$hh \rightarrow 4b$ with Transformer

Yongcheng Wu Nanjing Normal University LHC Mini-Workshop 2024.12.13-2024.12.16



- Higgs Milestone for the SM
- Current Measurement @LHC
 - Production
 - Decay Width



 \leq 140 fb⁻¹ (13 TeV)

 $2\ell 2\nu + 4\ell$ off-shell + 4ℓ on-shell

 $2\ell 2\nu$ off-shell + 4ℓ on-shell

CMS

14

- Higgs Milestone for the SM
- Current Measurement @LHC
 - Couplings
 - Higgs Mechanism









- Higgs Pair Production
- JHEP 10 (2012) 112 Optional: Recoil against hard jet
 - Higher order corrections
 - Single Higgs production
 - NLO correction depends on Higgs self-coupling JHEP 12 (2016) 080
 - Probe Higgs coupling w/o Higgs
 - V_L Scattering Goldstone equivalent

Phys.Rev.Lett. 123 (2019) 181801 JHEP 10 (2021) 099 Phys.Rev.D 105 (2022) 053009

• Electroweak precision measurement

$$\mathcal{O}_{BSM} = \mathcal{O}_{SM}(1 + (\kappa_{\lambda} - 1)C_1 + (\kappa_1^2 - 1)C_2)$$

JHEP 04 (2017) 155 Phys. Lett. B 817 (2021) 136307



Current Status of Higgs Pair



Particle Transformer

- ParT SOTA for Jet Tagging ICML 2022/arXiv: 2202.03772
 - Used by CMS in $h(b\overline{b})h(\tau^+\tau^-)$ to reconstruct $m_{\tau\tau}$ Nuovo Cim.C 47 (2024) 5, 250
 - Better than commonly used CPU consuming algorithm
 - Quark Flavor tagging in ILC project arXiv: 2410.11322
 - Evolution:
 - ParMAT arXiv: 2406.06638
 - MIParT arXiv: 2407.08682

| | Accuracy | # params | FLOPs |
|--------------|----------|----------|--------|
| PFN | 0.772 | 86.1 k | 4.62 M |
| P-CNN | 0.809 | 354 k | 15.5 M |
| ParticleNet | 0.844 | 370 k | 540 M |
| ParT | 0.861 | 2.14 M | 340 M |
| ParT (plain) | 0.849 | 2.13 M | 260 M |

EvenTransformer (ET)

- Based on ParT ICML 2022/arXiv: 2202.03772
- Extensions
 - The whole event as extremely fat Jet
 - Tagging the whole event
- Signal vs Background ⇒ Process Tagging



EvenTransformer (ET)

- Input Features
 - Low-level information From detector simulation (Delphes)
 - Particle Flow Objects
 - Tracks Displaced Vertices
 - Energy Deposition in Ecal and Hcal
 - Auxiliary Higher-level information Testing on Trying not using these
 - Slim Jet $\Delta R = 0.5$
 - Fat Jet $\Delta R = 1.0$
 - No B-tagging information

Higgs Pair to 4b

• Processes – Non-resonant Higgs Pair

- $g g \rightarrow h h \rightarrow b \overline{b} b \overline{b}$
- $b\overline{b} + 2jets$
- *bbbb*
- *bbh*
- *tt*
- *tth*
- $t\overline{t}b\overline{b}$
- *ZZ*
- *zh*
- Multi-class Classification
 - Training on 2.5M/process data

| Process | Cross Section [fb] | k-factor |
|---------|----------------------|----------|
| hh | 1.45×10^{1} | 2.4 |
| bbjj | 5.67×10^{8} | 1.6 |
| bbbb | 2.48×10^{6} | 1.3 |
| bbh | 4.73×10^{1} | - |
| tt | 5.06×10^{5} | 1.4 |
| tth | 4.00×10^{2} | - |
| ttbb | 1.36×10^{4} | - |
| ZZ | 9.34×10^{3} | - |
| zh | 5.80×10^{2} | - |

 $p_T^{b,j} > 20 \text{ GeV}, |\eta_{b,j}| < 4$

Model Results

- Confusion Matrix on validation set
 - Cross Entropy Loss Function



Major QCD Background For Higgs Pair to 4b

Model Results

- Confusion Matrix on validation set
 - Weighted Cross Entropy Overcome the large 2b2j background



Model Results

- Efficiency and Relative Error on $\sigma(HH)$ vs. Threshold
 - For $\kappa_{\lambda} = 1$ on test set



Model Results on σ and κ_{λ}

• With different κ_{λ} - Not optimized for other κ_{λ}



Compare with other Results

- Studies on 4b channel
 - DNN

JHEP 12 (2020) 115; 2004.04240

- 2 hidden layers, each with 200 nodes
- Inputs: information from reconstructed objects
- Comprehensive backgrounds

- SPA-NET (Symmetry preserving attention Network)
 - Compared with Cut-based and DNN-based analysis
 - Two networks
 - Pairing of the 4b into two Higgs
 - Classification of Signal/background
 - Inputs: List of jets (p_T, η, ϕ, m) and b-tag
 - Only QCD 4b background

JHEP 09 (2024) 139; 2401.14198

Limits on κ_{λ}

• Compared with Other Results



Summary

- Higgs Self Coupling Measurement
- Higgs Pair in 4b channel
 - Attention Mechanism ParT
 - Train on full event event classification
- Improved over current phenomenology studies
 - Depends on different working points w.r.t signal efficiency
- On progress Results might change
 - Improve on different κ_{λ}
 - Experiments on the model
 - Tuning for other tasks
 - Fully end-to-end
 - Interpretability Attention Map

Thanks for the

Thanks for your **Attention!**