

# $B \rightarrow X_{s\gamma}$ study with the hadronic tagging method at <u>Belle</u>

### Hanjin Kim

Institute of Physics and Applied Physics Yonsei Univ. <u>hanjin@yhep.yonsei.ac.kr</u>



## INDEX

### I. 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 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 $E_{\gamma}$ 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,  $A_I$ .

- 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.

Comparable study from Babar using "the recoil
$B.F(B \rightarrow X_s \gamma) = (3.97 \pm 0.97 \pm 0.64) \times 10^{-4}$
210 fb <sup>-1</sup> used



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

 $B.F = (3.55 \pm 0.32^{+0.30+0.11}_{-0.31-0.07}) \times 10^{-4}$  with  $140 f b^{-1}$ P. Koppenburg(2004), PRL 061803

Inclusive results  $B.F = (3.45 \pm 0.15_{stat} \pm 0.40_{syst}) \times 10^{-4}$  with  $605 f b^{-1}$ A. Limosani(2009), PRL 103, 241801

# 2. SIMULATION



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

(2)B $\rightarrow$ X<sub>s</sub> $\gamma$  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  $m_b = 4.574 \text{ GeV}, \ \mu_{\pi}^2 = 0.459 \text{ GeV}^2$ 

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

Generation Results		Gen & Gsim	Random selection	Eff.	
	Va ma a da a	Mixed	I2 mil.	6369933	53.08%
	Xs modes	Charged	I2 mil.	6370224	53.09%
	K*	Mixed	868632		otroomo of
K* modes		Charged	868674	signal	MC sample

Table 1. The number of signal MC events

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

SYW2014, Jan. 14th, 2014

#### **Pre-selection Best-Btag Selection** NBRank = 1(if a multiple choices, the one with higher NBout is chosen.) 7000 $5.24 < M_{bc} < 5.29 \text{ GeV/c}^2$ 6000 $|\Delta E| < 0.06 \text{ GeV}$ 5000 NBout > 0.14000 **Candidate Selection** 3000 2000 Most energetic ( in B rest frame ) gamma among sig-side gammas. 1000 $E_{\gamma B_{candi}} > 1.3 \text{ GeV}$ Table 2. Pre-selection criteria

	Simulated	After fullrecon & pre-selection	Efficiency
Signal	14477463	25287	0.17%
Generic	5 streams	414812	
Continuum	5 streams	266364	

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

5 Streams of Generic & Continuum MC50 streams of RareB, 20 streams of Ulnu MC employed.



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

# 3. EVENT SELECTION

### Selection Criteria

### $\pi^{0}$ & $\eta$ and addbg(911, off-timing) particles are dominant.



### Selection Criteria



\* In each step, tag-correction is applied.

### Analysis Regions

Region I :  $1.3 < E_{\gamma}^{B} < 1.8 \text{ GeV}$ 

Region 11 :  $1.8 < E_{\gamma}^{B} < 2.0 \text{ GeV}$  (Optimization regions)

Region III : 2.0 GeV <  $E_{\gamma B}$  < 2.8 GeV

# 3. EVENT SELECTION

## (I) 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) $P(\pi^0) < 0.3 \& P(\eta) < 0.3 Cut$

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  $P(\pi^0) < 0.3 \& P(\eta) < 0.3 Cut$ 

## 3. EVENT SELECTION (3) $|\cos \theta_{thrust}| < 0.8 Cut$

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

Thrust Axis = 
$$\frac{\sum_{i=1}^{n} \frac{p_i}{p_i}}{\frac{p_i}{p_i}}$$

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

$$\cos\theta_{\rm thrust} = \vec{A}_{\rm sig} \cdot \vec{A}_{\rm tag}$$

### (4) ECL variables, MVA



Hadronic shower (e.g. anti-n, anti-p) shows widely spread shower shape compared to the real gamma shower inside calorimeters. 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 5.  $|\cos\theta_{thrust}|$  distribution of sig/BB/qq

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



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) $\cos\theta_{\rm e} < 0.8$ 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, $1.3 < E\gamma^B < 1.8 \text{ GeV}$

Region I	Sig	BB	рр	FoM
911 veto	99.48%	93.02%	95.74%	0.21
Ρ(π0)&Ρ(η)	87.10%	48.59%	50.32%	0.26
lcos0thrustl	80.27%	79.01%	22.79%	0.29
PDERS	86.56%	63.59%	64.22%	0.31
cosθe	97.46%	93.01%	96.24%	0.31
tot-efficiency	59%	21.12%	6.79%	
tot-cutoff	41%	78.88%	93.21%	

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

Region 111	Sig	BB	рр	FoM
911 veto	99.48%	38.81%	62.94%	8.71
Ρ(π0)&Ρ(η)	97.78%	69.4%	68.35%	10.00
lcos0thrustl	79.27%	80.03%	18.43%	10.61
PDERS	99.39%	77.25%	80.67%	11.43
cosθe	99.26%	97.76%	97.78%	11.45
tot-efficiency	76.07%	16.28%	6.25%	
tot-cutoff	23.93%	83.72%	93.75%	

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

Region 11	Sig	BB	рр	FoM
911 veto	99.48%	83.22%	89.16%	1.01
Ρ(π0)&Ρ(η)	95.17%	54.79%	59.68%	1.27
lcos <del>0</del> thrustl	77.35%	79.43%	20.45%	1.32
PDERS	96.87%	79.9%	79.13%	1.43
cosθe	99.17%	96.22%	97.38%	1.44
tot-efficiency	70.64%	27.84%	8.39%	
tot-cutoff	29.65%	72.16%	91.61%	



Table4. Signal Efficiency of selection criteria

# 4. EXPECTED YIELD

SYW2014, Jan. 14th, 2014



Fig 12. M<sub>bc</sub> Fit result in the bin 1.9-2.0 GeV

# 4. EXPECTED YIELD

Bin#	range	Subt. Yield	error	Signifi- cance	True Significance
0	1.6-1.7	8.70	45.12	0.193	0.212
1	1.7-1.8	15.36	38.77	0.396	0.410
2	1.8-1.9	23.82	32.85	0.725	0.774
3	1.9-2.0	37.84	27.78	1.362	1.447
4	2.0-2.1	51.73	22.56	2.293	2.427
5	2.1-2.2	69.00	16.70	4.131	4.321
6	2.2-2.3	82.68	12.35	6.693	6.917
7	2.3-2.4	81.37	10.31	7.889	7.981
8	2.4-2.5	60.93	8.87	7.276	7.349
9	2.5-2.6	40.18	6.59	6.096	6.104

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



Fig 13. bkg-subtracted yield of each bin btw  $1.6 < E_{r^B} < 2.6 \text{ GeV}$ 

# 4. EXPECTED YIELD



# 5. SUMMARY & PLANS

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

# 6. BACK UP

### **ECL** variables

OK, apparently, it seems hard to generally apply MVA(distances, cosθ<sub>e</sub>) 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 ~ 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)



		PDERS	Rect. Cuts	Improvement(PDERS to Cuts)
FoM <sub>MAX</sub>		25.5169	25.0428	I.89%
at	Sig eff.	0.8641	0.8814	-1.96%
FoM <sub>MAX</sub>	Bkg Cut-off	0.7173	0.6426	11.62%

### 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. <u>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</u>





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 1	$, 1.3 < E\gamma^{B}$	<sup>3</sup> < 1.8 GeV
----------	-----------------------	------------------------

Region I	Sig	BB	рр	rare	ulnu	FoM
911 veto	99.48%	93.02%	95.74%	97.99%	95.65%	0.21
Ρ(π0)&Ρ(η)	87.1%	48.59%	50.32%	41.58%	47.4%	0.26
lcos0thrustl	80.27%	79.01%	22.79%	78.28%	80.39%	0.29
PDERS	86.56%	63.59%	64.22%	69.38%	75.44%	0.31
cosθe	97.46%	93.01%	96.24%	94.69%	78.97%	0.31
tot-efficiency	59%	21.12%	6.79%	20.95%	21.71%	
tot-cutoff	41%	78.88%	93.21%	79.05%	78.29%	

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

Region 11	Sig	BB	рр	rare	ulnu	FoM
911 veto	99.48%	83.22%	89.16%	97.28%	92.36%	1.01
Ρ(π0)&Ρ(η)	95.17%	54.79%	59.68%	46.30%	55.10%	1.27
lcos0thrustl	77.35%	79.43%	20.45%	79.29%	81.22%	1.32
PDERS	96.87%	79.9%	79.13%	86.59%	95.65%	1.43
cosθe	99.17%	96.22%	97.38%	96.33%	85.66%	1.44
tot-efficiency	70.64%	27.84%	8.39%	29.79%	33.87%	
tot-cutoff	29.65%	72.16%	91.61%	70.21%	66.13%	

### Region 111 , 2.0 < $E\gamma^{B}$ < 2.8 GeV

Region 111	Sig	BB	рр	rare	ulnu	FoM
911 veto	99.48%	38.81%	62.94%	93.92%	79.85%	8.71
Ρ(π0)&Ρ(η)	97.78%	69.4%	68.35%	55.87%	61.45%	10.00
lcos0thrustl	79.27%	80.03%	18.43%	79.64%	79.73%	10.61
PDERS	99.39%	77.25%	80.67%	91.93%	98.33%	11.43
cosθe	99.26%	97.76%	97.78%	97.81%	93.55%	11.45
tot-efficiency	76.07%	16.28%	6.25%	37.58%	35.99%	
tot-cutoff	23.93%	83.72%	93.75%	62.42%	64.01%	



### Bkg subtracted yield Procedure for each bin



(5) Subtracted yield =  $N_{whole,true} - N_{bkg,MC}$  error =  $sqrt(\sigma_{whole,true}^2 + \sigma_{bkg,MC}^2)$ 

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,

$$\begin{array}{l} \textbf{I employed True Significance} = & \frac{N_{sig,true}}{\sigma_{whole}} \end{array}$$

## M<sub>bc</sub> Fit results - Signal



## M<sub>bc</sub> Fit results - Signal



## M<sub>bc</sub> Fit results - BB/qq



## M<sub>bc</sub> Fit results - BB/qq



# M<sub>bc</sub> Fit results - Sig + BB/qq



# M<sub>bc</sub> Fit results - Sig + BB/qq



## Bkg subtracted yield

### Significances

Bin#	range	TRUE		Subtract.
0	1.3-1.4	0.046	<	0.048
1	1.4-1.5	0.053	>	0.047
2	1.5-1.6	0.122	>	0.112
3	1.6-1.7	0.212	>	0.193
4	1.7-1.8	0.410	>	0.396
5	1.8-1.9	0.774	>	0.725
6	1.9-2.0	1.447	>	1.362
7	2.0-2.1	2.427	>	2.293
8	2.1-2.2	4.321	>	4.131
9	2.2-2.3	6.917	>	6.693
10	2.3-2.4	7.981	>	7.889
11	2.4-2.5	7.349	>	7.276
12	2.5-2.6	6.104	>	6.096
13	2.6-2.7	1.784	<	1.851



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

 $S_{true}$ : True Significance =

 $\frac{N_{sig,true}}{\sigma_{whole}}$ 

### Fitted Signal Yield

E	1.3-	1.4-	1.5-	1.6-	1.7-	1.8-	1.9-	2.0-	2.1-	2.2-	2.3-	2.4-	2.5-	2.6-
Yield	3.01	3.00	5.99	9.01	14.98	23.99	38.02	51.93	69.01	83.00	80.99	61.00	40.03	3.99
err	1.73	1.72	2.44	3.00	3.86	4.89	6.10	7.18	8.13	8.77	8.70	7.78	6.26	1.98

### Fitted Total Yield

E	1.3-	1.4-	1.5-	1.6-	1.7-	1.8-	1.9-	2.0-	2.1-	2.2-	2.3-	2.4-	2.5-	2.6-
Yield	4,302.0	3,254.0	2,406.0	1,808.0	1,336.0	961.0	690.0	458.0	255.0	144.0	103.0	68.9	43.0	5.0
err	65.6	57.0	49.0	42.5	36.5	31.0	26.3	21.4	16.0	12.0	10.1	8.3	6.6	2.2

### Fitted bkg MC yield

E	1.3-	1.4-	1.5-	1.6-	1.7-	1.8-	1.9-	2.0-	2.1-	2.2-	2.3-	2.4-	2.5-	2.6-
Yield	4298.7	3251.2	2400.2	1799.3	1320.6	937.2	652.2	406.3	186.0	61.3	21.6	8.0	2.8	0.8
err	23.4	20.4	17.5	15.1	12.9	10.9	9.0	7.1	4.9	2.9	1.8	1.1	0.7	0.3

### Subtracted Yield

Е	1.3-	1.4-	1.5-	1.6-	1.7-	1.8-	1.9-	2.0-	2.1-	2.2-	2.3-	2.4-	2.5-	2.6-
Yield	3.33	2.85	5.81	8.70	15.36	23.82	37.84	51.73	69.00	82.68	81.37	60.93	40.18	4.19
err	69.65	60.58	52.06	45.12	38.77	32.85	27.78	22.56	16.70	12.35	10.31	8.37	6.59	2.26