Institute of High Energy Physics
Chinese Academy of Sciences

CEPC

The 2024 International Workshop on the High Energy Circular Electron Positron Collider

Top Quark EW Coupling Precision Measurement

Mustapha BIYABI On Behalf of

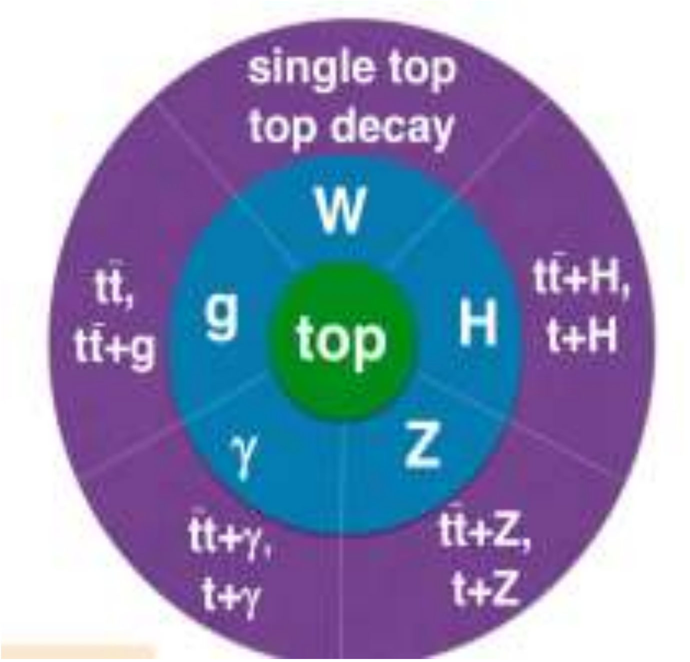
LI Gang , Xiaoxu Zhang, Hongbo Liao, Shudong WANG, JIARONG LI, Yaquan FANG.

24 Oct 2024

mustapha.biyabi@cern.ch

Outline :

- **Introduction &Motivations**
- **Circular Electron-Positron Collider**
- **Analysis Strategy**
- **Top Quark Study**
- **Background Analysis**
- **Summary**

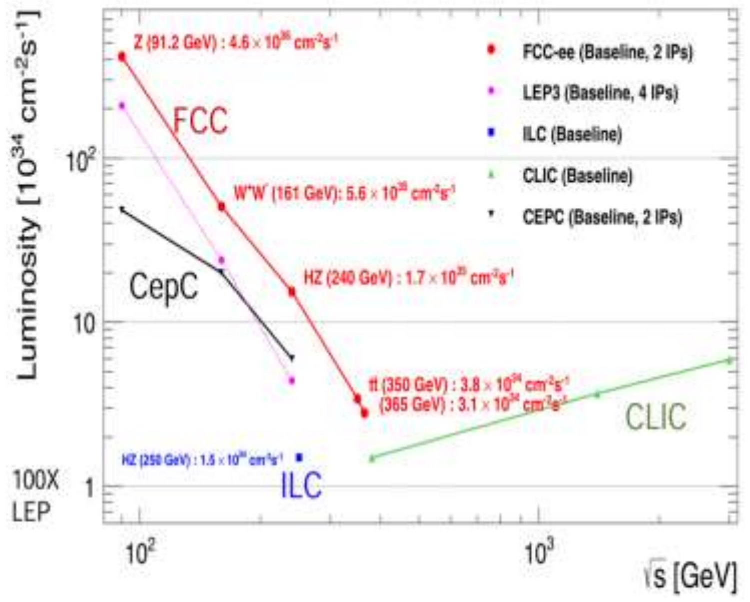


Physics Motivations

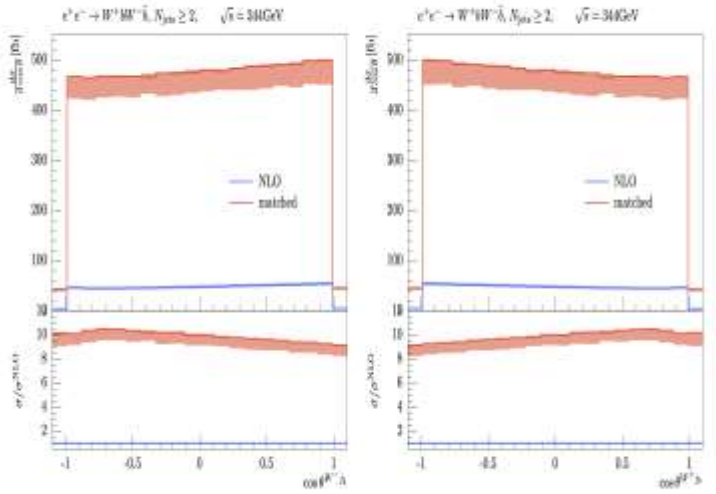
- The top quark : a great candidate to search for new physics, precise properties and interactions are a big part of the high energy physics programme of the coming decades.

An electron-positron collider such as **CEPC** can precisely measure the electroweak couplings of the top quark to the level of a few percent.

$$\Gamma_{\mu}^{t\bar{t}X}(k^2,q,\bar{q}) = -ie\left\{\gamma_\mu\left(F_{1V}^X(k^2)+\gamma_5 F_{1A}^X(k^2)\right)+\frac{\sigma_{\mu\nu}}{2m_t}(q+\bar{q})^\mu\left(iF_{2V}^X(k^2)+\gamma_5 F_{2A}^X(k^2)\right)\right\}$$



- **The determination of top antitop EW couplings at the LHC reached an interesting precision level ! Improvements beyond this level can be foreseen at proposed ee colliders.**

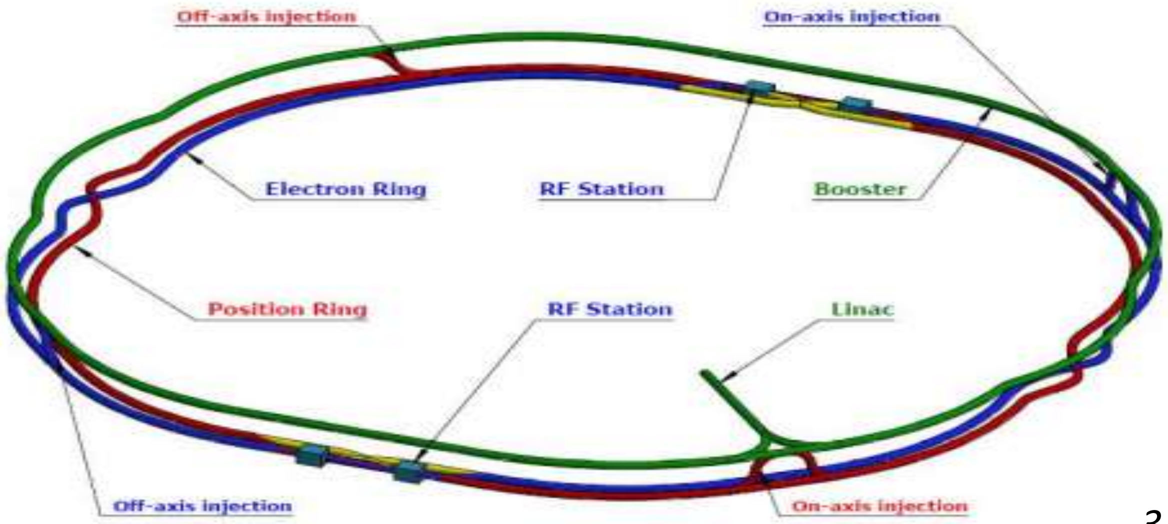


- **Advantages : Lepton colliders have the ability to produce tt pairs with low background, which is crucial for precise measurements of the tt coupling.**

- **Validation of the differential distribution (cross section) is not finished.**

Circular Electron Positron Collider

The CEPC is a proposed future lepton collider project hosted by China. It is designed to operate at a center-of-mass energy of 360 GeV, which will allow it to produce top quark-antiquark pairs for detailed studies of top quark properties using unpolarized beams.



Operation mode			ZH	Z	W^+W^-	$t\bar{t}$
$\sqrt{s}$ [GeV]			240	91	160	360
Run time [years]			7	2	1	-
CDR (30 MW)		L / IP [$\times 10^{34}$ $cm^{-2}s^{-1}$]	**3**	**32**	**10**	-
		$\int L\, dt$ [ab^{-1}, 2 IPs]	5.6	16	2.6	-
		Event yields [2 IPs]	1×10^6	7×10^{11}	2×10^7	-
Run Time [years]			10	2	1	5
TDR (Latest)	30 MW	L / IP [$\times 10^{34}$ $cm^{-2}s^{-1}$]	5.0	115	16	0.5
		$\int L\, dt$ [ab^{-1}, 2 IPs]	13	60	4.2	0.65
		Event yields [2 IPs]	2.6×10^6	2.5×10^{12}	1.3×10^8	4×10^5
	50 MW	L / IP [$\times 10^{34}$ $cm^{-2}s^{-1}$]	8.3	192	26.7	**0.8**
		$\int L\, dt$ [ab^{-1}, 2 IPs]	21.6	100	6.9	1.0
		Event yields [2 IPs]	4.3×10^6	4.1×10^{12}	2.1×10^8	6×10^5

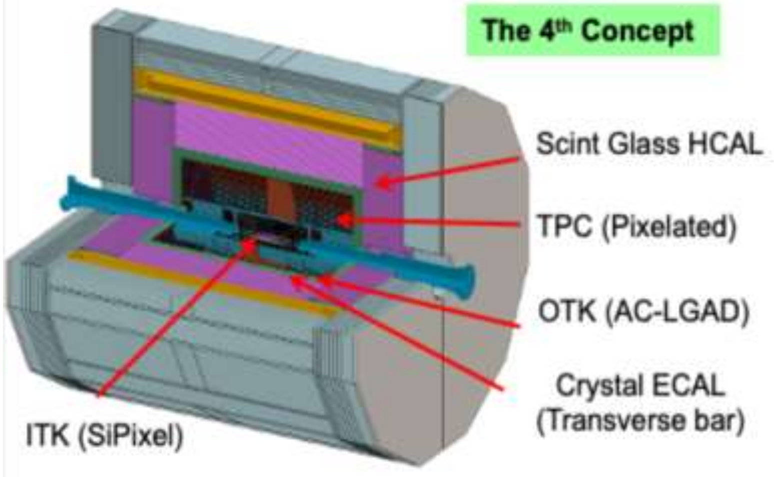
CEPC Detector

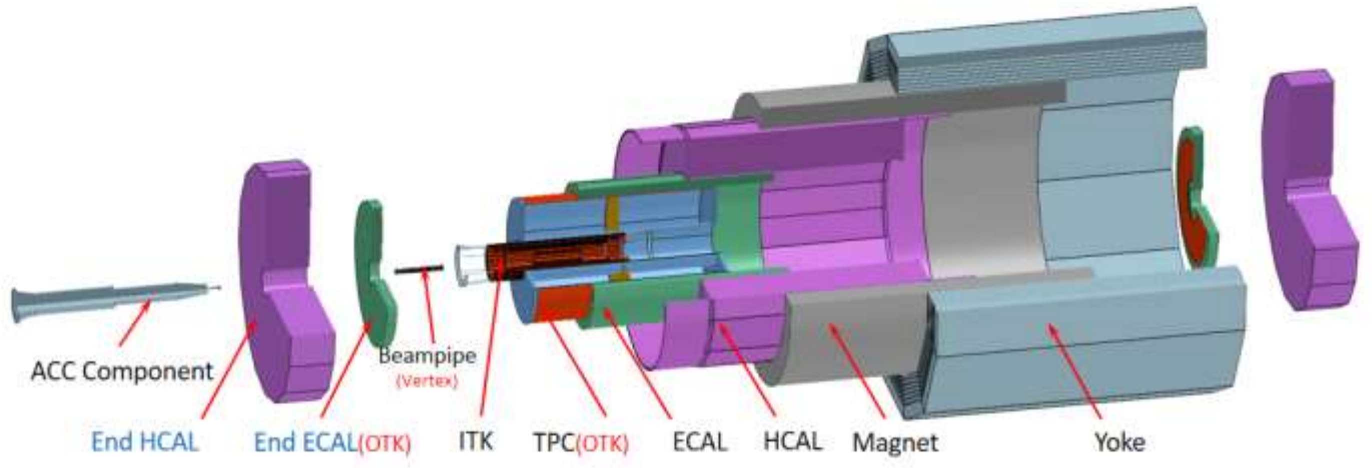
CEPC Conceptual Detector Design :

System	Technologies	
	Baseline	Backup / Comparison
Beam pipe	Φ20 mm	
LumiCal	SiTrk+Crystal	
Vertex	CMOS+Stitching	CMOS Pixel
Tracker	CMOS Si Pixel ITK	
		CMOS Si Strip ITK
	Pixelated TPC	PID Drift Chamber
	AC-LGAD OTK	SSD / SPD OTk
		LGAD ToF
ECAL	4D Crystal Bar	Stereo Crystal Bar, GS+SiPM, PS+SiPM+W, SiDet+W
HCAL	GS+SiPM+Fe	PS+SiPM+Fe, RPC+Fe
Magnet	LTS	HTS
Muon	PS bar+SiPM	RPC
TDAQ	Conventional	Software Trigger
BE electr.	Common	Independent

Radius

- ❑ **Excellent e/gamma energy resolution;**
- ❑ **PID capability;**
- ❑ **Better hadronic energy resolution;**





Analysis Strategy

- Generate Monte Carlo Data that correspond to ttbar semi-leptonic channel Using @Madgraph5.(*QQbarThreshold* is used for signal Xsection Calculation);

- Object reconstruction is done using the particle flow algorithm, Arbor, ue the ee-kT algorithm is used for jet clustering based on their distance :

$$d_{ji} = p_{t_j}^2 \frac{\Delta R_{ij}^2}{R^2}$$

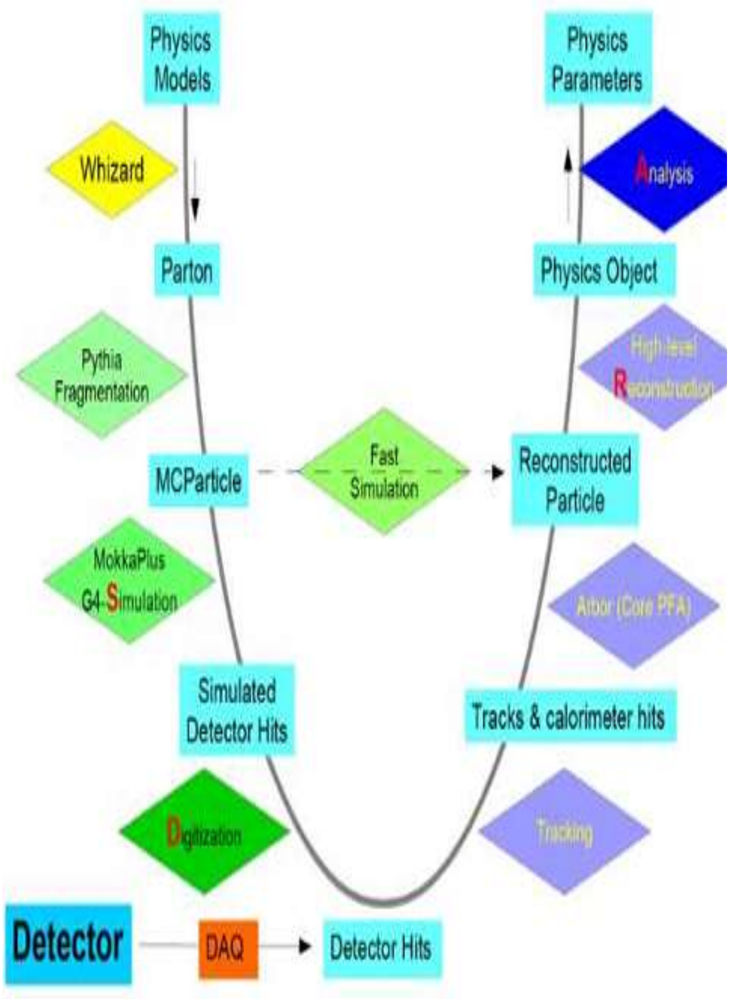
- Employ LCFIPLUS to perform jet tagging. Assuming that there are only 4 jets in each event, all particles are forced into 4 jets (performance of b-tagging is given by LCFIPlus);

- Lepton isolation for muon/electron;

- Reconstruct the Hadronic Top Mass and W Mass at the reconstructed level using a χ^2 Method.

- Background MC Samples generation, apply two techniques to reject the background: cut-based selection and machine learning methods.

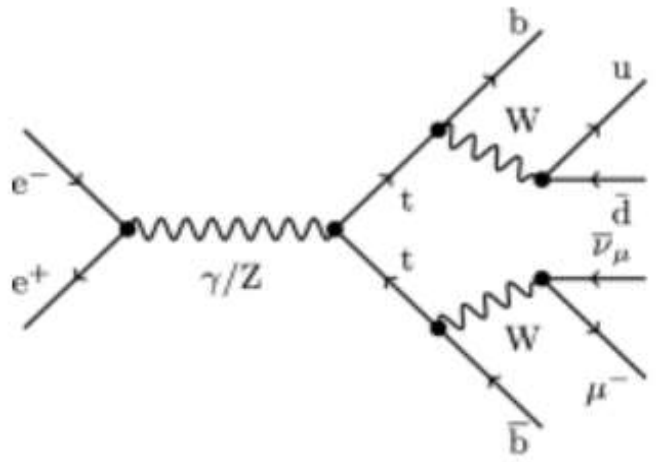
- Measure the top quark's electroweak couplings in the context of the Standard Model By inspectig the Top polar angle.



SemiLeptonic Channel

The are three differents channels for ttbar study :

- Semileptonic
- Dileptonic
- Hadronic



- For the **semi-leptonic channel :**

The data is generated using MG5 and passed through Pythia8 for showering and hadronization.

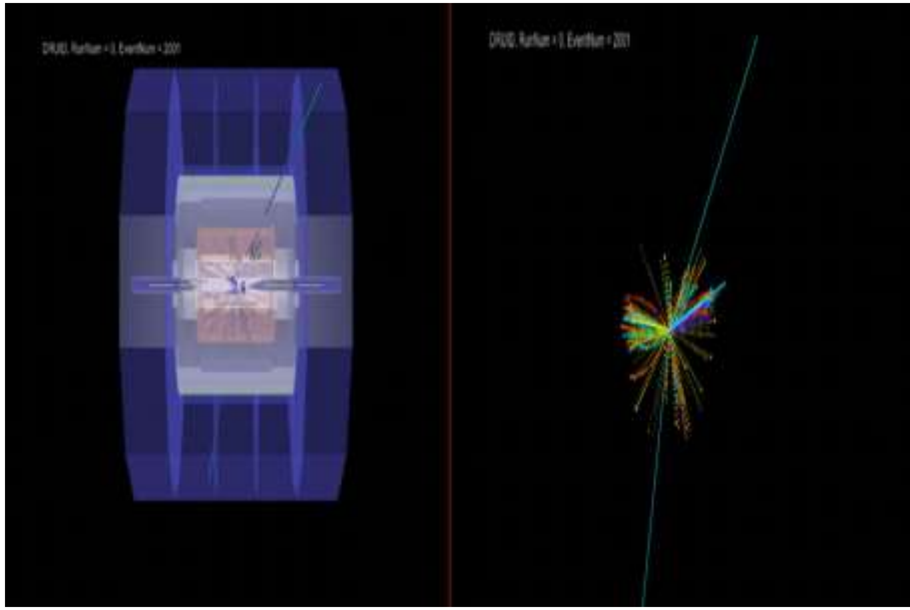
- CEPC_v4 (√s =360GeV, Bfield = 3.0 T)
- Luminosity is 1000 ab-1.

- Generated data is then passed through CEPC reco & sim.

- **There are two simulation methods for CEPC detector:**

- Fast Simulation : uses Delphes to perform a fast multipurpose detector response simulation.

- Full Simulation is a method using Geant4-based tools to simulate the particle transport and the detector response and then reconstruct the responses back to original signals.



Lepton Isolation

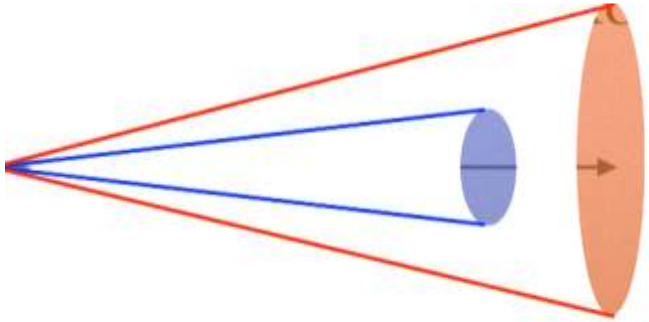
- ***Full simulation :***

Different Parameters were used for Lepton isolation :

❖ Specifically for identifying e and μ based on their energy deposits in various parts of a particle detector, such as:

- The maximum allowed ratio of energy deposited in the Ecal to the sum of energies deposited in both the Ecal and Hcal.

- Use rectangular cuts on track and cone energy, like : (Econe > 0.98)
  The cosine of the half-angle of the cone used in the isolation criteria.
  The maximum allowed energy within the isolation cone.
  The maximum allowed energy for any track within the isolation cone.



For Fast Simulation:

E_{lepton} and IPS (Impact Parameter Significance) are used to select the lepton (μ/e) :

$$IPS = \sqrt{(|D_0|/\sigma(D_0))^2 + (|D_Z|/\sigma(D_Z))^2} > 3.3$$

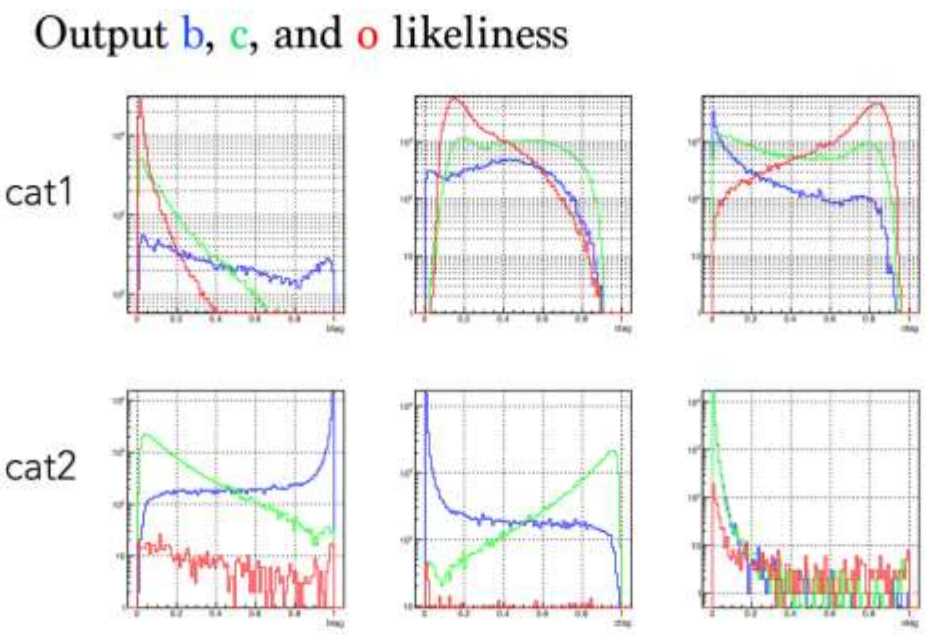
$$E_{Lepton} < 12\ GeV$$

Jet Tagging

LCFIplus :

- Use of BDT for multi-class classification
- 3 classes: b, c, o(uds, g assumed as same as uds)
- 4 categories: #vtx = 0, 1, 1+single, 2
- 20-30 variables for each category

- 2 independent outputs: b-likeliness, c-likeliness



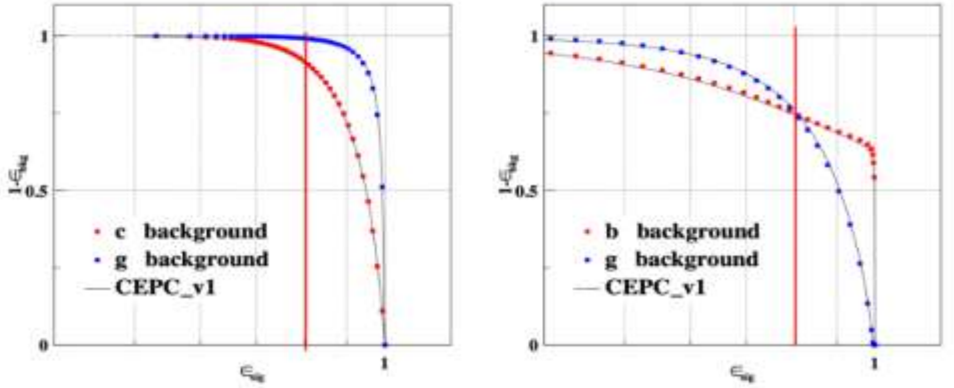
btag+ctag+otag=1: 2 outputs

❑ Receiver Operating Characteristic Curve (ROC)

❖ 80% b-tagging eff : Reject 90% c and 99% o jets

❖ 80% c-tagging eff. : Reject 75% b and 75% o jets

CEPC_v4 shows similar results :



For fast simulation:

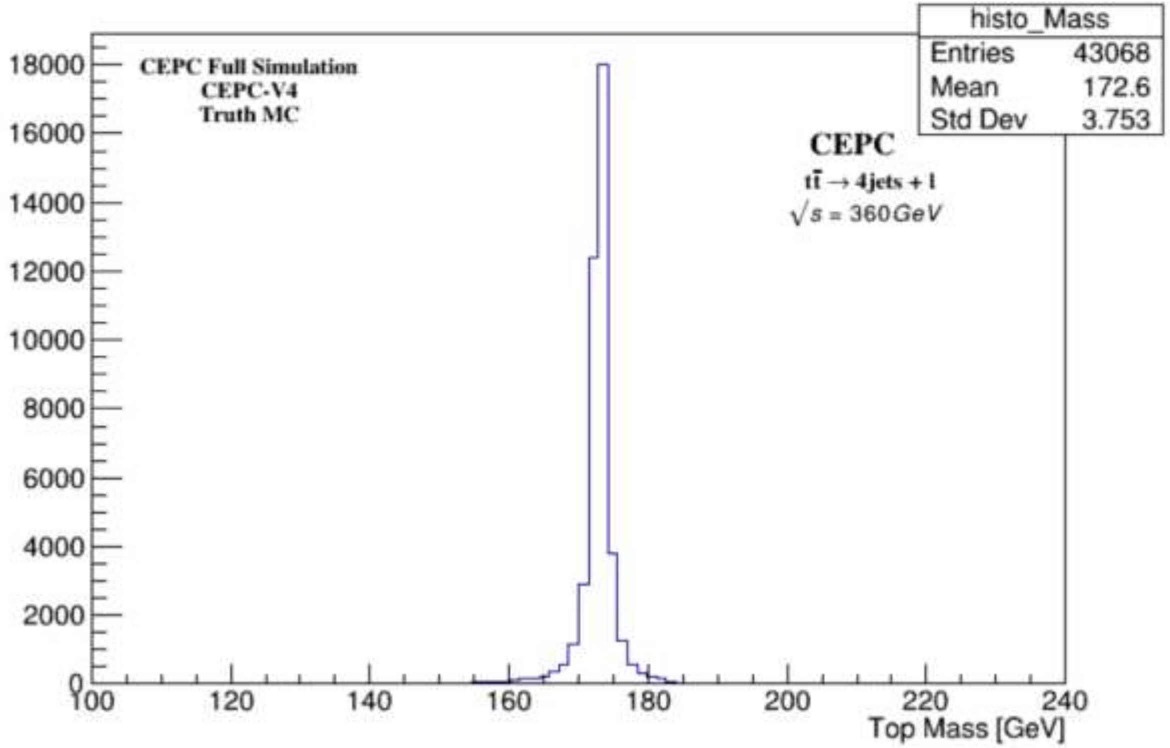
➢ B-tagging simulated according to a confusion matrix with accuracy ~ **90%.**

➢ B-tagging has 2 integer values: **1** (b quark) or **0** (other quark)

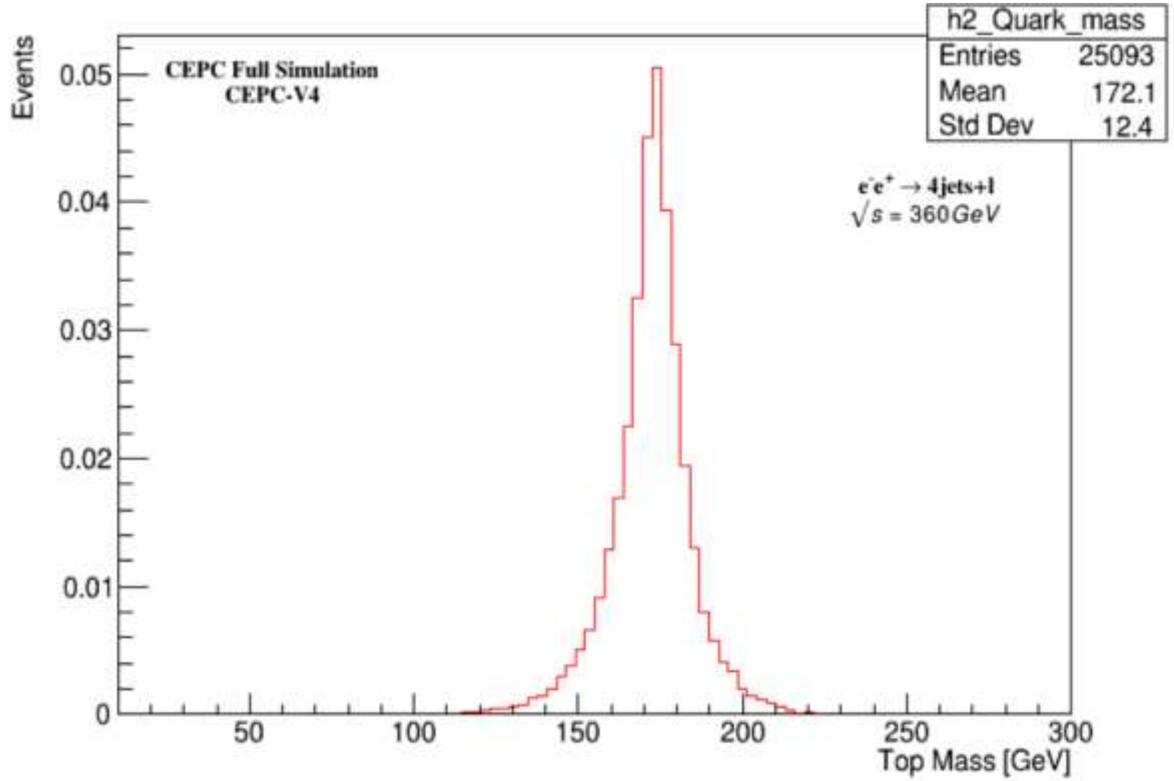
Top Mass Determination

- The reconstructed hadronic Top quark mass and truth level :
- Using a χ^2 method and requiring one jet to be tagged as a b-jet (80%).

Truth-Level



reco-Level



- Good Mass Reconstruction results with a reasonable standard deviation at 360GeV.

Top Mass Determination

➢ ***For better top mass resolution, we used kinematic fitting with the following χ² function:***

$$\chi^2 = \left(\frac{\sum_{b_H b_L jjl\nu} E_i - 360}{\sigma_E}\right)^2 + \left(\frac{\sum_{b_H b_L jjl\nu} P_{x_i}}{\sigma_{P_x}}\right)^2 + \left(\frac{\sum_{b_H b_L jjl\nu} P_{y_i}}{\sigma_{P_y}}\right)^2 + \left(\frac{\sum_{b_H b_L jjl\nu} P_{z_i}}{\sigma_{P_z}}\right)^2$$

$$+ \left(\frac{M_{b_H jj} - M_t}{\sigma_{M_{t_H}}}\right)^2 + \left(\frac{M_{b_L l\nu} - M_t}{\sigma_{M_{t_L}}}\right)^2 + \left(\frac{M_{jj} - M_W}{\sigma_{M_{W_H}}}\right)^2 + \left(\frac{M_{l\nu} - M_W}{\sigma_{M_{W_H}}}\right)^2$$

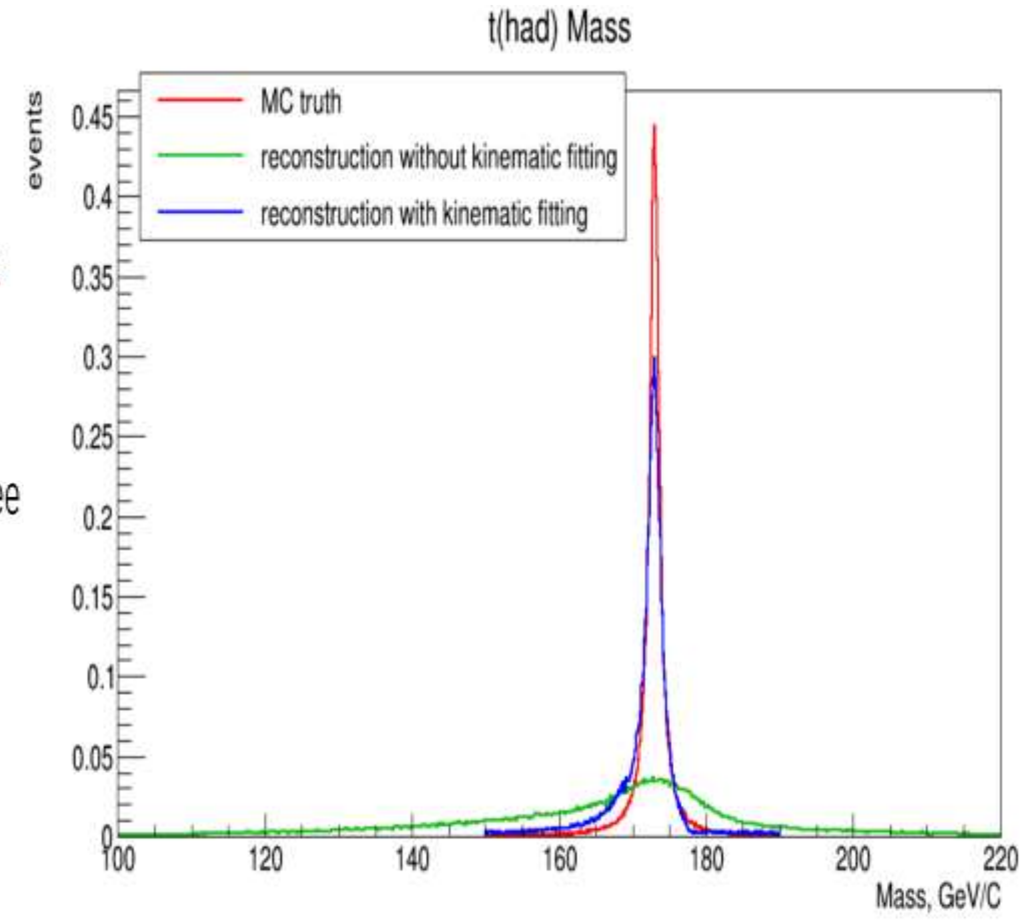
$$+ (sf_{b_H} - 1)^2 + (sf_{b_L} - 1)^2 + (sf_{jj} - 1)^2 + (sf_l - 1)^2$$

8 Kinematic Consrtraints

1 Degree

7 Unknows: 4 scalefactors(sf_{b_H}, sf_{b_L}, sf_{jj}, sf_l) and 3-momentum of neutrino(P_{x_ν}, P_{y_ν}, P_{z_ν}).

where M_t, M_W, σ_E, σ_{P_x}, σ_{P_y}, σ_{P_z}, $\sigma_{M_{t_H}}$, $\sigma_{M_{t_L}}$, $\sigma_{M_{W_H}}$, $\sigma_{M_{W_H}}$ are the parameters to be determined.



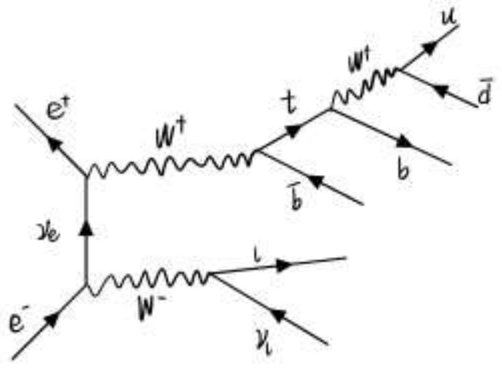
Background MC Samples

- **The background processes considered :**
  - **- diMuons Production ,**
  - **- qqbar Production ,**
  - **- diboson Production(WW/ZZ),**
  - **- Tri-Bosons Production (ZWW/ZZZ),**
  - **- Single Top Production .**

Background	Cross Section [pb]
$\mu\mu$	0.8163 ± 0.000475
$b\bar{b}$	0.7674 ± 0.00064
$q\bar{q}$	4.265 ± 0.00314
W^+W^-	10.94 ± 0.02035
ZZ	0.656 ± 0.001273
ZZZ	$0.0007322 \pm 7.52 \times 10^{-7}$
ZW^+W^-	$0.01484 \pm 2.226 \times 10^{-5}$
$SingleTop$	$0.007934 \pm 9.82610^{-6}$

- The background samples were generated using MG5 and simulated a full simlation and Fast Simulations.
- Table shows each Process with their corresponding cross-sections.

Background Analysis



- ***Preselection*** : Request 4jets + one Lepton (muon/electron).
- Cut Based Analysis :

Variables	TTbar	ZZZ	ZWW	WW	ZZ	qqbar	Muons	Single Top	bbar
Events	100000	50000	50000	50000	50000	100000	50000	100000	50000
nPFOS > 60	76.86%	13.65%	12.46%	0.22%	1.61%	27.12%	0.0%	71.91%	30.52%
ntrks > 15	76.86%	13.65%	12.40%	0.22%	1.50%	27.12%	0.0%	71.91%	30.50%
TotalP > 30	71.20%	2.95%	9.04%	0.10%	0.66%	5.70%	0.0%	58.45%	9.64%
Pmax < 45	68.75%	2.60%	2.20%	0.088%	0.27%	4.07%	0.0%	55.40%	4.88%
log(y34) > -10	68.41%	2.60%	2.05%	0.08%	0.27%	4.07%	0.0%	54.77%	4.80%
Thrust < 0.92	68.22%	2.68%	2.05%	0.08%	0.27%	2.46%	0.0%	54.60%	2.88%
Aplanarity > 0.01	68.22%	2.68%	2.05%	0.08%	0.27%	2.46%	0.0%	54.60%	2.88%
Log(Major) > -1	59.15%	2.30%	1.80%	0.009%	0.13%	0.75%	0.0%	48.04%	0.68%
Survived Events	16364	168	428	8	30	7	0	10573	21

BKG Process	BKG Rejections Efficiency
qqbar	99.25%
bbar	99.32%
$\mu\mu$	100%
ZZ	99.80%
WW	99.99%
ZZZ	97.7%
ZWW	98.2%
Single top	52.15%

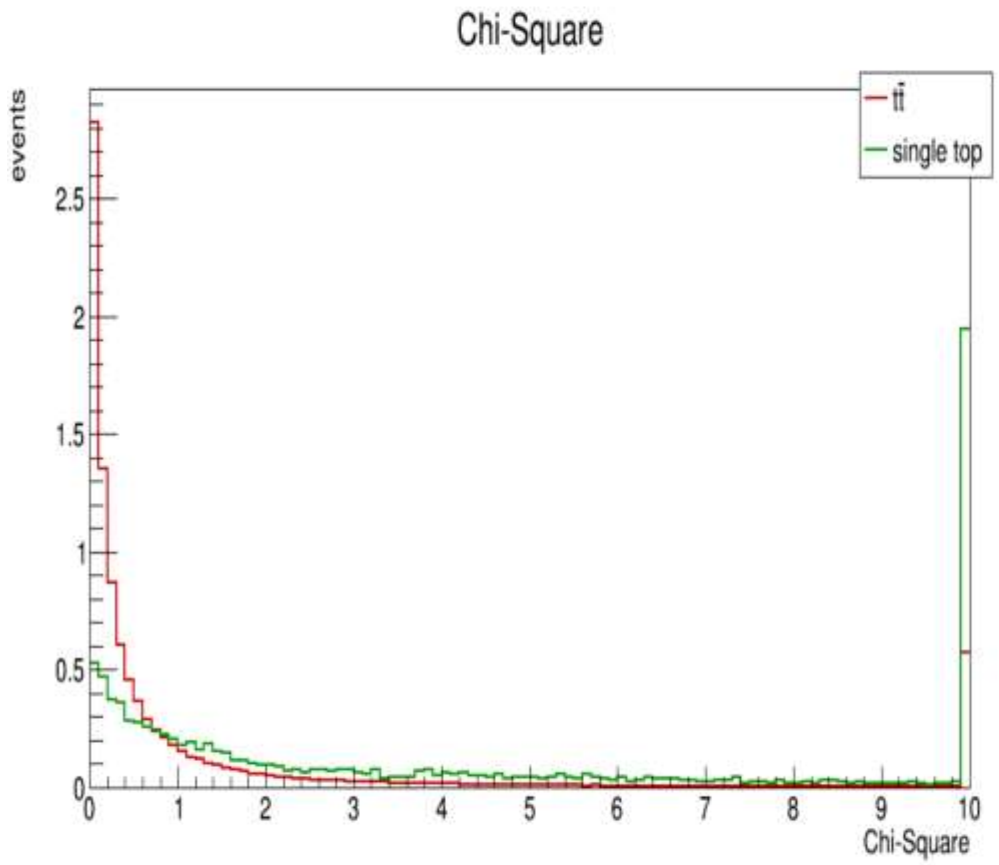
- ❑ All SM backgrounds are nearly removed, except for the Single top process.
- ❑ Different techniques were used to suppress the single top and enhance the signal efficiency.

* **Thrust quantifies how aligned the momenta of the particles are in a given direction.**

Background Analysis

❖ **Using Cut-based analysis and kinematic fitting (Fast simulation) :**

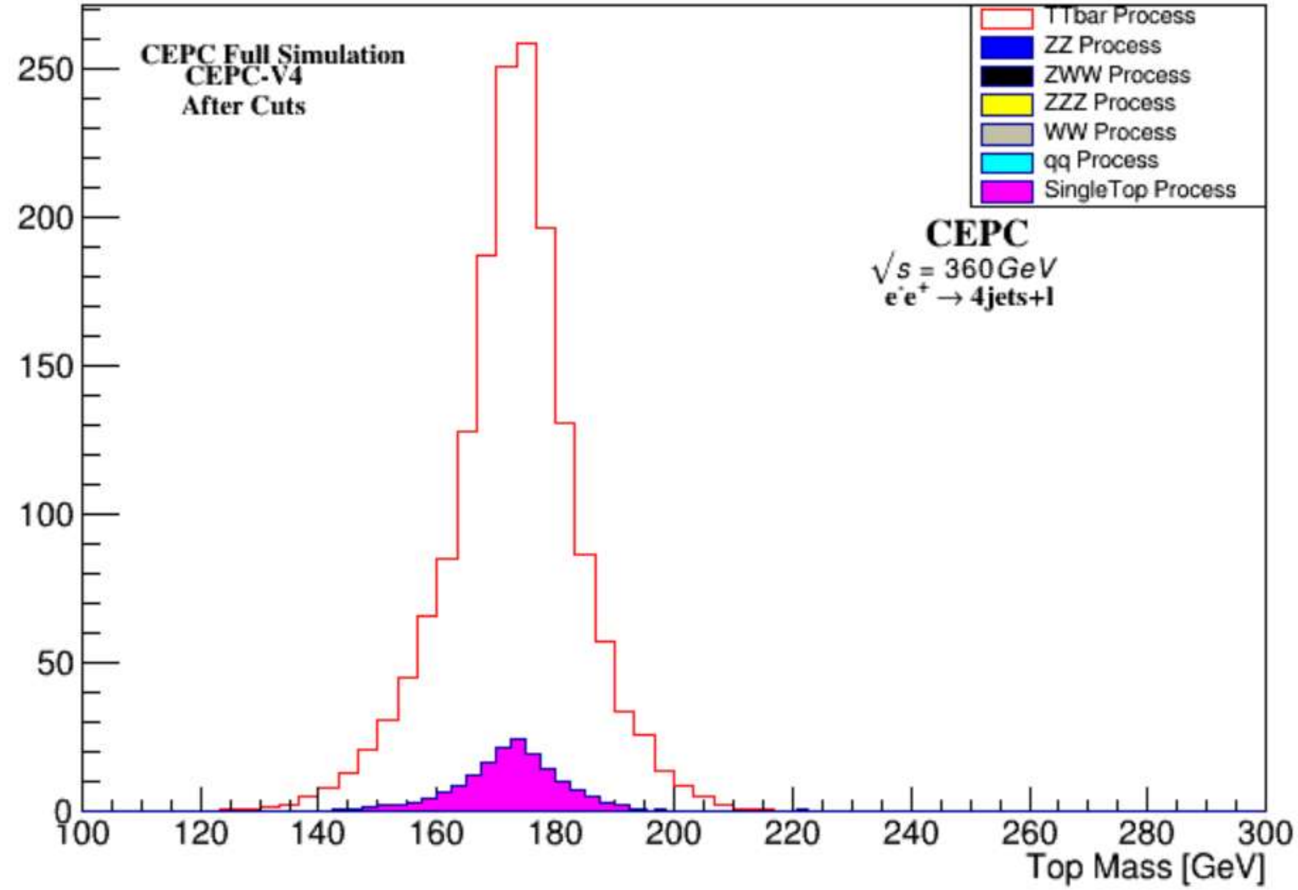
Process	$TTbar$	Single Top	$q\bar{q}$	W^+W^-	$b\bar{b}$	ZW^+W	ZZZ	ZZ
Events	1,000,000	20,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
$W_L \rightarrow \mu/e$ directly	66.67%	66.49%	0.72%	21.81%	0.60%	27.7%	16.6%	12.23%
4 jets	66.67%	66.49%	0.72%	15.66%	0.60%	26.19%	15.18%	9.24%
$\sum_{4jets} BTag = 2$	46.48%	44.73%	0.019%	0.096%	0.21%	2.18%	1.7%	0.58%
$\log(y_{34}) >$ -3.2	46.38%	44.49%	0.0097%	0.033%	0.13%	2.05%	1.64%	0.38%
$\log(y_{45}) >$ -4.1	46.36%	44.46%	0.0094%	0.033%	0.13%	2.05%	1.64%	0.38%
PFOs > 38	46.33%	44.42%	0.0094%	0.024%	0.13%	1.9%	1.52%	0.30%
Charged PFOs > 18	46.31%	44.40%	0.0094%	0.023%	0.13%	1.89%	1.51%	0.29%
PMax < 105	46.25%	43.27%	0.0092%	0.0091%	0.13%	1.69%	1.39%	0.076%
200 < TotalE < 344	46.14%	42.97%	0.0051%	0.0078%	0.094%	1.6%	0.56%	0.034%
Sphericity < 0.23	46.02%	42.82%	0.0021%	0.0056%	0.034%	1.59%	0.56%	0.034%
Thrust > 0.92	46.02%	42.82%	0.0021%	0.0056%	0.034%	1.59%	0.56%	0.034%
$\chi^2 < 1$	34.06%	14.14%	0.0001%	0.00%	0.0022%	0.23%	0.093%	0.0024%
Survived Events	59766	2239	4	0	15	34	1	16



- ❑ Using the χ² method helped further suppress the single top contribution.
- ❑ Different techniques were used to suppress the single top and enhance the signal efficiency.

Background Analysis

- The signal and background shapes after all cuts;
- Scaling each process to its cross section :



Background	Cross Section [pb]
$\mu\mu$	0.8163 ± 0.000475
$b\bar{b}$	0.7674 ± 0.00064
$q\bar{q}$	4.265 ± 0.00314
W^+W^-	10.94 ± 0.02035
ZZ	0.656 ± 0.001273
ZZZ	$0.0007322 \pm 7.52 \times 10^{-7}$
ZW^+W^-	$0.01484 \pm 2.226 \times 10^{-5}$
$SingleTop$	$0.007934 \pm 9.82610^{-6}$

The single top is still considered to be negligible here due to its small cross section(signal cross section ~ 0.087[pb]).

Summary

- ***Determination of the top quark mass for the semileptonic channel near top pair production at the CEPC using Full and Fast Simulations.***

- ***Background study conducted using cut-based analysis and kinematic fitting.***

- ***Validation of the differential distribution (cross section) is still under investigation.***

- ***Investigate other machine learning techniques to further suppress single top production, i.e., Particle Transformer and BDT.***

Thanks for your attention!

mustapha.biyabi@cern.ch