Rb measurement at CEPC MC Level

Bo Li

Outline

- Introduction
  - Rb
  - Double tagging method
  - B-tagging performance

- Method
  - I/O test
  - Effect of $\boldsymbol{\varepsilon_c}$ **,** $\boldsymbol{\varepsilon_{uds}}$ **and** $\boldsymbol{C_b}$

Introduction

- R_b : the relative decay width of Z into b quarks

$$\frac{\Gamma(\mathrm{Z} \to \mathrm{b\bar{b}})}{\Gamma(\mathrm{Z} \to \mathrm{had})}$$

- Based on the CEPC MC samples: Zbb, Zcc, Zuds
  - **Produced from FSClasser** **~110k events**

- By using the Double tagging method

Introduction

The two jets are divided into two hemispheres

Double tagging method :

- The ratio of one jet tagged as b jet
- The ratio of both jets tagged as b jets

$$\frac{N_t}{2N_{had}} = R_b\ \varepsilon_b + R_c\varepsilon_c + (1 - R_b - R_c)\varepsilon_{uds}$$

$$\frac{N_{tt}}{N_{had}} = C_b R_b \varepsilon_b^2 + C_c R_c \varepsilon_c^2 + C_{uds}(1 - R_b - R_c)\varepsilon_{uds}^2$$

$$C_b = \frac{\varepsilon_{2jet-tagged}}{(\varepsilon_{1jet-tagged})^2}$$

- The R_{uds} , R_c , ε_{uds} , ε_c , C_b , C_c , C_{uds} are gotten from MC samples

- The N_t , N_{tt} , N_{had} are gotten from DATA samples

Rb method

Double Tagging Method

Get From Mixed MC Sample

$$\frac{N_t}{2N_{had}} = R_b\varepsilon_b + R_c\varepsilon_c + (1 - R_b - R_c)\varepsilon_{uds}$$

$$\frac{N_{tt}}{N_{had}} = C_b R_b \varepsilon_b^2 + C_c R_c \varepsilon_c^2 + C_{uds}(1 - R_b - R_c)\varepsilon_{uds}^2$$

R_c, ε_c, ε_{uds}
C_b, C_c, C_{uds}
Get from MC

$$C_b = \frac{\varepsilon_{2jet-tagged}}{(\varepsilon_{1jet-tagged})^2}$$

- LEP measurement 0.21594 ±0.00066
  - Syst error : ~0.2%
    - Major systematics is hemisphere tag correlations
- CEPC
  - Expected Syst error (0.02%)
    - hemisphere tag correlations depends on b tagging efficiency
    - with a high b-tagging efficiency above 80% and rejection of charm and light jet above 90%

Rb method

Double Tagging Method

Get From Mixed MC Sample

$$\frac{N_t}{2N_{had}} = R_b\varepsilon_b + R_c\varepsilon_c + (1 - R_b - R_c)\varepsilon_{uds}$$

$$\frac{N_{tt}}{N_{had}} = C_b R_b \varepsilon_b^2 + C_c R_c \varepsilon_c^2 + C_{uds}(1 - R_b - R_c)\varepsilon_{uds}^2$$

R_c, ε_c, ε_{uds}
C_b, C_c, C_{uds}
Get from MC

$$C_b = \frac{\varepsilon_{2jet-tagged}}{(\varepsilon_{1jet-tagged})^2}$$

$$\frac{N_t}{N_{had}} = R_b\varepsilon^I{}_b + R_c\varepsilon^I{}_c + (1 - R_b - R_c)\varepsilon^I{}_{uds}$$ jet1 Btag work point I

$$\frac{N_t}{N_{had}} = R_b\varepsilon^J{}_b + R_c\varepsilon^J{}_c + (1 - R_b - R_c)\varepsilon^J{}_{uds}$$ jet2 Btag work point J

$$\frac{N_{tt}}{N_{had}} = C_b R_b \varepsilon^I{}_b \varepsilon^J{}_b + C_c R_c \varepsilon^I{}_c \varepsilon^J{}_c + C_{uds}(1 - R_b - R_c)\varepsilon^I{}_{uds}\varepsilon^J{}_{uds}$$

$$C_b = \frac{\varepsilon_{2jet-tagged}}{\varepsilon^I{}_b \varepsilon^J{}_b}$$

B-tagging method

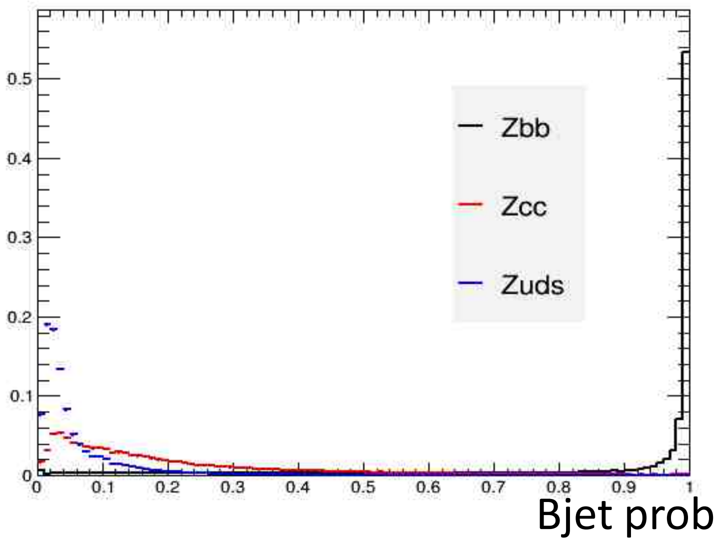
- The default algorithm is based on the LCFIPlus: combines more than 60 discriminant variables to calculate the b jet probability by BDT method

- B-tagging method for hemisphere 1:
  - Calculating the b jet probability based on the number of vertex

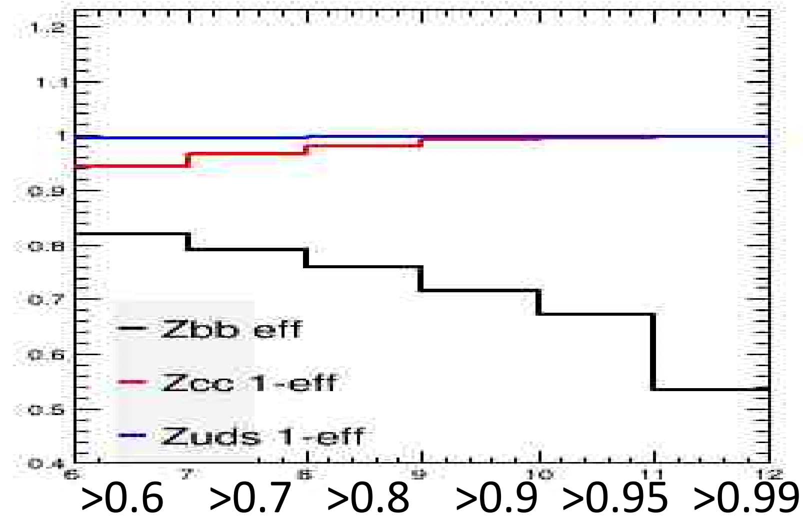
- B-tagging method for hemisphere 2:
  - Calculating the b jet probability based on the number of lepton(we set >=1 lepton)

B-tagging method

- B-tagging method for hemisphere 1:
  - Calculating the b jet probability based on the number of vertex



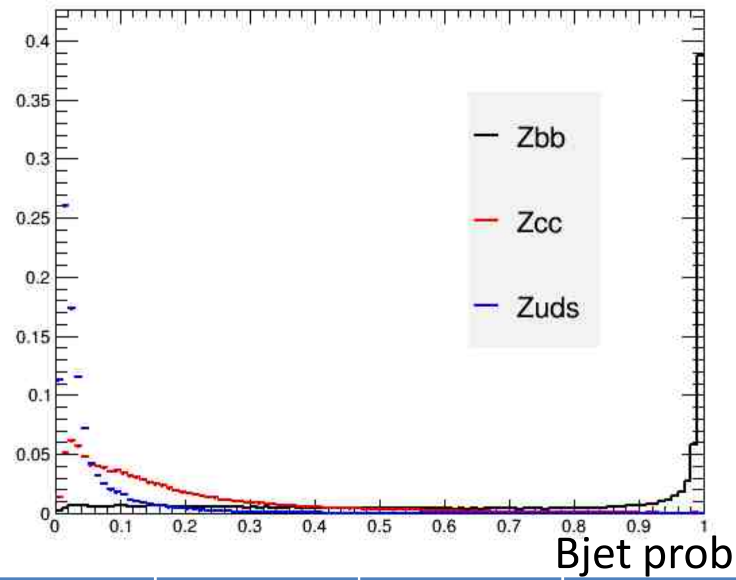
b jet efficiency and c, light jet rejection



	Prob>0.6	Prob>0.7	Prob>0.8	Prob>0.9	>0.95	>0.99
Zbb_eff	0.8202	0.7915	0.7593	0.7144	0.6722	0.5336
Zcc_Rej:	0.9437	0.9666	0.9817	0.9920	0.9957	0.9990
Zll_Rej	0.9941	0.9961	0.9973	0.9985	0.9990	0.9995

B-tagging method

- B-tagging method for hemisphere 2:
  - –Calculating the b jet probability based on the number of lepton



	Prob>0.6	Prob>0.7	Prob>0.8	Prob>0.9	Prob>0.95	Prob>0.99
Zcc_Rej:	0.9921	0.9955	0.9978	0.9994	0.9997	0.9999
Zll_Rej	0.9991	0.9995	0.9998	0.9999	1.000	1.000

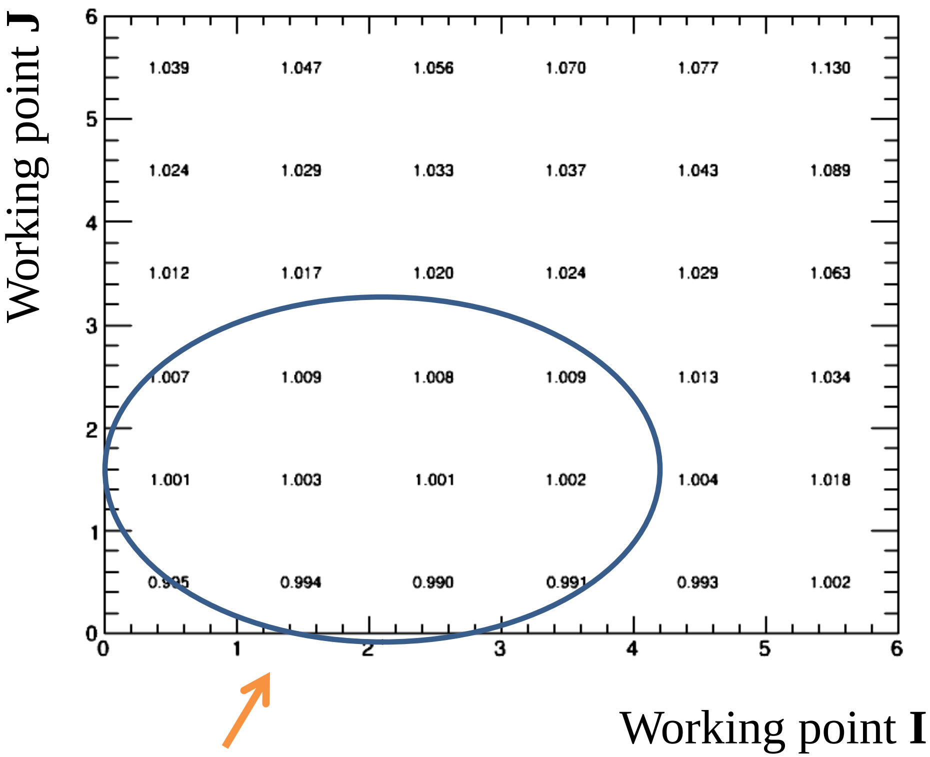
Correlation Factor

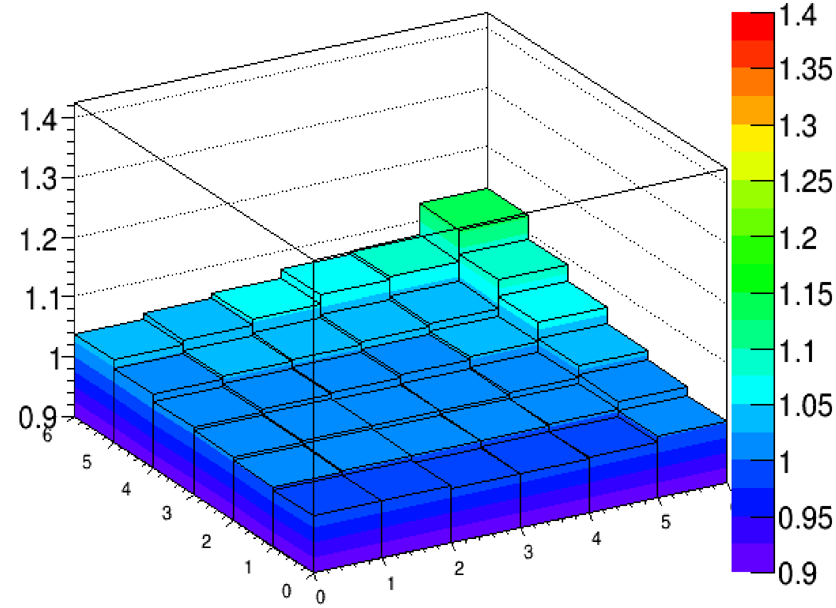
zbb

$$C_b = \frac{\varepsilon_{2jet-tagged}}{\varepsilon^{I}{}_{b}\varepsilon^{J}{}_{b}}$$

hemisphere 1 I : Prob>0.6 >0.7 >0.8 >0.9 >0.95 >0.99

hemisphere 2 J : Prob>0.6 >0.7 >0.8 >0.9 >0.95 >0.99





Correlation Factors ~ 1

Rb method

I/O test

$$\frac{N_t}{N_{had}} = R_b\varepsilon^I{}_b + R_c\varepsilon^I{}_c + (1 - R_b - R_c)\varepsilon^I{}_{uds}$$ jet1 Btag work point I

$$\frac{N_t}{N_{had}} = R_b\varepsilon^J{}_b + R_c\varepsilon^J{}_c + (1 - R_b - R_c)\varepsilon^J{}_{uds}$$ jet2 Btag work point J

$$\frac{N_{tt}}{N_{had}} = C_b R_b\varepsilon^I{}_b\varepsilon^J{}_b + C_c R_c\varepsilon^I{}_c\varepsilon^J{}_c + C_{uds}(1 - R_b - R_c)\varepsilon^I{}_{uds}\varepsilon^J{}_{uds}$$

$$C_b = \frac{\varepsilon_{2jet-tagged}}{\varepsilon^I{}_b\varepsilon^J{}_b}$$

Following this procedure, we can meausere the R_b, ε_b

The Z hadronic pseudo'**DATA**' is mixed by MC samples: Zbb **sample**, Zcc **sample**, Zll sample

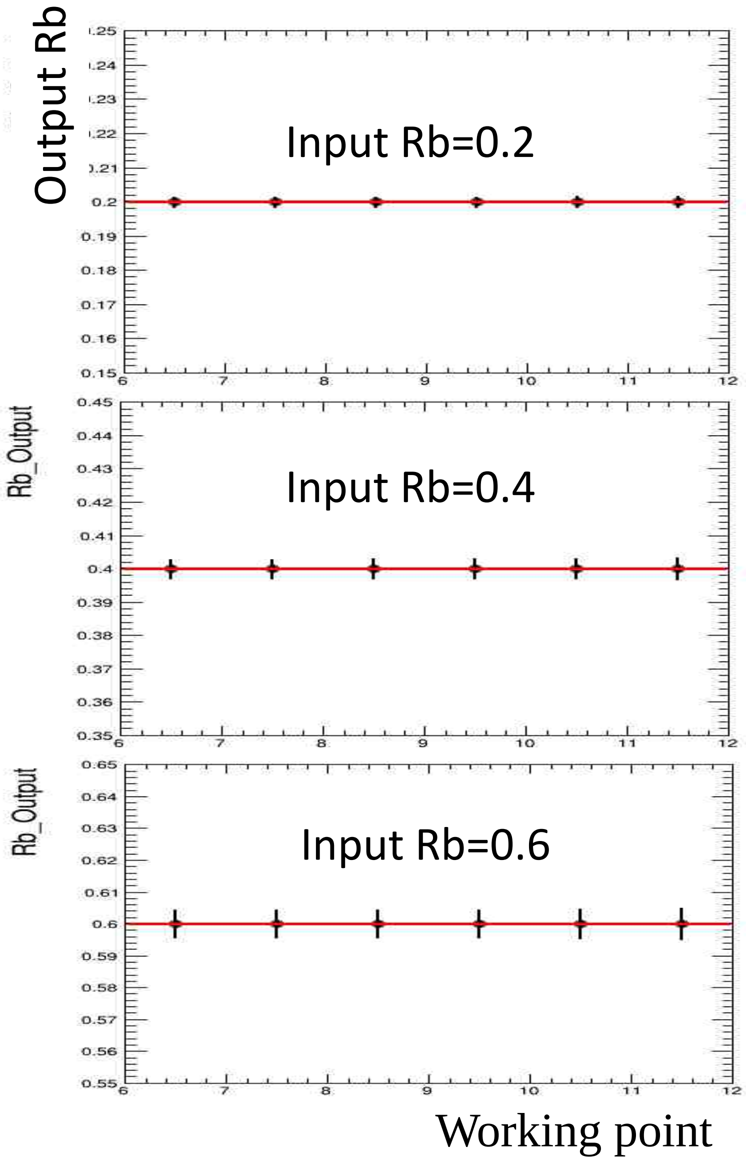
We set Rb=0.2, Rb=0.4, Rb=0.6 as the Input Rb to mix the 'DATA'

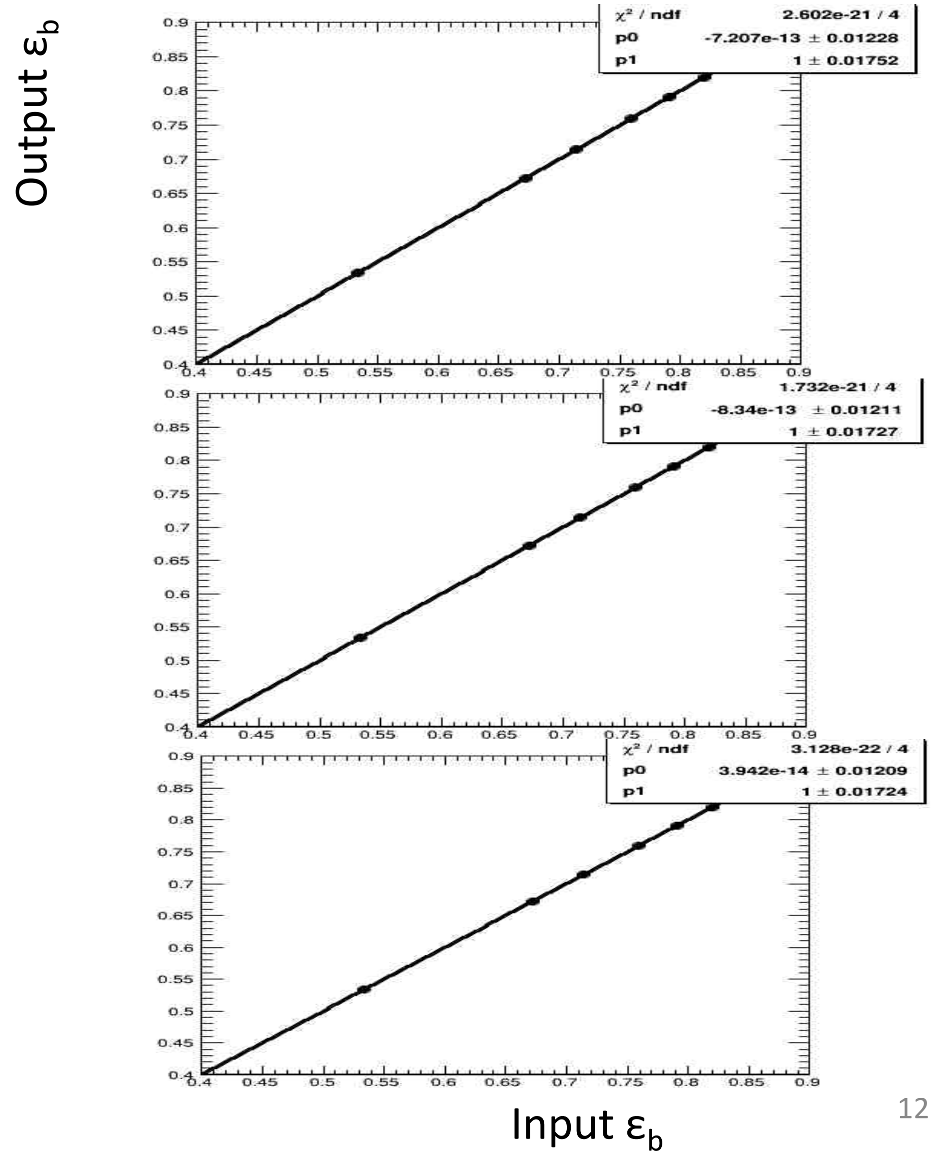
Method

I/O test

Input Rb=0.2, 0.4, 0.6.

Btag work points: **Prob>0.6, >0.70, >0.80, >0.90, >0.95, >0.99 >0.95,>0.99**





Method

Effect of ε_c , ε_{uds} and C_b

Input theory Rb=0.2158, Btag Prob work points: **Prob>0.6, >0.70, >0.80, >0.90, >0.95,>0.99**

I/O test with ε_c ±10%, ε_{uds} ±10%, C_b ±10%

	(Measured Rb-0.2158)/0.2158					
	Prob>0.6	**Prob>0.70**	**Prob>0.80**	**Prob>0.90**	**Prob>0.95**	**Prob>0.99**
ε_c ±10%	0.55%	0.34%	0.19%	0.09%	0.05%	0.01%
ε_{uds} ±10%	0.21%	0.14%	0.10%	0.06%	0.04%	0.02%
C_b ±10%	10.12%	10.09%	10.08%	10.06%	10.06%	10.05%

Plan

- Both the B-tagging efficiency for Z->bb and rejection for Z->cc, Z->uds are good for method

- The double tagging method procedure works well as shown in I/O test

- The b tagging correlation factors ~ 1

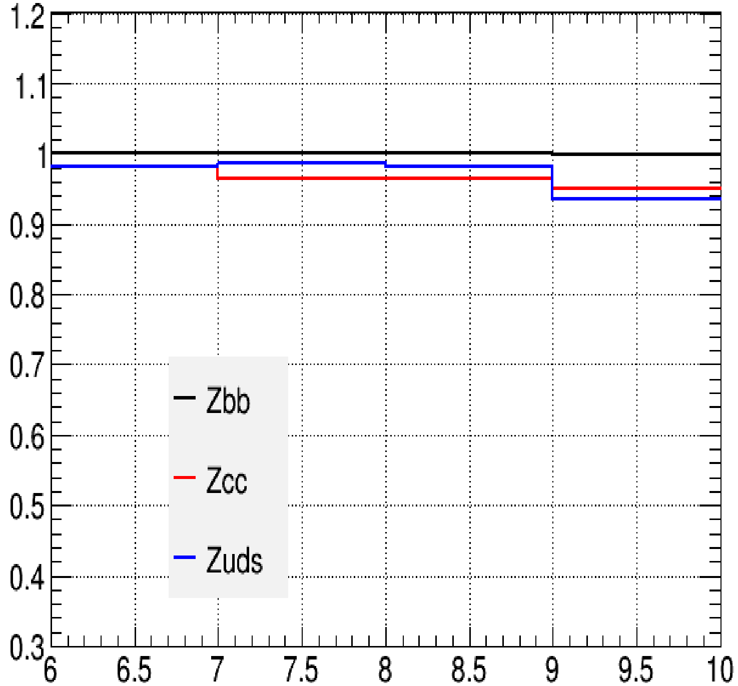
- Next step :
  - Increase the statistic to reduce the effect from $\boldsymbol{\varepsilon_c}$ **,** $\boldsymbol{\varepsilon_{uds}}$ **and** $\mathbf{C_b}$
  - Study on the systematic errors such as the gluon splitting, charm physics modeling

Check

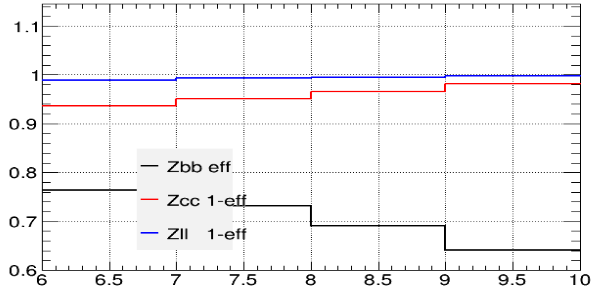
the measured Rb and effb in DATA are different from the Input Truth Rb and effb at Prob>0.9

The R_c, ε_c, C_b, C_c, C_{uds} is got by MC samples: Zbb sample2, Zcc sample2, Zllsample2
So if DATA **sample1≠** sample2, which means the MC R_c, ε_c, C_b, C_c, C_{uds} is different from the 'DATA'

The difference as a Ratio: Eff in 'DATA'/ Eff in MC



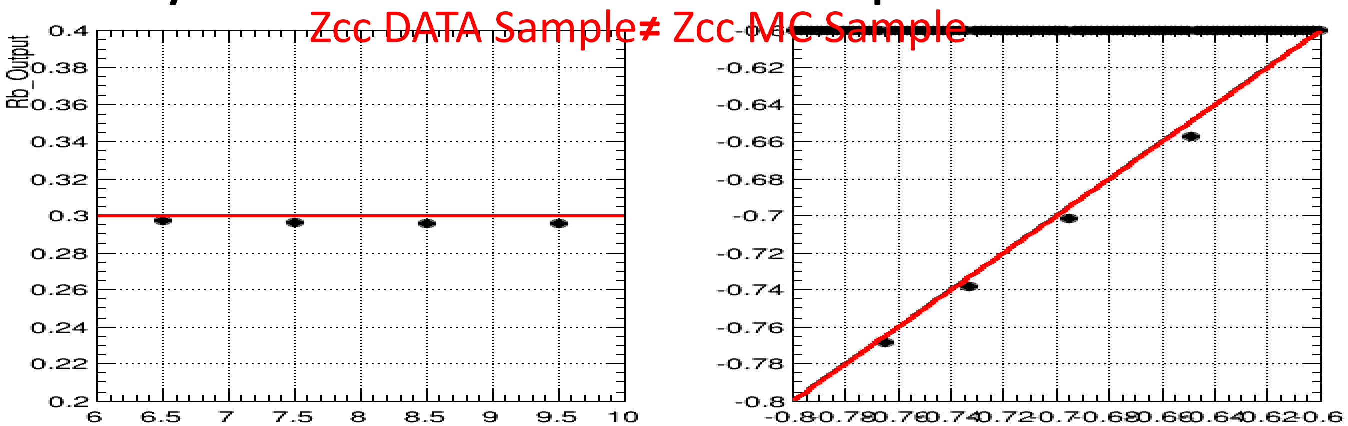
1. ε_b difference between DATA and MC are very small
2. ε_c and ε_{uds} differences are big at Prob>0.9 :
   - which may come from the very low statistics after B-tagging
   - which will lead to the difference in the IO test
3. ε_{uds} effect is very small, as **The Zll rejction at four work point are ~100%**

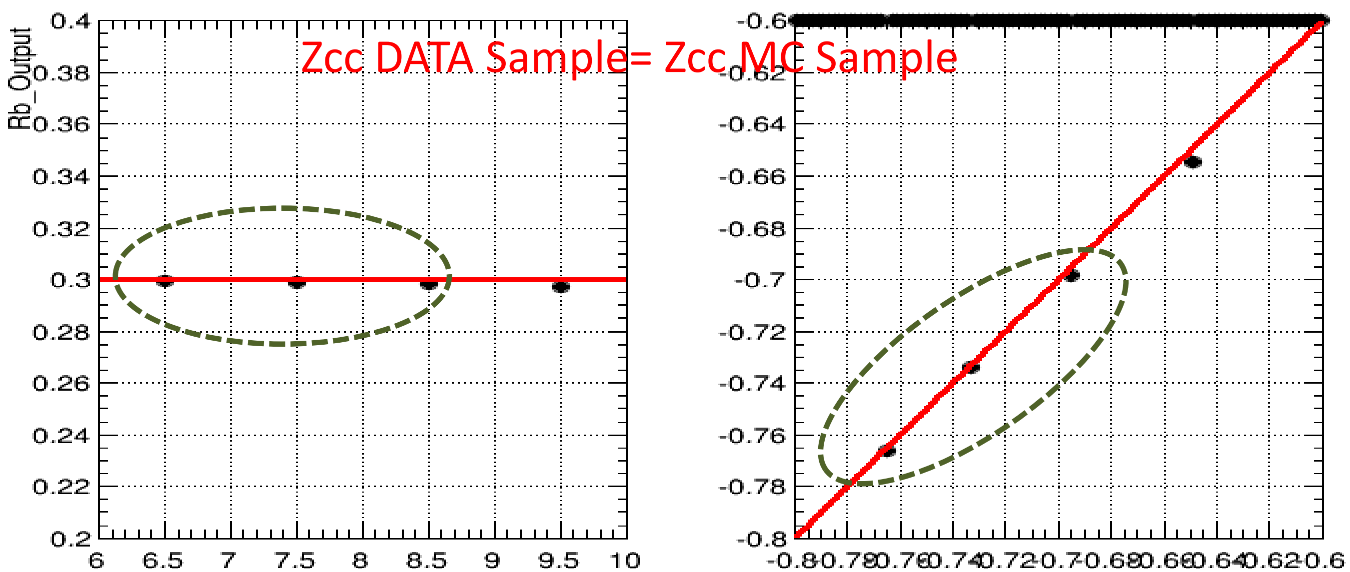


Check

Input Rb=0.3

We redo the IO test by 'DATA' and MC with same Zcc sample

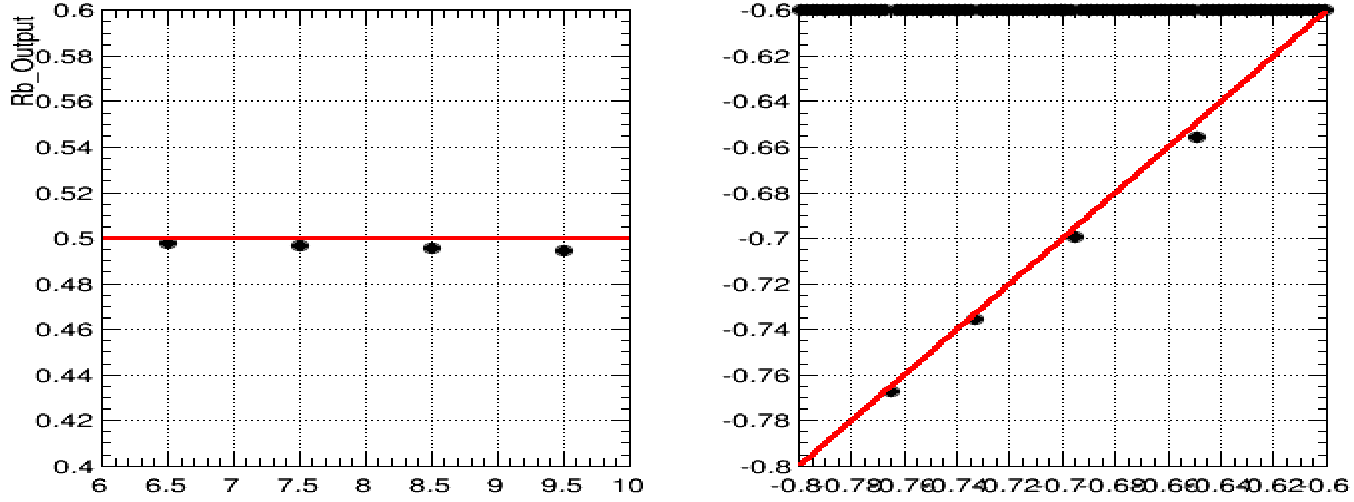


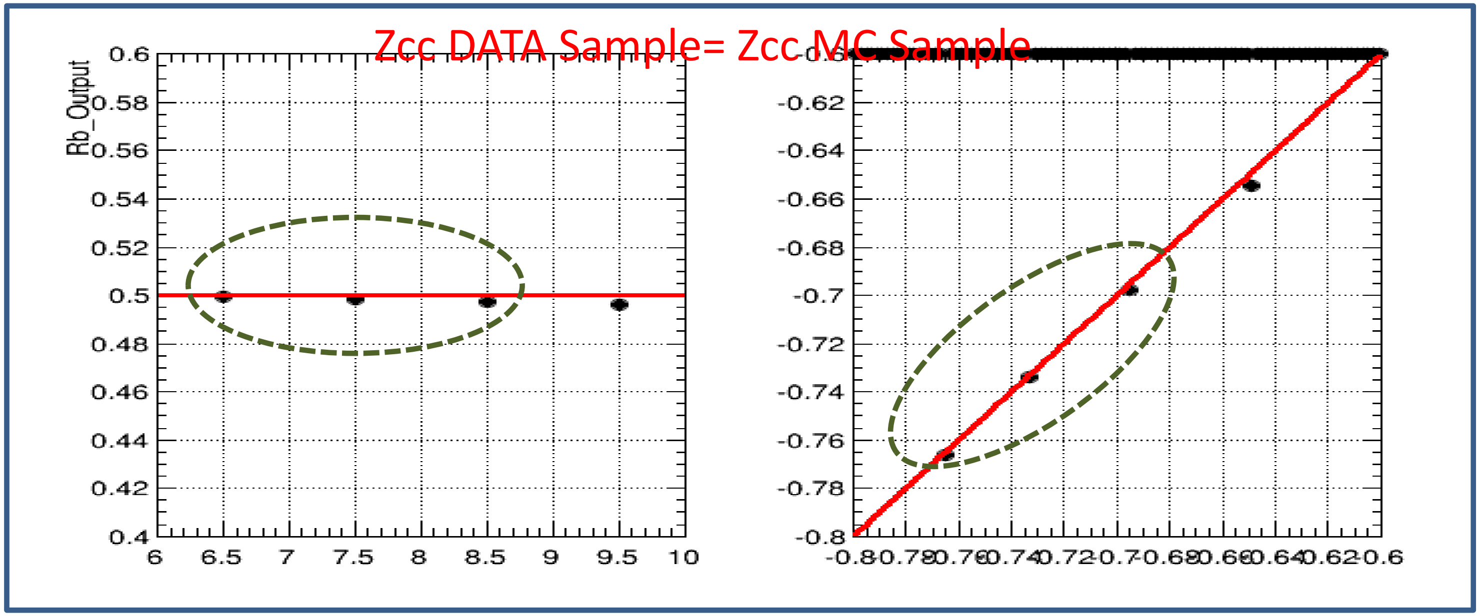


We can see the differences of measured Rb and effb between DATA and MC are smaller

Check

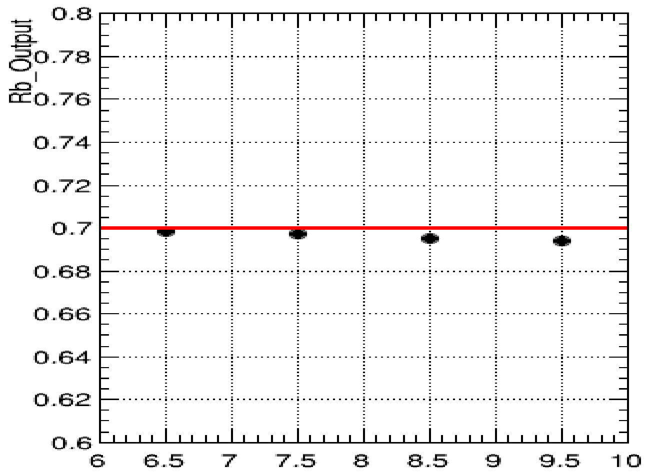
Input Rb=0.5

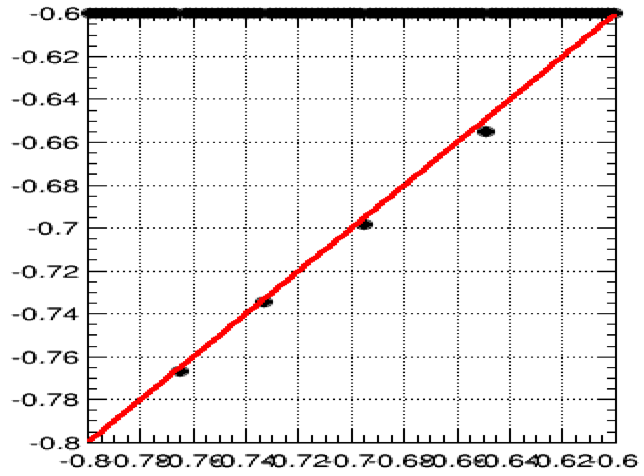




Check

Input Rb=0.7





Zcc DATA Sample= Zcc MC Sample