
Measuring the Trilinear Higgs Self-Coupling at Multi-TeV Muon Colliders with Machine Learning

Kingman Cheung, Jae Sik Lee, **Soojin Lee**, Chen Wang

arXiv:2608.01888



The 7th Joint Workshop on the Standard Model and Beyond
Nanjing Normal University · October 31, 2026



Contents

01

Introduction

02

Signal and event selection

03

Topological Data Analysis

04

Machine learning analysis

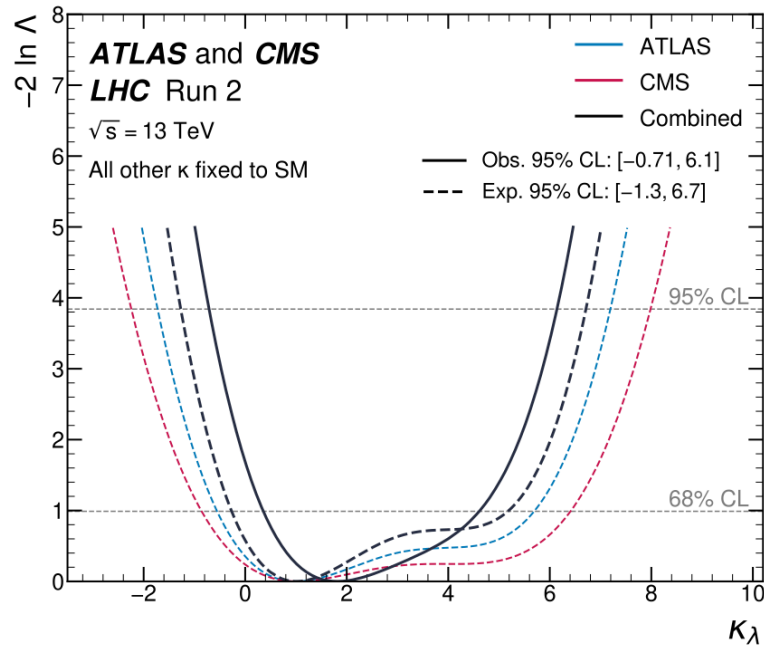
05

Results

01 Introduction

The Higgs self-coupling

Since its discovery in 2012, how precisely have we measured the properties of the Higgs boson?

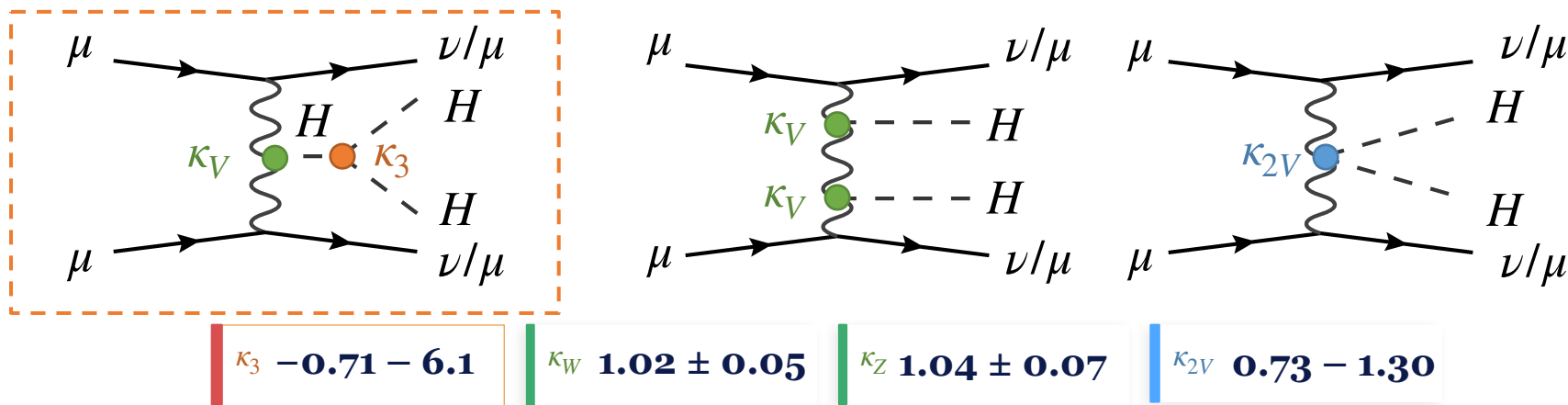


ATLAS+CMS combined
arXiv:2602.23991

- $V(H) = -\mu^2|H|^2 + \lambda|H|^4$ sets the **shape of the electroweak vacuum**
- SM trilinear Higgs Coupling, $\mathcal{L} \supset -\lambda_3 h^3$, $\lambda_3 = m_H^2/2v$
- $\kappa_3 \equiv \lambda_3 / \lambda_3(\text{SM})$, probed directly in HH production
- LHC Run 2 (ATLAS+CMS): **$-0.71 < \kappa_3 < 6.1$** at 95% CL
- HL-LHC, 3 ab^{-1} : **$-42\% / +48\%$** at 68% CL

Precise measurement of κ_3 calls for a new collider !

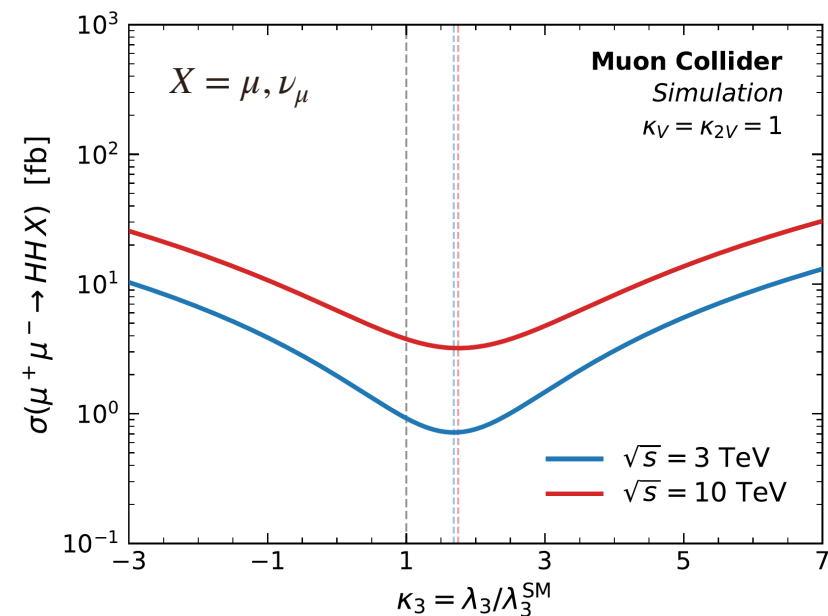
VBF Higgs pair production at a muon collider



Why a multi-TeV muon collider

- ▶ VBF HH grows with $\sqrt{s}$
- ▶ No QCD multijet background and strongly suppressed $t\bar{t}$
- ▶ Full $\sqrt{s}$ goes into the collision, no PDF dilution

- $\mu^+\mu^- \rightarrow HH \nu\bar{\nu}$ (W fusion) + $HH \mu^+\mu^-$ (Z fusion), $HH \rightarrow b\bar{b}b\bar{b}$
- We fix $\kappa_V = \kappa_{2V} = 1$
- Benchmarks: 3 TeV, 1 ab^{-1} and 10 TeV, 10 ab^{-1}
- **Destructive interference:** σ minimum near $\kappa_3 \approx 1.7$



This work at a glance

01 Resolved $\oplus$ Boosted

- 4 small-R jets, $p_T(H) < 200$ GeV
- 2 large-R jets, $p_T(H) > 200$ GeV
- **Combined in one likelihood**

02 Topological data analysis

- Persistent homology of the energy flow
- Compact and interpretable

03 Machine learning chain

- **SPANet** jet-to-Higgs pairing
- **D_{HH}**: signal vs background
- **D _{κ_3}** : κ_3 shape

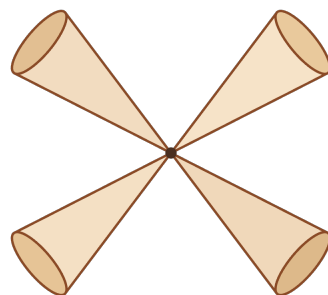
02 Signal and event selection

Two complementary regions

Resolved: $p_T(H_{1,2}) < 200$ GeV

- 4 jets, $R = 0.5$
- 78% (67%) of signal

Resolved

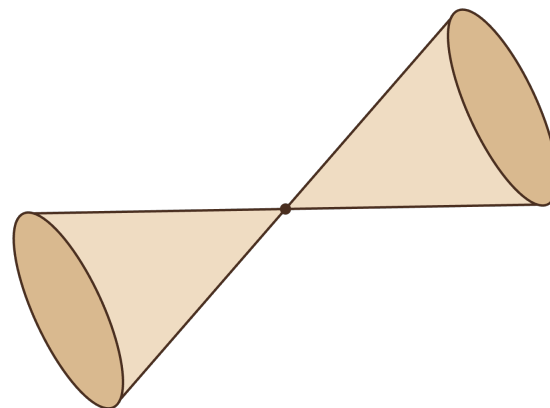


$p_T(H) < 200$ GeV
4 jets, $R = 0.5$

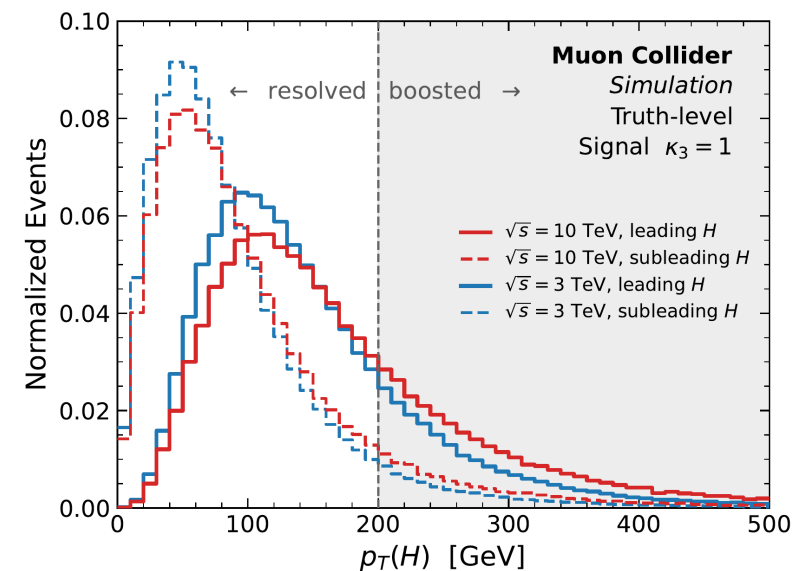
Boosted: $p_T(H_{1,2}) > 200$ GeV

- 2 jets, $R = 1.0$
- 8% (12%) of signal

Boosted



$p_T(H) > 200$ GeV
2 jets, $R = 1.0$



Truth $p_T(H)$, SM signal: resolved | boosted at 200 GeV

Simulation and pre-selection

MadGraph5_aMC@NLO



Pythia 8



Delphes 3, MuC card

RESOLVED

σ_{gen} [fb]	3 TeV	10 TeV
HHX ($\kappa_3 = 1$)	0.312	1.273
Hq$\bar{q}$v$\bar{\nu}$	2.597	6.879
ZZv $\bar{\nu}$	12.16	23.93
W $^+$ W $^-$ v $\bar{\nu}$	19.52	33.97
t $\bar{t}$ v $\bar{\nu}$	2.228	6.982
ZZ	0.755	0.0004
W $^+$ W $^-$	9.191	0.0026
t $\bar{t}$	7.792	0.704
Total bkg	54.24	72.47

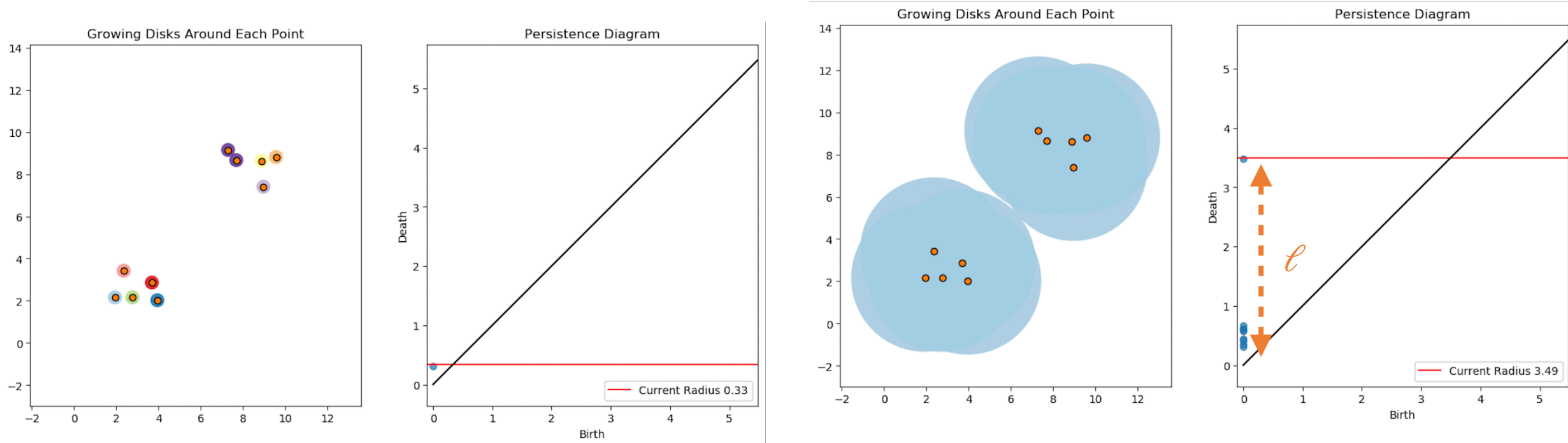
BOOSTED

σ_{gen} [fb]	3 TeV	10 TeV
HHX ($\kappa_3 = 1$)	0.0208	0.128
HZv$\bar{\nu}$	0.520	2.864
ZZv $\bar{\nu}$	6.469	39.84
W $^+$ W $^-$ v $\bar{\nu}$	12.23	70.54
ZZ	9.617	1.336
W $^+$ W $^-$	150.7	22.01
t $\bar{t}$	4.424	0.0354
b $\bar{b}$	2.326	0.0187
q $\bar{q}$ (q $\neq$ b)	13.45	0.108
Total bkg	199.7	136.8

- b tagging from **full simulation** (arXiv:2511.23273)
: $\epsilon_b \approx 0.58$, $\epsilon_c \approx 0.22$, $\epsilon_{\text{light}} \approx 0.007$
- Large-R jets: ϵ^2 as a **double-b tag**
- **Beam-induced background not simulated**

03 Topological Data Analysis

Topological data analysis of the energy flow



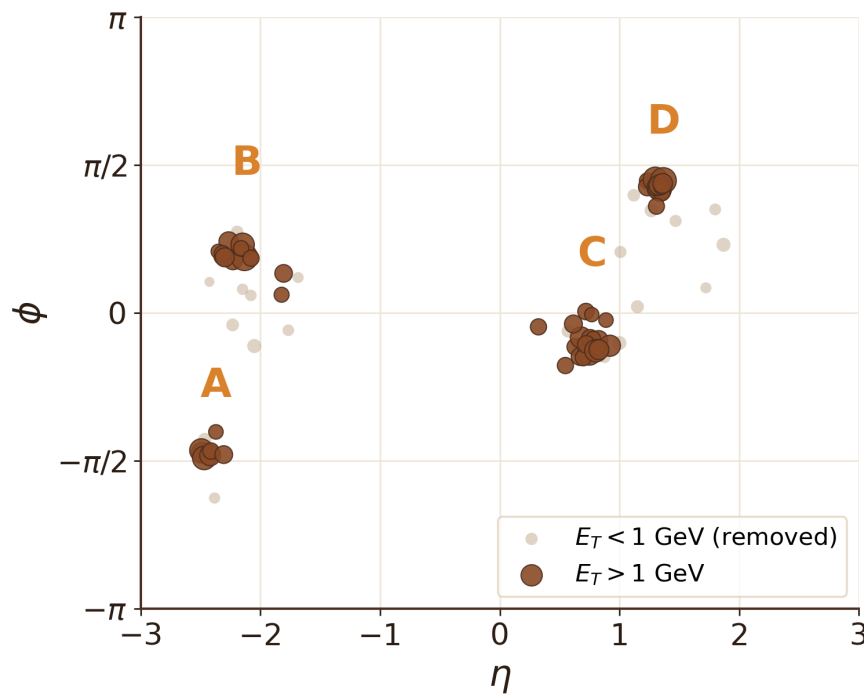
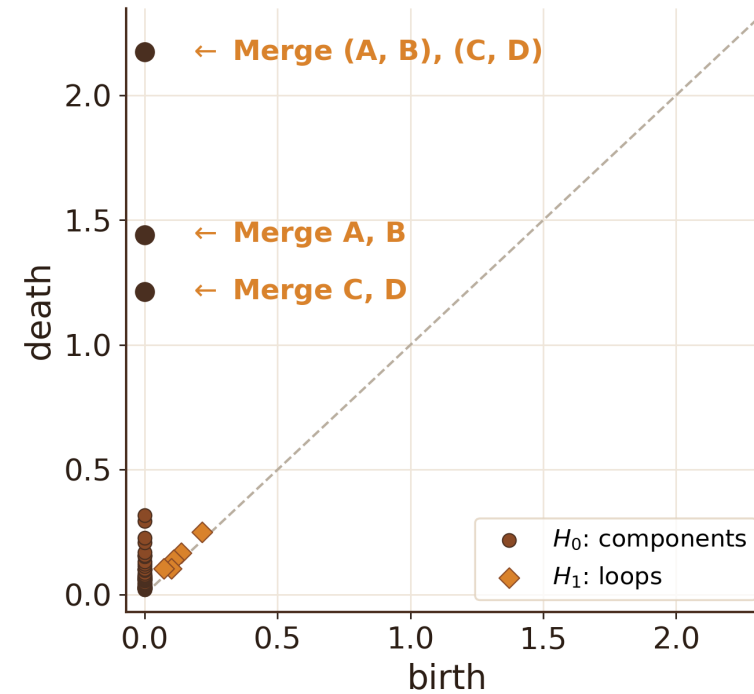
$$\ell_k \equiv d_k - b_k$$

$$H_0 = \sum_{k \in \text{dgm}_0} \ell_k$$

- Grow a ball around each point: clusters merge (H_0)
- H_0 is large when energy spread widely

Topological data analysis of the energy flow

Resolved: one SM HH event, 3 TeV

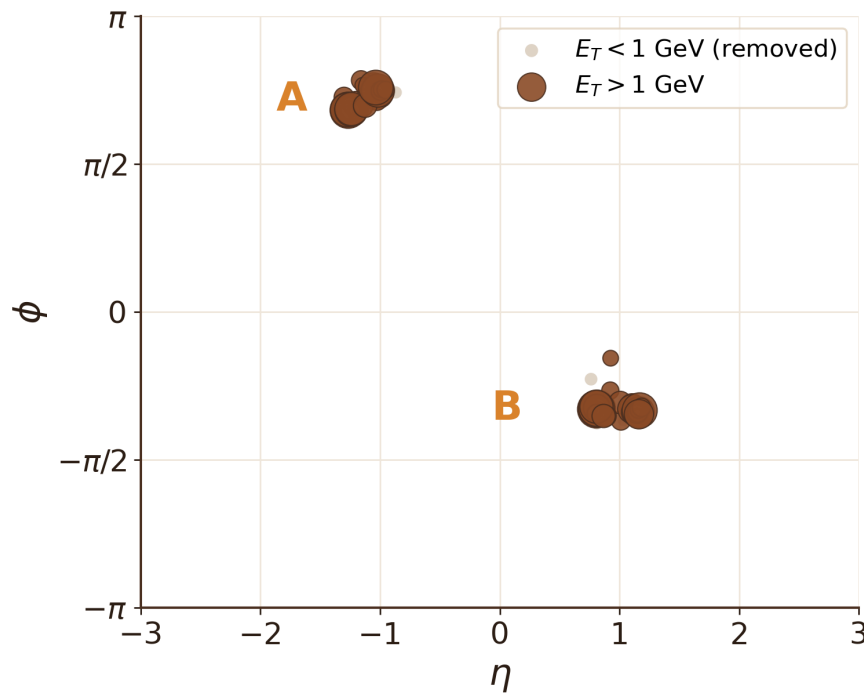
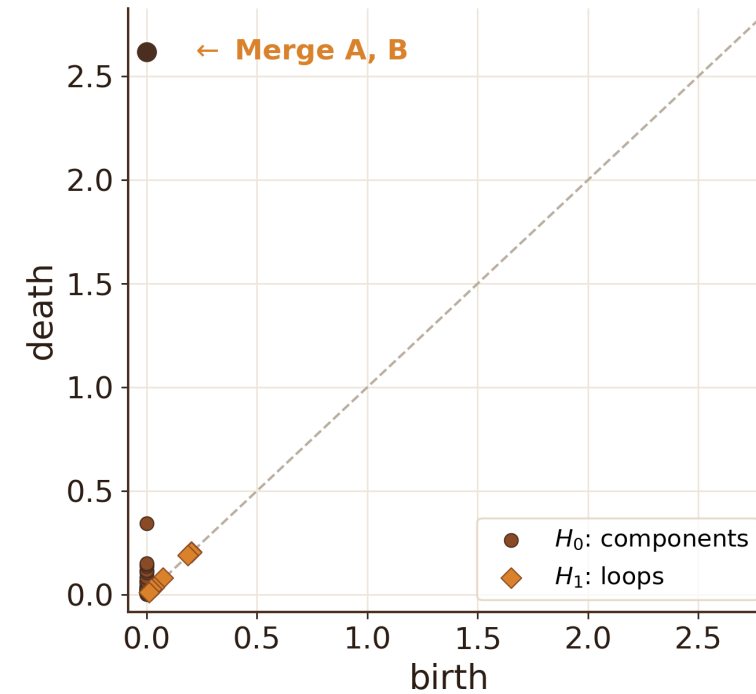


$$d_{ij} = \sqrt{(\Delta\eta_{ij})^2 + \left(2 \sin \frac{\Delta\phi_{ij}}{2}\right)^2}$$

Marker size $\propto E_T$; pale: $E_T < 1$ GeV, removed (BIB mostly below 1 GeV)

Topological data analysis of the energy flow

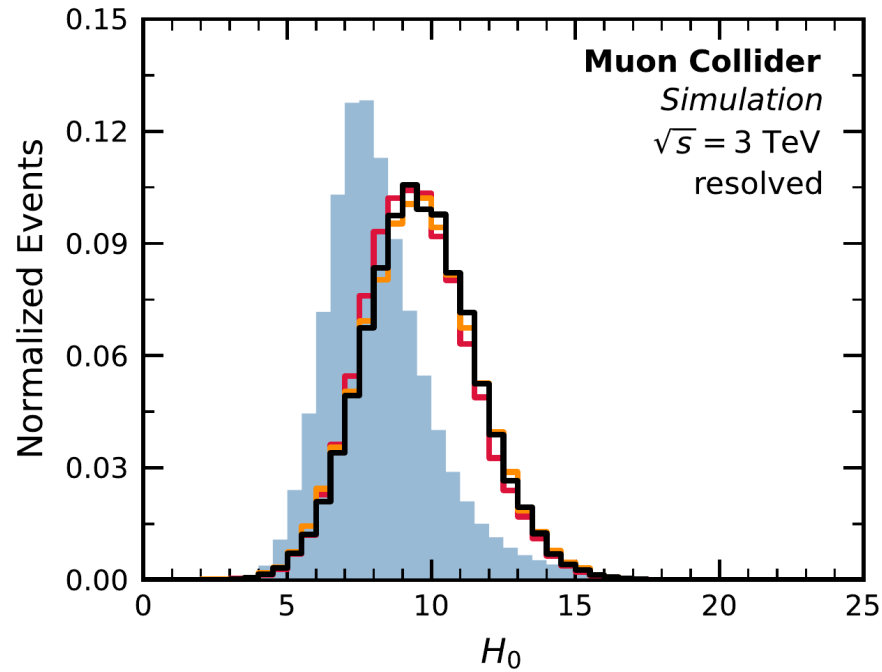
Boosted: one SM HH event, 3 TeV



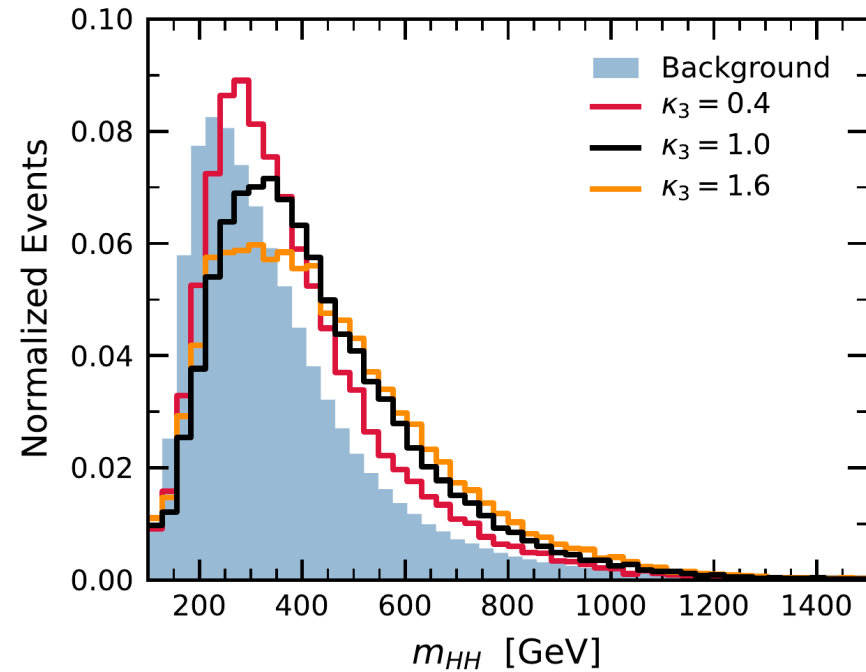
$$d_{ij} = \sqrt{(\Delta\eta_{ij})^2 + \left(2 \sin \frac{\Delta\phi_{ij}}{2}\right)^2}$$

Marker size $\propto E_T$; **pale: $E_T < 1$ GeV, removed** (BIB mostly below 1 GeV)

H_0 finds the signal, m_{HH} reads κ_3



- **Signal vs background**
- Same for all κ_3 : same 4b topology



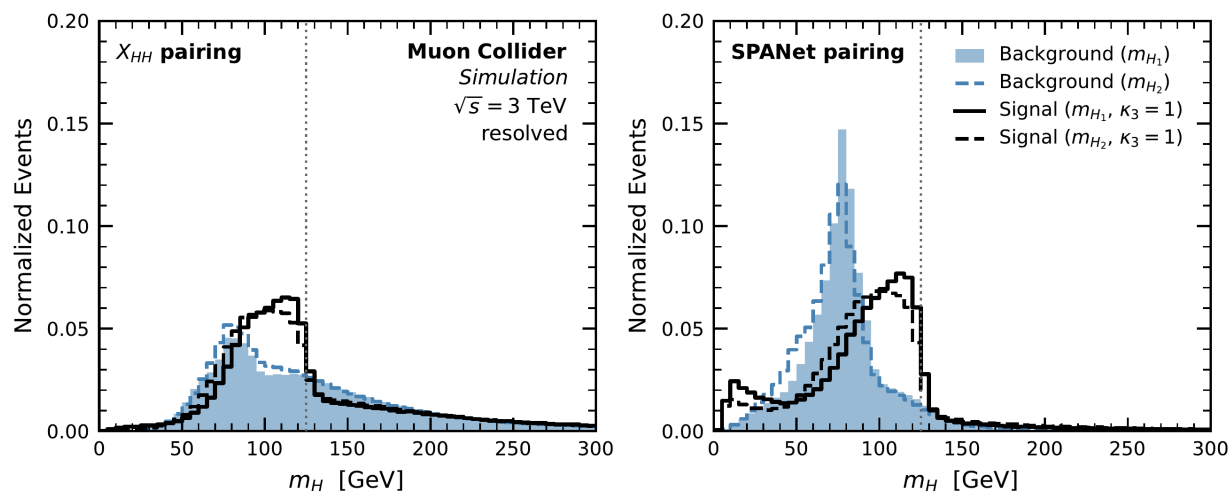
- **Threshold peak changes with κ_3**
- Overlaps with the background

04 Machine learning analysis

SPANet jet-to-Higgs pairing

Classical X_{HH} Pairing

- Minimises $X_{HH} = \sqrt{(m_{H_1} - m_H)^2 + (m_{H_2} - m_H)^2}$
- Forces background to peak near m_H by construction



- X_{HH} picks the pairing closest to m_H : **background is pushed to m_H**
- SPANet: **background stays broad, signal peak stays sharp**

SPANet

Symmetry-Preserving Attention Network

Role of attention

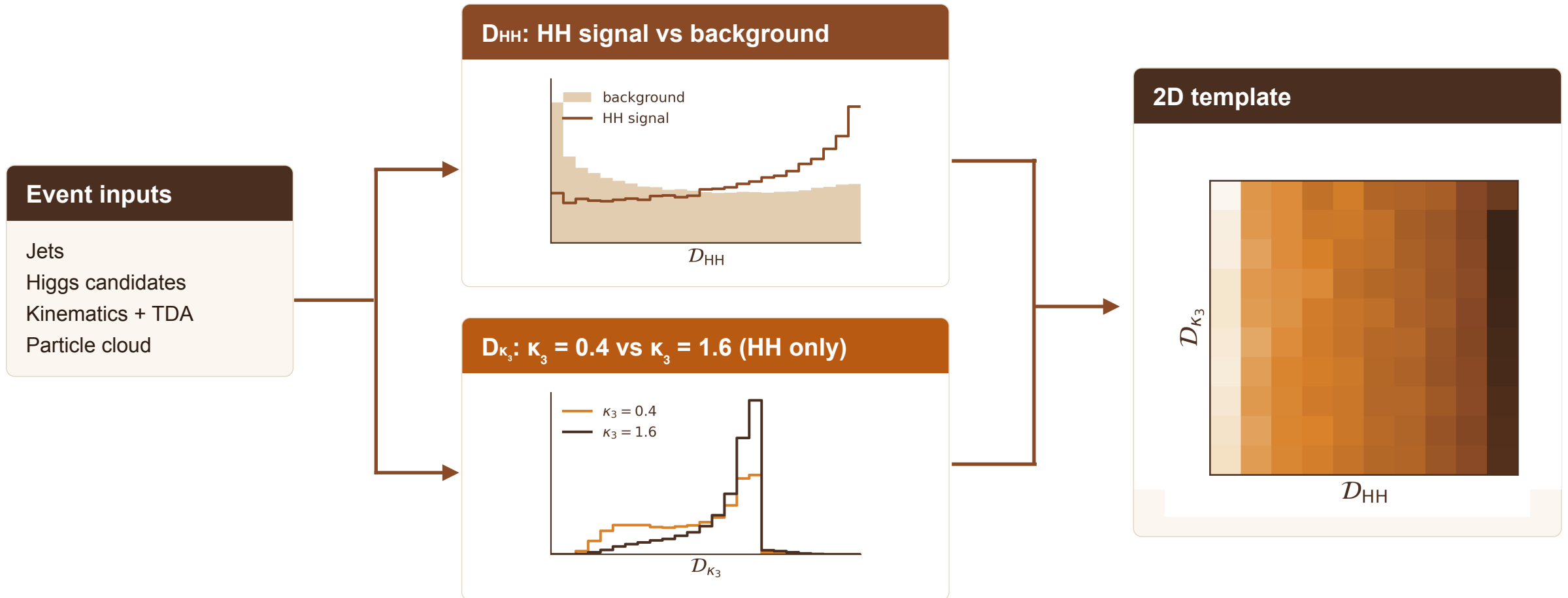
Each jet weighs all other jets, so the pairing is decided with the **full event context**, not one jet pair at a time

Built-in symmetry

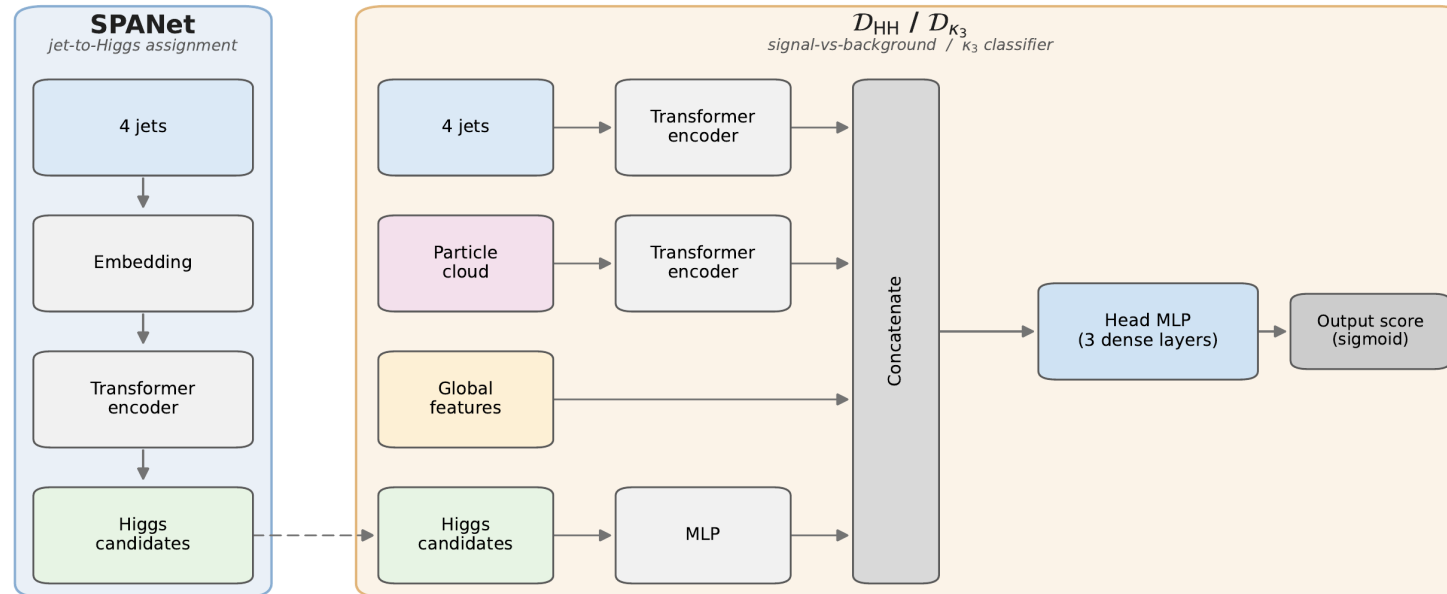
Invariant under jet order, swap within a pair and $H_1 \leftrightarrow H_2$

➡ **95% correct pairings!**

Two classifiers, one 2D template



Resolved classifiers: inputs



ONE ARCHITECTURE, TWO TARGETS. **DHH**: HH vs background · **D κ_3** : $\kappa_3 = 0.4$ vs 1.6

Jets · 4 × 7

$\ln p_T, \eta, \sin \varphi, \cos \varphi,$
 $\ln(1 + m/m_0), |\Delta\varphi(j, \text{MET})|, \text{b-tag}$

Higgs candidates · 2 × 7

$\ln p_T, \eta, \sin \varphi, \cos \varphi, \ln m, \text{Nb},$
 ΔR_{jj} from SPANet

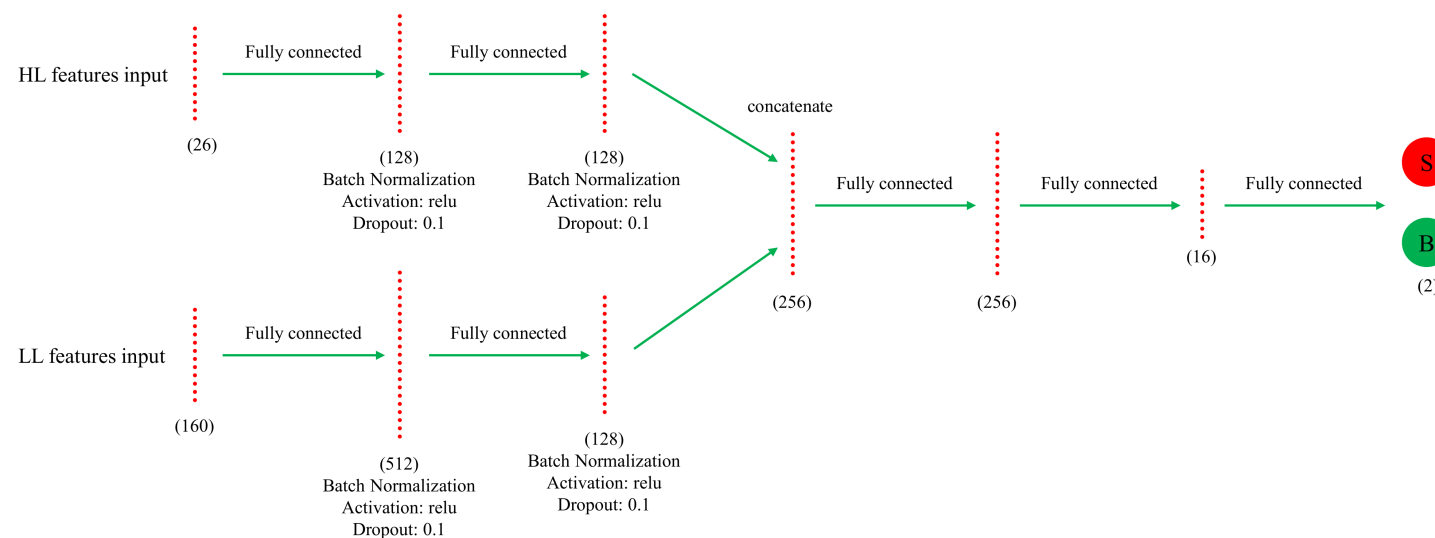
Global · 8 + 5 TDA

$m_{HH}, p_T(HH), X_{HH}, \Delta\varphi(HH, \text{MET}),$
 $\Delta R(H_1, H_2), \text{Nb}, \text{MET}, \Sigma E, \mathbf{H}_0, \mathbf{H}_1, \mathbf{S}_0, \mathbf{S}_1, \mathbf{LB}_1$

Particle cloud · 40 × 4

10 leading EFlow per jet:
 $p_T/p_T(\text{jet}), \Delta\eta, \Delta\varphi, \text{particle type}$

Boosted classifiers: inputs and network



DHH: Two-branch DNN, HH vs background · **D κ_3 : XGBoost**, $\kappa_3 = 0.4$ vs 1.6

Fat jets · 2 × 8

M, pT, double-b tag, **MSD**, **τ_{21}** , **τ_{32}** , **D2**, **N2**

Global · 5 + 5 TDA

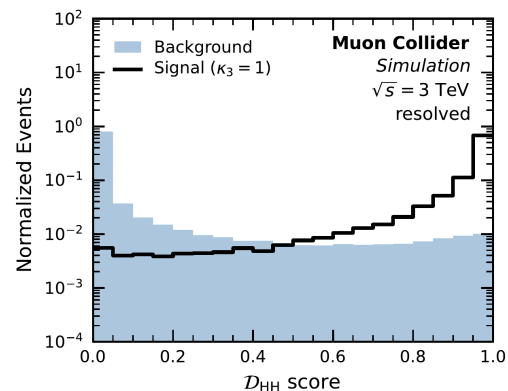
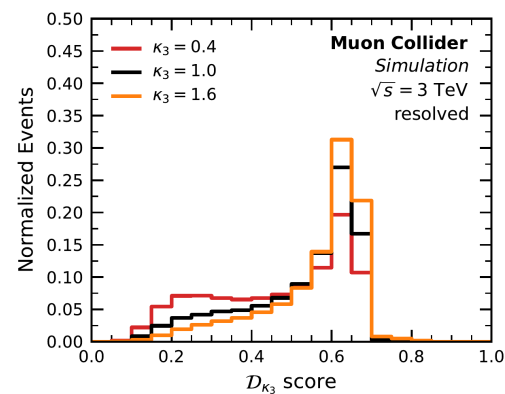
MET, XHH, MJJ, PTJJ, $\Delta\eta(J_1, J_2) + H_0, H_1, S_0, S_1, LB_1$

Particles · 2 × 20 × 4

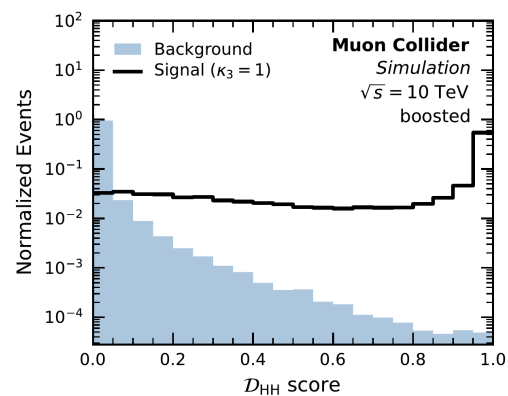
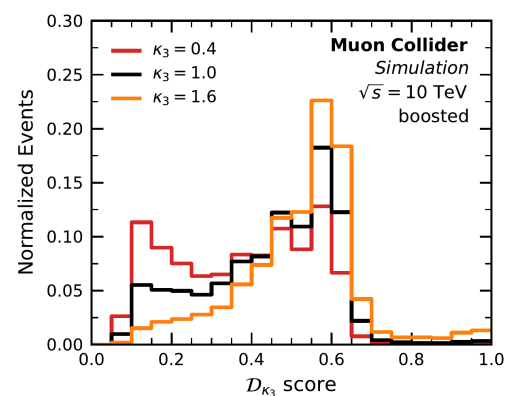
20 leading per fat jet: pT, η , ϕ , particle type

Scores: resolved and boosted classifiers

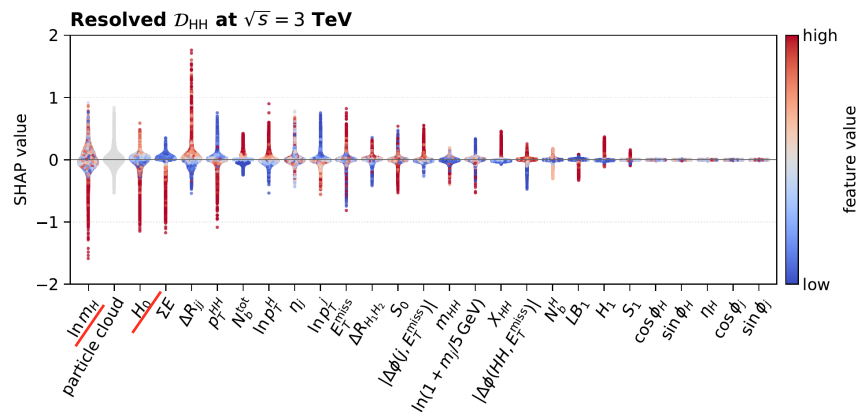
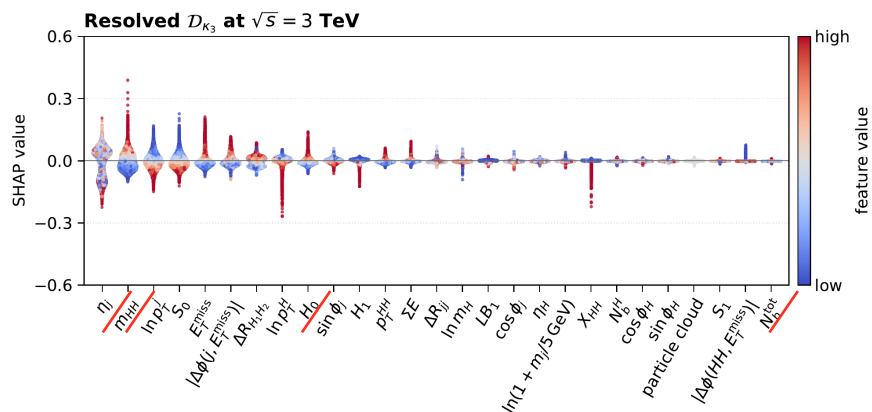
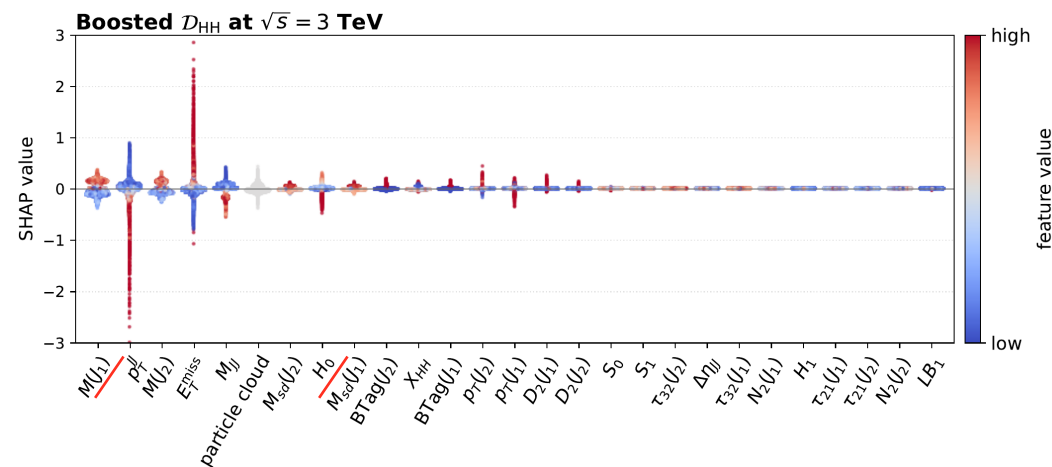
Resolved



Boosted



What the classifiers learn

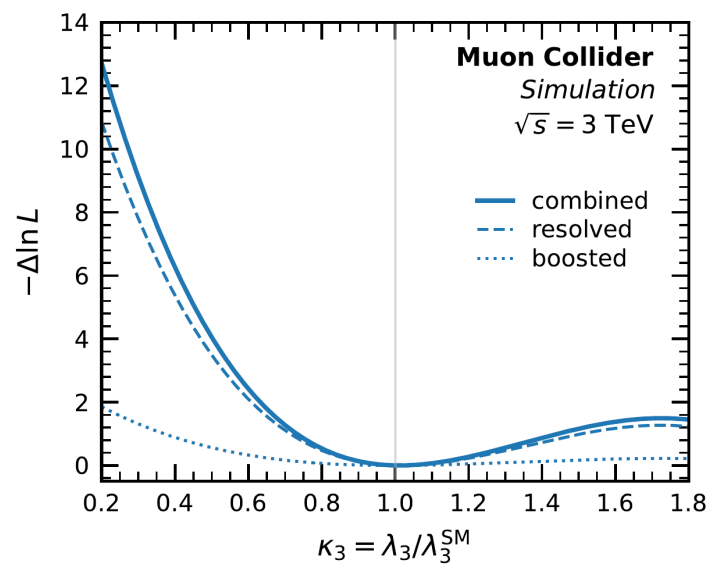
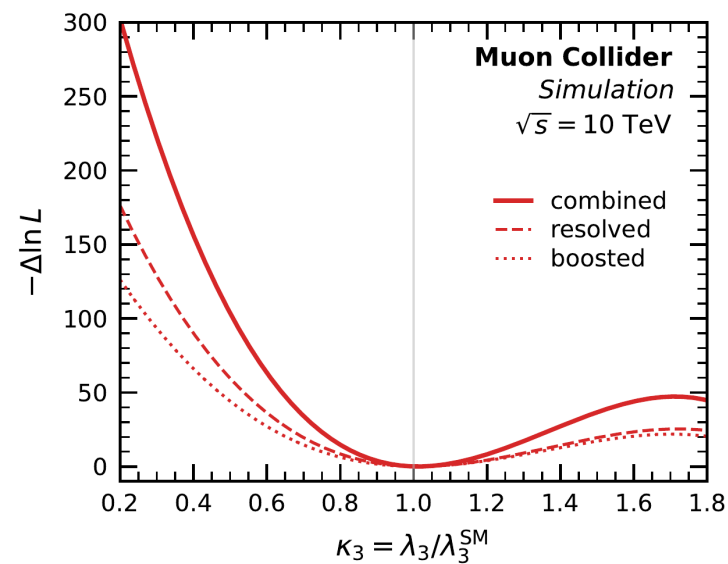


SHAP RANKING, 3 TeV

- Resolved DHH: Higgs mass first, **H_0 third**
- Resolved D_{K_3} : jet η and **m_{HH}** lead
- Boosted DHH: fat-jet mass and MET lead

05 Results

Sensitivity to κ_3



68% CL	3 TeV, 1 ab ⁻¹	10 TeV, 10 ab ⁻¹
Resolved	[0.79, 1.31]	[0.94, 1.06]
Boosted	[0.52, open]	[0.94, 1.08]
Combined	[0.81, 1.28]	[0.96, 1.05]

Summary

01 Combining the **resolved and boosted regions** in one likelihood tightens the κ_3 sensitivity

02 **Topological descriptors** of the energy flow give a compact, interpretable summary of each event

03 **SPANet** pairs the jets correctly, **D_{HH}** finds the HH events and **D κ_3** reads the κ_3 shape.

At 68% CL

3 TeV, 1 ab⁻¹

$0.81 < \kappa_3 < 1.28$

10 TeV, 10 ab⁻¹

$0.96 < \kappa_3 < 1.05$

vs HL-LHC, $0.58 < \kappa_3 < 1.48$

Thank you

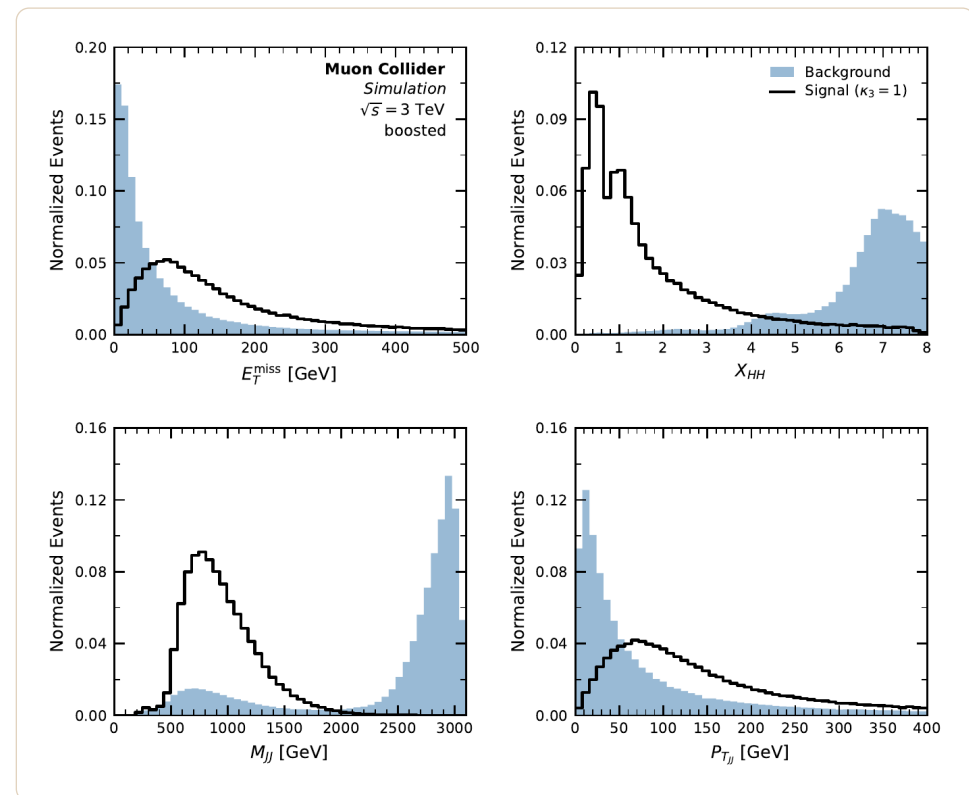
[arXiv:2608.01888](https://arxiv.org/abs/2608.01888) · github.com/LSJ957/muc-hh-kappa3

Backup

κ_3 sensitivity across colliders

Experiment	κ_3 , 68% CL	Note (arXiv)
LHC Run 2	$[-0.3, 4.4]$	13 TeV, 140 fb ⁻¹ , ATLAS single H + HH (2608.07332). ATLAS+CMS HH: $-0.71 < \kappa_3 < 6.1$ at 95% (2602.23991)
HL-LHC	$[0.74, 1.29]$	14 TeV, 3 ab ⁻¹ per experiment, ATLAS+CMS HH (2504.00672)
FCC-ee	$\delta\kappa_3 \approx 17\text{--}20\%$	indirect via single-H loops, with HL-LHC (2511.03883)
CLIC	$[0.92, 1.11]$	1.4 + 3 TeV, 5 ab ⁻¹ at 3 TeV, ZHH + HHv $\bar{\nu}$ (1901.05897)
FCC-hh	$\delta\kappa_3 \approx 3\text{--}4\%$	85 TeV, 30 ab ⁻¹ , HH $\rightarrow b\bar{b}\gamma\gamma$, $b\bar{b}\tau\tau$ (2511.03883)
MuC 3 TeV, Andreetto et al.	$[0.81, 1.44]$	1 ab ⁻¹ , Geant4, no BIB, assumed b-tag efficiency (2405.19314)
MuC 10 TeV, Andreetto et al.	$[0.94, 1.08]$	10 ab ⁻¹ , MUSIC full simulation with BIB (2609.08392)
MuC 3 TeV, this work	$[0.81, 1.28]$ (−19% / +28%)	1 ab ⁻¹ , Delphes, no BIB, resolved $\oplus$ boosted
MuC 10 TeV, this work	$[0.96, 1.05]$ (−4% / +5%)	10 ab ⁻¹ , Delphes, no BIB, resolved $\oplus$ boosted

Boosted region: signal vs background



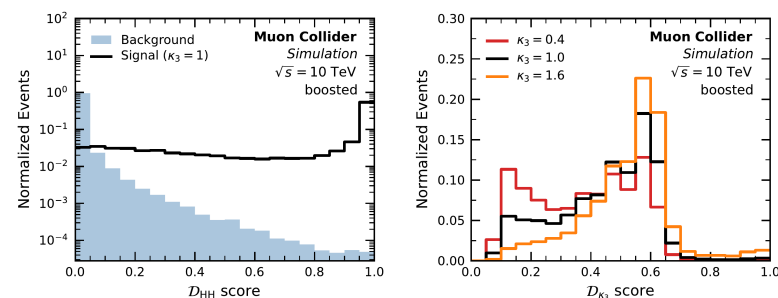
3 TeV after pre-selection: MET, XHH, MJJ, PTJJ

VBF signal

- **Large MET** from the forward neutrinos
- Small XHH: both jet masses near m_H
- MJJ peaks near 800 GeV

Two-body background

- Full $\sqrt{s}$ in two jets: **MJJ near $\sqrt{s}$**
- Back-to-back: small MET and PTJJ



10 TeV scores: DHH (left), D_{κ_3} (right)

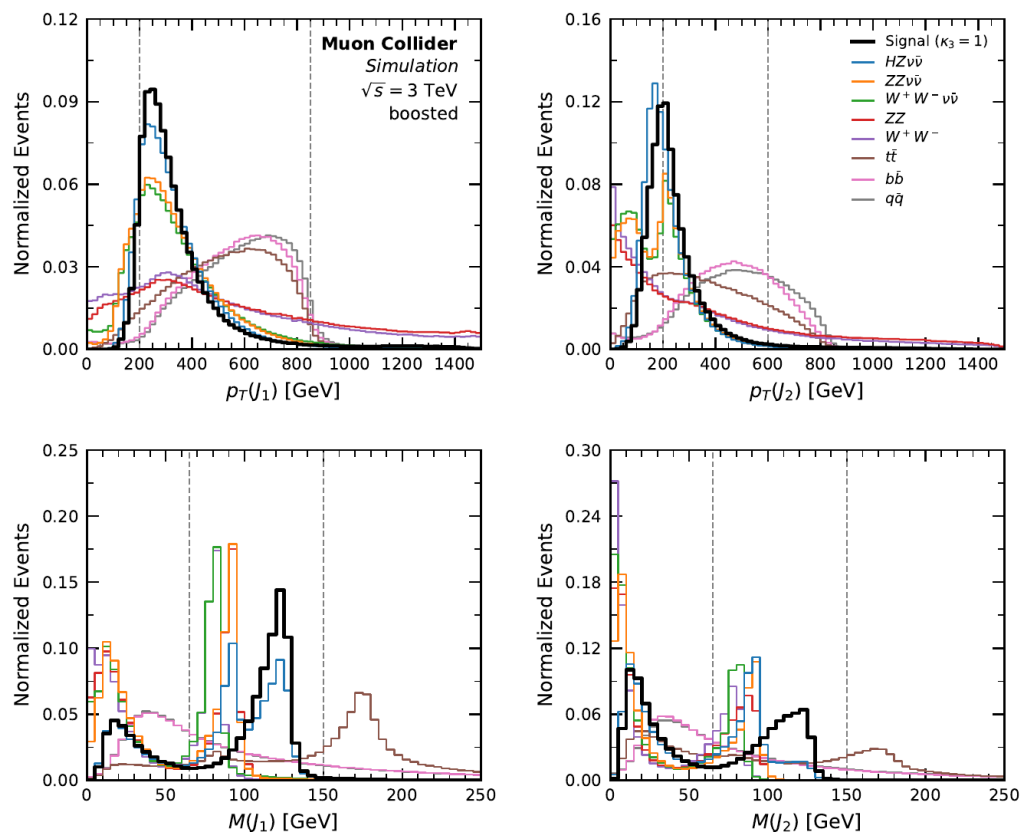
Resolved pre-selection yields

Process	σ_{gen} [fb]	ϵ_{pre} [%]	Npre	σ_{gen} [fb]	ϵ_{pre} [%]	Npre
HHX ($\kappa_3 = 1$)	0.312	42.4	132	1.273	27.2	3,467
Hq $\bar{q}$ v $\bar{\nu}$	2.597	44.7	1,160	6.879	32.1	22,055
ZZv $\bar{\nu}$	12.16	45.3	5,514	23.93	36.7	87,860
W ⁺ W ⁻ v $\bar{\nu}$	19.52	47.6	9,292	33.97	40.6	137,855
t $\bar{t}$ v $\bar{\nu}$	2.228	58.8	1,310	6.982	45.4	31,726
ZZ	0.755	14.4	109	0.0004	3.7	0.1
W ⁺ W ⁻	9.191	22.2	2,043	0.0026	4.4	1.1
t $\bar{t}$	7.792	1.6	122	0.704	0.3	23.4
Total bkg	54.24	–	19,550	72.47	–	279,521

Boosted pre-selection yields

Process	σ_{gen} [fb]	ϵ_{pre} [%]	Npre	σ_{gen} [fb]	ϵ_{pre} [%]	Npre
HHX ($\kappa_3 = 1$)	0.0208	29.6	6.15	0.128	23.0	293.4
HZ $\nu\bar{\nu}$	0.520	26.6	138.2	2.864	13.8	3,953
ZZ $\nu\bar{\nu}$	6.469	24.2	1,567	39.84	14.8	59,022
W ⁺ W ⁻ $\nu\bar{\nu}$	12.23	20.0	2,451	70.54	11.6	81,832
ZZ	9.617	12.2	1,168	1.336	0.0036	0.48
W ⁺ W ⁻	150.7	10.1	15,287	22.01	0.0014	3.08
t $\bar{t}$	4.424	7.2	316.3	0.0354	3.7	12.97
b $\bar{b}$	2.326	7.8	181.8	0.0187	2.1	3.87
q $\bar{q}$ (q $\neq$ b)	13.45	7.1	949.8	0.108	1.9	20.50
Total bkg	199.7	–	22,059	136.8	–	144,848

Boosted pre-selection



PRE-SELECTION (DASHED LINES)

- 2 fat jets, $|\eta(J)| < 2.5$
- $200 < p_T(J_1) < 850 \text{ GeV}$
- $200 < p_T(J_2) < 600 \text{ GeV}$
- $65 < M(J_{1,2}) < 150 \text{ GeV}$

GENERATION CUTS

- Signal: $p_T(H) > 200 \text{ GeV}$
- Di-boson: $p_T(V) > 200 \text{ GeV}$
- $t\bar{t}$, $b\bar{b}$, $q\bar{q}$: $200 < p_T < 850 \text{ GeV}$

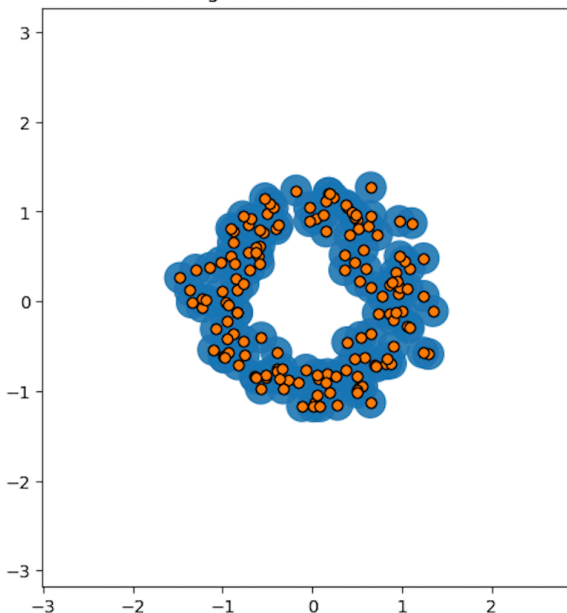
What each ingredient adds

68% CL	3 TeV	10 TeV
Resolved: XHH pairing, DHH only	[0.72, 1.44]	[0.90, 1.07]
+ SPANet pairing	[0.72, 1.42]	[0.91, 1.07]
+ TDA	[0.74, 1.38]	[0.92, 1.07]
+ $D\kappa_3$	[0.79, 1.31]	[0.94, 1.06]
Boosted: DHH only	[0.46, -]	[0.92, 1.10]
+ TDA	[0.52, -]	[0.92, 1.10]
+ $D\kappa_3$	–	[0.94, 1.08]
Combined	[0.81, 1.28]	[0.96, 1.05]

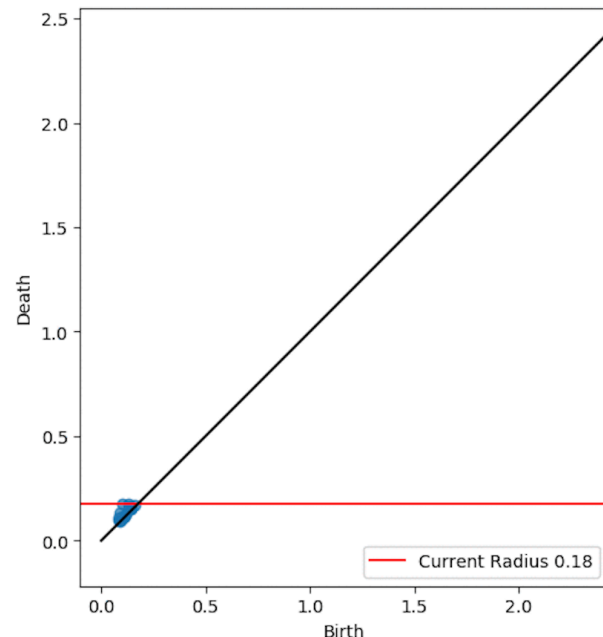
DHH retrained at each step. Dash: edge beyond the scanned range.

Topological data analysis of the energy flow

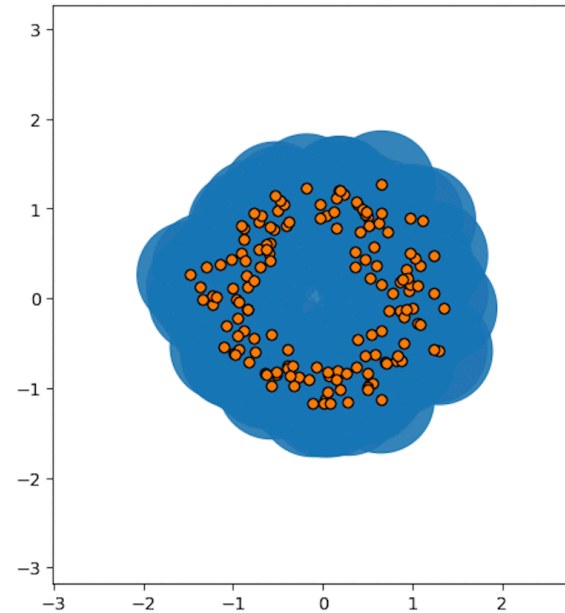
Growing Disks Around Each Point



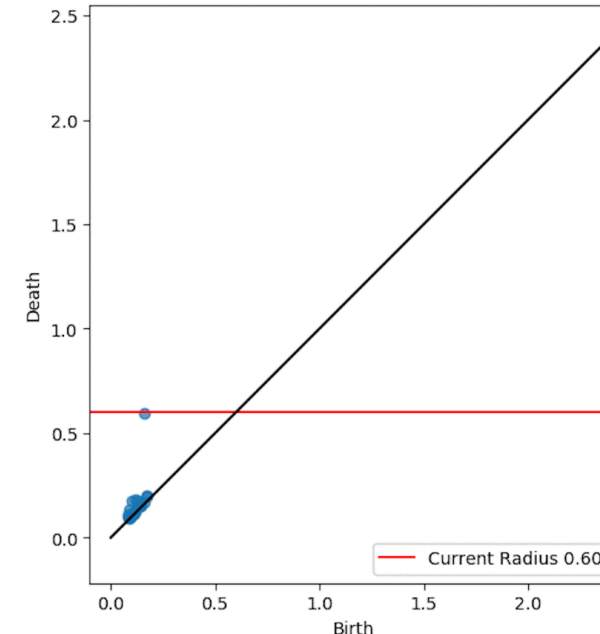
Persistence



Growing Disks Around Each Point



Persistence



$$H_i = \sum_{k \in \text{dgm}_i} \ell_k, \quad S_i = - \sum_k \frac{\ell_k}{H_i} \log_2 \frac{\ell_k}{H_i}, \quad LB_1 = \sum_{k \in \text{dgm}_1} b_k \ell_k$$

TDA	Definition	Large when
H_0	$\sum \ell$, components	energy spread widely
H_1	$\sum \ell$, loops	voids between clusters
S_0, S_1	entropy of ℓ	many similar features
LB_1	$\sum \text{birth} \times \ell$, loops	large late loops

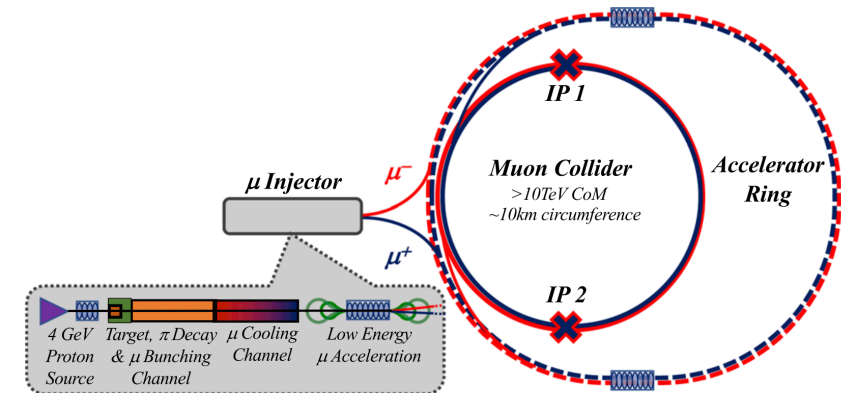
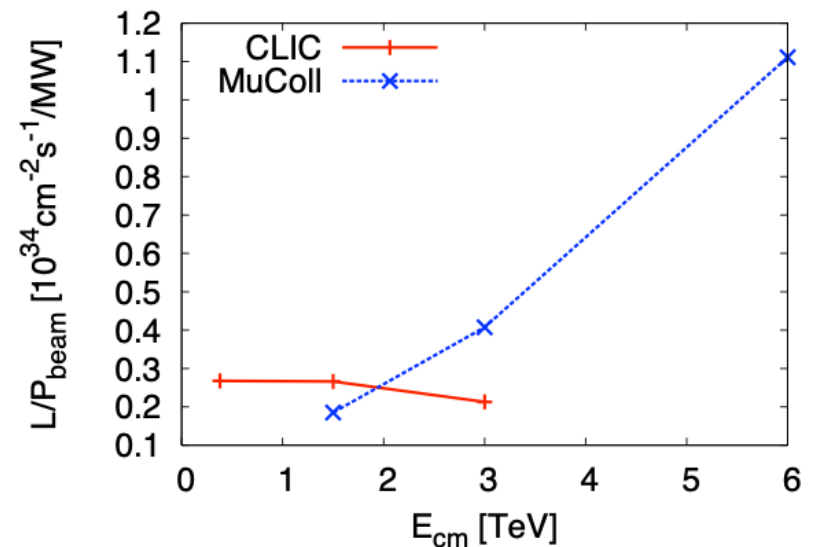
Why a muon collider

- Point-like beams: full $\sqrt{s}$ available
- No underlying event, negligible pile-up
- Electroweak backgrounds, no QCD multijet flood
- 200 $\times$ heavier than e: little synchrotron loss
- 3 TeV: ring $\approx$ 4.5 km; 10 TeV: ring $\approx$ 10 km

CHALLENGES

- Muon production and ionization cooling
- High-field magnets for 10 TeV
- Beam-induced background in the detector
- Neutrino radiation along the ring plane

arXiv:2407.12450



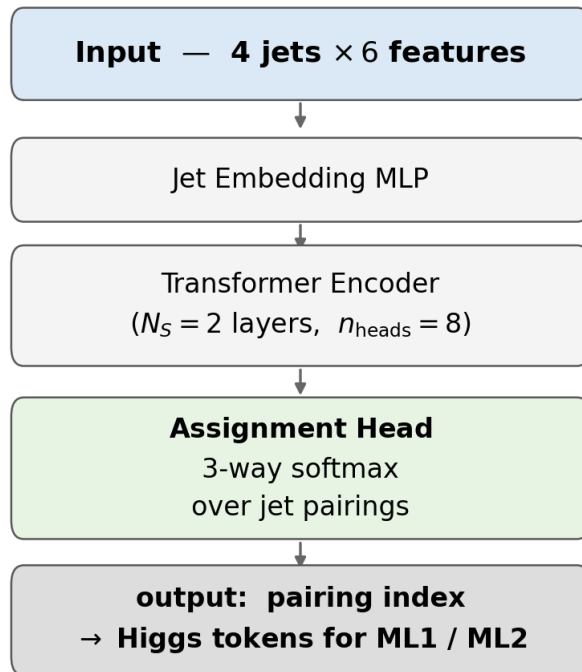
Profiling κ_{2V}

3 TeV, 68% CL	$\kappa_{2V} = 1$	$\sigma = 5\%$	$\sigma = 10\%$
Resolved	[0.79, 1.31]	[0.72, 1.37]	[0.64, 1.49]
Combined	[0.81, 1.28]	[0.73, 1.37]	[0.70, 1.41]

10 TeV, 68% CL	$\kappa_{2V} = 1$	$\sigma = 0.5\%$	$\sigma = 1\%$
Resolved	[0.94, 1.06]	[0.93, 1.06]	[0.93, 1.08]
Boosted	[0.94, 1.08]	[0.90, 1.12]	[0.86, 1.15]
Combined	[0.96, 1.05]	[0.94, 1.06]	[0.93, 1.07]

SPANet Network

SPANet — jet-to-Higgs pairing



Symmetry-Preserving Attention Network

- Job: **choose which 2 jets form H1, which 2 form H2** — 3 pairings to score
- Input: **4 jets × 6 features** (log pT, η , sin φ , cos φ , log mass, b-tag)
- Backbone: per-jet MLP + small Transformer (2 layers, 8 heads)
- Assignment head: enumerate 3 pairings → MLP scorer → softmax(3)
- Built-in symmetry: mean-pool inside a Higgs candidate;
H1 $\leftrightarrow$ H2 absorbed by the enumeration
- Classification accuracy of SPANet: **95% for both 3 TeV and 10 TeV**
- Output: argmax pairing index {0,1,2} → forwarded as **Higgs tokens to ML1/ML2**