11th CLHCP2025, Xinxiang, Oct 29-Nov 2, 2025

河南大學
Henan University

Prospects for a 95 GeV Higgs Boson at Future Higgs Factories with Transformer Networks

Yabo Dong（董亚博）

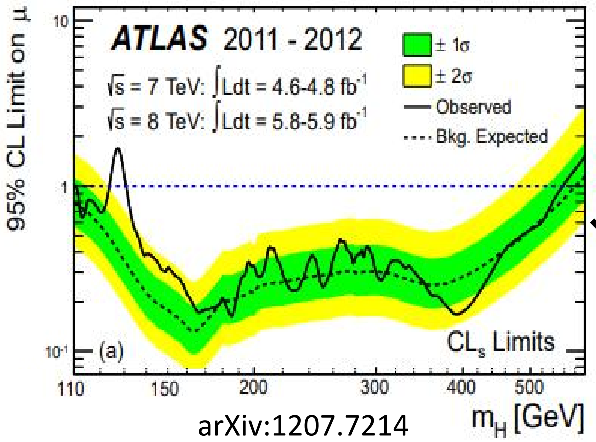
Based on arXiv: 2510.24662

on behalf of Kun Wang, Haijun Yang and Jingya Zhu

Outline

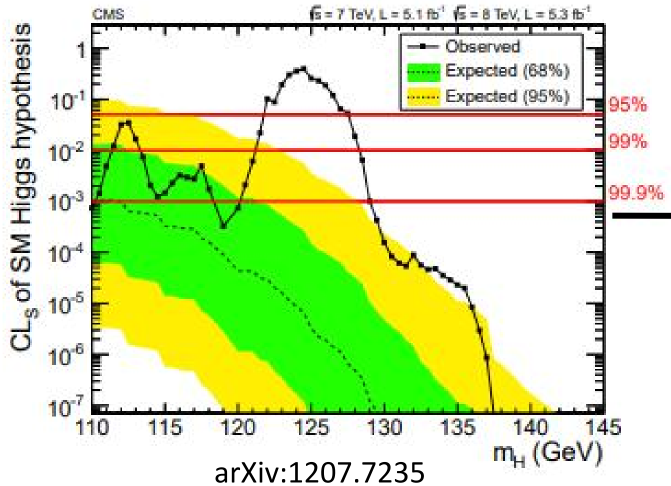
Background and motivation

The puzzle of SM



arXiv:1207.7214

add the final piece

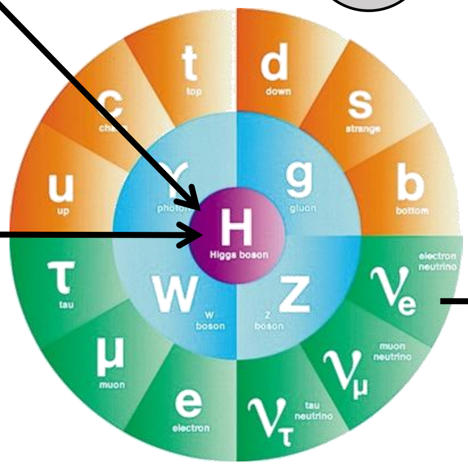


arXiv:1207.7235

Dark matter

Baryon asymmetry

Neutrino mass



extend

Additional Higgs

additional C and CP Violation

First-order electroweak phase transition

solve

Shaposhnikov conditions:

1. Baryon Number Violation

2. strong C and CP Violation

3. Departure from Thermal Equilibrium

Background and motivation

Recent excesses

a. 95.3 GeV dipohoton channel 2.8σ CMS 2018.
arXiv: 1811.08459.

b. 95.4 GeV dipohoton channel 2.9σ CMS 2024.
arXiv: 2407.07546.

c. 95.4 GeV dipohoton channel 1.7σ ATLAS 2024.
arXiv: 2405.18149.

d. 100 GeV ditau channel 3.1σ CMS 2022.
arXiv: 2208.02717

Signal strength

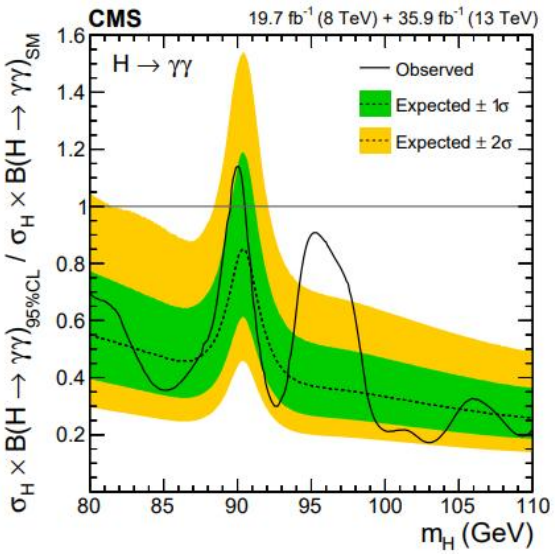
$$\mu_{\tau\tau}^{exp} = 1.2 \pm 0.5$$

$$\mu_{\gamma\gamma}^{exp} = 0.24^{+0.09}_{-0.08}$$

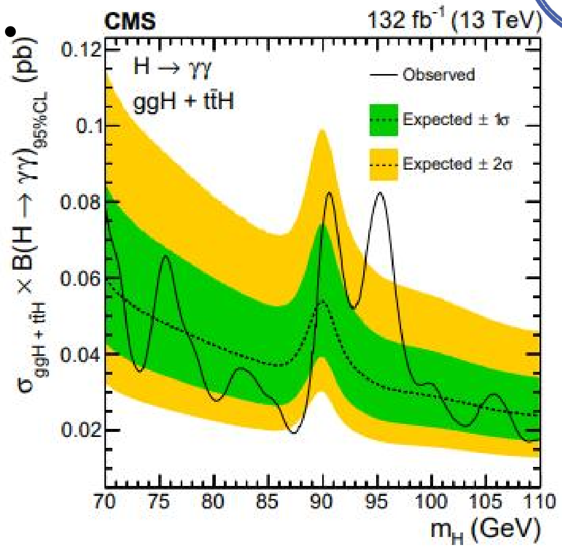
$$\mu = \frac{\sigma_{\text{exp}}}{\sigma_{\text{SM}}}$$

95 GeV Higgs ← favor

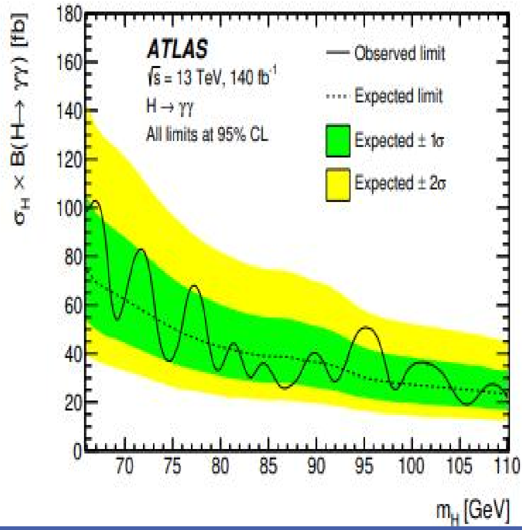
a.



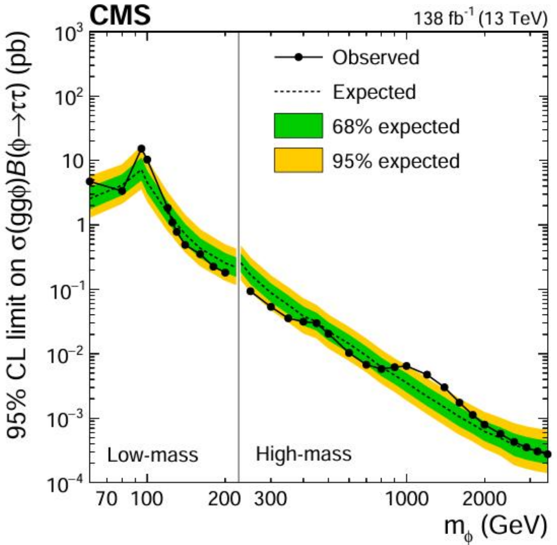
b.



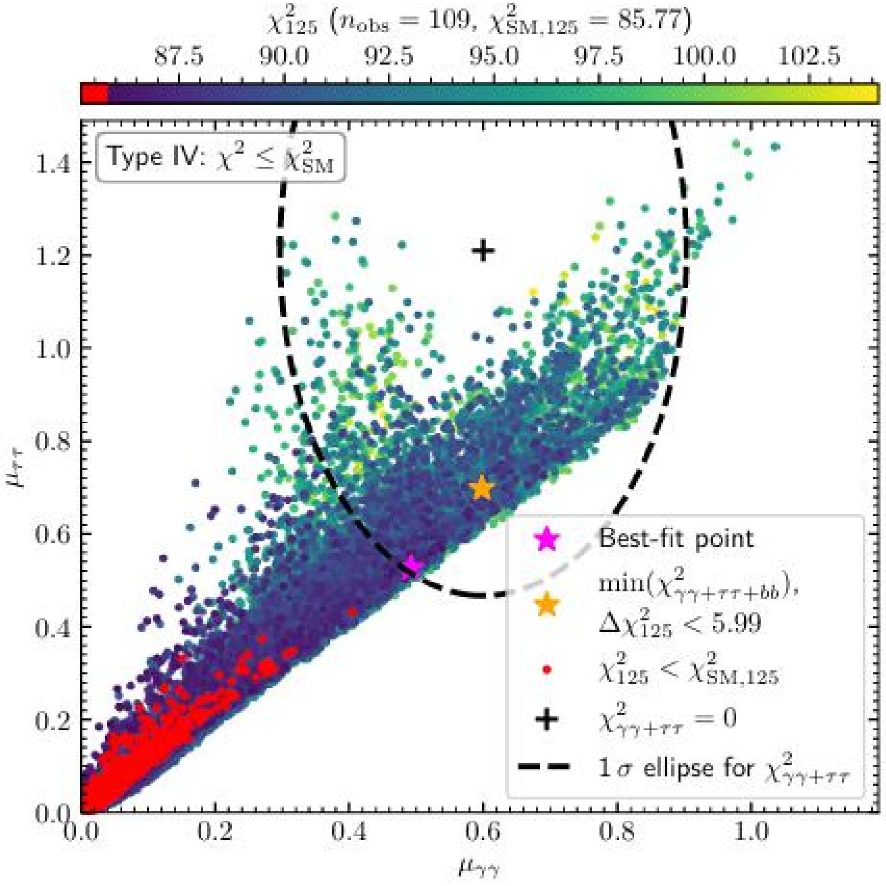
c.



d.

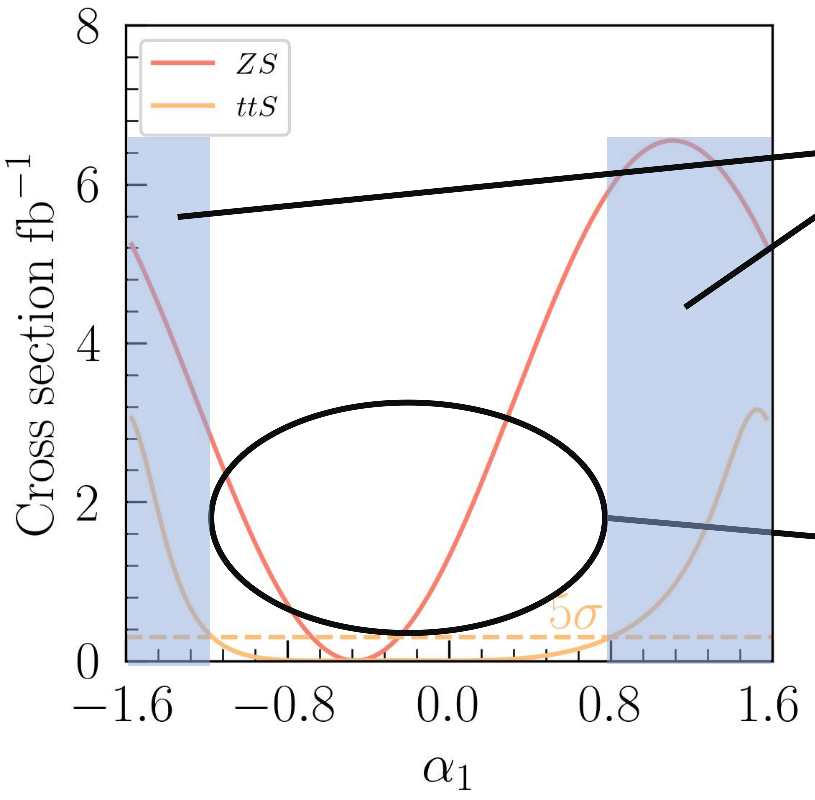


N2HDM-F and Collider phenomenology



Surviving samples of flipped N2HDM[1]

ttS > 0.3 fb can be covered at 5 σ on HL-LHC[2]



$\tan\beta = 2, \cos\alpha_2 = \sqrt{2}/2$

HL-LHC (L=3 ab^{-1})

ttS:
$pp \to t\bar{t}S(\to \gamma\gamma)$

CEPC (L=20 ab^{-1})

ZS:
$ee \to Z(\to \mu\mu)S$

1. arXiv: 2203.13180
2. arXiv: 2410.13636

The N2HDM

The Scalar potential

Doublet part

$$V_{\mathrm{N2HDM}} = \boxed{m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - m_{12}^2(\Phi_1^\dagger \Phi_2 + h.c.) + \frac{\lambda_1}{2}(\Phi_1^\dagger \Phi_1)^2 + \frac{\lambda_2}{2}(\Phi_2^\dagger \Phi_2)^2 + \lambda_3 \Phi_1^\dagger \Phi_1 \Phi_2^\dagger \Phi_2 + \lambda_4 \Phi_1^\dagger \Phi_2 \Phi_2^\dagger \Phi_1 + \frac{\lambda_5}{2}[(\Phi_1^\dagger \Phi_2)^2 + h.c.]}$$

$$\boxed{+\frac{m_S^2}{2}\Phi_S^2 + \frac{\lambda_6}{8}\Phi_S^4 + \frac{\lambda_7}{2}\Phi_1^\dagger \Phi_1 \Phi_S^2 + \frac{\lambda_8}{2}\Phi_2^\dagger \Phi_2 \Phi_S^2}$$

Singlet part

6 scalars H_a, H_b, H_c CP even Higgs; A pseudoscalar; $H^\pm$ charged Higgs

12 parameters

$$m_{H_1}, m_{H_2}, m_{H_3}, m_A, m_{H^\pm},$$
$$sign(R_{23}), R_{12}, tan\beta,$$
$$C_{H_2tt}^2,\ C_{H_2VV}^2$$
$$m_{12}^2, v_s$$

effective couplings and mixing matrix elements representation

The Coupling

Higgs to massive gauge bosons

$$C_{H_iVV} = \begin{cases} c_{\alpha_2} c_{\beta-\alpha_1}, & i = 1, \\ -c_{\beta-\alpha_1} s_{\alpha_2} s_{\alpha_3} + c_{\alpha_3} s_{\beta-\alpha_1}, & i = 2, \\ -c_{\beta-\alpha_1} s_{\alpha_2} c_{\alpha_3} - s_{\alpha_3} s_{\beta-\alpha_1}, & i = 3. \end{cases}$$

Higgs to up type quarks and leptons

$$C_{H_iff} = \begin{cases} s_{\alpha_1} c_{\alpha_2}/s_\beta, & i = 1, \\ (c_{\alpha_1} c_{\alpha_3} - s_{\alpha_1} s_{\alpha_2} s_{\alpha_3})/s_\beta, & i = 2, \\ -(c_{\alpha_1} s_{\alpha_3} + s_{\alpha_1} s_{\alpha_3} c_{\alpha_2})/s_\beta, & i = 3. \end{cases}$$

Higgs to down type quarks

$$C_{H_iq_dq_d} = \begin{cases} c_{\alpha_1} c_{\alpha_2}/c_\beta, & i = 1, \\ -(c_{\alpha_1} s_{\alpha_2} s_{\alpha_3} + c_{\alpha_3} s_{\alpha_1})/c_\beta, & i = 2, \\ -(c_{\alpha_1} s_{\alpha_2} c_{\alpha_3} - s_{\alpha_1} s_{\alpha_3})/c_\beta, & i = 3. \end{cases}$$

$c_x = \cos(x)$ and $s_x = \sin(x)$

N2HDM-Flipped

Parameter space scan

The selection of parameters

mass of excess candidate H_1	$95 < m_{H_1} < 96$ GeV
mass of SM-like H_2	$m_{H_2} = 125.09$ GeV
mass of CP-even H_3, pseudoscalar A	$300 < m_{H_3,A} < 1500$ GeV
mass of charged $H^\pm$	$580 < m_{H^\pm} < 1500$ GeV
coupling of H_2 to boson, top quark	$0.7 < C^2_{H_2VV} < 1.0,\ 0.7 < C^2_{H_2tt} < 1.2$
mixing matrix elements	$\mathrm{sign}(R_{23}) = \pm 1, -1 < R_{12} < 1$
VEV ratio of two doublets	$0.8 < \tan\beta < 10$
VEV of singlet	$1 < v_s < 3000$ GeV
Z_2 breaking parameter	$1 \times 10^{-3} < m^2_{12} < 5 \times 10^5\ \mathrm{GeV}^2$

The constraints

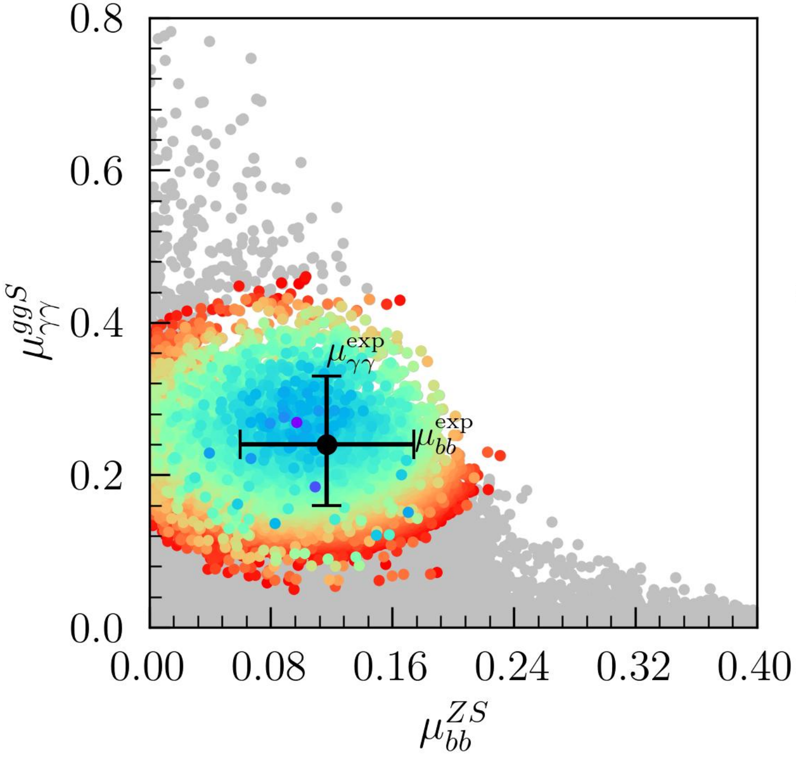
1. Perturbative unitarity.
2. Vacuum stability.
3. B physic.
4. Electroweak precision observables.
5. Direct search for Higgs.
6. SM-like Higgs Measurements.

N2HDM-Flipped

Surviving samples

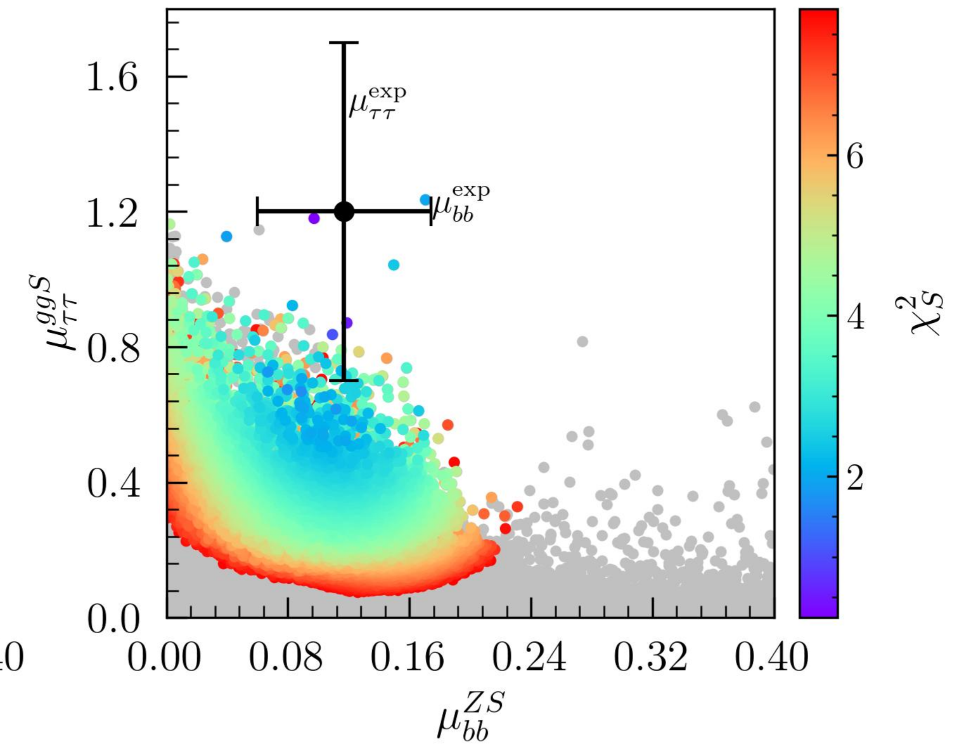
Gluon fusion process

$$\mu_{xx}^{ggS} = \frac{\sigma(gg \to S) \times \mathcal{B}(S \to xx)}{\sigma_{SM}(gg \to H_{95})\mathcal{B}_{SM}(H_{95} \to xx)}$$



Higgsstrahlung process

$$\mu_{xx}^{ZS} = \frac{\sigma(ee \to ZS) \times \mathcal{B}(S \to xx)}{\sigma_{SM}(ee \to ZH_{95})\mathcal{B}_{SM}(H_{95} \to xx)}$$



Three channel excesses

$$\mu_{bb}^{exp} = 0.117 \pm 0.057$$

$$\mu_{\tau\tau}^{exp} = 1.2 \pm 0.5$$

$$\mu_{\gamma\gamma}^{exp} = 0.24_{-0.08}^{+0.09}$$

Fit value

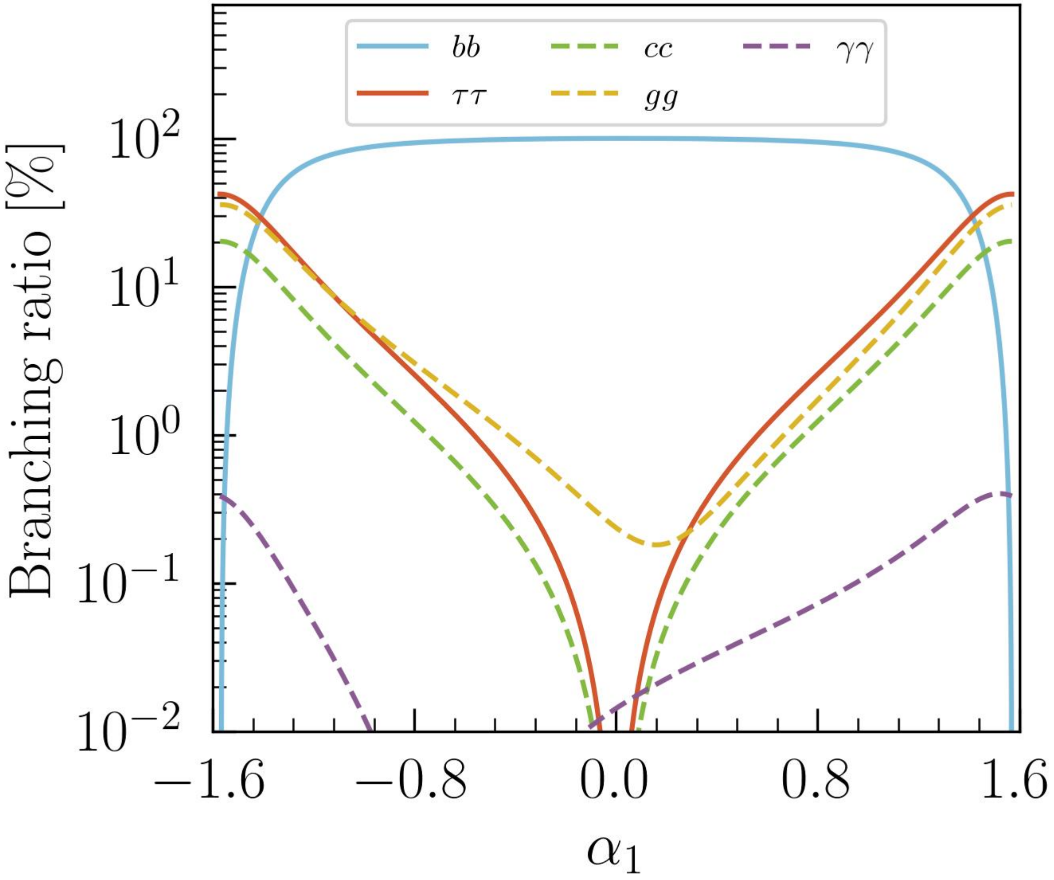
$$\chi_S^2 = \sum_{xx=\tau\tau,\gamma\gamma,bb} \frac{\left(\mu_{xx} - \mu_{xx}^{exp}\right)^2}{\left(\Delta\mu_{xx}^{exp}\right)^2}$$

< 7.82 at 95 C. L.

MC Simulation

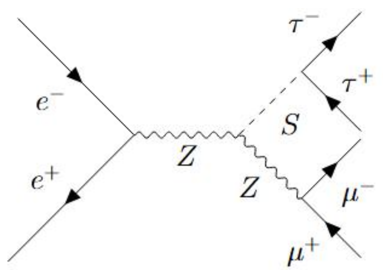
Signal and background

dominant decay channels

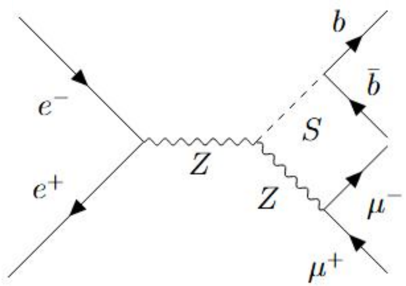


$tan\beta = 2,\ cos\alpha_2 = \sqrt{2}/2$

Signal

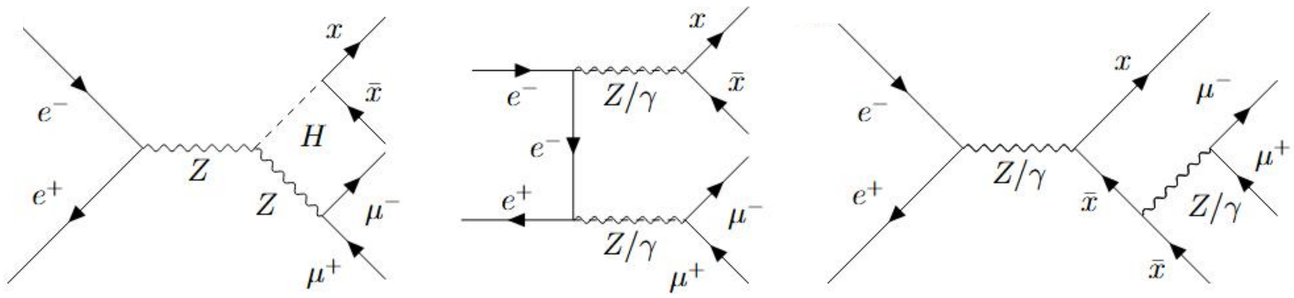


$e^+e^- \to Z(\to \mu^+\mu^-)S(\to \tau^+\tau^-)$



$e^+e^- \to Z(\to \mu^+\mu^-)S(\to bb)$

Background



$e^+e^- \to \mu^+\mu^- xx,\ x = u, d, c, s, e, \mu, \tau, b$

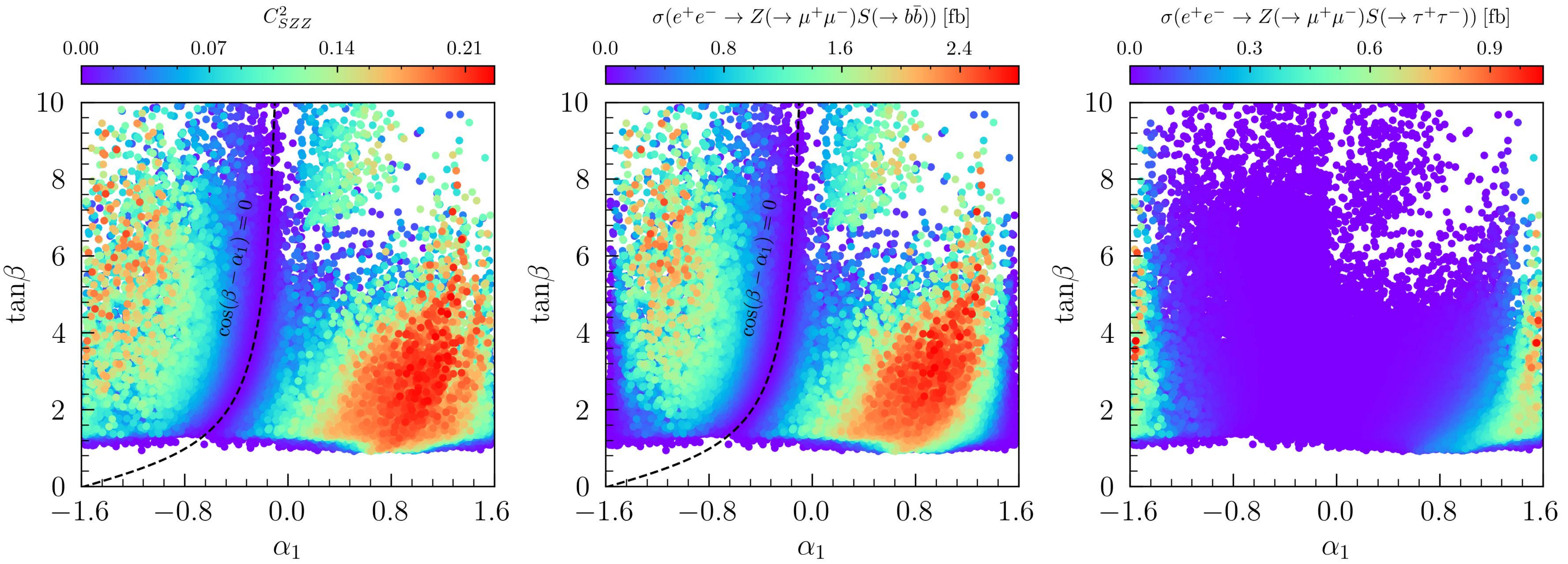
MC Simulation

Cross section in two channels

Madgraph5_aMC@NLO   SM value

Effective coupling

$$\sigma = \sigma_{SM}(ee \to ZS)|_{m_H=95.5\ \mathrm{GeV}} \cdot c_{SZZ}^2 \cdot \mathrm{Br}(Z \to \mu\mu) \cdot \mathrm{Br}(S \to xx)$$



$C_{H_1VV} = cos\alpha_2 \cdot cos(\beta - \alpha_1)$ $C_{H_1bb} = cos\alpha_1 \cdot cos\alpha_2 / cos\beta$ $C_{H_1\tau\tau} = sin\alpha_1 \cdot cos\alpha_2 / sin\beta$

MC Simulation

Event generation and pre cuts

generate 1 million signal and background

Event generation

Particle generate	Particle decays	Detector effect
MadGraph5_aMC@NLO_v3.6.2 →	PYTHIA_v8.2 →	Delphes_v3.5.0

CEPC 4th card

Pre selections

Basic cuts

$N(\mu) \geq 2,\ p_T^{\mu} \geq 10\ \text{GeV}$

$\eta(\mu) < 2, \Delta R(\mu_1, \mu_2) > 0.4$

$M_{\mu\mu}$ cut

$M(\mu_1, \mu_2) > 50\ \text{GeV}$

calculate

$$M_{\text{recoil}} = \sqrt{s - M^2{}_{\mu^+\mu^-} - 2\sqrt{s}\left(E_{\mu^+} + E_{\mu^-}\right)}$$

MC Simulation

Cut based analysis on $\tau\tau$ channel

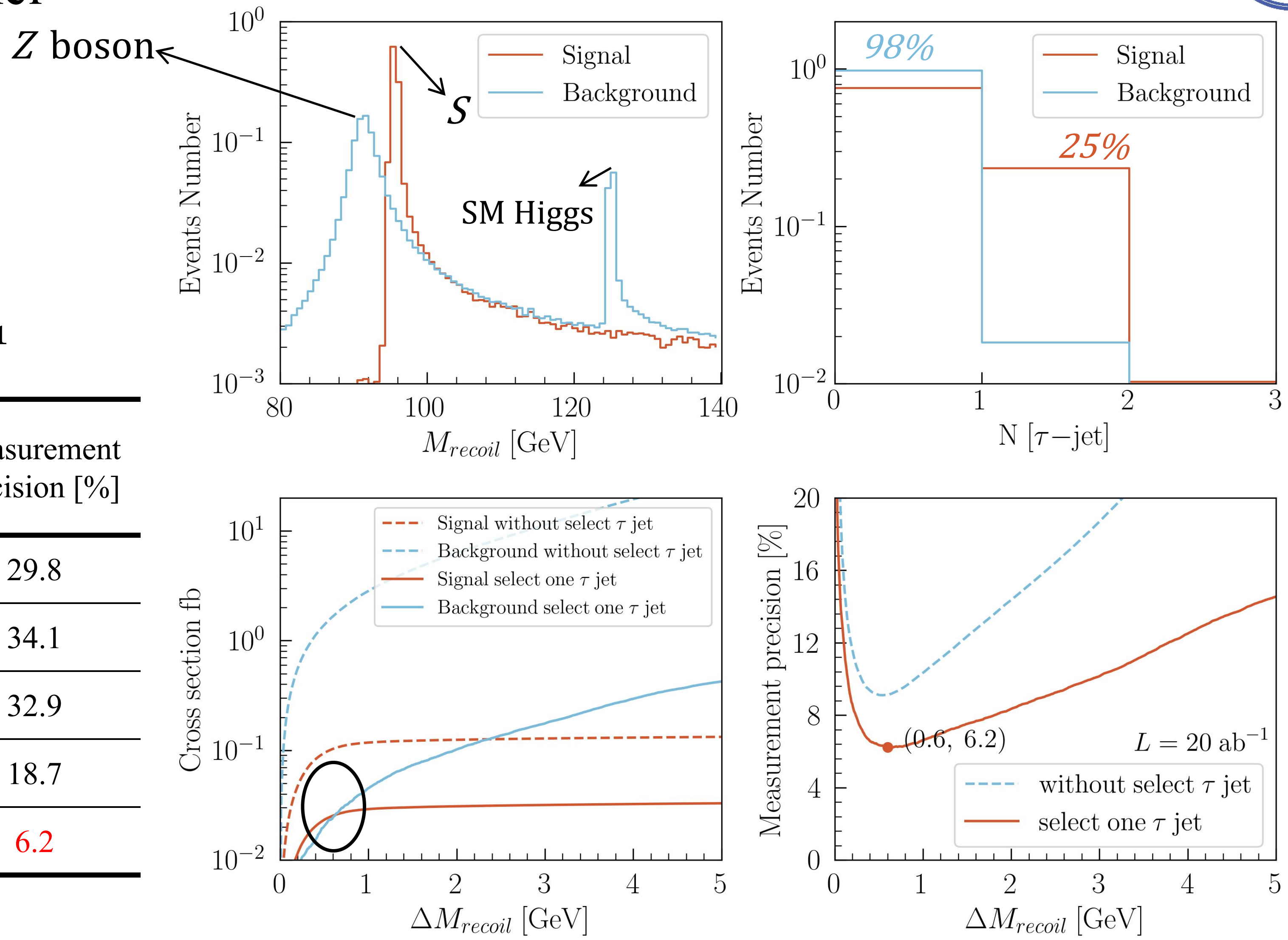
Measurement precision

$$\Delta = \frac{\sqrt{S+B}}{S}$$

The cut flow with $L = 20\ \mathrm{ab}^{-1}$

Cuts	Cross section [fb]		Measurement precision [%]
	Signal	Background	
Initail	0.218	84.250	29.8
Basic	0.177	72.202	34.1
$M_{\mu_1\mu_2} > 50$ GeV	0.165	58.617	32.9
N[τ-jet] ≥ 1	0.040	1.094	18.7
$\Delta M_{recoil} = 0.6$ GeV	0.026	0.026	6.2

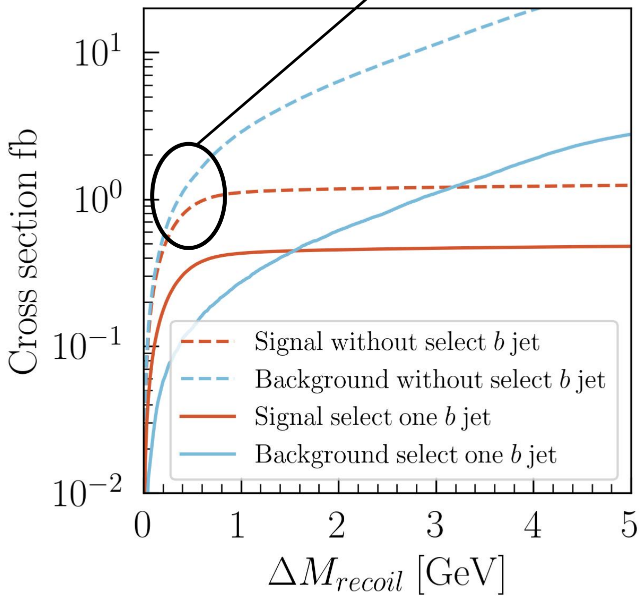
xsec is set to 0.2 SM value for signal

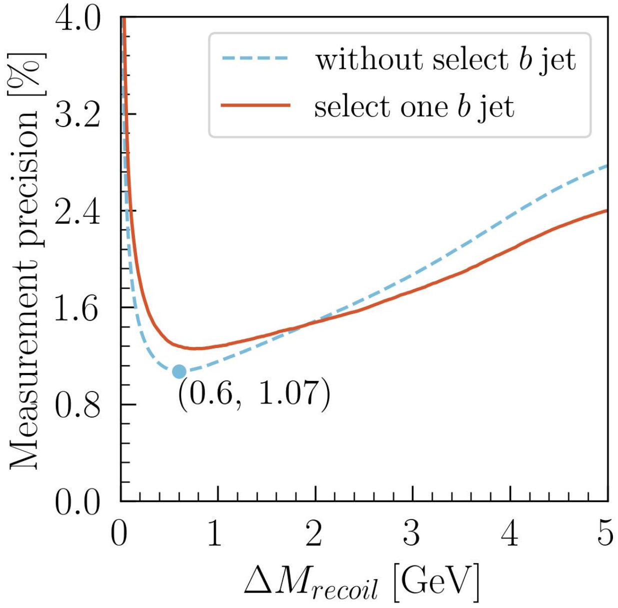


MC Simulation

Cut based analysis on bb channel

The cross section can comparable





The cut flow with $L = 20\ \text{ab}^{-1}$

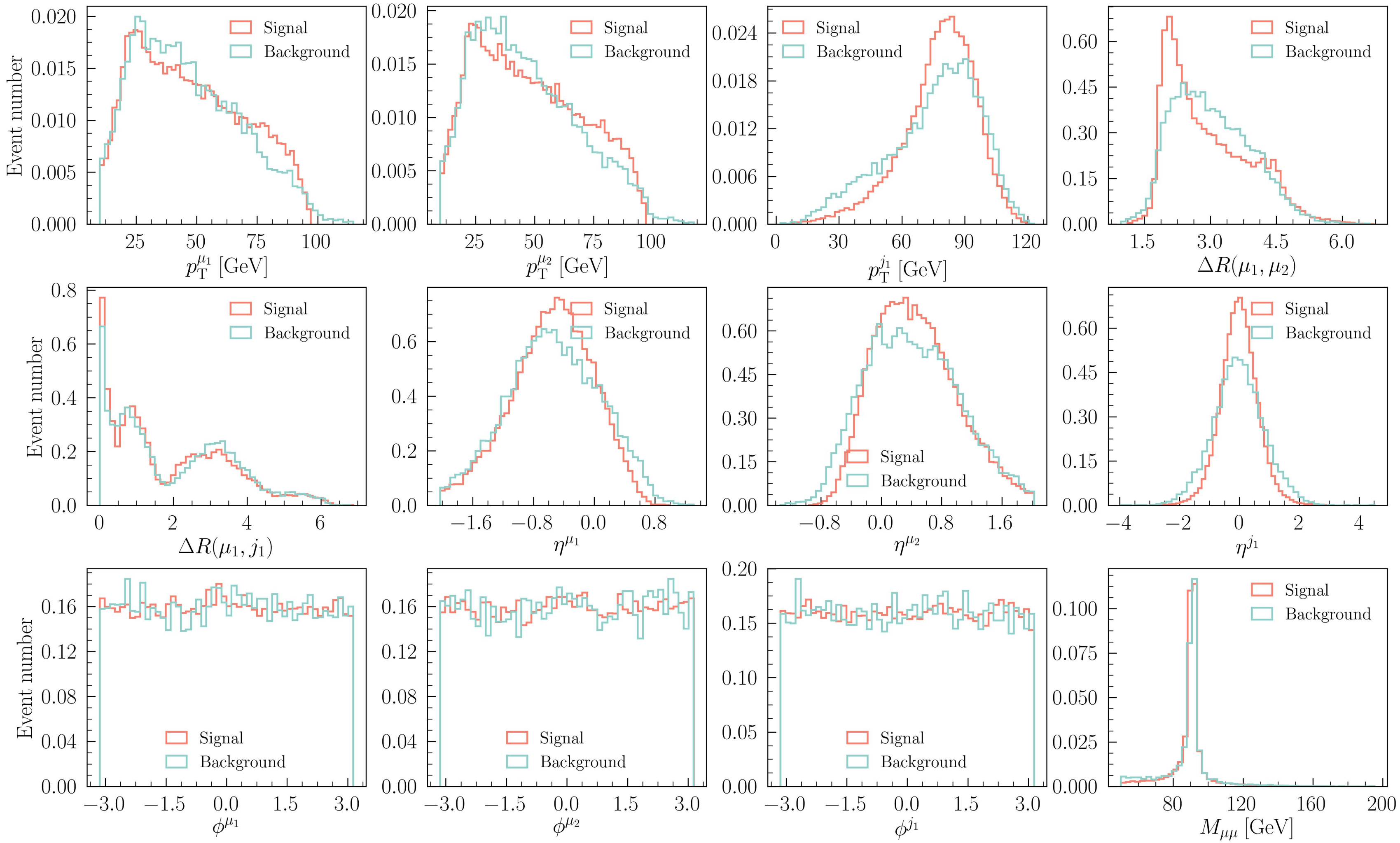
Cuts	Cross section [fb]		Measurement precision [%]
	Signal	Background	
Initail	2.099	84.250	3.13
Basic	1.700	72.244	3.58
$M_{\mu_1\mu_2} > 50$ GeV	1.633	58.552	3.36
$\Delta M_{recoil} = 0.6$ GeV	1.094	1.633	1.07

How to improve precision？

xsec is set to 0.2 SM value for signal

MC Simulation

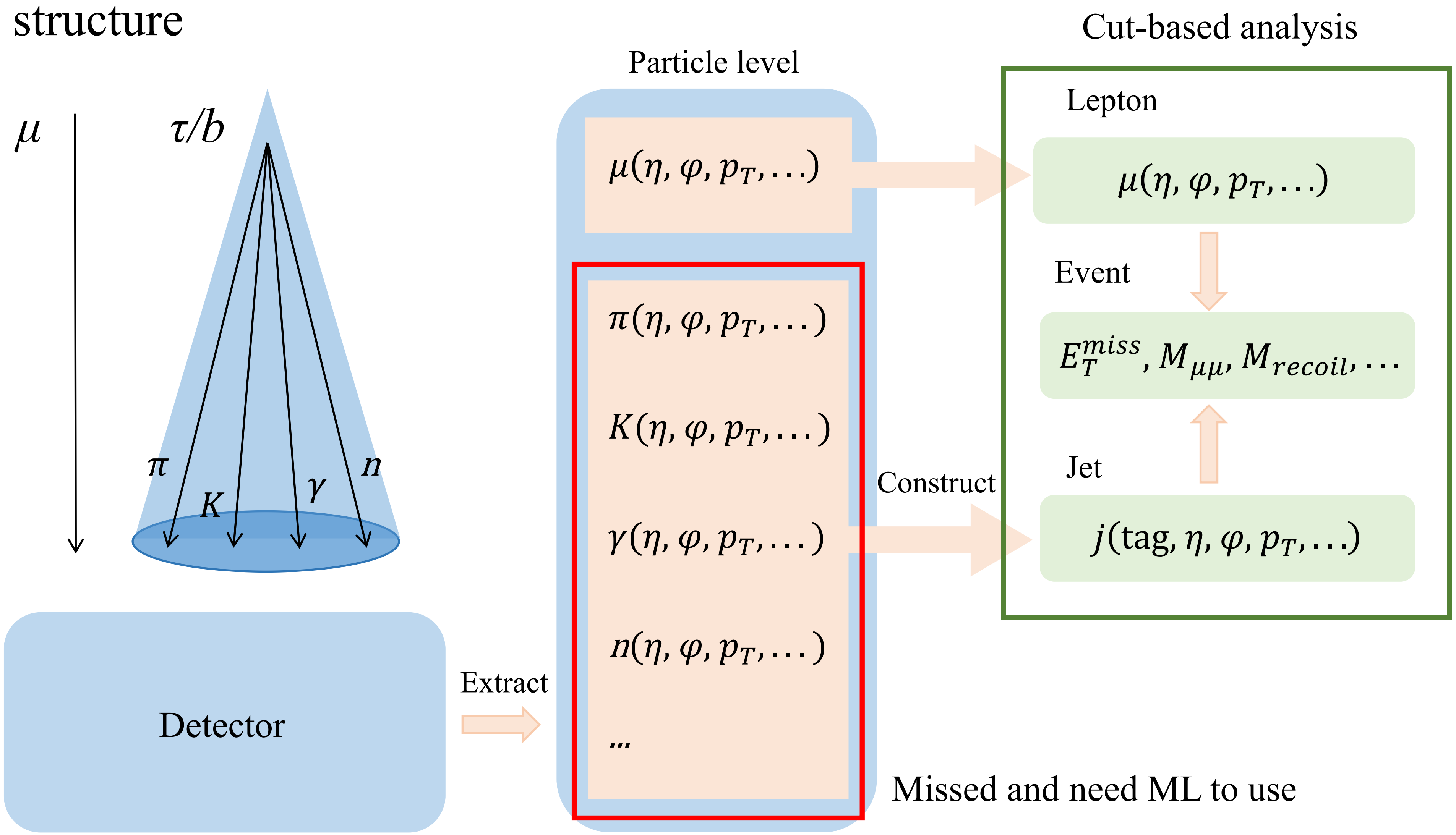
Event distributions for other observables



The precision can hardly significance be improved by these observables!

The Application of ML

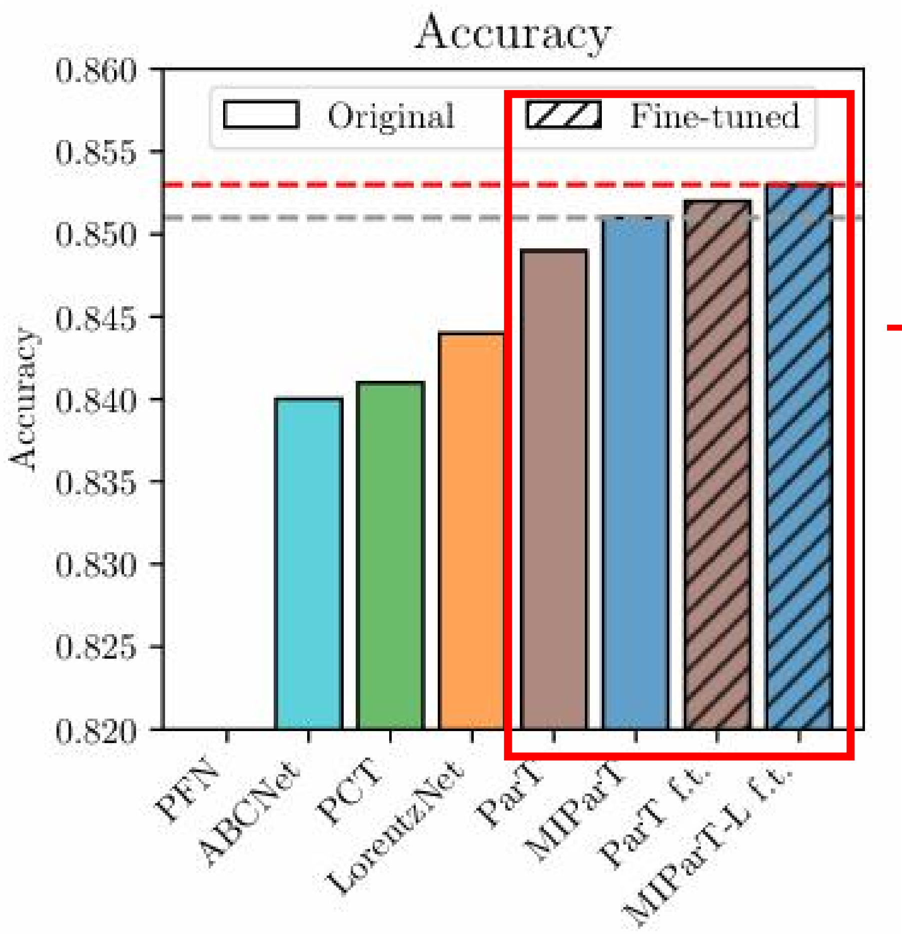
Feature structure



ML performance in jet tagging

Advanced ParT[1] and MIParT[2] + particle level feature

Performance in QuarkGluon tagging



Detailed performance

	ACC	AUC
ParT	0.849	0.9203
MIPaT	0.851	0.9215
ParT.f.t	0.852	0.9230
MIParT.f.t	0.853	0.9237

1. arXiv: 2202.03772
2. arXiv: 2407.08682

The Application of ML

ML in $\tau\tau$ channel

Input features

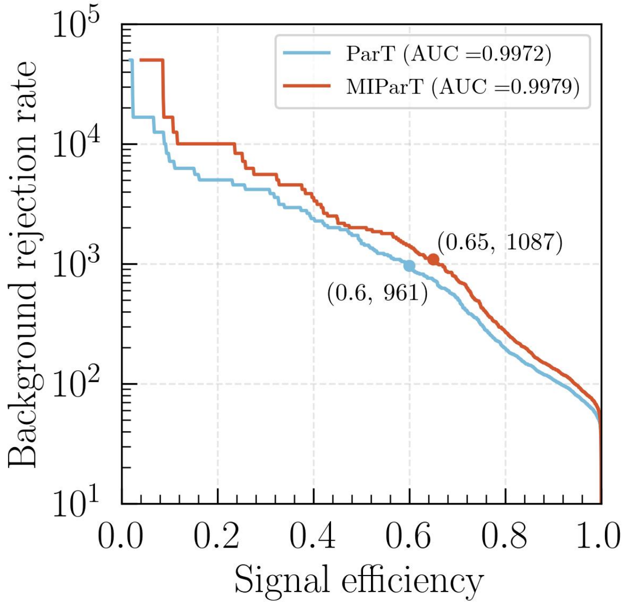
Event level:

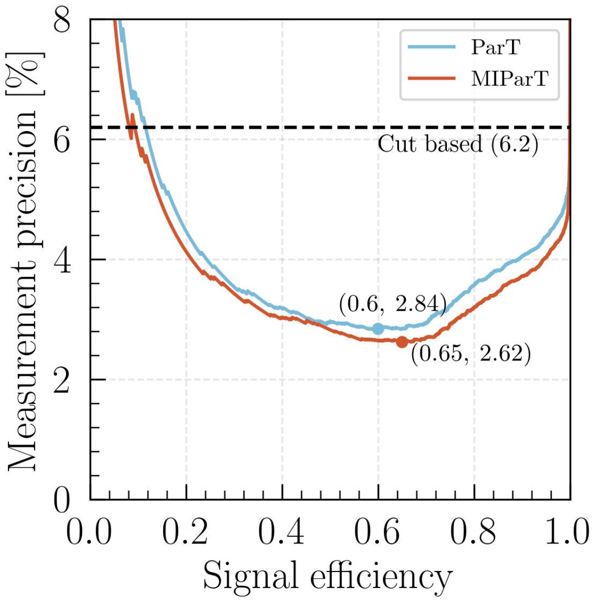
- Recoil mass: M_{recoil}
- Invariant mass: $M_{\mu\mu}$
- Average azimuthal angle: $\varphi = (\varphi_{\mu_1} + \varphi_{\mu_2})/2$
- Average pseudorapidity: $\eta = (\eta_{\mu_1} + \eta_{\mu_2})/2$

Particle level (max length=50):

- Transverse momentum: P_T^p
- Eneragy: E^P
- Renormalized P_T: $P_T^p/(P_T^{\mu_1} + P_T^{\mu_2})$
- Renormalized E: $E^P/(E^{\mu_1} + E^{\mu_2})$
- Regularization φ: $\Delta\varphi = \varphi^p - \varphi$
- Regularization η:$\Delta\eta = \eta^p - \eta$
- Angular separartion: $\sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}$
- Charge: C
- Partilce kind.

The ML performance





Performance comparison

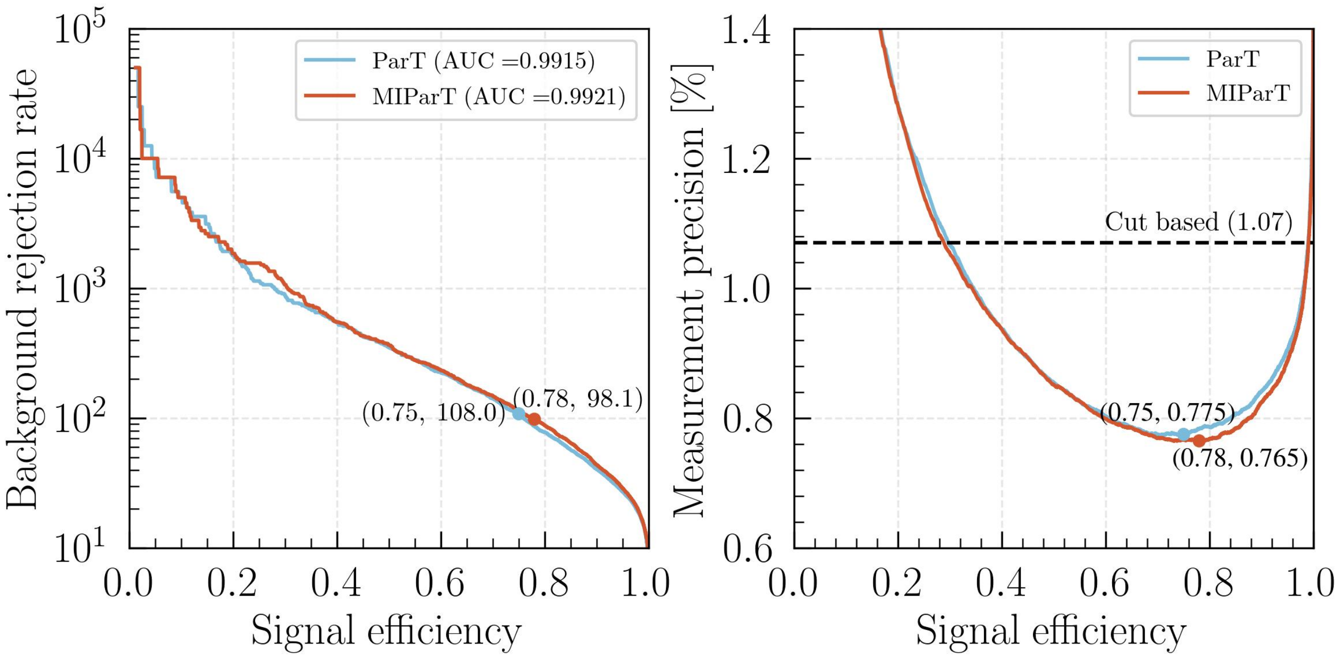
Method	Signal efficiency	background refection rate	Measurement precision [%]
Cut based	0.15	2777	6.2
ParT	0.60	916	2.8
MIParT	0.65	1087	2.6

↑ 58%

The Application of ML

ML in *bb* channel

The ML performance



Performance comparison

Method	Signal efficiency	background refection rate	Measurement precision [%]
Cut based	0.64	43.3	1.07
ParT	0.75	108	0.78
MIParT	0.78	98.1	0.77

↑ 28%

The Application of ML

CEPC detect range for surviving samples

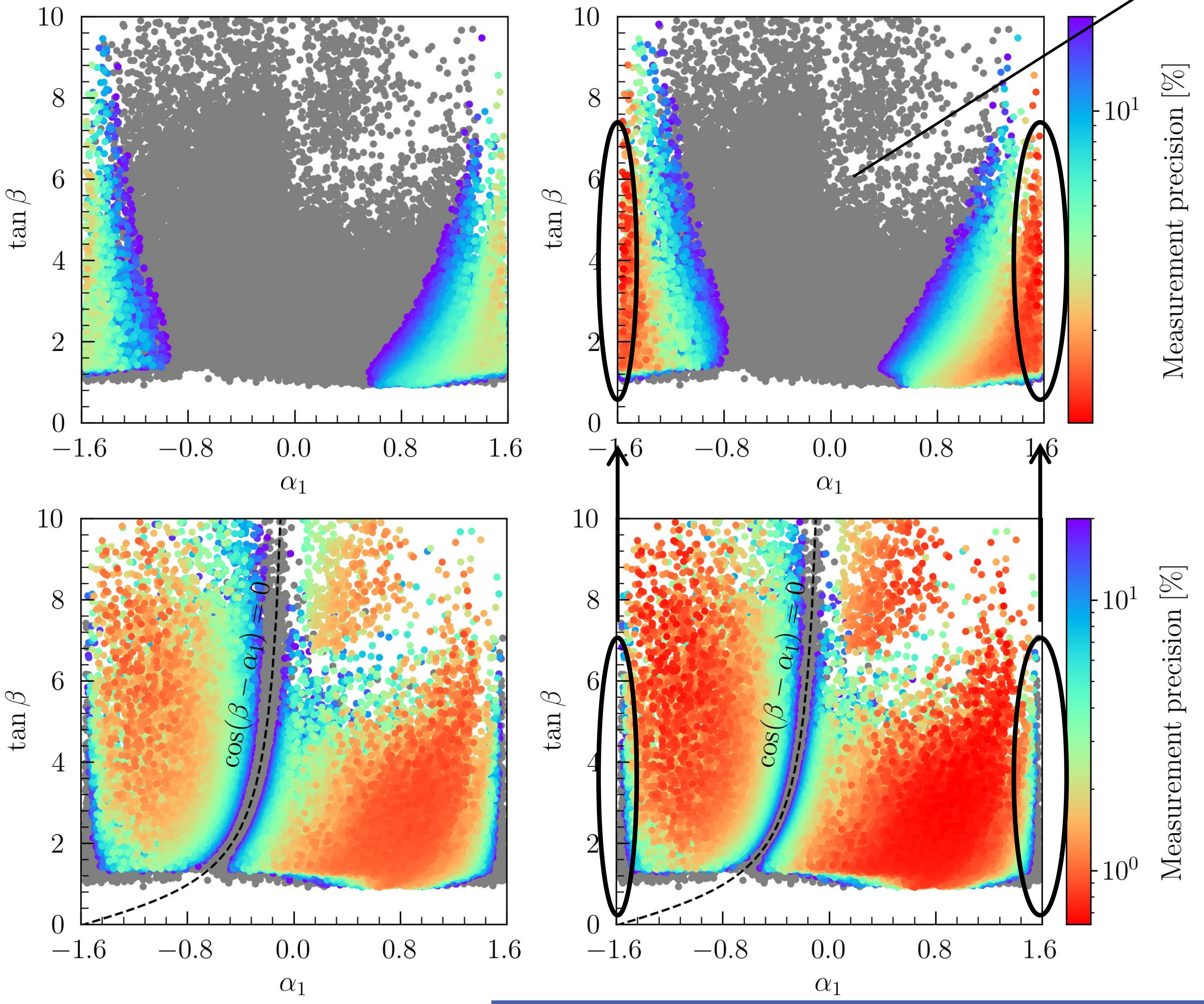
without ML

with MIParT

$< 5\sigma$

$\tau\tau$ channel

bb channel



Signal significance

$$\mathcal{S} = \sqrt{2\left[(S+B)ln\left(1+\frac{S}{B}\right)-S\right]}$$

Measurement precision

$$\Delta = \frac{\sqrt{S+B}}{S}$$

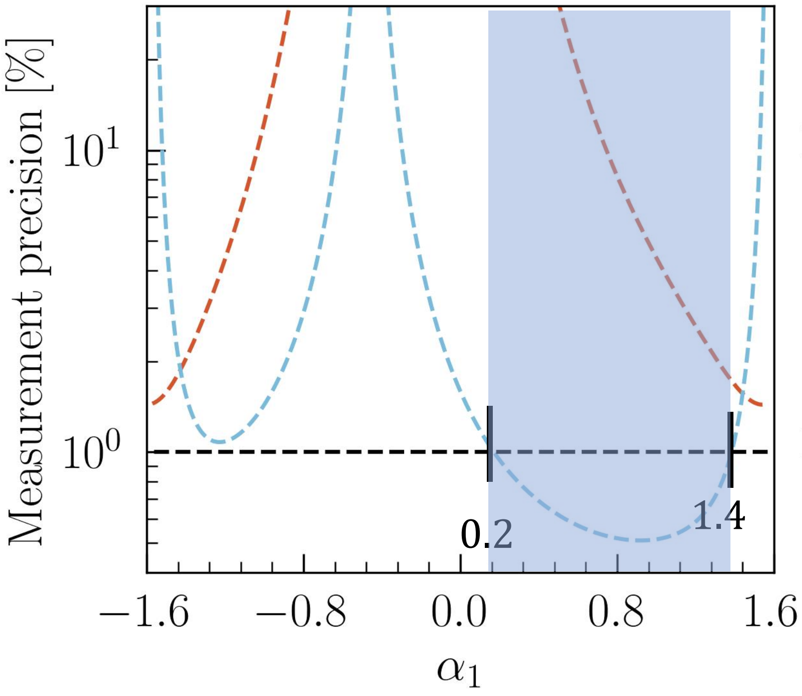
For the surviving samlples

Channel	Max cross section [fb]	Max precision [%]	
		Without ML	MIParT
$\tau\tau$	1.03	2.3	1.0
bb	2.8	0.86	0.63

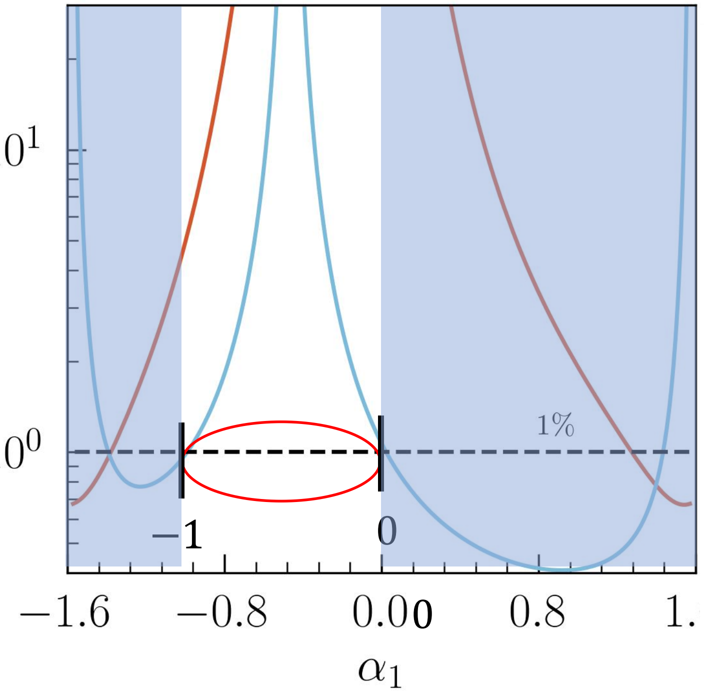
CEPC detect range for benchmark

High-precision detection range has been significantly increased

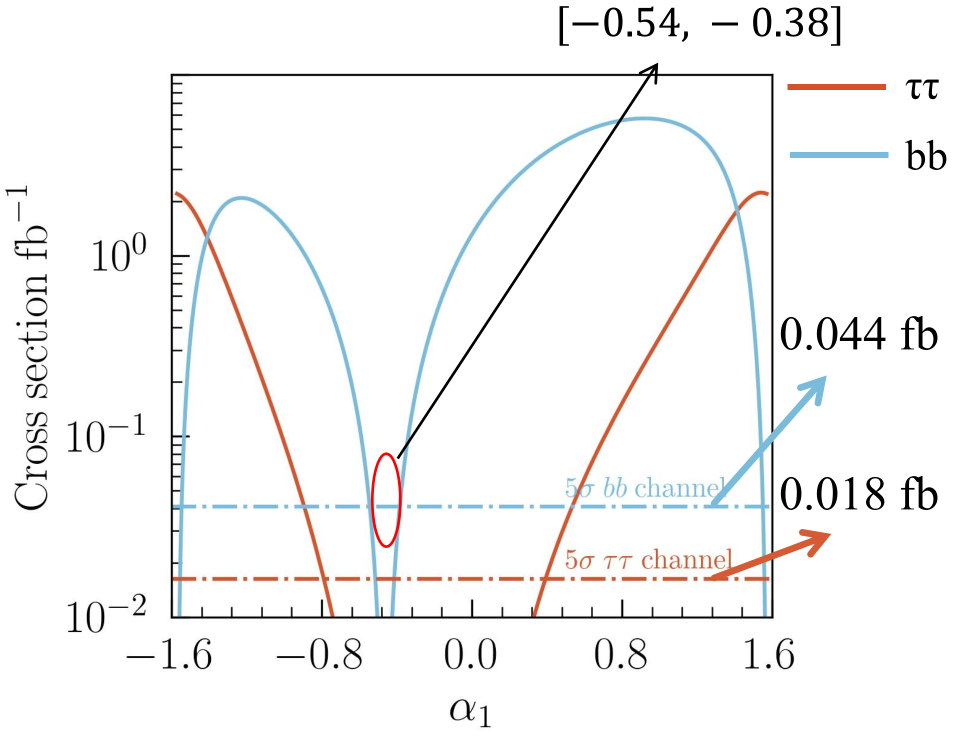
Cut based



MIParT



MIParT

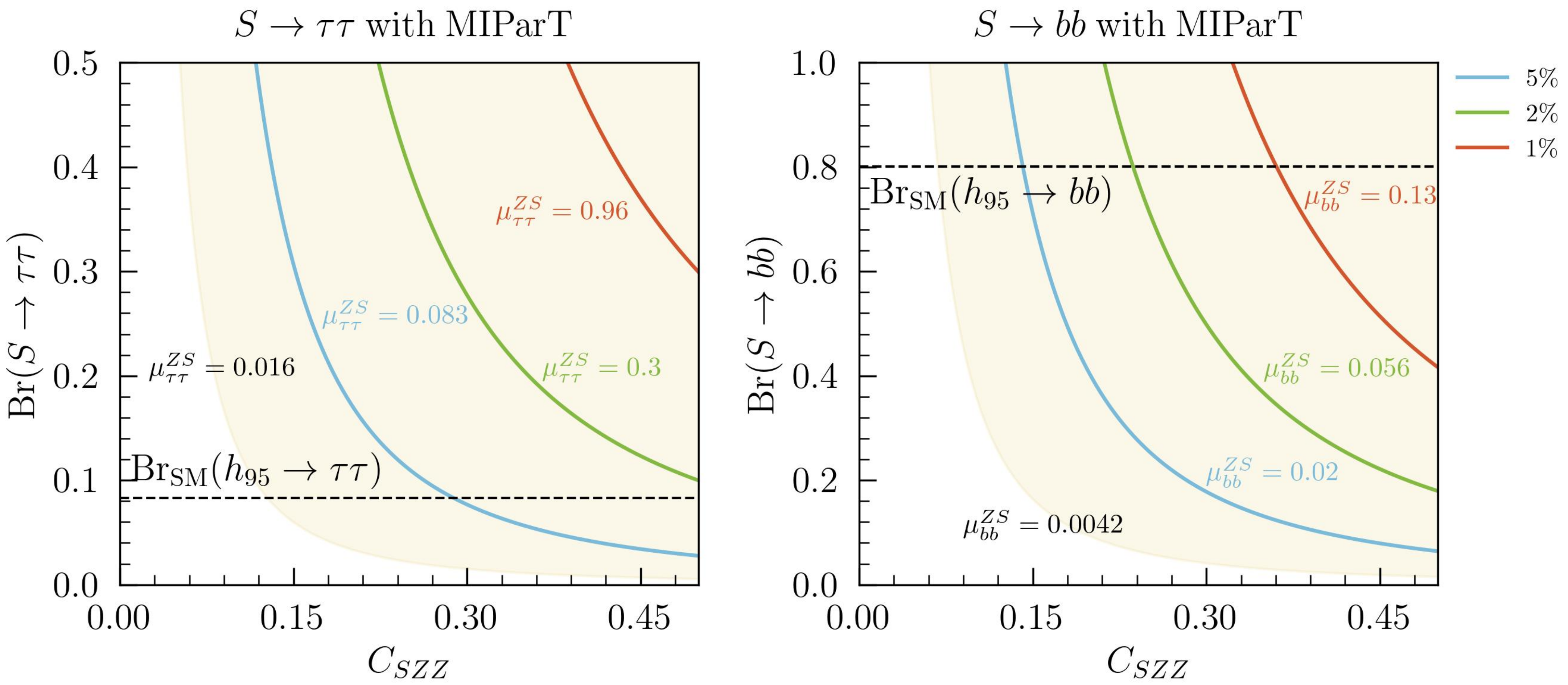


$\tan\beta = 2, \cos\alpha_2 = \sqrt{2}/2$, and $\cos(\beta - \alpha_1) = 0$ for $\alpha_1 =- 0.46$

The Application of ML

CEPC detect range in signal strength

$$\mu_{xx}^{ZS} = C_{SZZ}^2 \cdot \frac{Br(S \to xx)}{Br_{SM}(H_{95} \to xx)}$$



Conclusion

MC result in $e^+e^- \to Z(\to \mu^+\mu^-)S$ process at CEPC with L=20 ab^{-1} shows:

Particle level features and MIParT network can improve measurement precision

- $S\to\tau\tau$: 6.2% → 2.6% (↑58%)
- $S\to bb$: 1.07% → 0.77% (↑28%)

CEPC detect range in our method

Channel	5σ significance covered	1% detcted precison
$S \to \tau\tau$	μ =0.016	μ =0.96
$S \to bb$	μ =0.0042	μ =0.13

Based on arXiv: 2510.24662

Thanks!

Thanks

N2HDM-Flipped

- Constraints from perturbative unitarity: The constraint require the eigenvalues $\mathcal{M}^i_{2\to 2}$ of the $2 \to 2$ scattering matrix $\mathcal{M}_{2\to 2}$ satisfy

$$|\mathcal{M}^i_{2\to 2}| < 8\pi \tag{8}$$

at tree level [39].

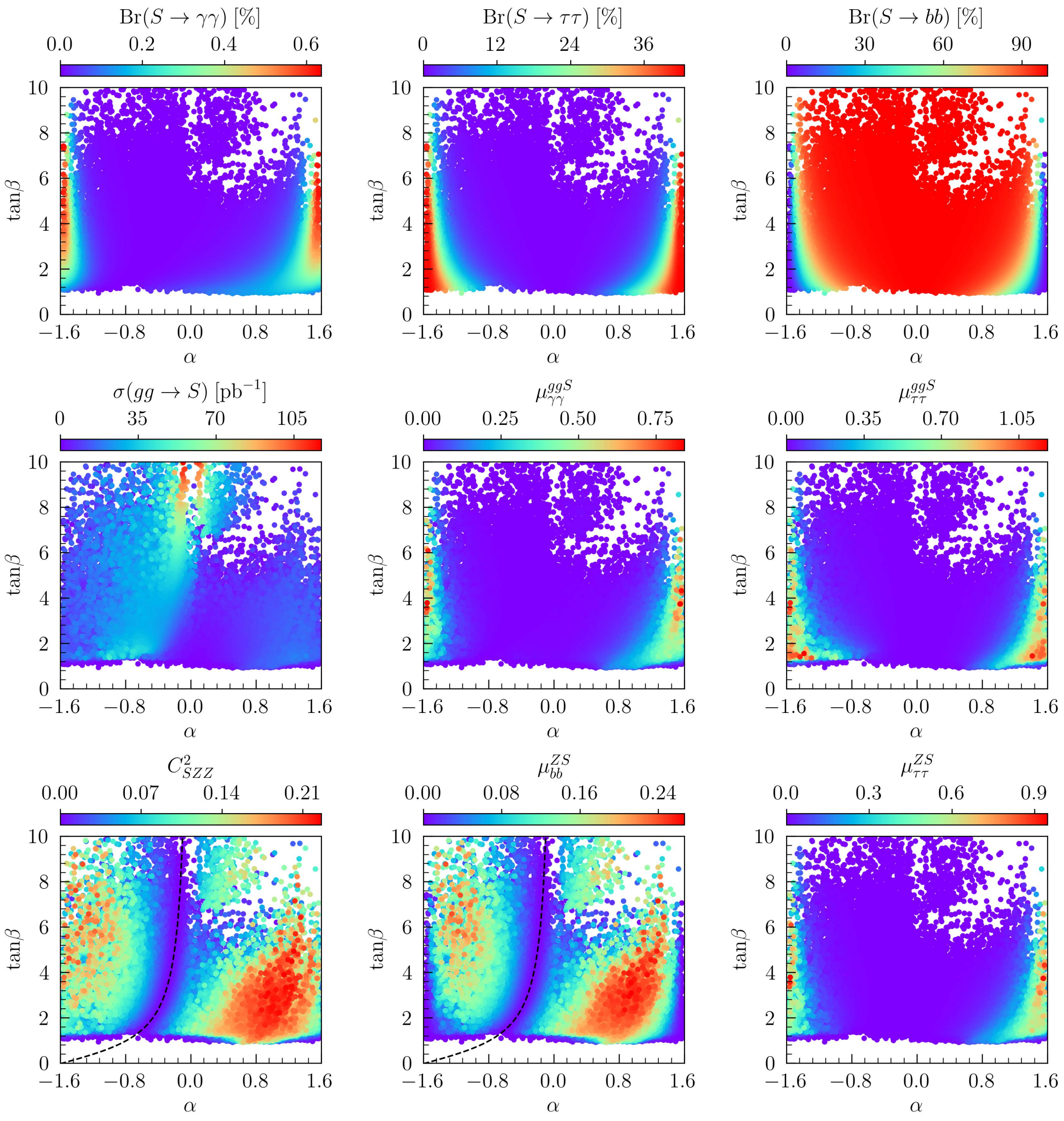
- Constraints from vacuum stability: The package of EVADE [40, 41] is used to test the stability at tree-level.

- Constraints from B physics and electroweak precision observables: The global fit result of Ref. [42] on oblique parameters and flavour constraints at 2σ level are used.

- Constraints from Higgs Searches and SM-like Higgs Measurements: For the additional scalars direct search constraints, the effective coupling approximation is exact to interface to the HiggsBounds-5.10.0 package [43] at the 95% confidence level using. For the SM-like Higgs Measurements, HiggsSignal-2.6.0 package [44] use 109 degrees of freedom data to calculate the χ of this model and. The SM χ^2 is 85.8 and

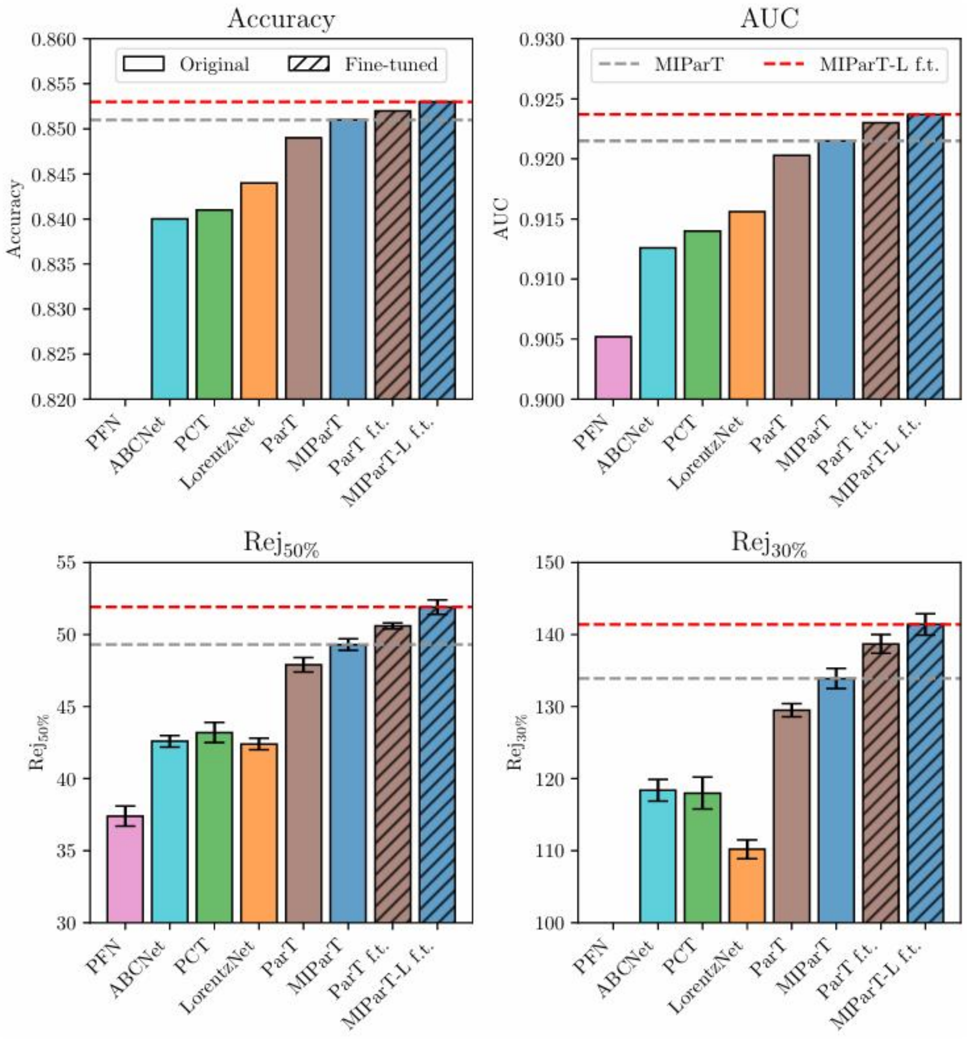
$$\Delta\chi^2 = \chi^2_{\rm N2HDM-F} - \chi^2_{\rm SM} < 6.18 \tag{9}$$

is required.



Background and motivation

ML in jet tagging



arXiv:2407.08682

Performance on QuarkGluon

	ACC	AUC
ParT	0.849	0.9203
MIPaT	0.851	0.9215
ParT.f.t	0.852	0.9230
MIParT.f.t	0.853	0.9237

ML can enhance the light Higgs search potential?

Advantages

1. Deal multiple features big data
2. Feature interaction extraction capability
3. Balance signal efficiency and background rejection rate
4. Automation, high efficiency

Disadvantages

Black box

The Application of ML

More features + advance ML = improve?

Our features

Event

$E_T^{miss}, M_{\mu\mu}, M_{recoil}, \ldots$

Lepton

$\mu(\eta, \varphi, p_T, \ldots)$

Jet

$j(\eta, \varphi, p_T, \text{tag}, \ldots)$

Cut-based features

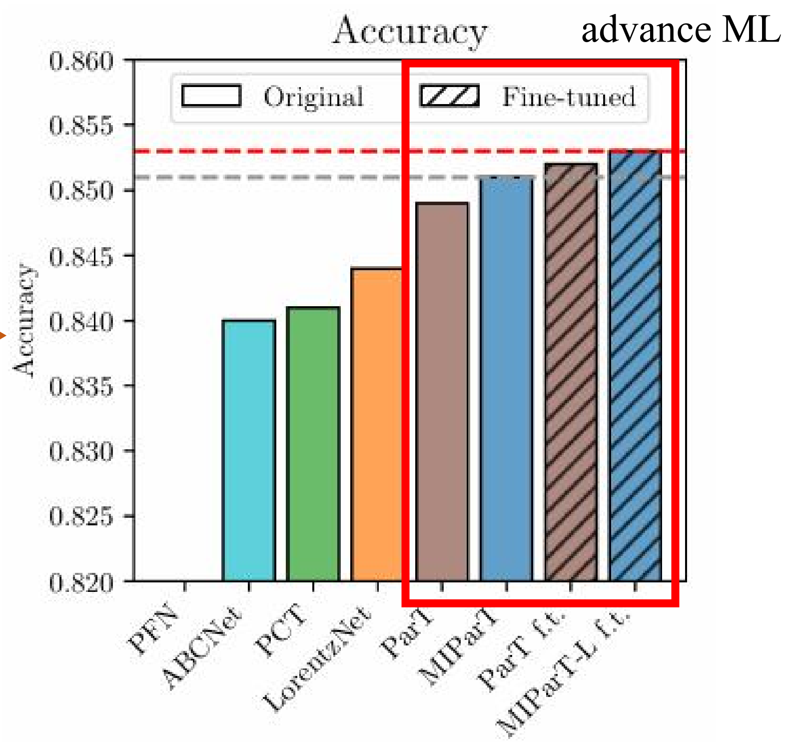
Particle

$\pi, K, \gamma, \ldots (\eta, \varphi, p_T, \ldots)$

Interaction

$\Delta R(p, j), E_\pi / E_j, \ldots\ldots$

ML features

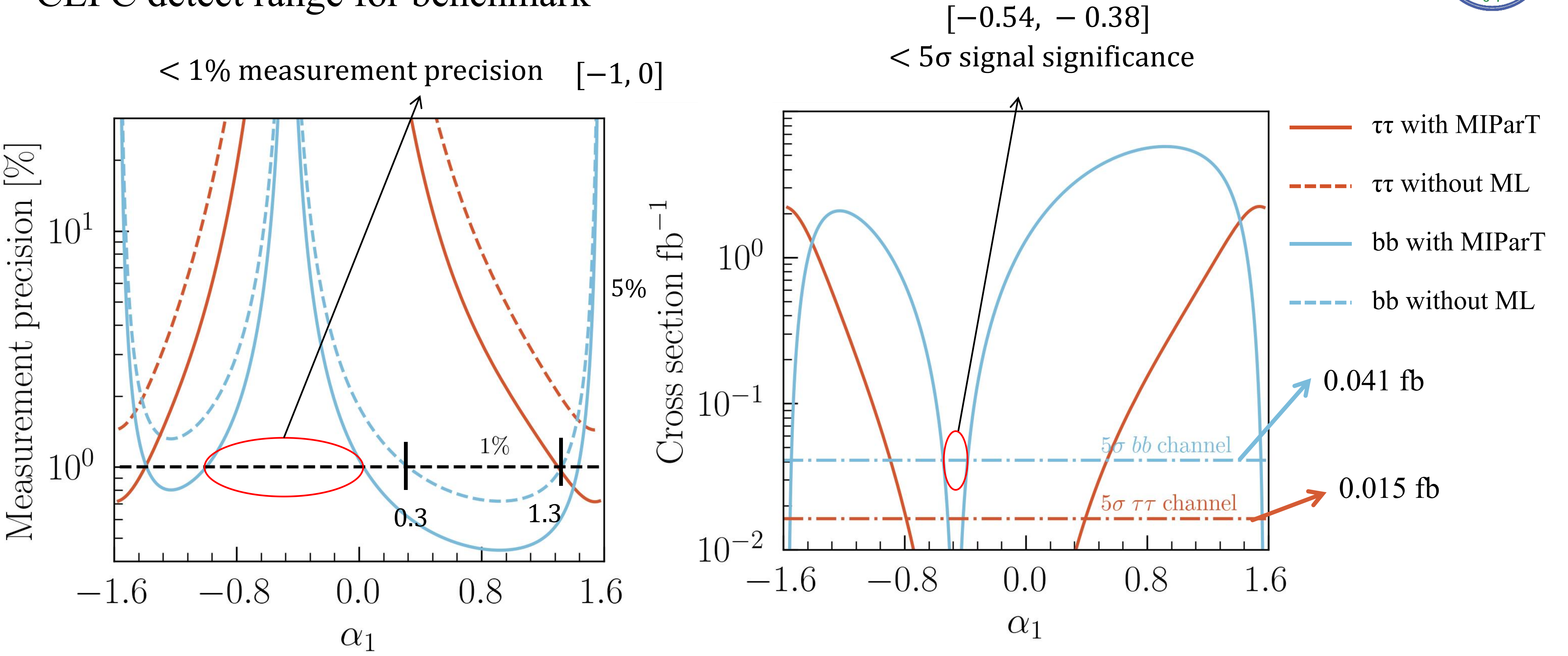


ML performance[1]

1. arXiv: 2407.08682

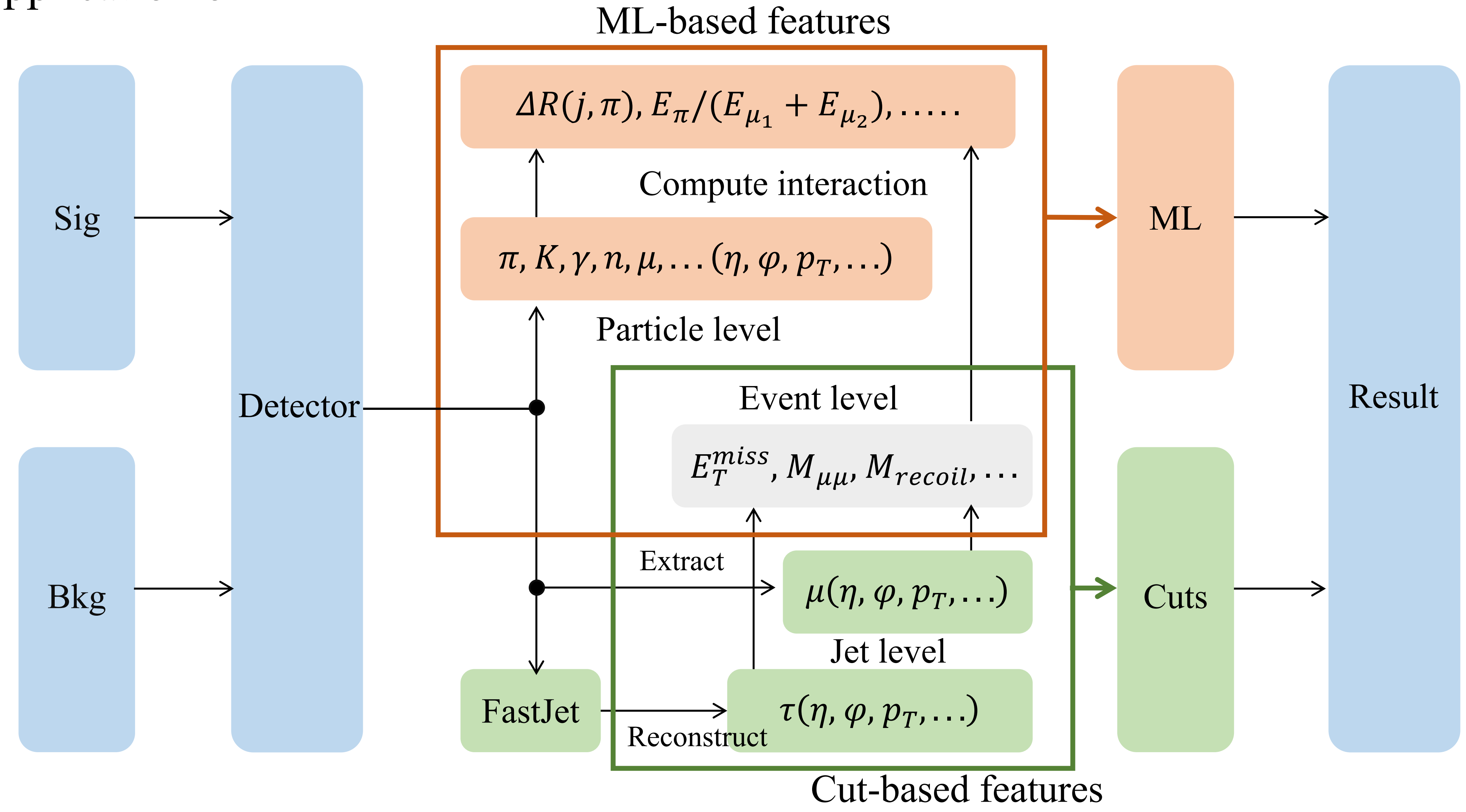
The Application of MIParT

CEPC detect range for benchmark



$\tan\beta = 2$, $\cos\alpha_2 = \sqrt{2}/2$, and $\cos(\beta - \alpha_1) = 0$ for $\alpha_1 = -0.46$

The Application of ML



MC Simulation

The Application of ML