Discussion follow-up

Danning Liu

Check Points

- Configuration setups for two Z(ll)gamma analysis
  - Both HGamGamStar Framework, h024 and h026 tag — Different CP recommendations
- Overflow check
  - Two different methods — add one extra overflow bin works better and close to optimised binning
- Add Full Background contributions
  - Add Z+jets contribution to the fit
  - Add bkg systematics and signal systematics

- Next steps

Production tag

- In this table, different configuration will be listed between two analysis : inclusive $Z(ll)\gamma$ and inclusive $Z(ll)\gamma + jets$

	Z(ll)+gamma	**Z(ll)+gamma+jets**
Production tag	h024	h026

- MxAOD files for both measurements are produced with HGamGamStar Framework, with Release 21
- Present $Z(ll)\gamma + jets$ use New CP recommendations
- Difference around 5%, more details need to check original code

Overflow Check

- In this slide, we will introduce two different overflow binnings and errors
    - Bin A : p_T^{γ} [200, 241, 500] GeV, add overflow events to the last bin
    - Bin B : p_T^{γ} [200, 241, 500, 2000] GeV, add overflow events to the last bin

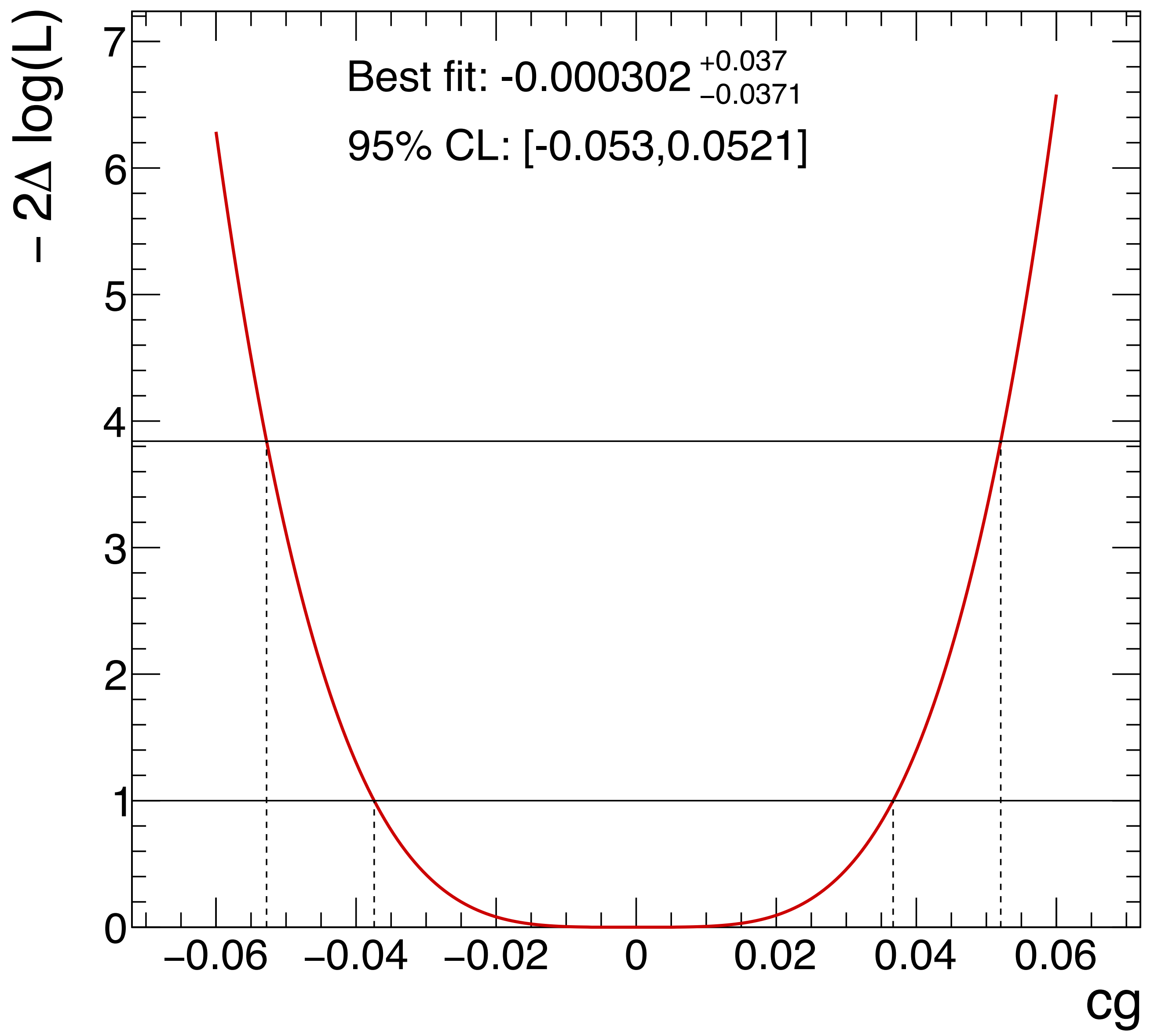
    - errors
    - final_error = TMath::Sqrt(TMath::Power(overflow_error, 2) + TMath::Power(lastbin_error, 2))
    - overflow_error : the error for the overflow bin of the distribution
    - lastbin_error : the error for the last bin of the distribution

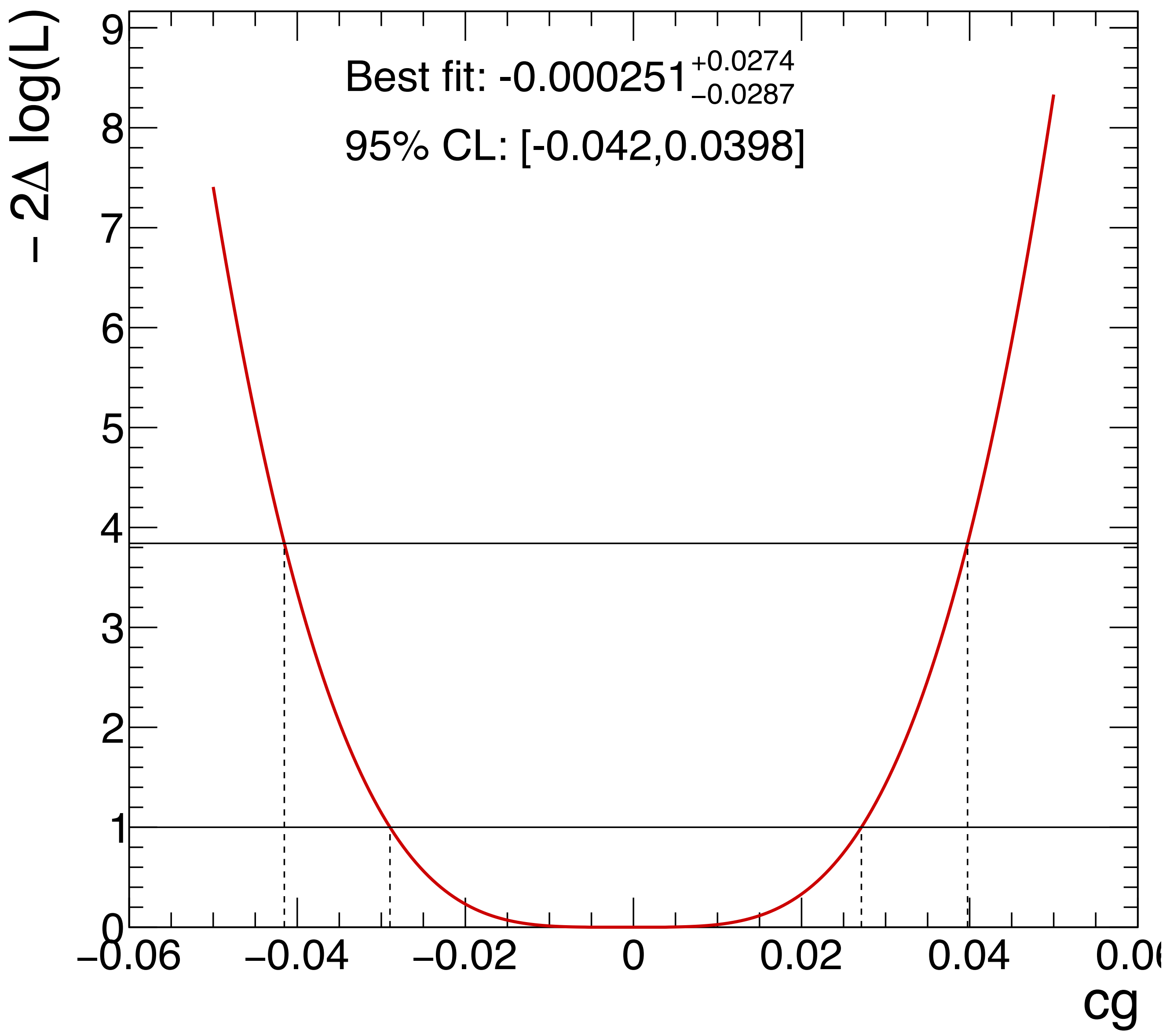
- Also added overflow events to the last bin for background samples

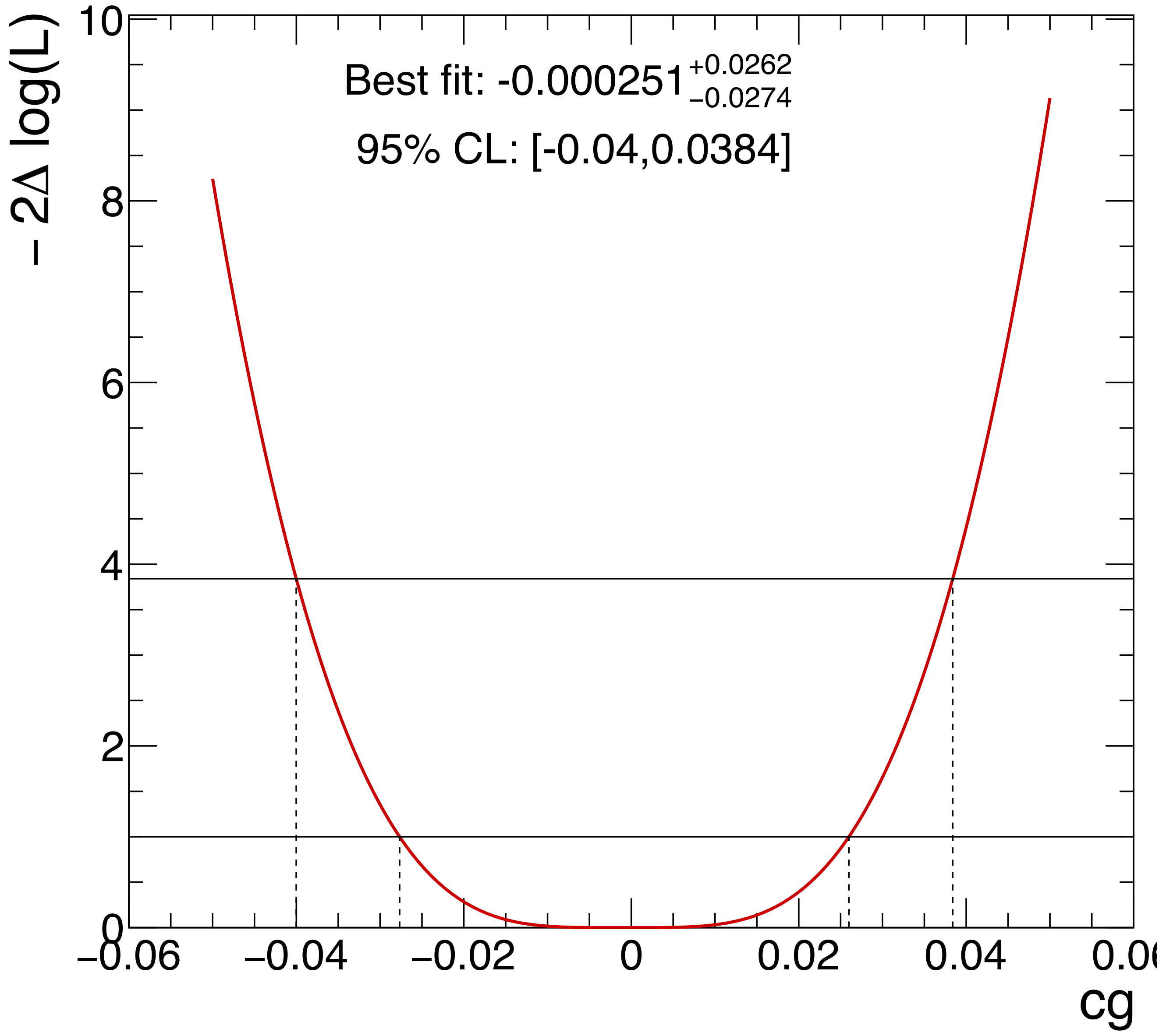
Overflow Check

- Constrained limits result

Binning	**Bin A (500GeV)**	**Bin B (2000GeV)**	**Optimised Bin**
Limits (stat - only)	[-0.053, 0.052]	[-0.042, 0.040]	[-0.040, 0.038]





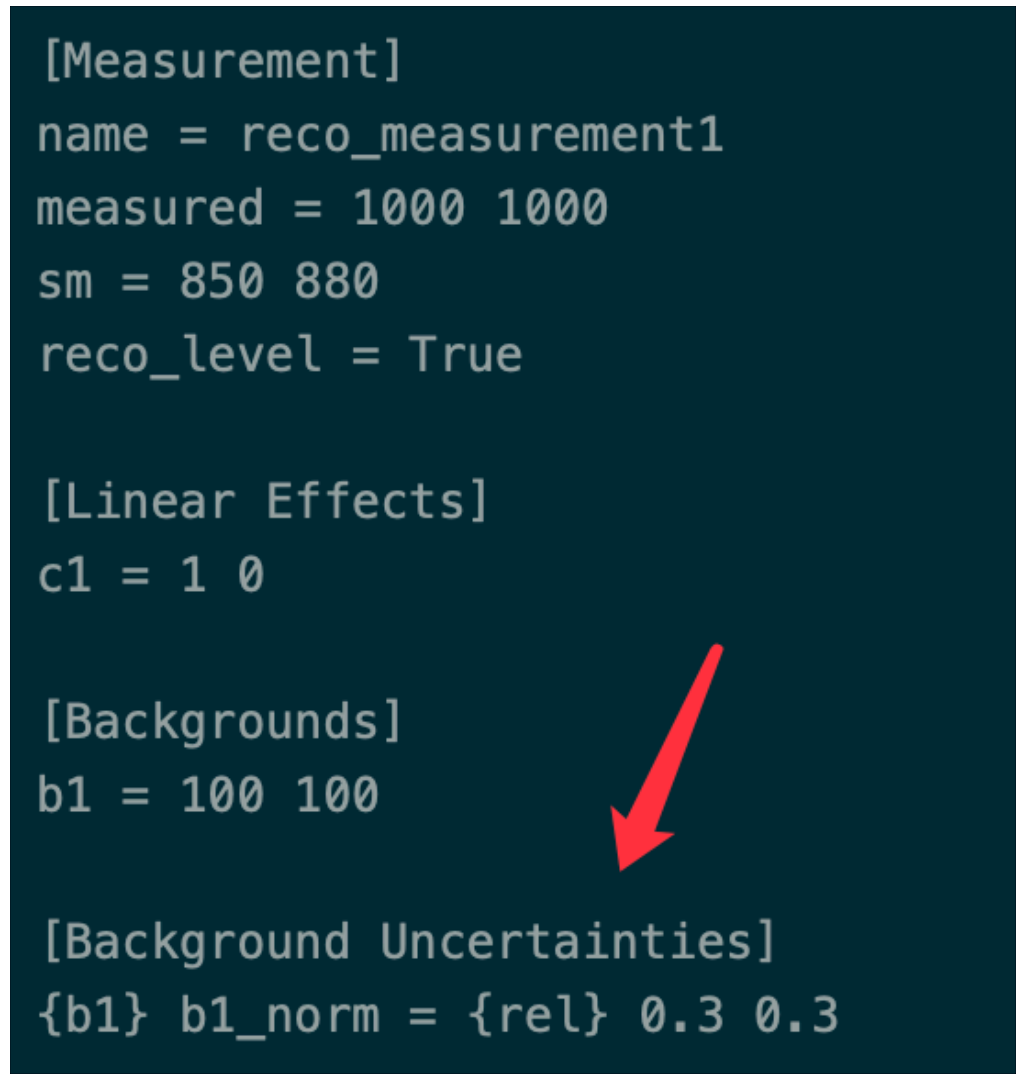


Recap : Optimised Bin (p_T^{γ} : [200, 230, 260, 300, 350, 430, 550, 2000] GeV

How to add Background Systematics

- Systematics account for luminosity uncertainties
  - 15% for $t\bar{t}\gamma$ simulation
  - 30% for all the other MC simulated sources ($WZ\gamma, WW\gamma, tW\gamma \ldots$)

  - Quadratic sum
  - $Unc_{tot} = \sqrt{Unc^2_{Zjets} + Unc^2_{pileup} + Unc^2_{t\bar{t}\gamma} + Unc^2_{tW\gamma} + Unc^2_{diboson} + Unc^2_{triboson}} \approx 0.807775$
  - $Unc_{Zjets} = Unc_{pileup} = 0$, data-driven method



```
[Measurement]
name = reco_measurement1
measured = 1000 1000
sm = 850 880
reco_level = True

[Linear Effects]
c1 = 1 0

[Backgrounds]
b1 = 100 100

[Background Uncertainties]
{b1} b1_norm = {rel} 0.3 0.3
```

How to add SM and BSM uncertainties

- How to calculate systematics ?
  - Experimental
    - Finite resolution of objects, modelling of the reconstructions (electrons, muons, jets, photons and so on)
    - $\sigma_{exp} = \frac{\sigma_{down} - \sigma_{up}}{2}$ (σ_{down} , σ_{up} : calculated from the difference between variated and nominal one)
  - Theoretical
    - Scale and PDF variations
    - $\sigma_{theory} = \frac{max - min}{2}$ (max, min : take maximum and minimum for each bin, envelope)

How to add SM and BSM uncertainties

- Key Point :
  - Uncertainties will vary with the change of coupling values (in order to get one conservative estimation, will take the envelope and not switch regard of the coefficients)
  - Illustration Photon uncertainties (from p_T^{γ} distribution, O_{G+} operator, coupling values are : 0.03, 0.05, 0.07, 0.08, 0.1)

Bin/coupling value	**Bin 1**	**Bin 2**	**Bin 3**	**Bin 4**	**Bin 5**	**Bin 6**	**Bin 7**
cg = 0.03	-0.473403	-0.314216	-0.357882	-0.226984	-0.18677	-0.0941598	-0.0644754
cg = 0.05	-0.442947	-0.327299	-0.346069	-0.236896	-0.197016	-0.0857605	-0.096733
cg = 0.07	-0.459166	-0.322781	-0.344047	-0.241693	-0.1714	-0.0960418	-0.152123
cg = 0.08	-0.45543	-0.314667	-0.362208	-0.227224	-0.182144	-0.105568	-0.191775
cg = 0.1	-0.447429	-0.325093	-0.353494	-0.237633	-0.187973	-0.0991581	-0.283357

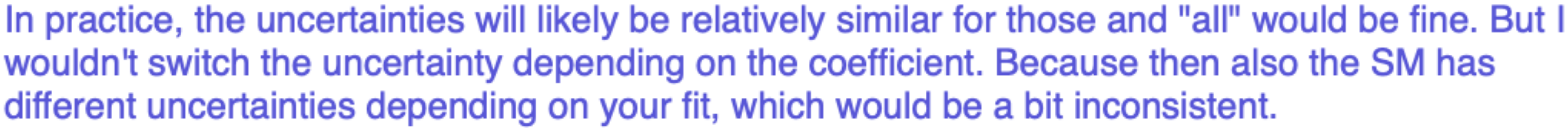
How to add SM and BSM uncertainties

- Key Point :
  - Uncertainties will vary with the dependence of coefficients (in order to get one conservative estimation, will take the envelope and not switch regard of the coefficients)
  - Illustration Photon uncertainties (from p_T^{γ} distribution, 6 operators, with similar X-Sec)

Bin/coupling value	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7
CBB = 1.1	-0.464964	-0.331438	-0.357601	-0.248526	-0.197257	-0.114902	-0.119815
CBtW = 1.4	-0.469689	-0.326359	-0.36301	-0.234285	-0.201883	-0.104064	-0.120964
CBW = 2.5	-0.451926	-0.320219	-0.372926	-0.237597	-0.190548	-0.111342	-0.107714
CWW = 5.5	-0.463189	-0.31923	-0.342997	-0.237487	-0.192507	-0.105631	-0.124554
cg = 0.05	-0.442947	-0.327299	-0.346069	-0.236896	-0.197016	-0.0857605	-0.096733
cgm = 0.9	-0.466689	-0.328365	-0.354337	-0.245857	-0.188888	-0.110012	-0.281408

How to add SM and BSM uncertainties

- According to the advice from Hannes



> In practice, the uncertainties will likely be relatively similar for those and "all" would be fine. But I wouldn't switch the uncertainty depending on the coefficient. Because then also the SM has different uncertainties depending on your fit, which would be a bit inconsistent.

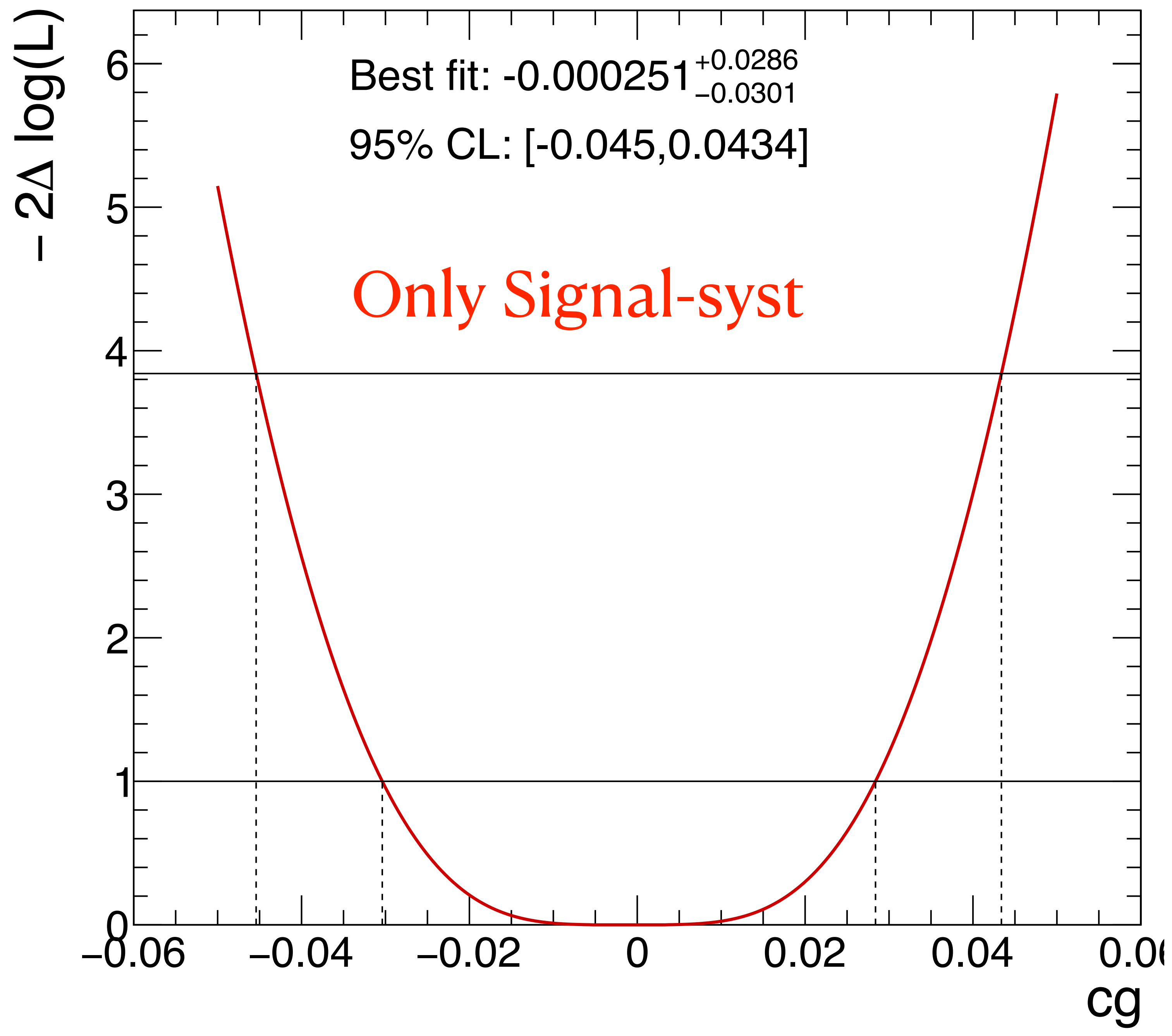
  - Solution :
    - take envelope for each term as an overall prediction uncertainty, and will not switch depending on the coefficients
    - In order to not to import extreme large predictions (which are related to the coupling values), will use estimated coupling value with similar X-sec (rough estimated at 95% C.L. during the official production)
    - CBB = 1.1, CBtW = 1.4, CBW = 2.5, CWW = 5.5, cg = 0.05, cgm = 0.9

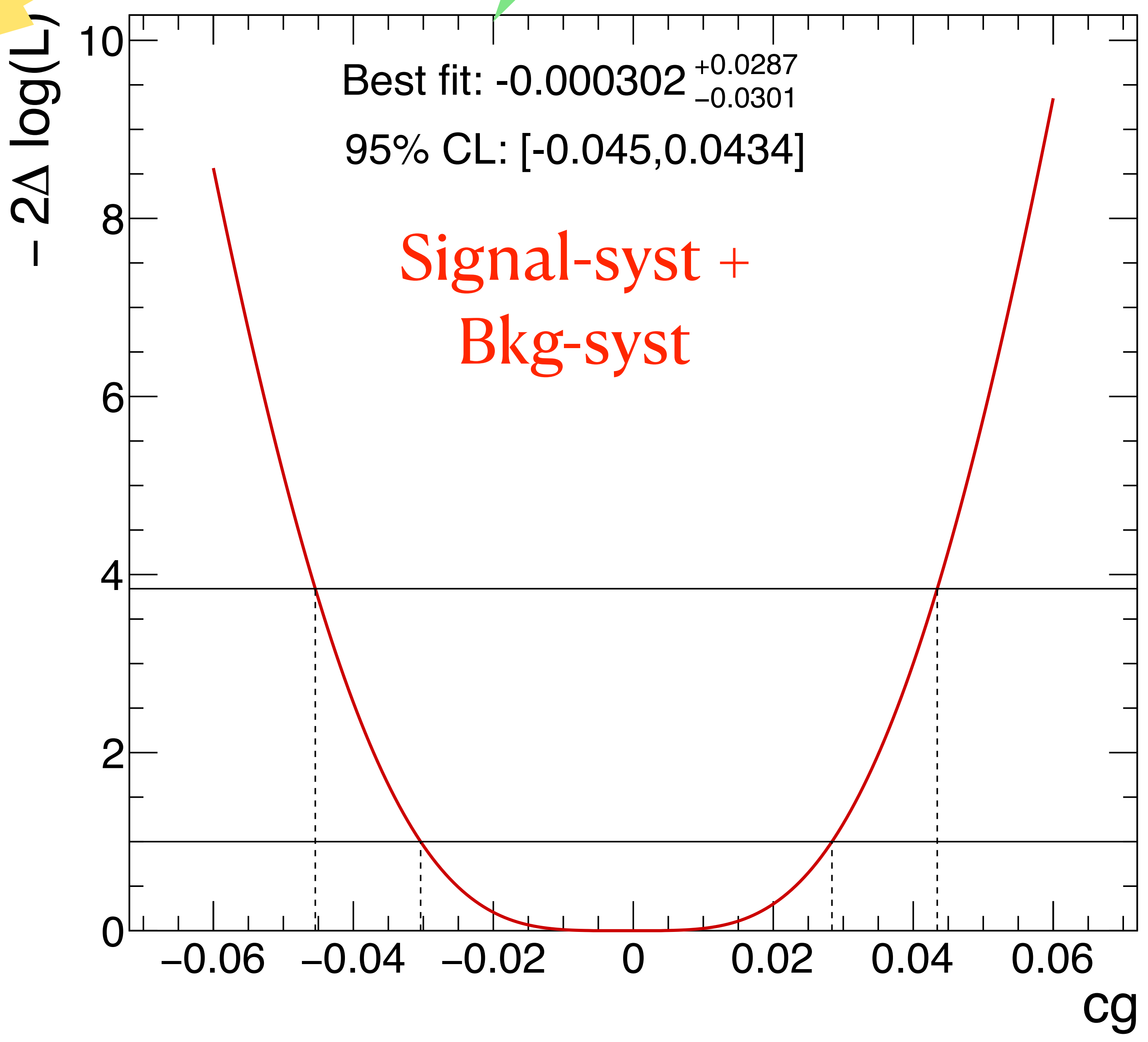
Limits comparison

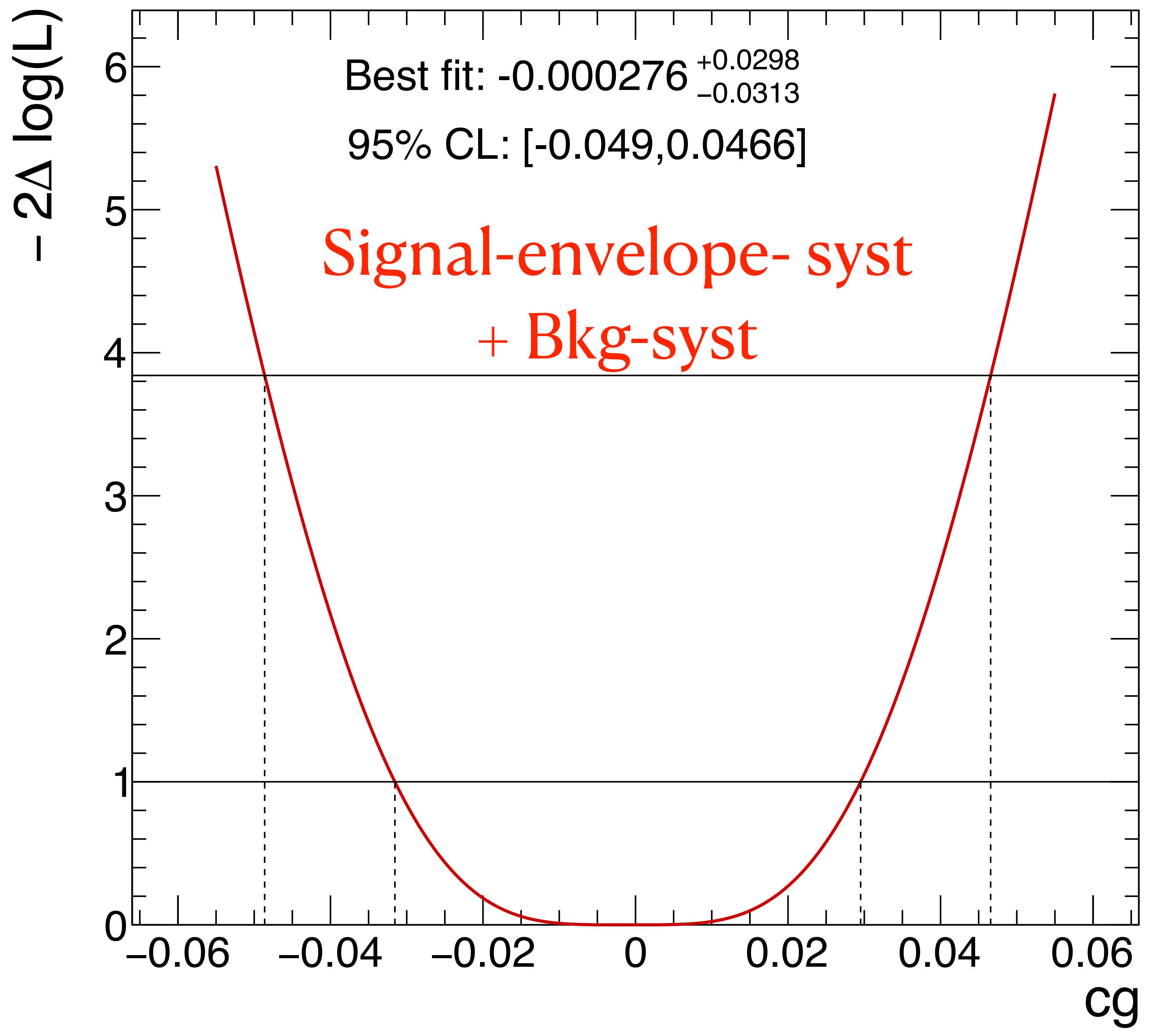
- Fitting results with different setups
  - Only Signal-syst
  - Signal-syst + Bkg-syst
  - Signal-envelope-syst + Bkg-syst

due to the small contribution of backgrounds except for SM Zgamma prediction

Almost no influence on limits, but small effects on best fit value







How to add SM and BSM uncertainties

- In order to get [Prediction Uncertainties] for both SM and BSM terms
- Here we use variated distributions and nominal distributions and take the differences as the uncertainties from different sources
  - For each variated distribution
    - histogram-SM : variated distribution with coupling values = 0
    - histogram-Interference : variated distribution with positive and negative coupling values (XS(a) - XS(-a)) / 2
    - histogram-Quadratic : variated distribution with positive , negative and zero coupling values (XS(a) + XS(-a)) /2 - XS(0)

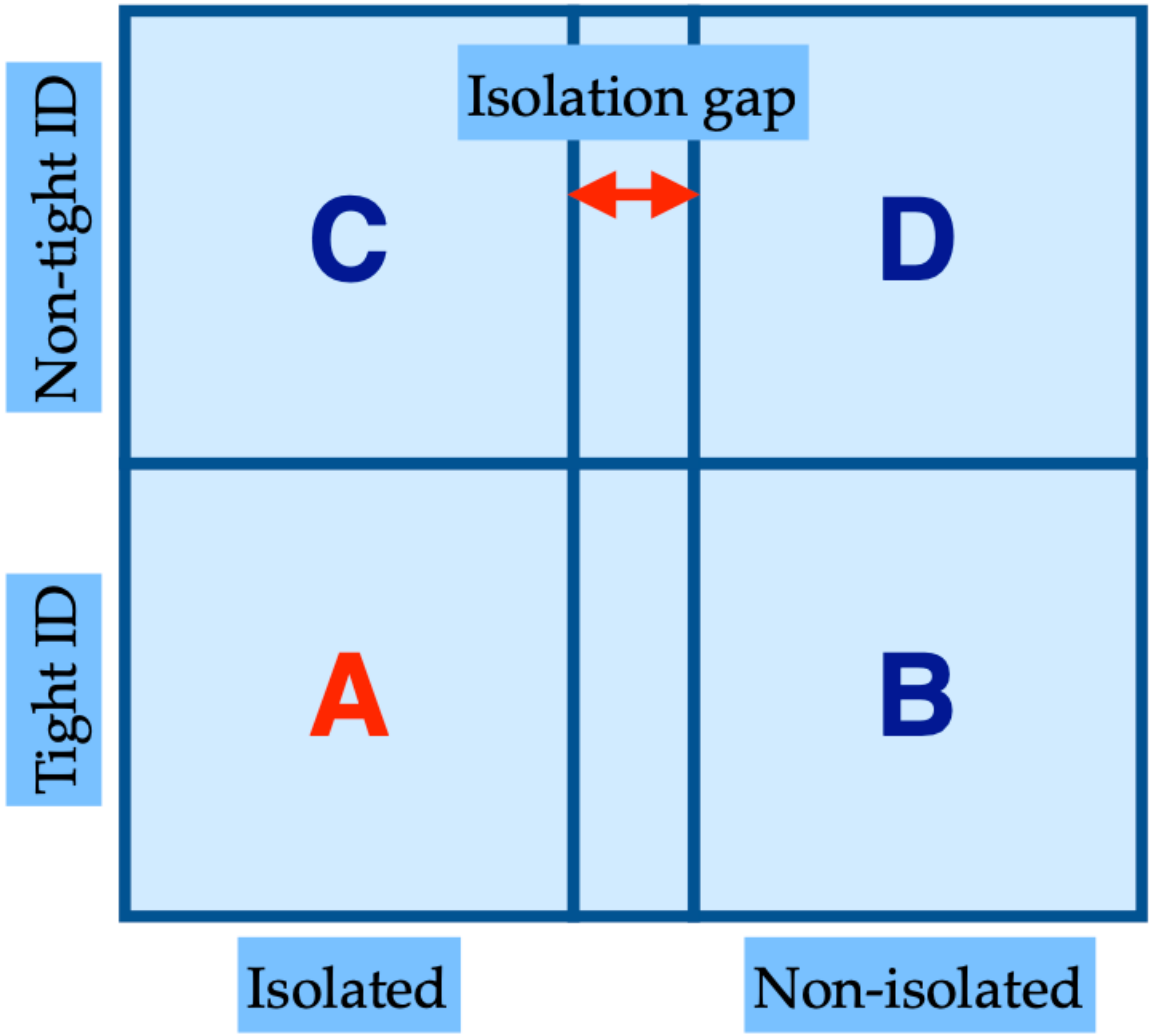
  - Same procedure with nominal distributions
  - Difference between variated and nominal distributions are considered as uncertainties

How to add SM and BSM uncertainties

- In principle, differences between the nominal ones and variated ones are considered as uncertainties
  - First use re-weighted distribution to extract SM, BSM (combine INT and QUA) terms with different sources of systematics (experimental or theoretical)
  - Then compare the variated distributions and nominal distributions
  - Use the differences as systematics for both SM and BSM

Background estimation — Z+jets

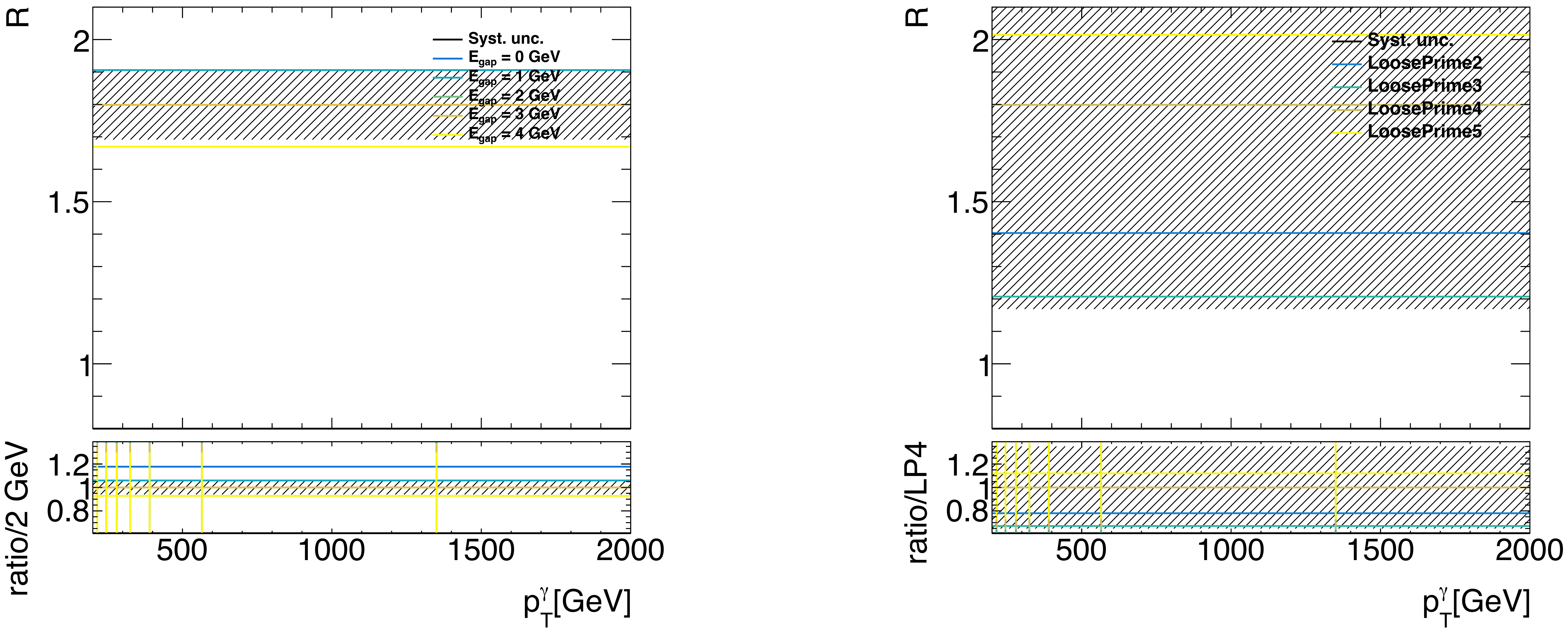
- Estimation code obtained from Vincent
  - Identification criteria : LoosePrime4
  - Isolation criteria : $E_T^\gamma > 0.065 p_T^\gamma$ ($E_{gap} = 2GeV$)
  - Apply event-selections in nTGC region
    - $p_T^\gamma > 200GeV$
    - $N_{jet} = 0$ (exclusive)



- Strategy :
  - Use larger bin until the purity computation
  - To preserve statistics and also distribution trends

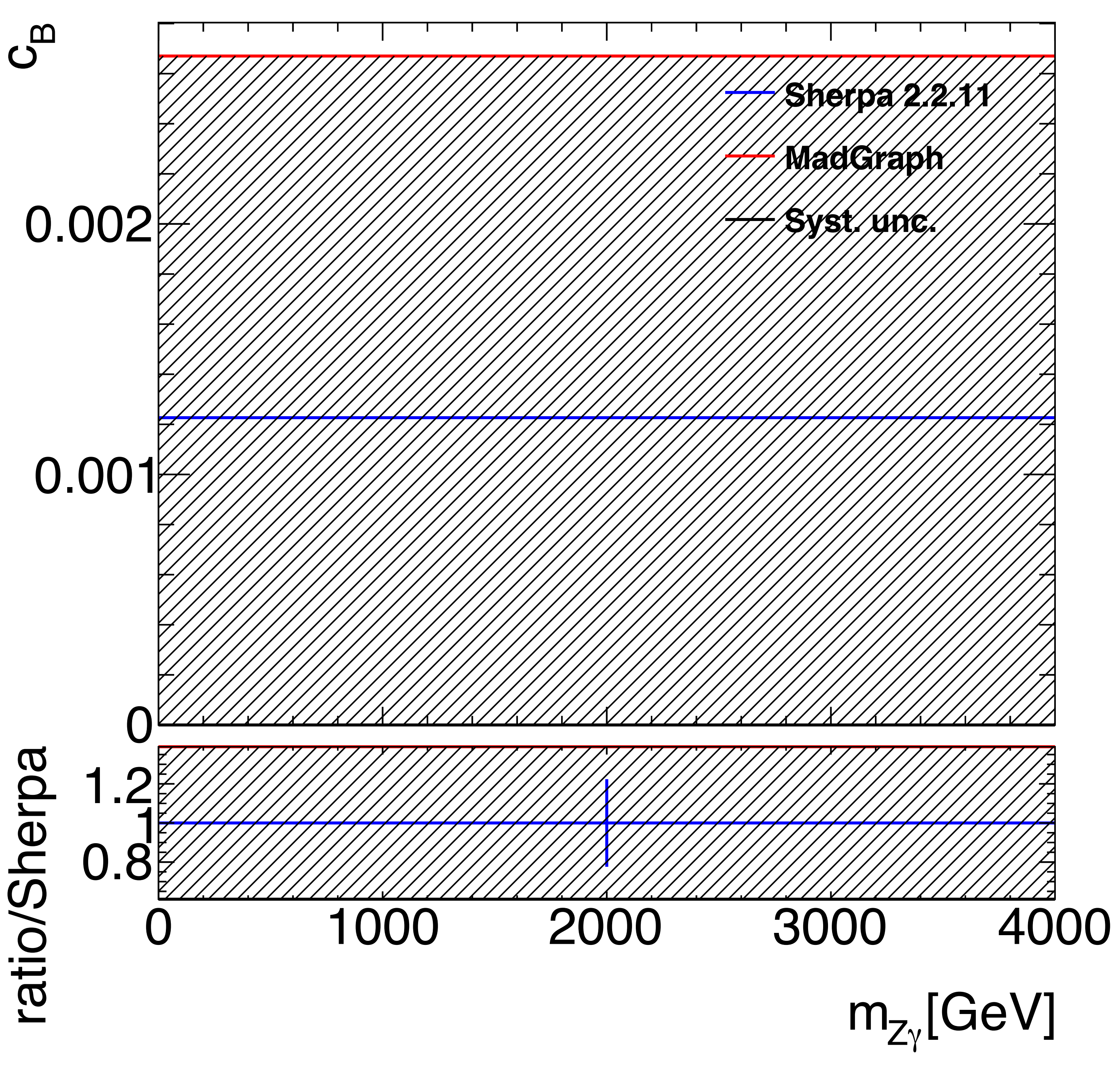
Background estimation — Z+jets

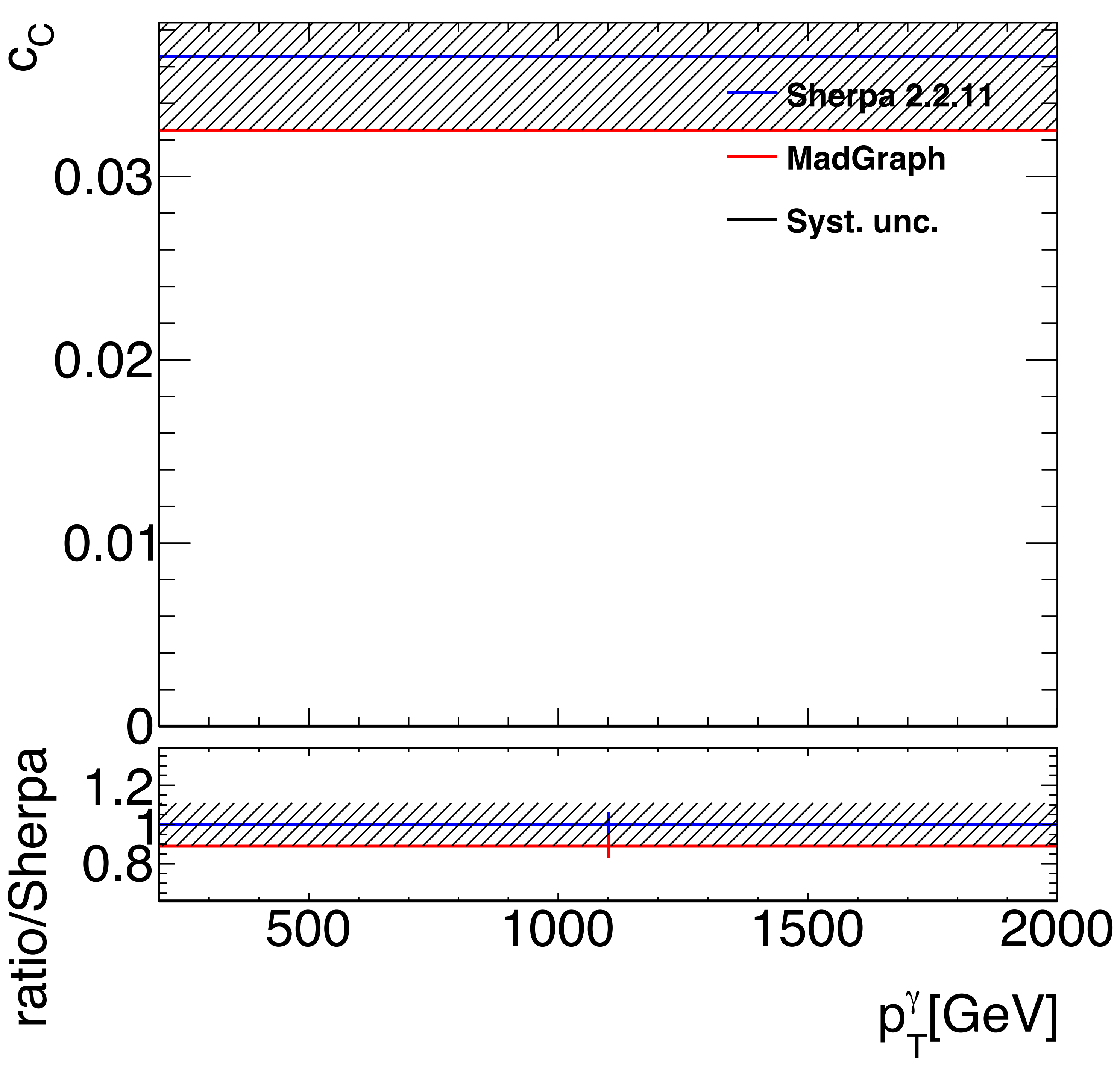
- Comparison of the R factor obtained with the different E_{gap} and Anti-tight definition for p_T^{γ} distribution
  - Nominal criteria : $E_{gap} = 2GeV$, anti-tight = LoosePrime4

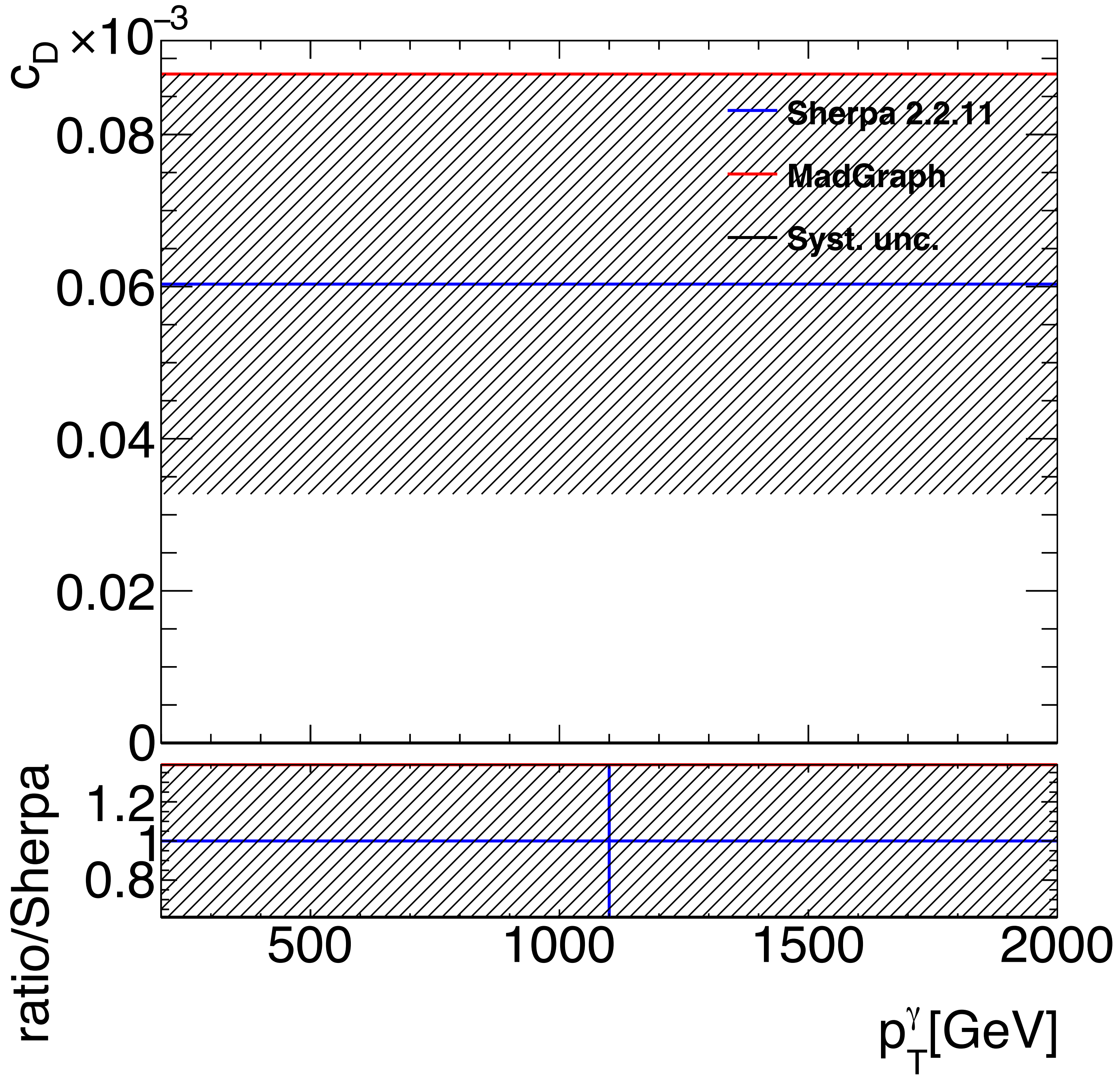


Background estimation — Z+jets

- Comparison of the leakage factors ($C_B/C_C/C_D$) for different generators, also as a function of p_T^γ distribution
  - From larger bin : one-bin used to calculate c_X due to lack of statistics







Background estimation — Z+jets

- Results
  - Table 1 : the parameters entering the ABCD method are summarised here :

Same issue with Pile-up backgrounds

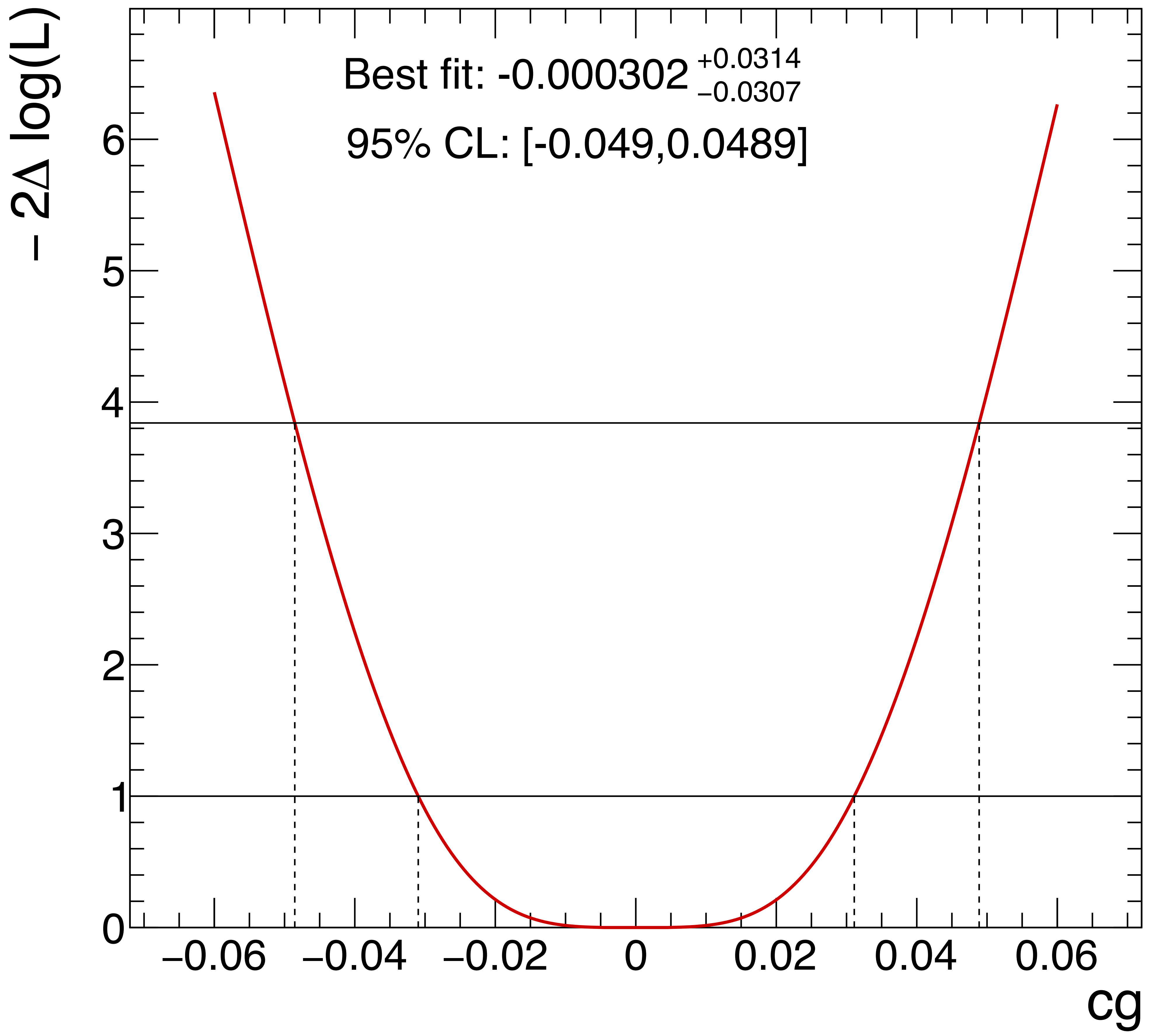
Lack of statistics with $N_{jets} = 0$

Apply correction factor : $\dfrac{N_{jets} = 0}{N_{jets} \geq 0}$

Anti-tight **Isolation gap**	**LoosePrime4** **2 GeV**
R	$1.80 \pm 0.65(stat) \pm 0.64(syst)$
$C_B(\times 10^{-3})$	$1.23 \pm 0.27(stat) \pm 1.44(syst)$
$C_C(\times 10^{-3})$	$36.57 \pm 2.26(stat) \pm 4.03(syst)$
$C_D(\times 10^{-3})$	$0.06 \pm 0.04(stat) \pm 0.03(syst)$
$N_A^{data} - N_A^{bkg}$	$412.78 \pm 10.37(stat) \pm 0.65(syst)$
$N_B^{data} - N_B^{bkg}$	$3.91 \pm 0.74(stat) \pm 0.02(syst)$
$N_C^{data} - N_C^{bkg}$	$52.50 \pm 3.13(stat) \pm 0.12(syst)$
$N_D^{data} - N_D^{bkg}$	$3.98 \pm 0.65(stat) \pm 0.003(syst)$

Fitting results with Background

- Add Z+jets background contribution to the fitting
  - With systematics (signal, bkg)
  - Converted limits for h_4 :
    - $[-9.0 \times 10^{-5}, 9.0 \times 10^{-5}]$



Fitting Results

- Add Z+jets background contribution to the fitting

- Summarised table for 6 operators :

Operator	**EFT**	**EVT**
O_{BB}	[-0.90, 0.90]	
$O_{\tilde{B}W}$	[-1.20, 1.16]	$[-7.16\times10^{-4}, 6.92\times10^{-4}]$
O_{BW}	[-2.00, 2.17]	
O_{WW}	[-3.90, 4.11]	
O_{G+}	[-0.05, 0.05]	$[-9.0\times10^{-5}, 9.0\times10^{-5}]$
O_{G-}	[-0.49, 0.50]	$[-1.64\times10^{-4}, 1.60\times10^{-4}]$

Next steps

- Double check the background estimation procedure, make sure everything is correct

- Currently, prediction uncertainty is “ALL” (contribution for SM and BSM are mixed up)
  - Next step, will try to derive uncertainties separately for SM and BSM effects

- Double check the background systematics
  - Especially combine systematics from data-driven backgrounds and MC background samples

- Starting to document everything in the note

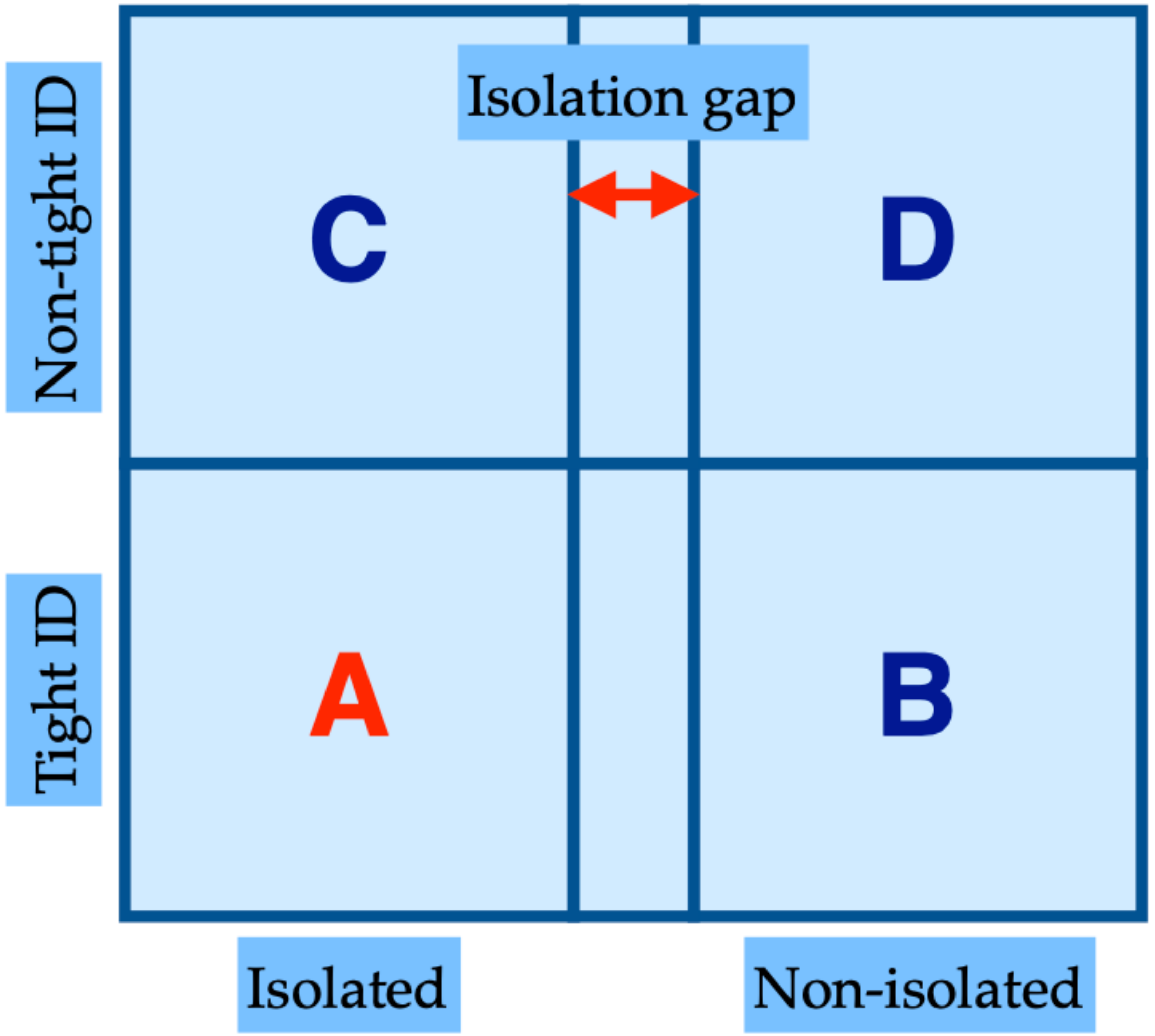
Backup

PDF systematic check

- PDF variation is around 20%
  - Now we need to check the compact of variation difference against the EFT signal strength in high p_T region

ABCD method

- ABCD : 2D sideband method



Two-dimensional sideband method is used for estimating Z+jets background

- Rely on the ratios of MC generated samples
- Region A : the photon passes the tight identification and isolation criteria ($E_T^{iso} < 0.065 p_T^{\gamma}$)

-->

- Region B : the photon passes the tight identification but fails the isolation criteria ($E_T^{iso} > 0.065 p_T^{\gamma} + E_{gap}$)
- Region C : the photon fails the tight identification but pass the isolation criteria ($E_T^{iso} < 0.065 p_T^{\gamma}$)
- Region D : the photon fails the tight identification (but still required to pass anti-tight ID criteria : LoosePrime4) and isolation gap ($E_T^{iso} > 0.065 p_T^{\gamma} + E_{gap}$)

Background estimation — Z+jets

- Correlation factor R： $R = \dfrac{N_A^{Z+jets} \times N_D^{Z+jets}}{N_B^{Z+jets} \times N_C^{Z+jets}}$
- Signal leakages in different control region : $C_B = \dfrac{N_B^{sig}}{N_A^{sig}}, C_C = \dfrac{N_C^{sig}}{N_A^{sig}}, C_D = \dfrac{N_D^{sig}}{N_A^{sig}}$
- The data-driven estimate number of signal events in the signal region can therefore be written as :
  - $N^{sig} = N_A^{data} - N_A^{bkg} - N_A^{Z+jets} = N_A^{data} - N_A^{bkg} - R\dfrac{[(N_B^{data} - N_B^{bkg}) - c_B N^{sig}][(N_C^{data} - N_C^{bkg}) - c_C N^{sig})]}{(N_D^{data} - N_D^{bkg}) - c_D N^{sig}}$
  - N_A^{bkg} is the number of background events obtained from MC simulations other than Z+jets
- Purity : $P = \dfrac{N^{sig}}{N_A^{data} - N_A^{bkg}}$
- $N^{Z+jets} = (N_A^{data} - N_A^{bkg}) \times (1 - P)$

Background estimation — Z+jets

- Strategy
  - Bin-by-Bin : the most intuitive way to compute correlation factor, purity and Z+jets within the bin, but lack of statistics
  - Fit function : some variables exhibit a trend, but lack of knowledge of fitting formula
  - Inclusive correlation factor or purity : use a unique number on all the events, but introduce an issue due to the trends within the distribution, will effect the overall Z+jets estimate shape, also a systematics was introduced
  - **Larger bins** : both combine a high enough statistics and a preservation of the distribution trends — used in the analysis
    - Kept until the purity computation
    - Then Z+jets estimate is computed Bin-by-Bin
    - Provides the best trade off between keeping the differential information and enough statistics

Background estimation — Z+jets

- Systematics : including both statistics and systematics
  - Background : systematics account for luminosity uncertainties (15% for $t\bar{t}\gamma$, and 30% for the other MC simulated sources of background ($WZ\gamma$, $WW\gamma$ . . .)
  - Experimental : due to the object reconstruction and modelling of the reconstruction
  - Signal leakage parameters : different choice on event generator
    $\delta^2_{generator} = (c_X(MadGraph) - c_X(Sherpa))^2$
  - Correlation factor R : the choice of identification and isolation criteria
    $\delta^2_{anti-tight} = (R(LoosePrime5) - R(LoosePrime4))^2 + (R(LoosePrime3) - R(LoosePrime4))^2$
    $\delta^2_{isolation-gap} = (R(E_{gap} = 3GeV) - R(E_{gap} = 2GeV))^2 + (R(E_{gap} = 1GeV) - R(E_{gap} = 2GeV))^2$
- Different systematic uncertainties sources are propagated to the Z+jets estimates by the ABCD method

Limits Comparison

	Z(ll)+gamma	**Z(vv)+gamma**
Operator	**EFT**	**EFT**
O_{BB}	[-0.90, 0.90]	[-0.28, 0.27]
$O_{\tilde{B}W}$	[-1.20, 1.16]	[-1.3, 1.3]
O_{BW}	[-2.00, 2.17]	[-0.74, 0.74]
O_{WW}	[-3.90, 4.11]	[-2.7, 2.7]
O_{G+}	[-0.05, 0.05]	
O_{G-}	[-0.49, 0.50]	

	Z(ll)+gamma	**Z(vv)+gamma**
h_4^Z	$[-9.0 \times 10^{-5}, 9.0 \times 10^{-5}]$	$[-5.3 \times 10^{-7}, 5.1 \times 10^{-7}]$
h_4^γ		$[-5.1 \times 10^{-7}, 5.0 \times 10^{-7}]$
h_3^Z	$[-7.16 \times 10^{-4}, 6.92 \times 10^{-4}]$	$[-3.8 \times 10^{-4}, 3.8 \times 10^{-4}]$
h_3^γ	$[-1.64 \times 10^{-4}, 1.60 \times 10^{-4}]$	$[-4.2 \times 10^{-4}, 4.3 \times 10^{-4}]$