PDF) **almost 1 : the same PDF(overtrained), almost 0 : no meaning of training, 0.1~0.9 : usual expectation**

**PDERS can be a better choice according to K-S test**





### **Region 11, 1.8 <**  $E\gamma^B$  **< 2.0 GeV**



### **Region 111**,  $2.0 < E$ <sub>*P*</sub><sup>B</sup> < 2.8 GeV





## **Bkg subtracted yield Procedure for each bin**



 $Subtracted yield = N<sub>whole, true</sub> - N<sub>bkg, MC</sub> error = sqrt( $\sigma$ <sub>whole, true</sub><sup>2</sup> +  $\sigma$ <sub>bkg, MC</sub><sup>2</sup>)$ **(5)**

$$
\text{True Significance} = \frac{N_{sig,true}}{\sqrt{\sigma_{sig,true}^2 + \sigma_{bkg,true}^2}}
$$

Due to the low bkg statistics in the ranges beyond 2.5 GeV,

Imployed True Significance = 
$$
\frac{N_{sig,true}}{\sigma_{whole}}
$$

# **Mbc Fit results - Signal**



# **Mbc Fit results - Signal**



# **Mbc Fit results - BB/qq**



# **Mbc Fit results - BB/qq**



# **Mbc Fit results - Sig + BB/qq**



# **Mbc Fit results - Sig + BB/qq**



# **Bkg subtracted yield**

### **Significances**





True Significance Points =  $N_{sub.}$  - ( $S_{true}$  x  $\sigma_{sub.}$ ) (Each point should be located under the zero line)

 $S<sub>true</sub>$ : True Significance =

*Nsig,true*  $\sigma_{whole}$ 

### Fitted Signal Yield



### Fitted Total Yield



### Fitted bkg MC yield



### Subtracted Yield

