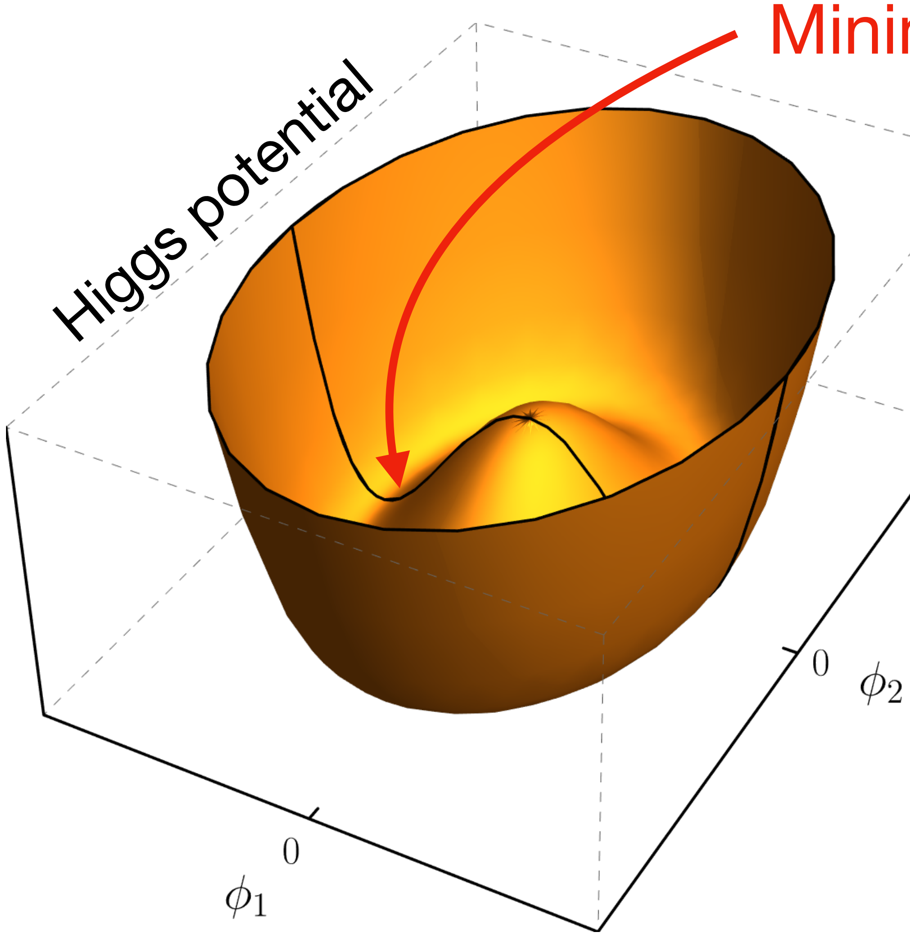
Investigating the Higgs self-couplings through HHH production

Panagiotis Stylianou
based on 2312.04646
in collaboration with Georg Weiglein

SUSY 2024 - IFT, Madrid
13 June 2024

Introduction: the Higgs potential

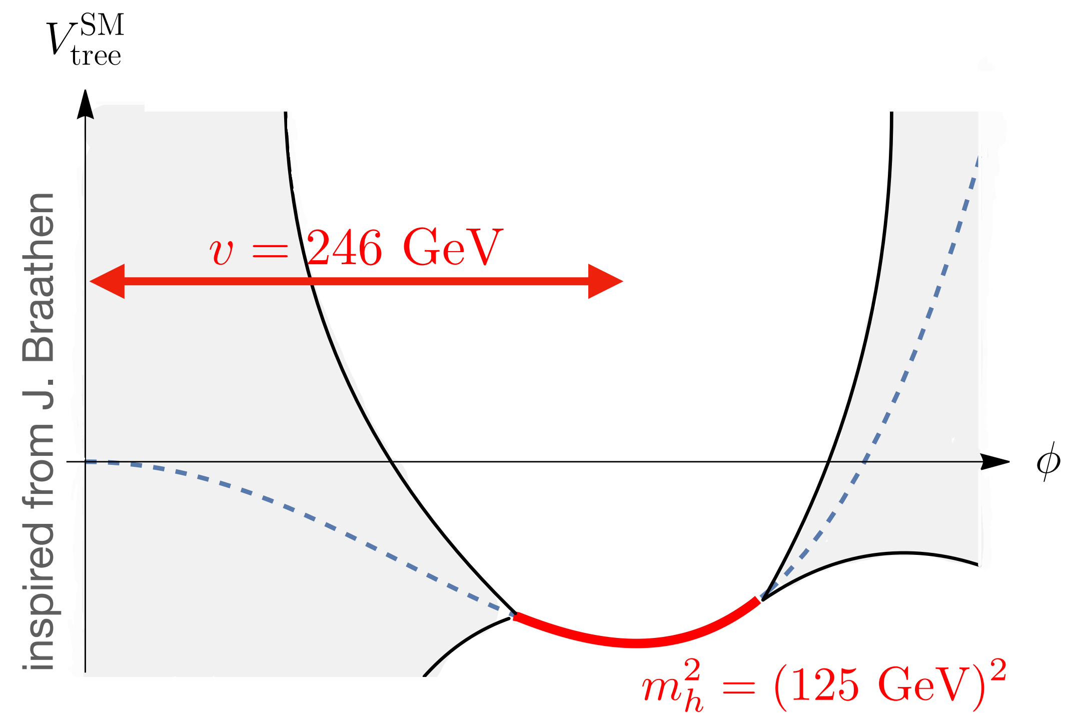
- Crucial questions about Electroweak Symmetry Breaking: What is the form of the Higgs potential?



So far we know:

→ location of minimum

→ curvature around minimum



SM Potential:

$$V(\Phi) = \lambda(\Phi^\dagger\Phi)^2 - \mu^2\Phi^\dagger\Phi$$

$$\supset -\lambda v H^3 - \frac{\lambda}{4}H^4$$

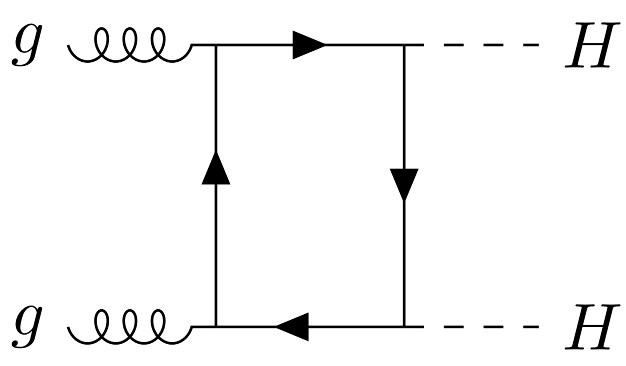
BSM theories → more complicated shapes

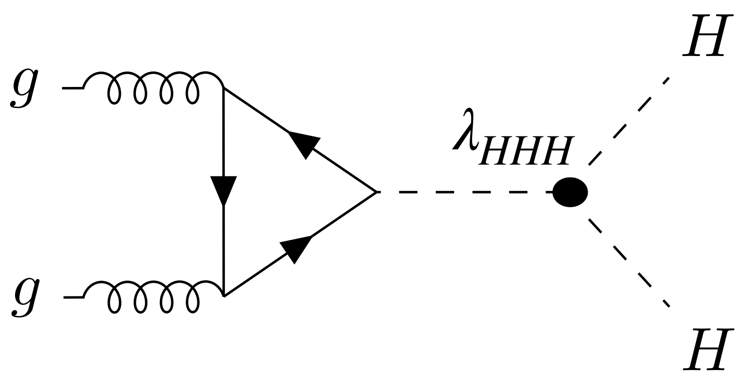
Very challenging experimentally

requires trilinear and quartic Higgs self-couplings

Trilinear Higgs coupling: experimental status

- Experimental bounds on signal strength from HH production: $\mu_{HH} < 2.4$ (ATLAS)





λ_{HHH} enters at LO order
→ most direct probe

- Signal strength translated to limit on $$\kappa_3 = \frac{\lambda_{HHH}}{\lambda_{HHH}^{\mathrm{SM},(0)}}$$

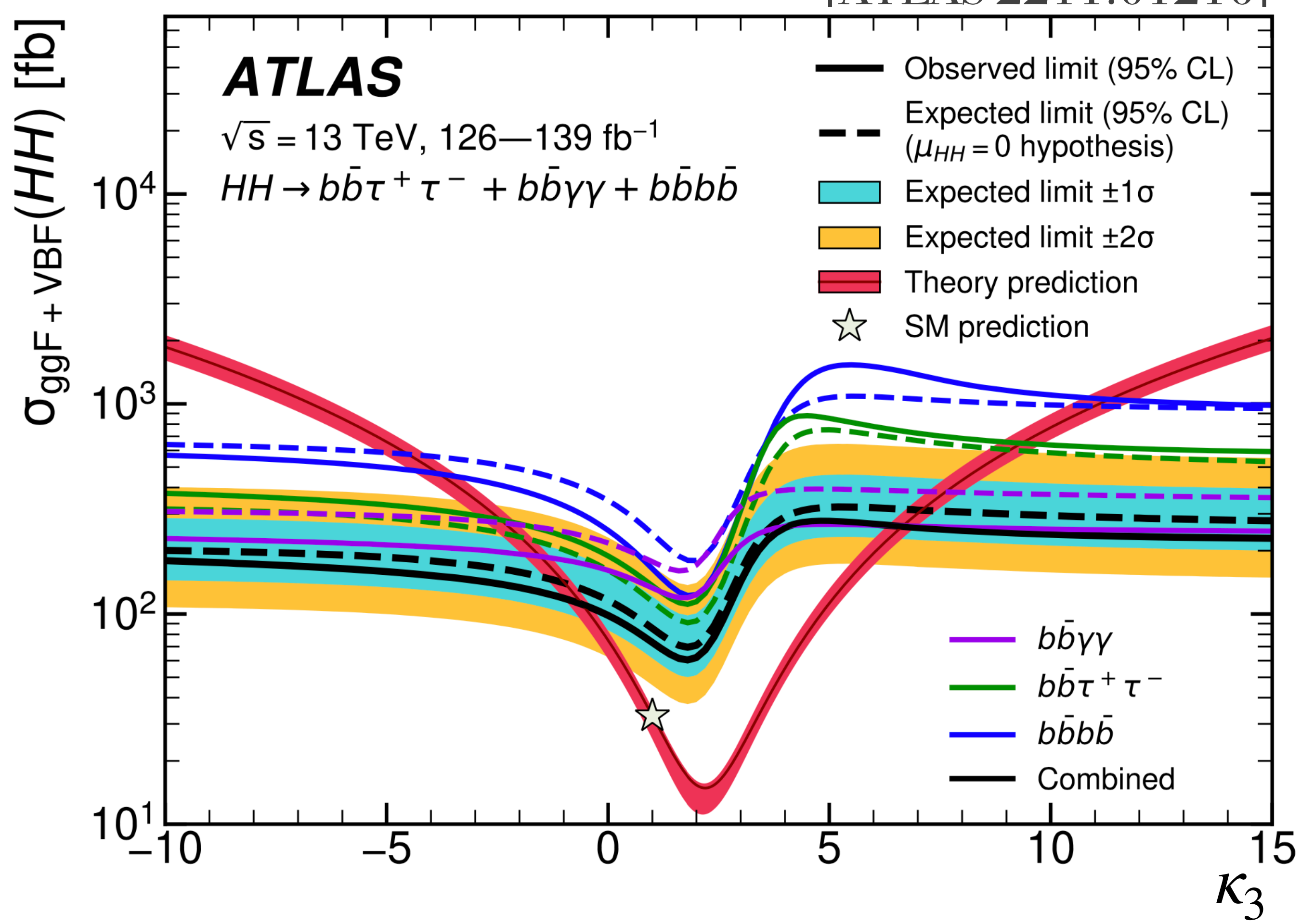
Fixing other couplings fixed to SM:

CMS: $[-1.2, 6.5]$
ATLAS: $[-0.4, 6.3]$
} @95 % CL

↓ Including single Higgs

[CMS 2207.00043]
[ATLAS 2211.01216]

[ATLAS 2211.01216]

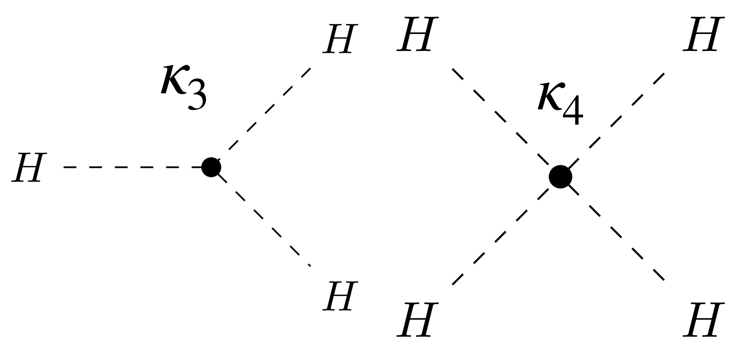


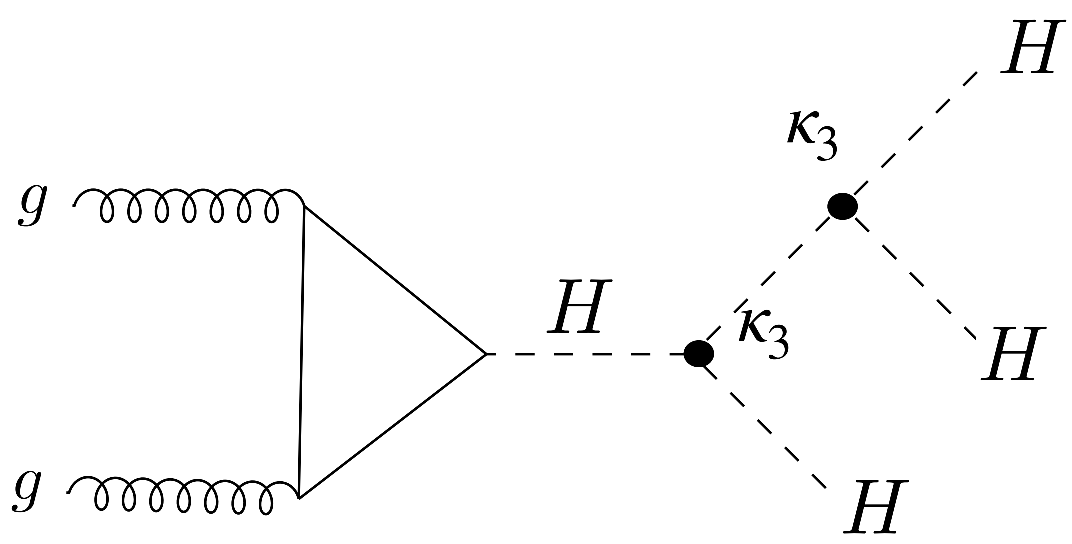
Triple Higgs production

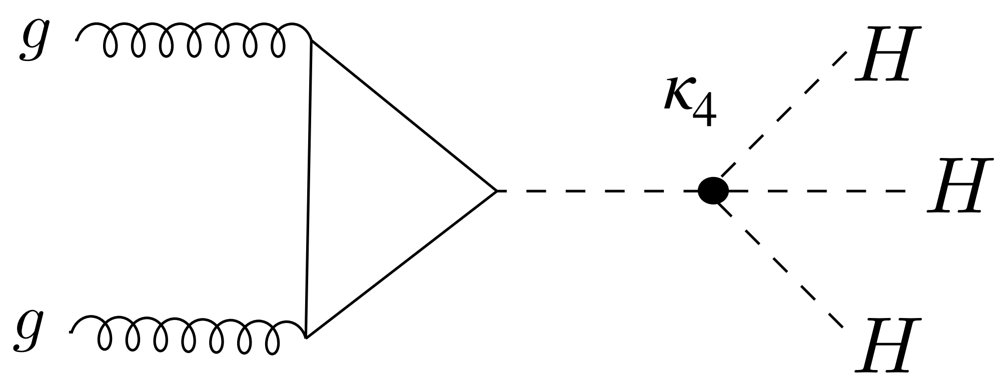
early works: [Plehn, Rauch, `05]
[Binoth, Karg, Kauer, Rückl, `06]

- Additional source of information → HHH production

- Dependence on both trilinear κ_3 and quartic κ_4







- Is it possible to obtain bounds on κ_3 and κ_4 from HHH production beyond theoretical bounds from perturbative unitarity?

- How big can deviations in κ_4 be in BSM theories from SM value ($= 1$)

- Is there potential to improve κ_3 constraints from HH production?

Perturbative unitarity and Higgs couplings

- Process relevant for κ_3, κ_4 is $HH \to HH$ scattering (see also [Liu et al `18])

- Jacob-Wick expansion allows to extract the zeroth partial wave:

$$a^0_{ii} = \frac{3M_H^2\sqrt{s^2-4M_H^2 s}}{32\pi s(s-M_H^2)v^2}\left[\kappa_4(s-M_H^2) - 3\kappa_3^2 M_H^2 + \frac{6\kappa_3^2 M_H^2(s-M_H^2)}{s-4M_H^2}\log\left(\frac{s}{M_H^2}-3\right)\right]$$

- Tree level unitarity:

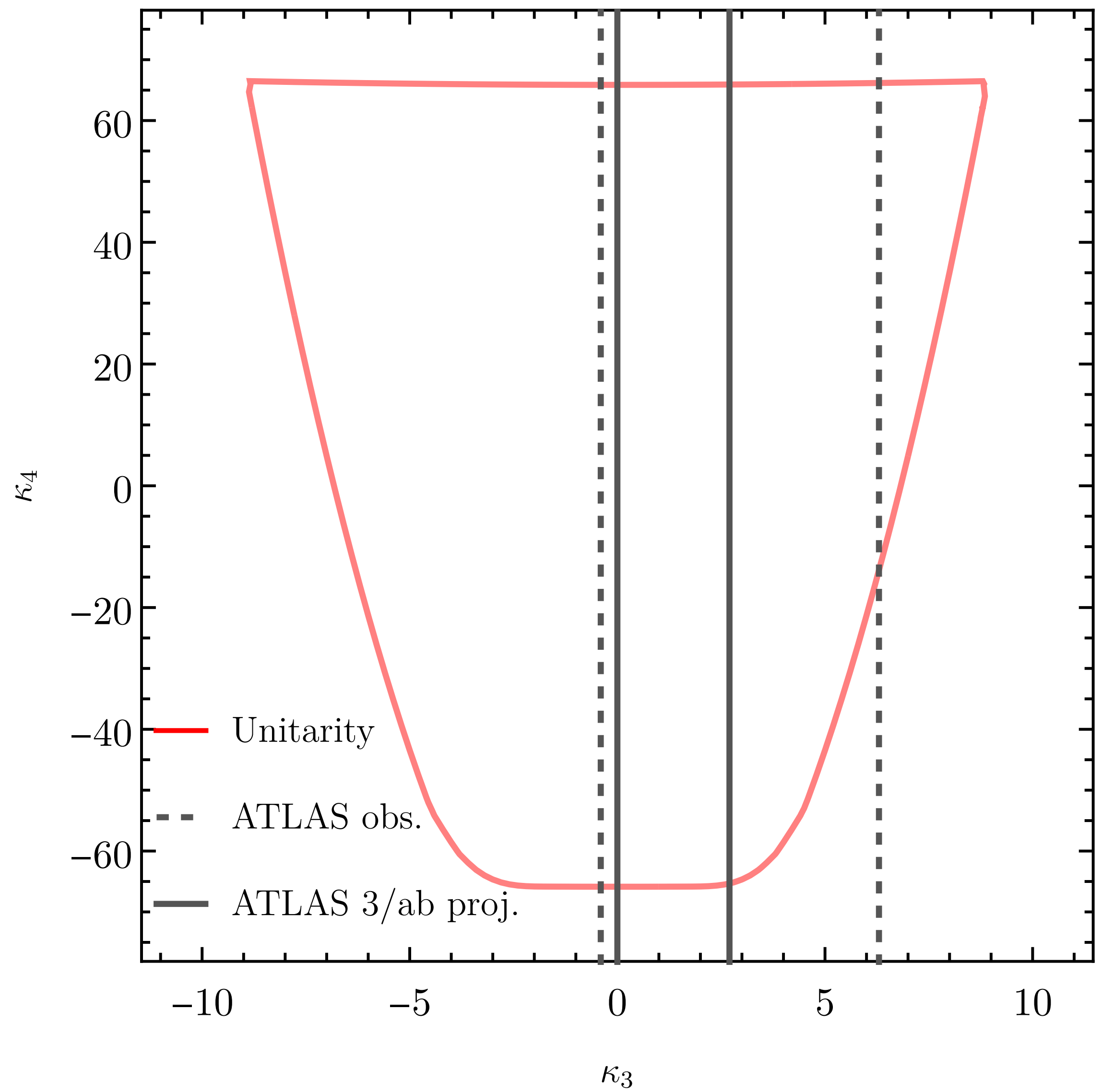
$$\mathrm{Im}a^0_{ii} \geq |a^0_{ii}|^2 \implies |\mathrm{Re}a^0_{ii}| \leq \frac{1}{2}$$

95 % CL

ATLAS current bounds: $[-0.4, 6.3]$

CMS & ATLAS HH projections: $[0.1, 2.3]$

[ATLAS 2211.01216]
[CERN Yellow Rep. 1902.00134]



Extension of SM potential by operators

Linear power expansion for higher order terms in Λ^{-1} orders:

[Boudjema, Chopin `96]
[Maltoni, Pagani, Zhao `18]

$$V_{\text{BSM}} = \frac{C_6}{\Lambda^2}\left(\Phi^\dagger\Phi - \frac{v^2}{2}\right)^3 + \frac{C_8}{\Lambda^4}\left(\Phi^\dagger\Phi - \frac{v^2}{2}\right)^4 + \ldots$$

Contributions to κ_3, κ_4:

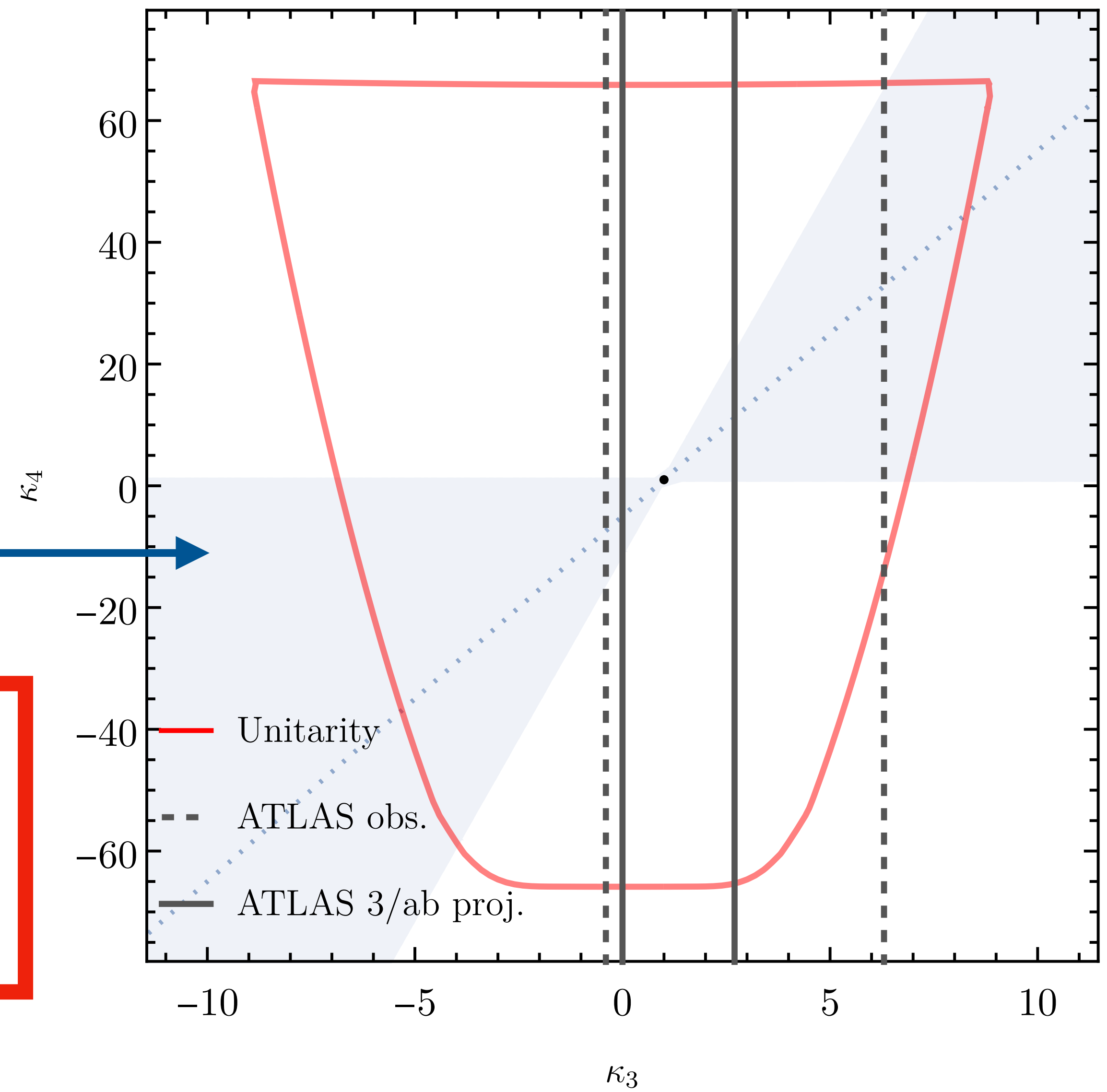
$$(\kappa_3 - 1) = \frac{C_6 v^2}{\lambda\Lambda^2}\,,$$

$$(\kappa_4 - 1) = \frac{6C_6 v^2}{\lambda\Lambda^2} + \frac{4C_8 v^4}{\lambda\Lambda^4}$$

vanishing dimension-8 →

$$\simeq 6(\kappa_3 - 1) + \mathcal{O}\left(\frac{1}{\Lambda^4}\right)$$

Shaded region: $\frac{4C_8 v^4}{\lambda\Lambda^4} < \frac{6C_6 v^2}{\lambda\Lambda^2}$ →



Electroweak Chiral Lagrangian (HEFT):
Higgs introduced as singlet and κ_3 and κ_4 are **free parameters** → probes **non-linearity**

Extension of SM potential by operators

Linear power expansion for higher order terms in Λ^{-1} orders:

[Boudjema, Chopin `96]
[Maltoni, Pagani, Zhao `18]

$$V_{\mathrm{BSM}} = \frac{C_6}{\Lambda^2}\left(\Phi^\dagger\Phi - \frac{v^2}{2}\right)^3 + \frac{C_8}{\Lambda^4}\left(\Phi^\dagger\Phi - \frac{v^2}{2}\right)^4 + \ldots$$

Contributio

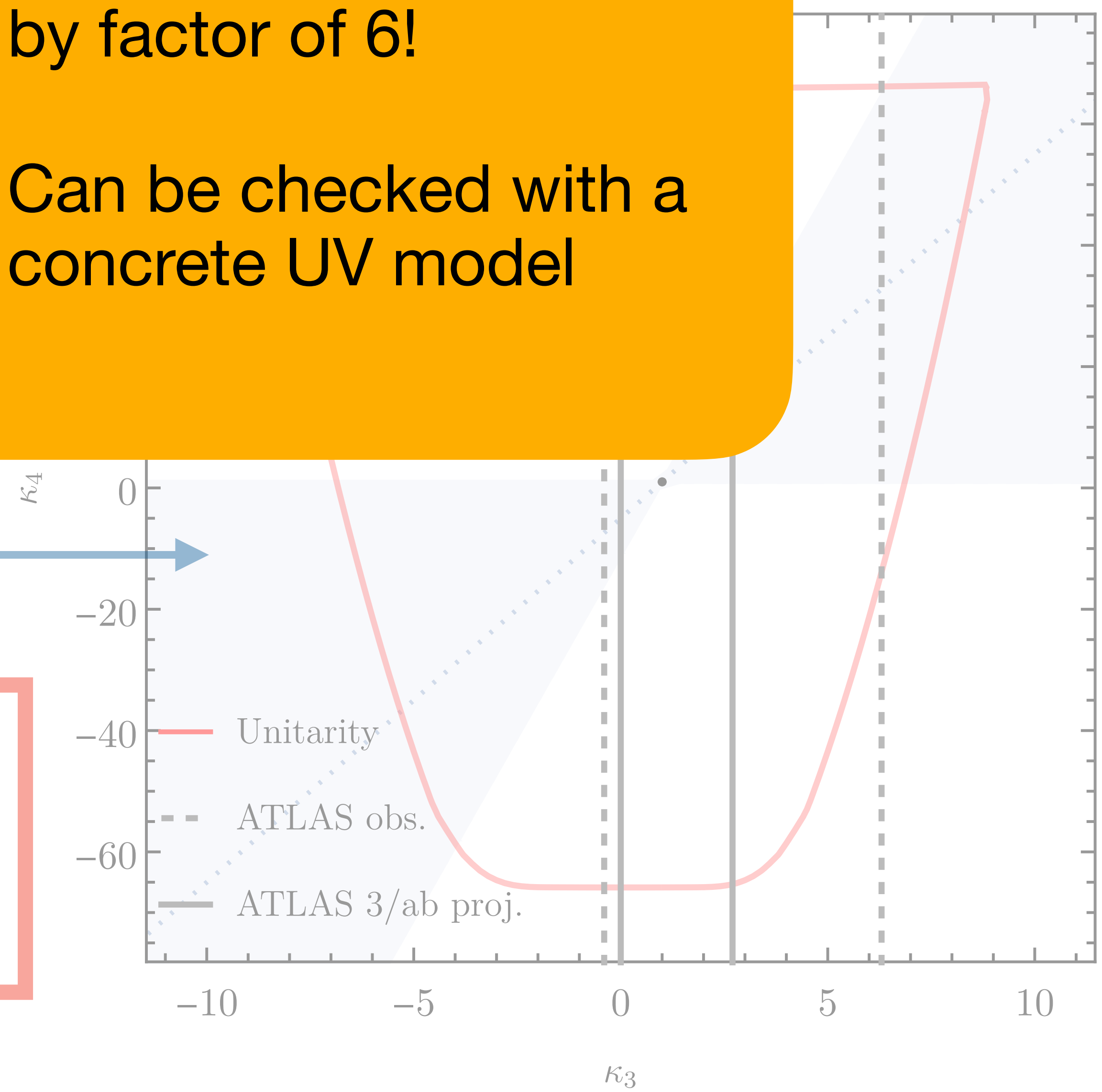
$$(\kappa_3 - 1) = \frac{C_6 v^2}{\lambda\Lambda^2}\,,$$

$$(\kappa_4 - 1) = \frac{6C_6 v^2}{\lambda\Lambda^2} + \frac{4C_8 v^4}{\lambda\Lambda^4} \simeq 6(\kappa_3 - 1) + \mathcal{O}\left(\frac{1}{\Lambda^4}\right)$$

- Deviation in κ_4 enhanced by factor of 6!
- Can be checked with a concrete UV model

vanishing dimension-8

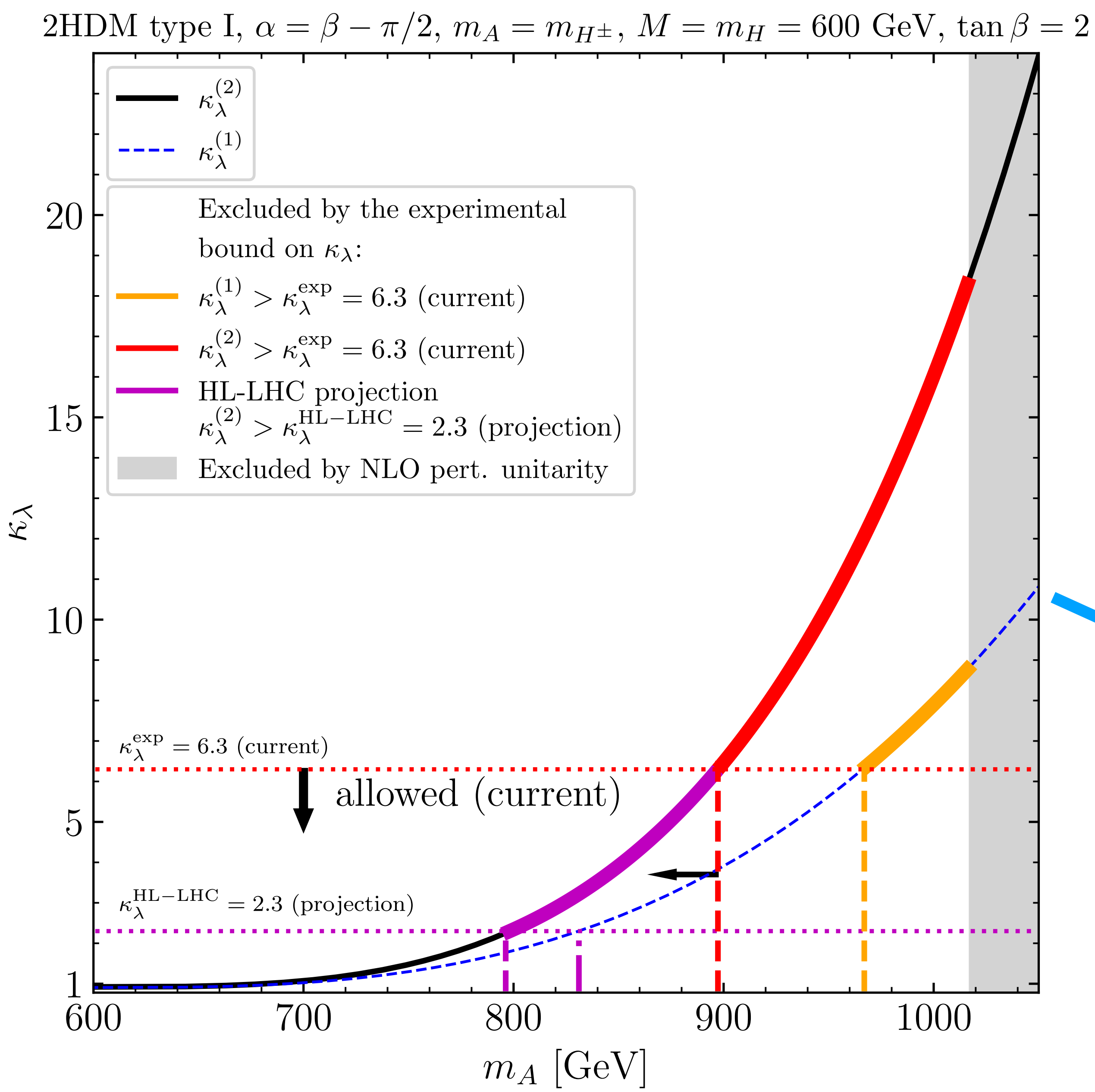
Shaded region: $\frac{4C_8 v^4}{\lambda\Lambda^4} < \frac{6C_6 v^2}{\lambda\Lambda^2}$



Electroweak Chiral Lagrangian (HEFT):
Higgs introduced as singlet and κ_3 and κ_4 are **free parameters** →probes **non-linearity**

BSM example: 2HDM

- Consider the 2HDM as an example
- Prediction for κ_3 up to two-loop level: [Bahl, Braathen, Weiglein `22]



Sizeable deviations for κ_3

→ even current experimental constraints can exclude otherwise allowed regions

Reproduce one-loop result (with FeynArts, FormCalc, LoopTools) and check the behaviour of κ_4

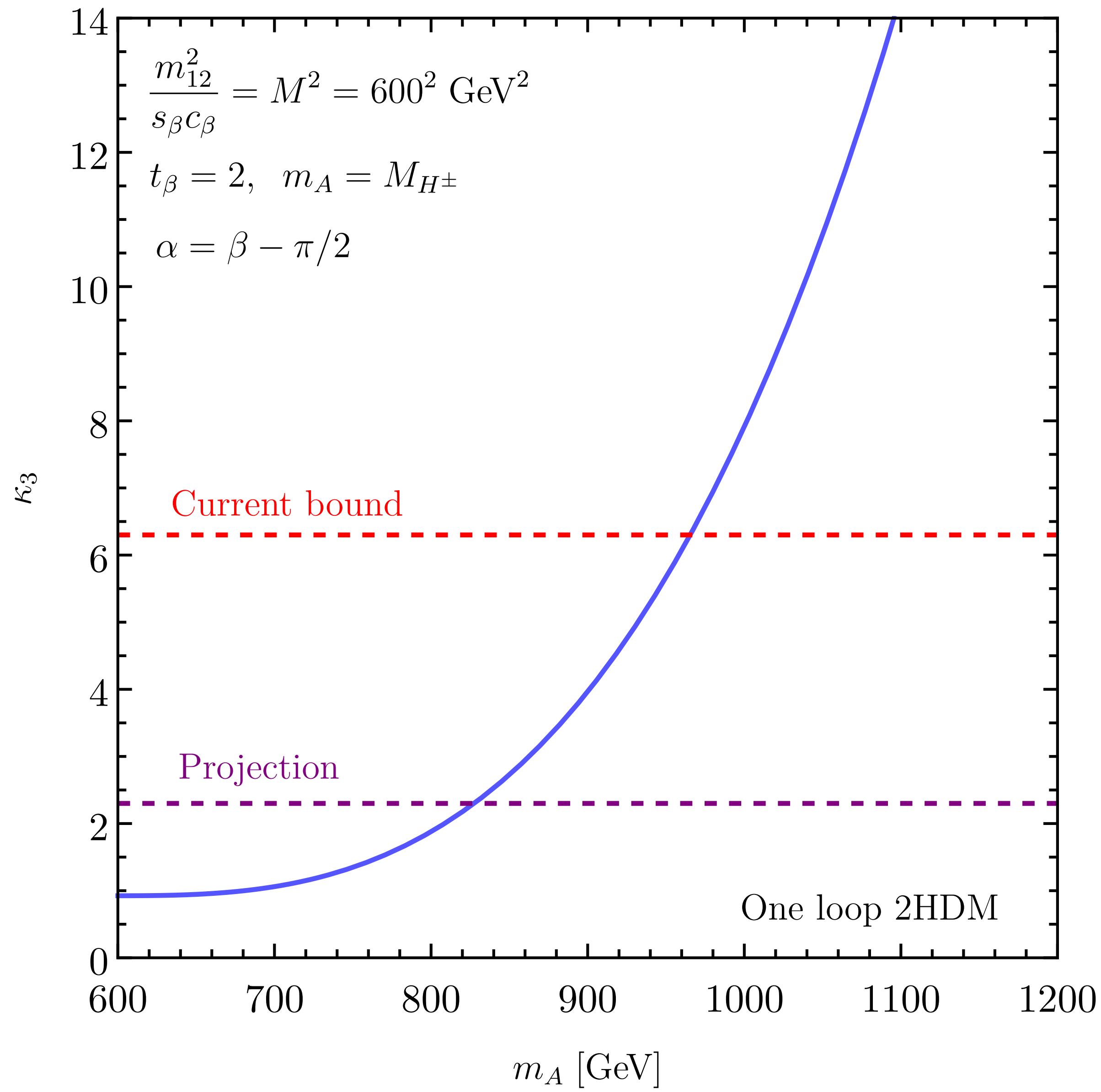
Model example: 2HDM - trilinear vs quartic

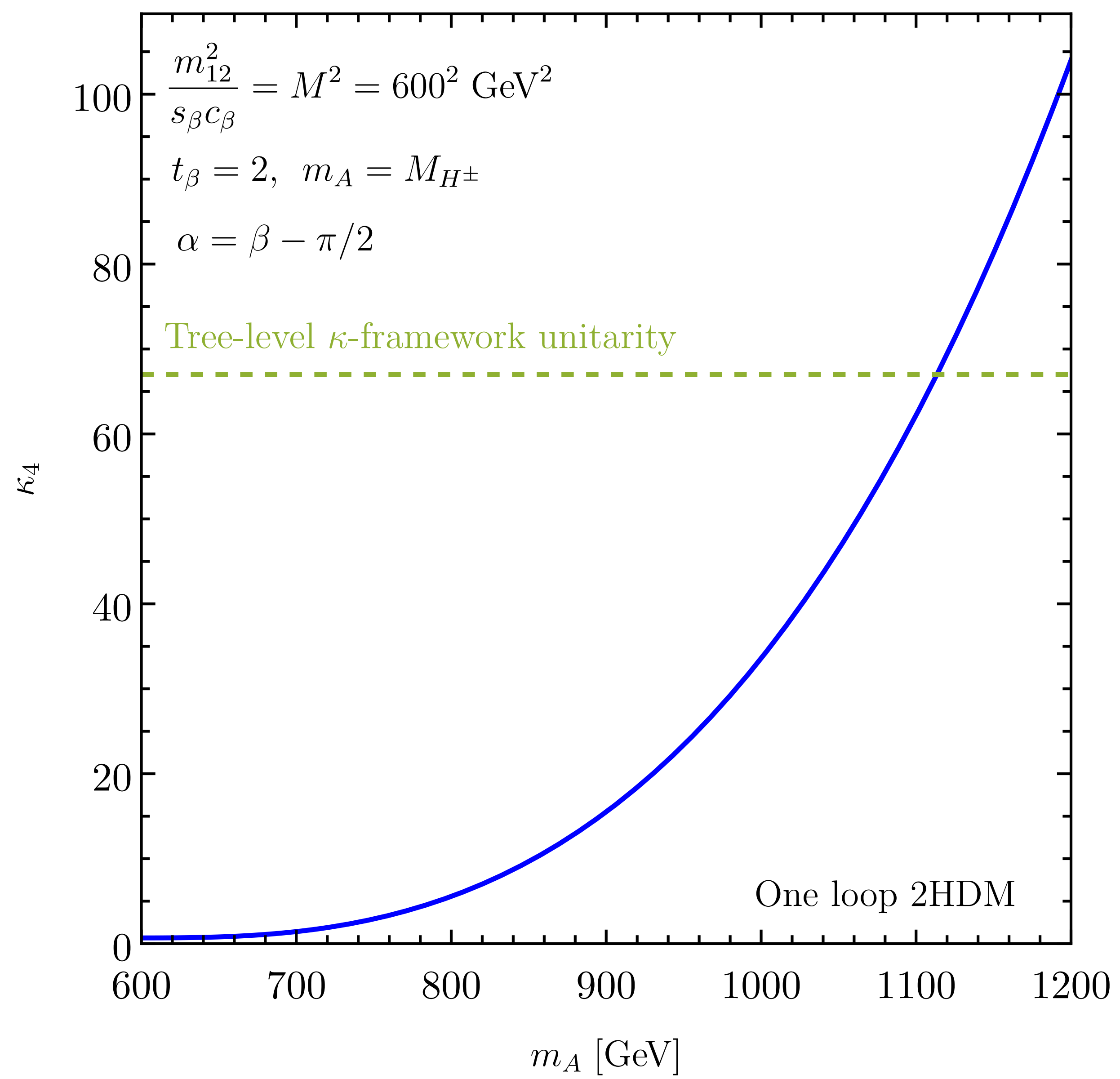
- Same benchmark Point of [Bahl, Braathen, Weiglein `22] → cross-check κ_3 result

- Expectedly deviations in κ_3 induce sizeable deviations in κ_4

$$\kappa_i = \frac{\Gamma_i^{(0)} + \hat{\Gamma}_i^{(1)}}{\Gamma_{\mathrm{SM},i}^{(0)}}$$

$$i \in \{3H, 4H\}$$





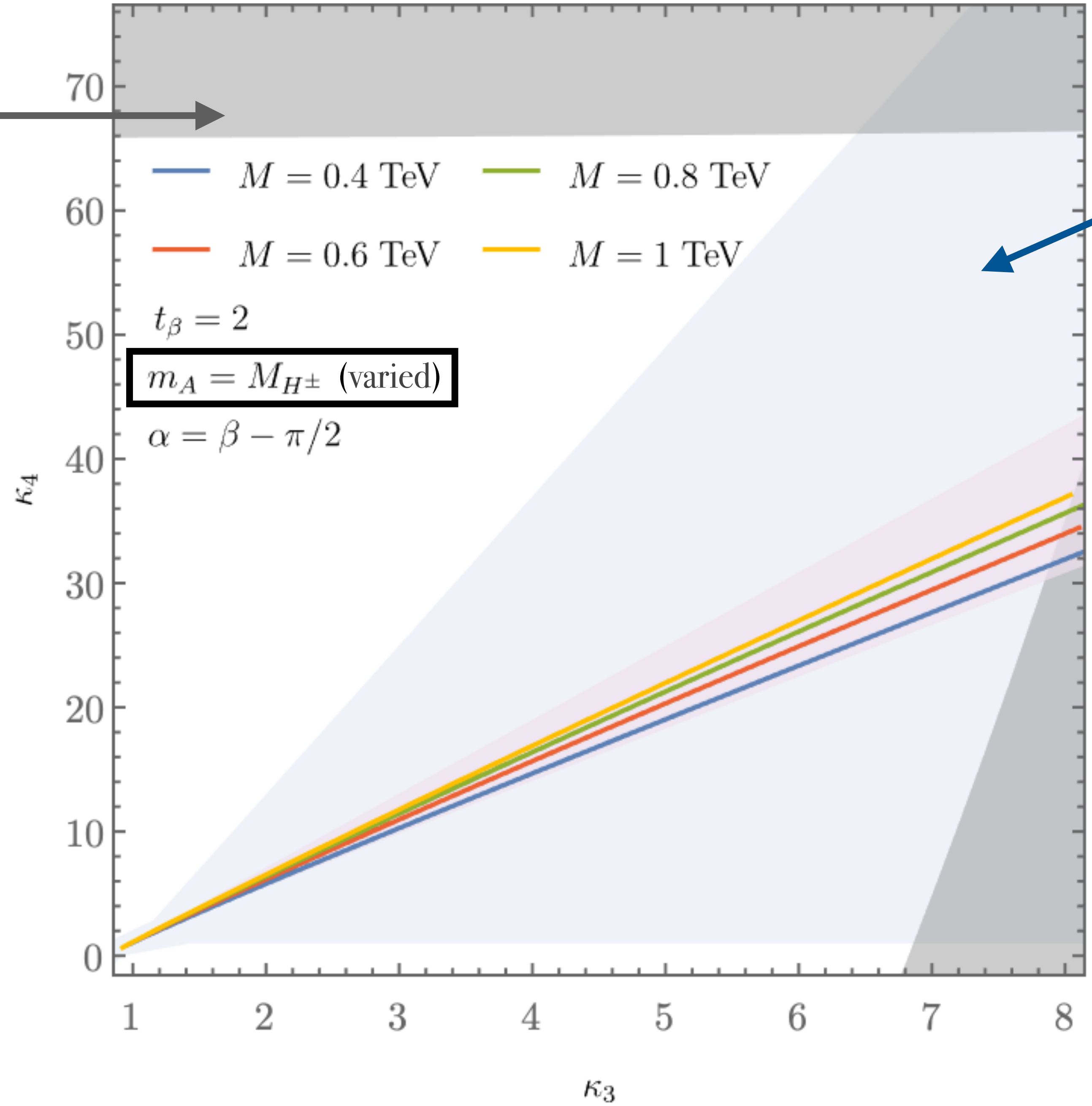
Model example: 2HDM - trilinear vs quartic

- Same benchmark Point of [Bahl, Braathen, Weiglein `22] → cross-check κ_3 result
- Expectedly deviations in κ_3 induce sizeable deviations in κ_4

$$\kappa_i = \frac{\Gamma_i^{(0)} + \hat{\Gamma}_i^{(1)}}{\Gamma_{\mathrm{SM},i}^{(0)}}$$

$$i \in \{3H, 4H\}$$

tree-level κ-framework perturbative unitarity



well-behaved perturbative expansion from before

Prospects for the HL-LHC

- Small rates at LHC

Need dominant production & decays

- gluon fusion
- BRs:

$$\mathrm{BR}(H \to b\bar{b}) = 0.584$$
$$\mathrm{BR}(H \to \tau^+\tau^-) = 6.627 \times 10^{-2}$$

- Focus on $6b$ and $4b2\tau$ final states { studied at FCC energies: [Fuks, Kim, Lee `17] [Papaefstathiou, Xolocotzi, Zaro `19]

- Assume 5 and 3 tagged b-quarks, respectively

Backgrounds:

$6b$: dominant QCD contributions (see also [Papaefstathiou, Robens, Xolocotzi `21])

$4b2\tau$: $W^+W^-b\bar{b}b\bar{b},\ Zb\bar{b}b\bar{b},$
$t\bar{t}(H \to \tau\tau),\ t\bar{t}(H \to b\bar{b}),$
$t\bar{t}(Z \to \tau\tau),\ t\bar{t}(Z \to b\bar{b}),\ t\bar{t}t\bar{t}$

Event generation and pre-selection

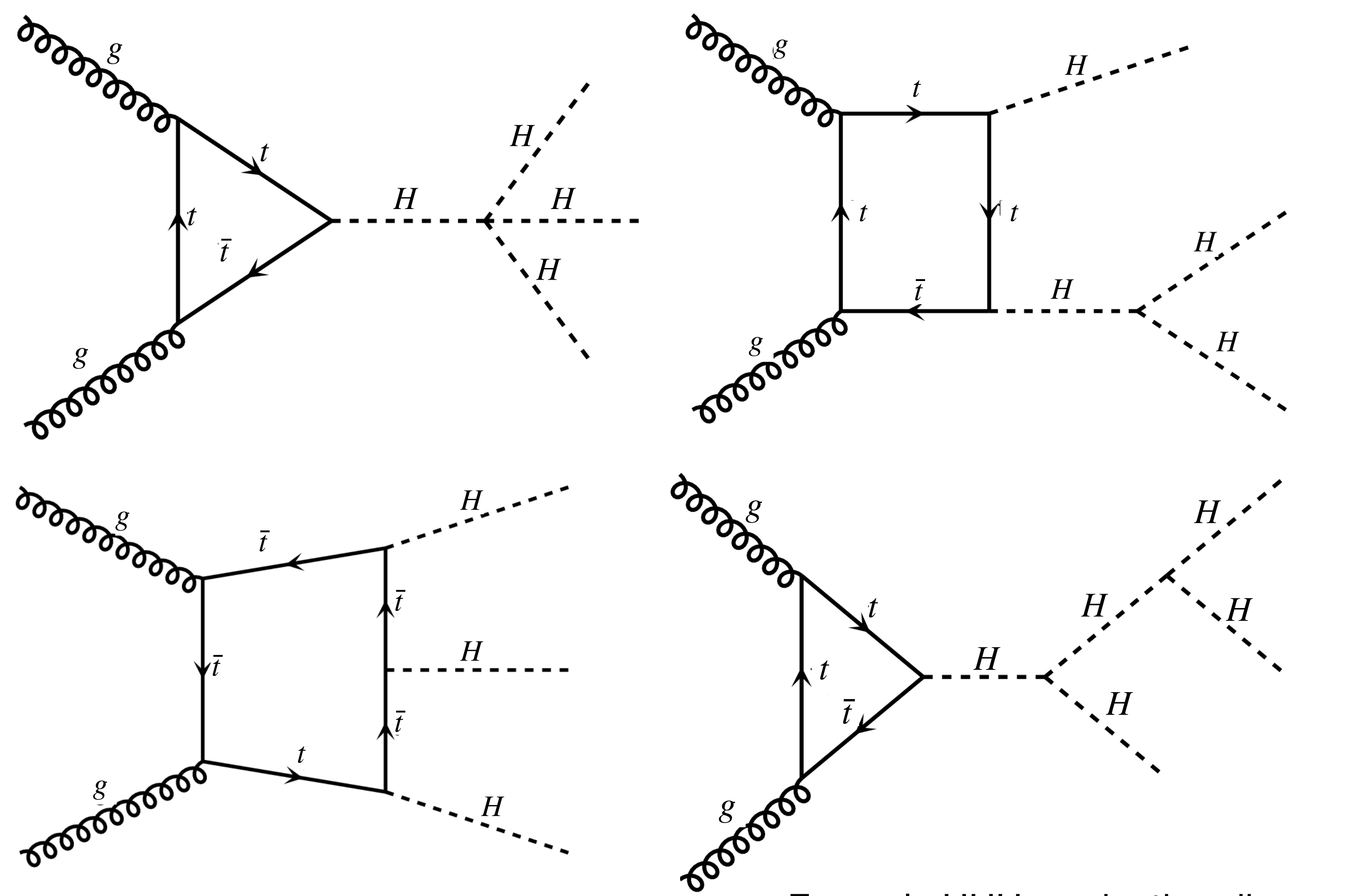
- Events generated with MadGraph5_aMC@NLO

- Higgs states decayed with MadSpin

(conservative) background K-factor of 2

signal K-factor of 1.7

[Florian, Fabre, Mazzitelli` 20]



Example HHH production diagrams

Pre-selection cuts:

Invariant mass of final states: $\gtrsim 350$ GeV

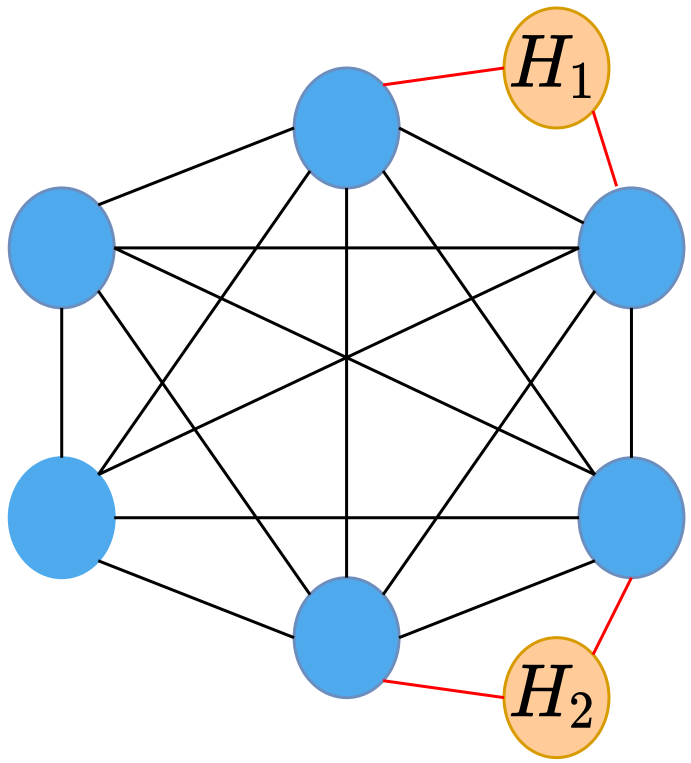
At least one pair of tagged states with $m_{ij} \in [110,140]$

$p_T(b) > 30$ GeV $\quad p_T(\tau) > 10$ GeV

$|\eta(\tau)| < 2.5 \quad |\eta(b)| < 2.5$

Graph Neural Network

- Add nodes for tagged b, τ and Missing Transverse Momentum
- Consider combinations of b-quarks and τ with reconstructed four-momentum $(p_i + p_j)$
- If $m_{ij} \in [100,150]$ (GeV) add extra node H_i
- Features for each node: $[p_T, \eta, \phi, m,$ PDGID$]$



Dataset with signal & background graphs → **GNN** (2 EdgeConv layers + linear) → $\left\{ P(\mathbf{S}), P(\mathbf{B}_1), \ldots, P(\mathbf{B}_N) \right\}$

- GNN trained on $(\kappa_3, \kappa_4) = (1,1)$ sample
- Signal regions selected with cuts on background scores

Graph Neural Network

- Consider combinations of b-quarks and τ with reconstructed four-momentum $(p_i + p_j)$

- If $m \in [100,150]$ (GeV) add extra node

- **<u>Assumption:</u>** Same GNN efficiency for other values of (κ_3, κ_4)
- Flat optimistic $80\,\%$ b-tagging and τ-tagging efficiency
- **<u>Significance</u>:**

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

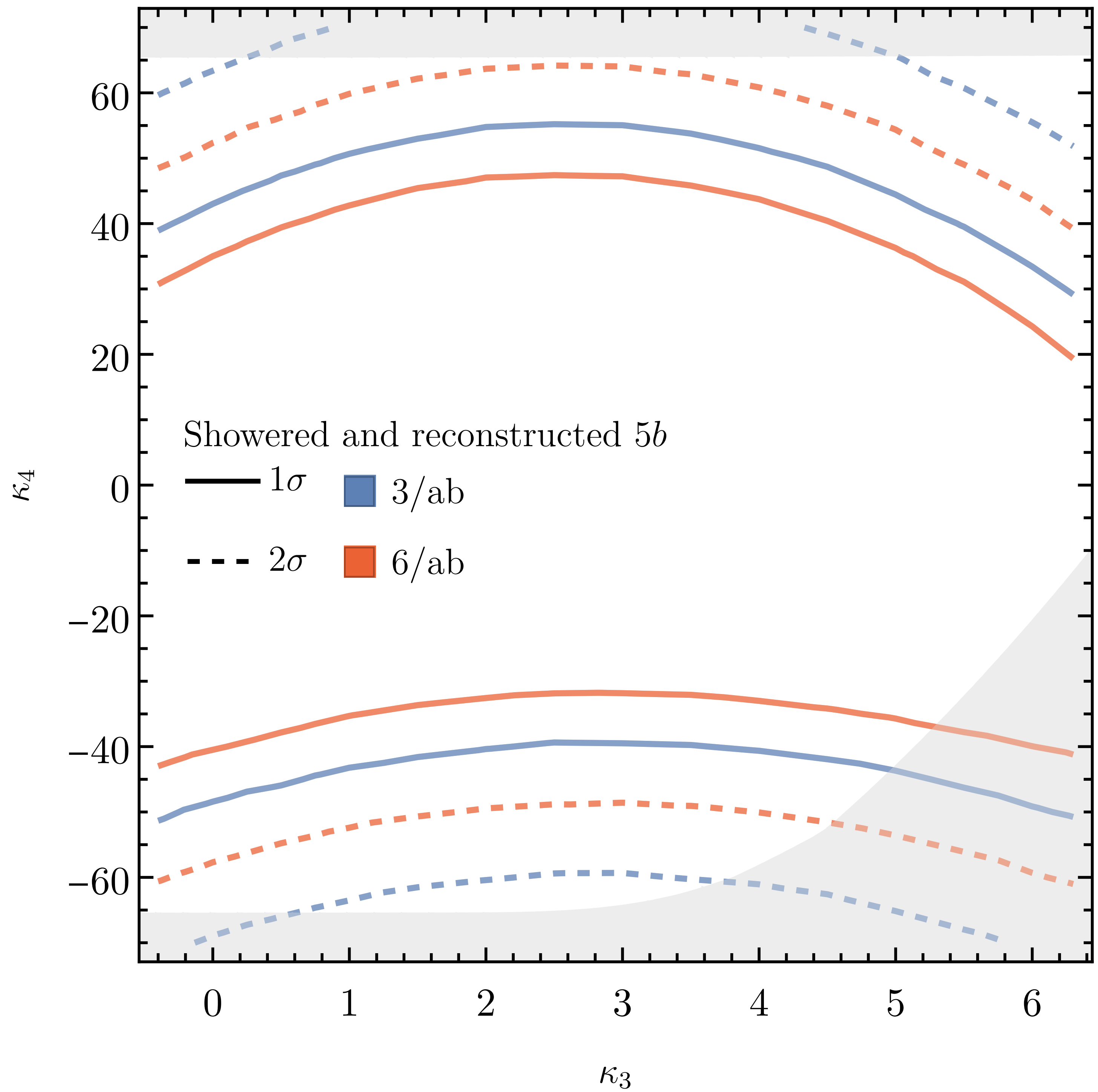
from [Cowan, Cranmer, Gross, Vitells ` 10]

- GNN trained on $(\kappa_3, \kappa_4) = (1,1)$ sample
- Signal regions selected with cuts on background scores

Results for $5b$

- Binary classification (only signal and QCD background)
- HL-LHC luminosity of 3/ab and ATLAS-CMS combined luminosity of 6/ab

Signal region selected with cut on background score $P[QCD] \lesssim 0.6\,\%$



Results for $3b2\tau$

- $3b2\tau$ more complicated due to multiple backgrounds → multi-class classification
- Train on backgrounds: $W^+W^-b\bar{b}b\bar{b},\ Zb\bar{b}b\bar{b},\ t\bar{t}(H \to \tau^+\tau^-)$

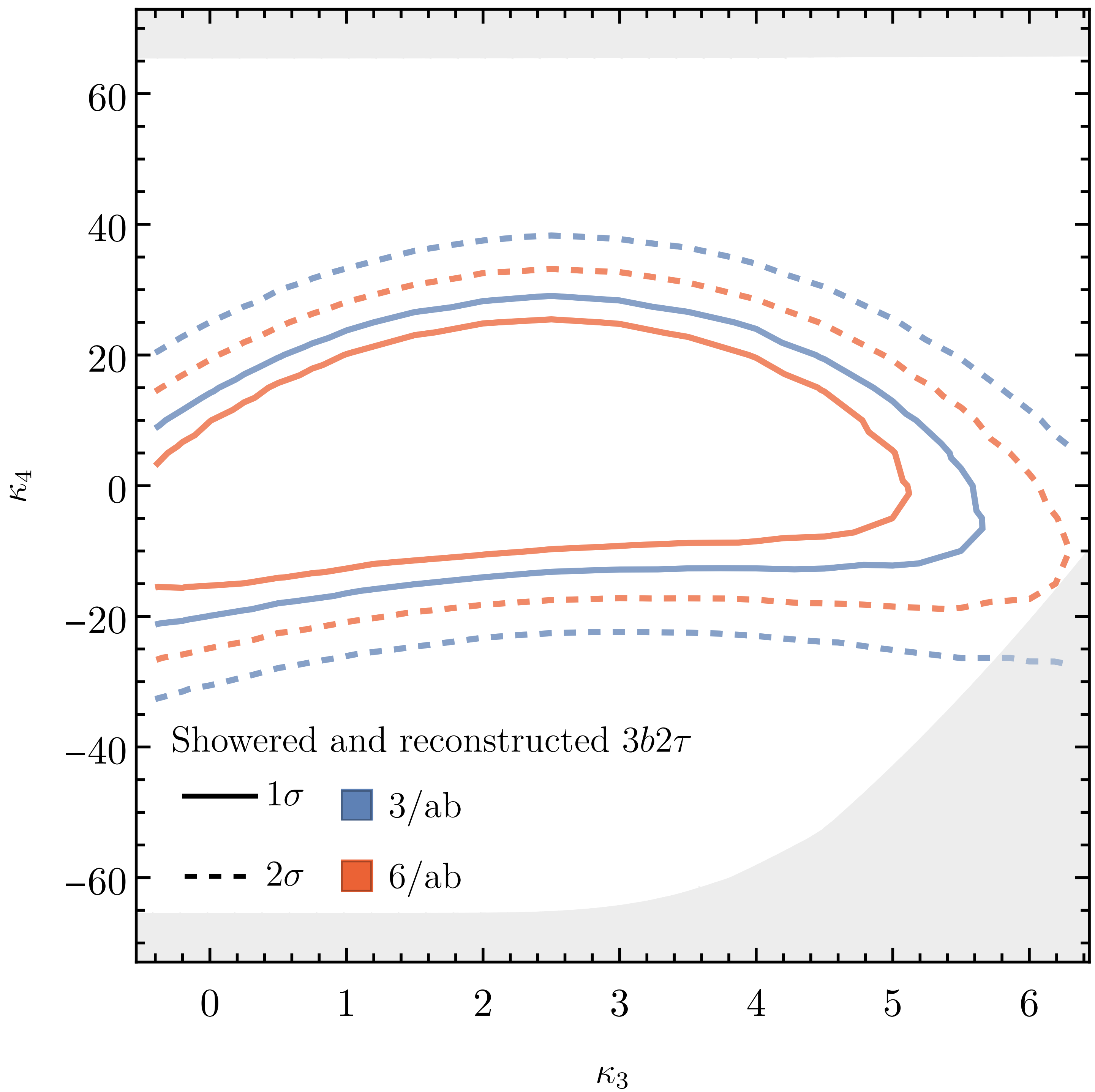
- Impose cuts on NN scores to reduce backgrounds:

$$P[W^+W^-b\bar{b}b\bar{b}] < 0.03, \quad P[Zb\bar{b}b\bar{b}] < 0.1, \quad P[t\bar{t}(H \to b\bar{b})] < 0.3$$

	σ(gen.)(fb)	σ(sel.)(fb)	σ(NN)(fb)
$tt(H \to \tau\tau)$	3.3	0.14	0.011
$WWbbbb$	27	4.0	7.1×10^{-3}
$tt(H \to bb)$	3.0	0.78	3.3×10^{-3}
$Zbbbb$	3.8	0.40	2.9×10^{-4}
$tt(Z \to bb)$	0.67	0.13	2.7×10^{-4}
$tttt$	0.33	0.080	1.8×10^{-4}
$tt(Z \to \tau\tau)$	4.1	0.073	9.0×10^{-5}

Results for $3b2\tau$

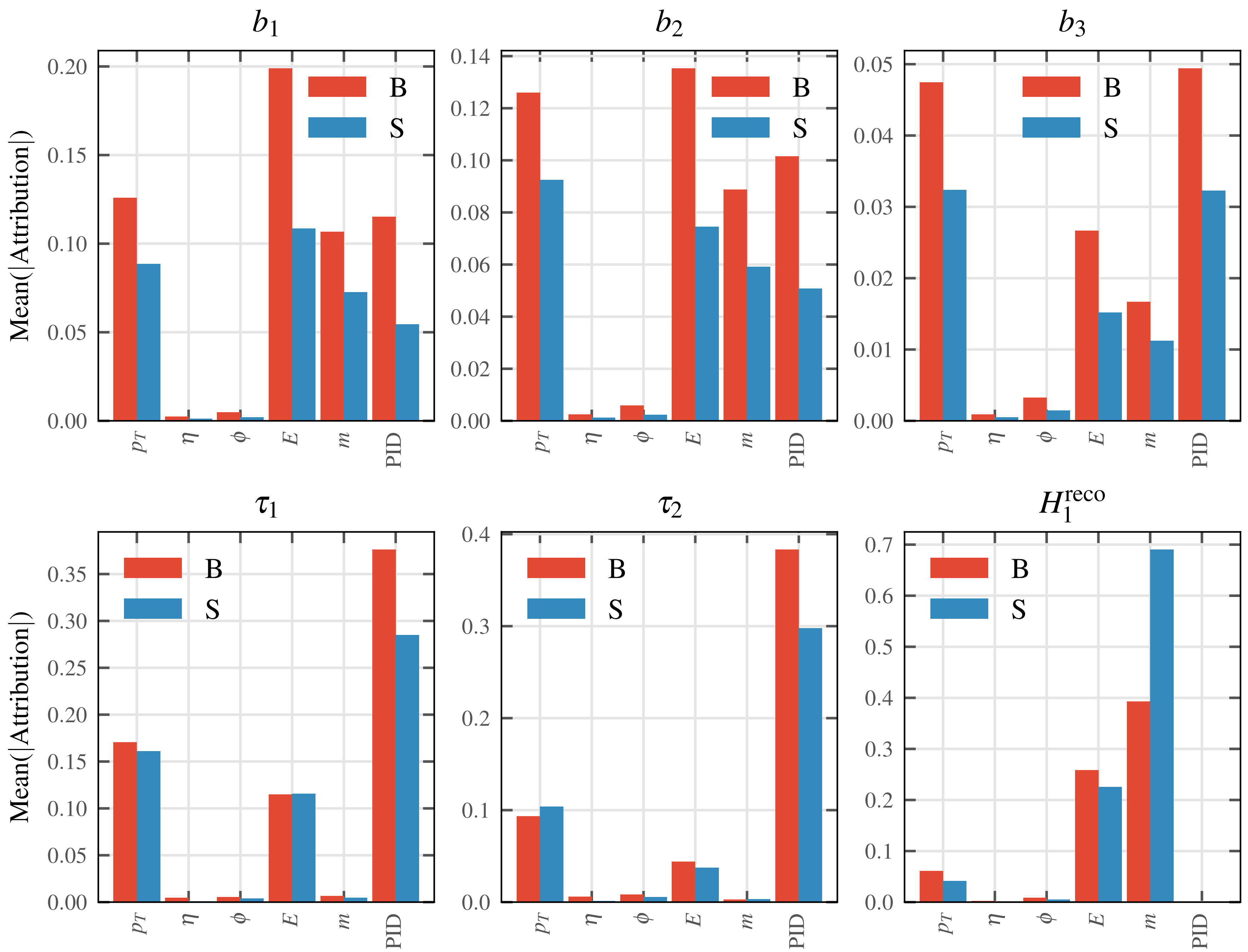
- $3b2\tau$ more complicated due to multiple backgrounds → multi-class classification
- Train on backgrounds: $W^+W^-b\bar{b}b\bar{b},\ Zb\bar{b}b\bar{b},\ t\bar{t}(H \to \tau^+\tau^-)$



Understanding the 'black box': NN interpretations

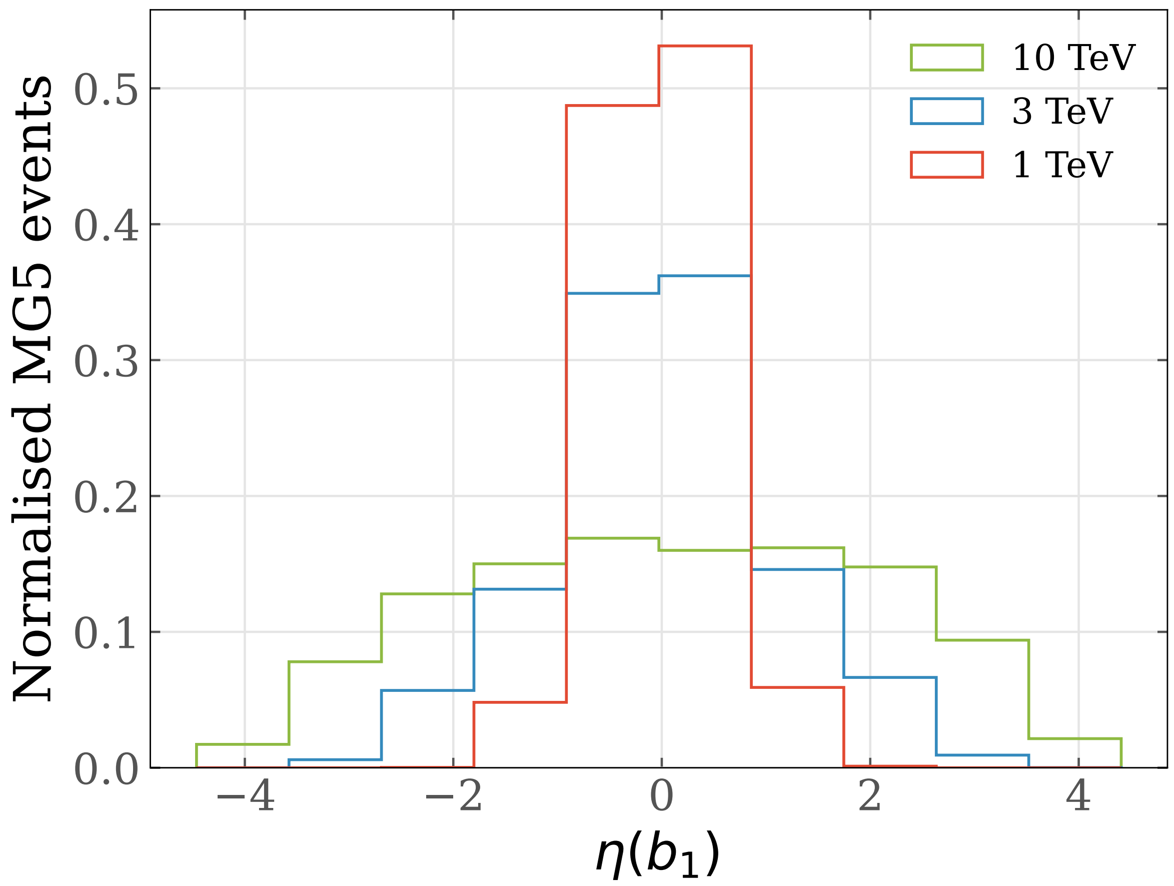
Which features are more important? Investigate with 'Integrated Gradients' method

- Tagged b-jets and τ nodes ordered by p_T
- 'Roughly' reconstructed Higgs nodes ordered by 'closeness' to 125 GeV
- p_T , E and PID more important than angular observables
- Higgs masses most important



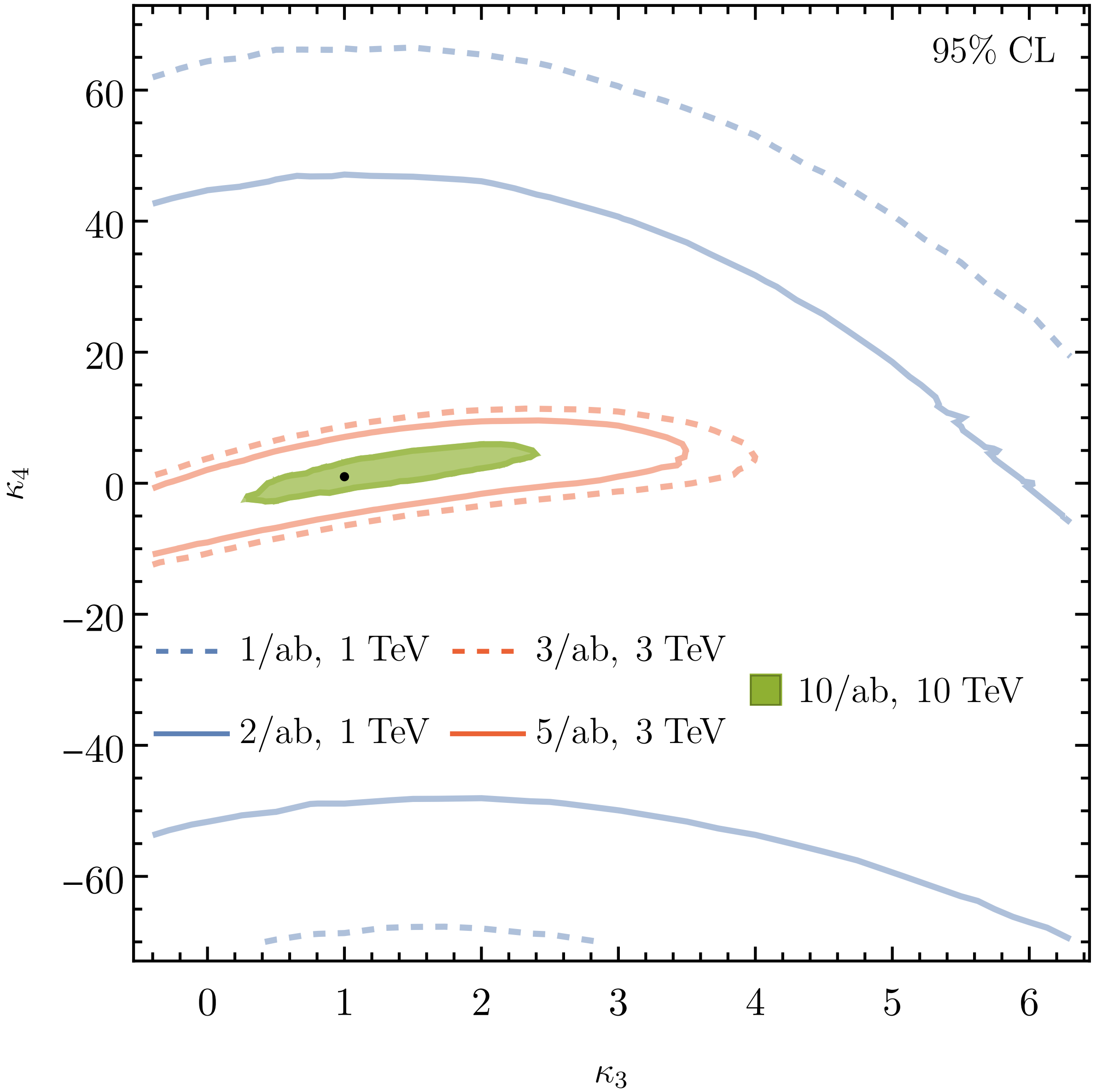
Lepton Colliders

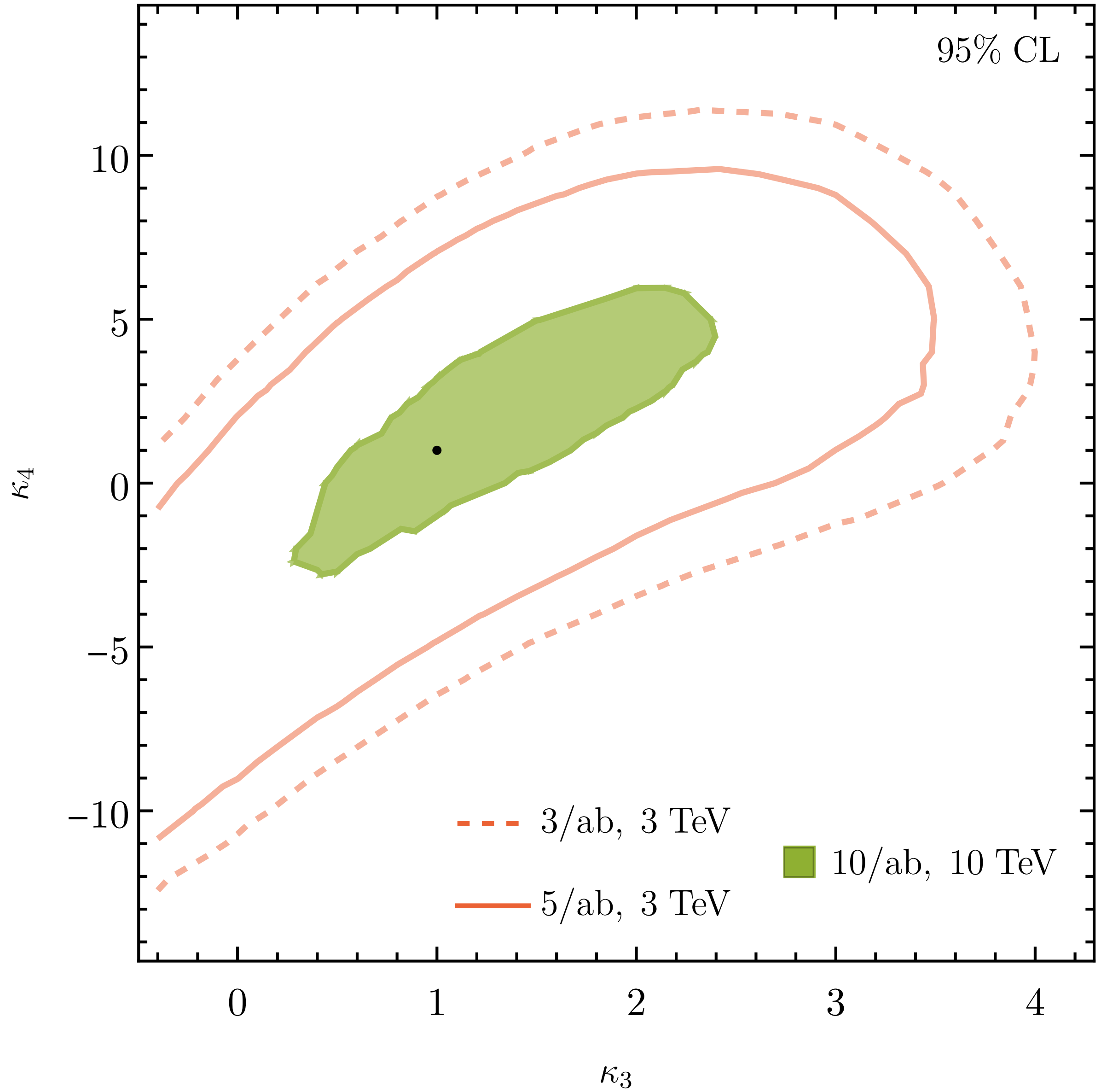
- Complete picture of $(\kappa_3, \kappa_4) \rightarrow$ lepton colliders?
- Inclusive $\ell\ell \rightarrow HHH + X$ analysis with $H \rightarrow b\bar{b}$
    - At least 5 tagged b-quarks with $p_T(b) > 30$ GeV
    - Tagging efficiency: $80\,\%$
- For high energies b-quarks are not only in the central part of detector $\rightarrow$ assume extended tagging capabilities: $|\eta| < 4$



Lepton Collider Results

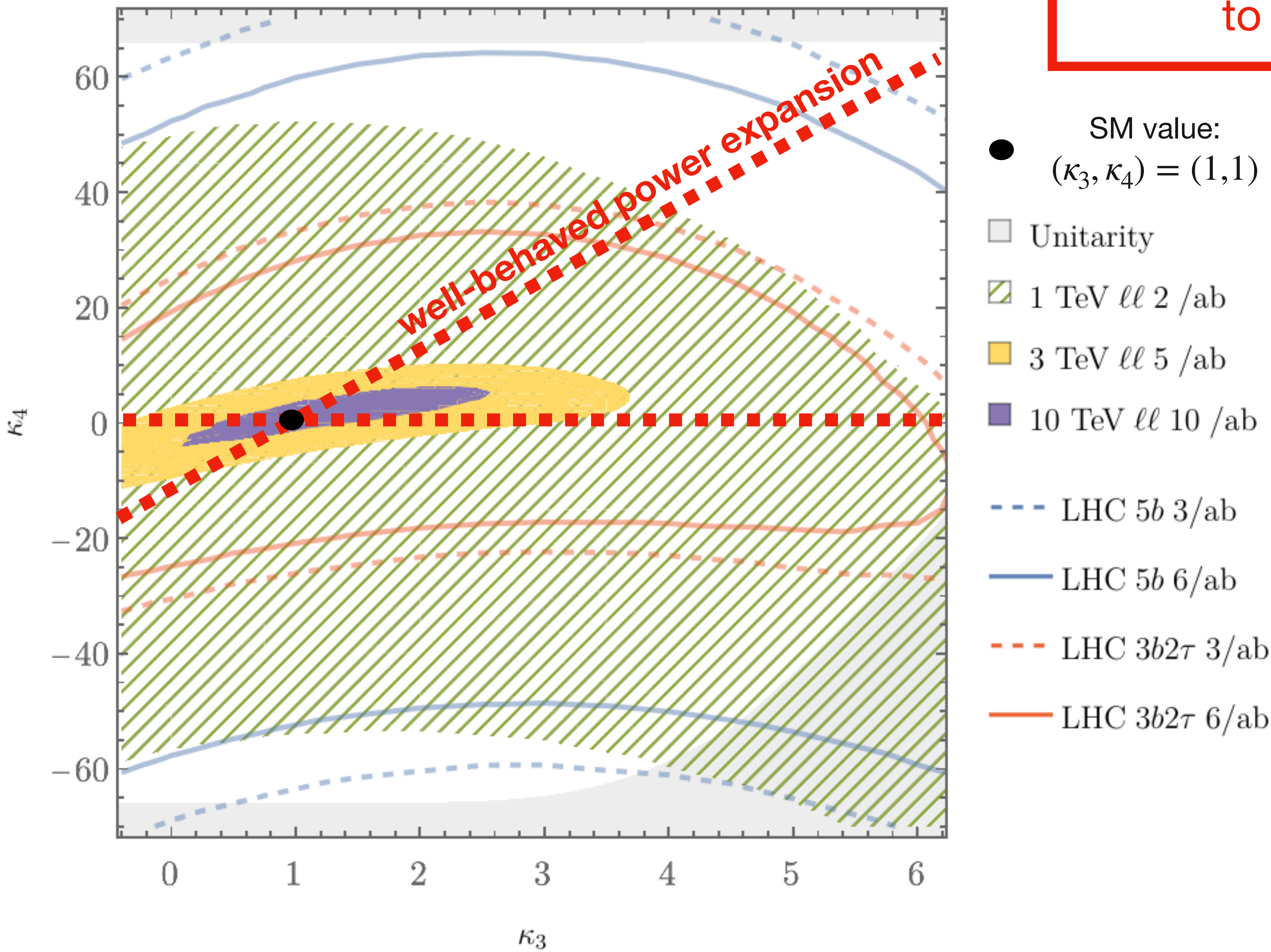
- Poissonian analysis: $\mu_{\rm up} = \frac{1}{2} F^{-1}_{\chi^2}\left[2(n+1); {\rm CL}\right]$ Assume no backgrounds!
- Results similar to other works with dedicated analyses for 1 and 3 TeV, e.g. [Maltoni, Pagani, Zhao ` 18]





HL-LHC vs. future lepton colliders

- HL-LHC can provide competitive results compared to 1 TeV collider
- High energy lepton collisions way more sensitive → **BUT** such machines more comparable to FCC



Conclusions

- If there is a sizeable deviation in κ_3, an even larger deviation in κ_4 is not unreasonable

  sizeable κ_4 deviations allowed by unitarity

- **GNNs** provide enhanced results at HL-LHC
  - HL-LHC should be able to probe regions allowed by unitarity
  - HHH not powerful enough to constrain κ_3 as well as di-Higgs bounds

  BUT can provide complementary information and be used in combination with di-Higgs

- HL-LHC competitive with 1 TeV lepton colliders but higher energies more sensitive

Thank you!

Backup

Experimental constraints on κ_3

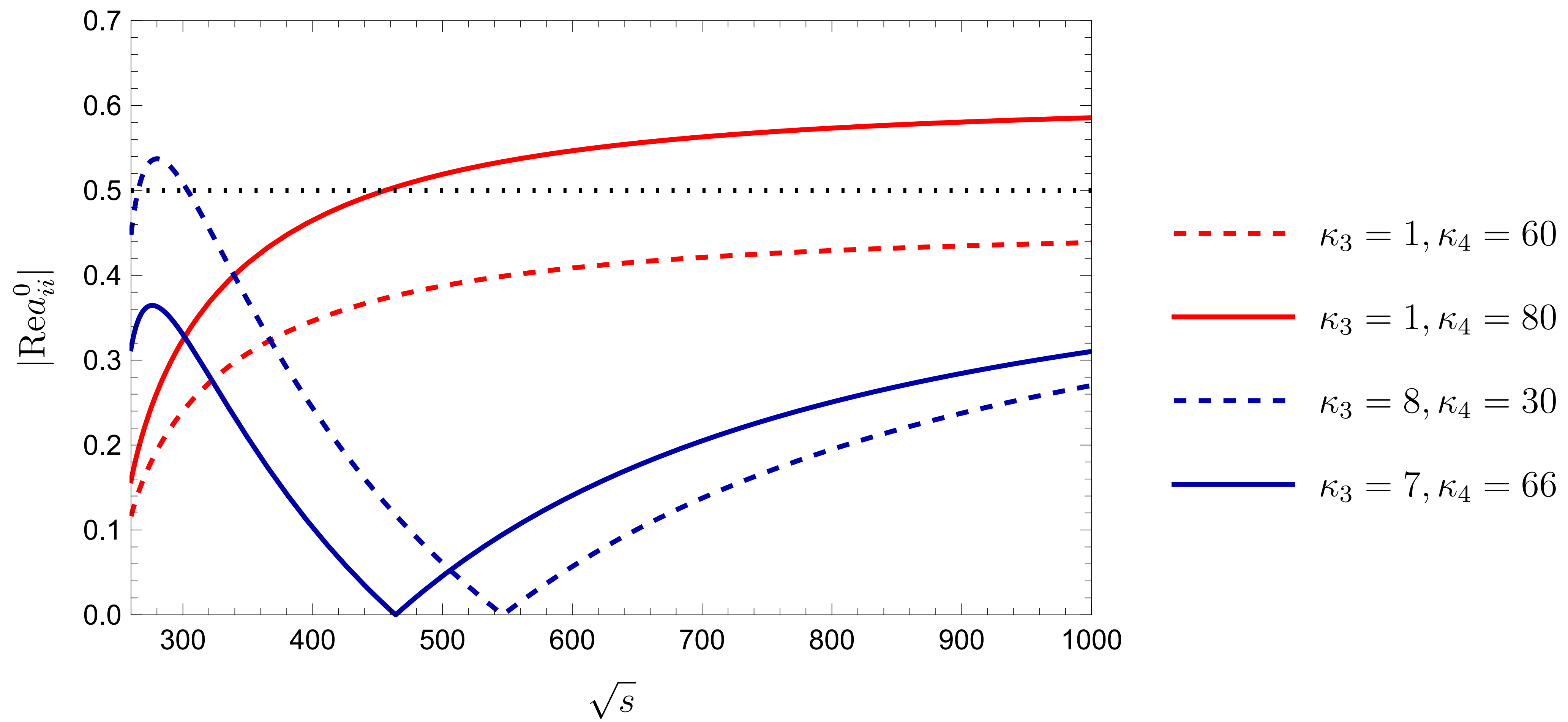
[ATLAS 2211.01216]

Combination assumption	Obs. 95% CL	Exp. 95% CL	Obs. value$^{+1\sigma}_{-1\sigma}$
HH combination	$-0.6 < \kappa_\lambda < 6.6$	$-2.1 < \kappa_\lambda < 7.8$	$\kappa_\lambda = 3.1^{+1.9}_{-2.0}$
Single-H combination	$-4.0 < \kappa_\lambda < 10.3$	$-5.2 < \kappa_\lambda < 11.5$	$\kappa_\lambda = 2.5^{+4.6}_{-3.9}$
HH+H combination	$-0.4 < \kappa_\lambda < 6.3$	$-1.9 < \kappa_\lambda < 7.5$	$\kappa_\lambda = 3.0^{+1.8}_{-1.9}$
HH+H combination, κ_t floating	$-0.4 < \kappa_\lambda < 6.3$	$-1.9 < \kappa_\lambda < 7.6$	$\kappa_\lambda = 3.0^{+1.8}_{-1.9}$
HH+H combination, κ_t, κ_V, κ_b, κ_τ floating	$-1.3 < \kappa_\lambda < 6.1$	$-2.1 < \kappa_\lambda < 7.6$	$\kappa_\lambda = 2.3^{+2.1}_{-2.0}$

Perturbative unitarity and Higgs couplings

- Process relevant for κ_3, κ_4 is $HH \to HH$ scattering (see also [Liu et al `18])
- Jacob-Wick expansion allows to extract zeroth partial wave:

$$a^0_{ii} = \frac{3M_H^2\sqrt{s^2-4M_H^2 s}}{32\pi s(s-M_H^2)v^2}\left[\kappa_4(s-M_H^2) - 3\kappa_3^2 M_H^2 + \frac{6\kappa_3^2 M_H^2(s-M_H^2)}{s-4M_H^2}\log\left(\frac{s}{M_H^2}-3\right)\right]$$



Two-Higgs Doublet Model (2HDM)

- Two-Higgs Doublet Model (2HDM) → a second doublet: $\Phi_1 = \begin{pmatrix} \phi_1^+ \\ \frac{1}{\sqrt{2}}(v_1 + \rho_1 + i\eta_1) \end{pmatrix}, \; \Phi_2 = \begin{pmatrix} \phi_2^+ \\ \frac{1}{\sqrt{2}}(v_2 + \rho_2 + i\eta_2) \end{pmatrix}$

$$V_{\rm 2HDM} = m_{11}^2(\Phi_1^\dagger\Phi_1) + m_{22}^2(\Phi_2^\dagger\Phi_2) - m_{12}^2(\Phi_1^\dagger\Phi_2 + \Phi_2^\dagger\Phi_1) + \frac{\lambda_1}{2}(\Phi_1^\dagger\Phi_1)^2 + \frac{\lambda_2}{2}(\Phi_2^\dagger\Phi_2)^2$$
$$+\lambda_3(\Phi_1^\dagger\Phi_1)(\Phi_2^\dagger\Phi_2) + \lambda_4(\Phi_1^\dagger\Phi_2)(\Phi_2^\dagger\Phi_1) + \frac{\lambda_5}{2}\left((\Phi_1^\dagger\Phi_2)^2 + (\Phi_2^\dagger\Phi_1)^2\right)$$

- Free parameters: $m_H, m_{H'}, \; m_A, \; m_{H^\pm}, \; m_{12}^2, \; v, \; \cos(\beta-\alpha), \; \tan\beta$

<u>Scalar Particle content:</u>

Neutral scalars: H', $H(m_H = 125$ GeV)

Neutral pseudoscalars: A

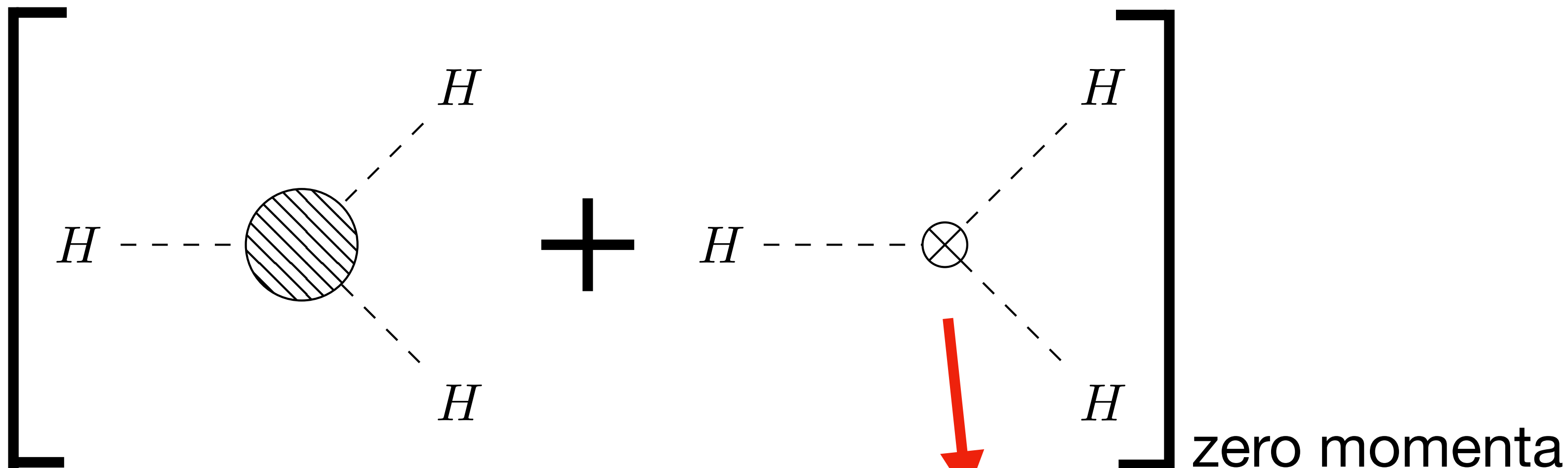
Charged scalars: $H^\pm$

Alignment limit → couplings of light Higgs same as SM
$\cos(\beta-\alpha) = 0$

Model example: 2HDM - calculation

- 1-loop calculation for κ_3, κ_4 with FeynArts, FormCalc, LoopTools in alignment limit

$\hat{\Gamma}^{(1)}_{3H} =$

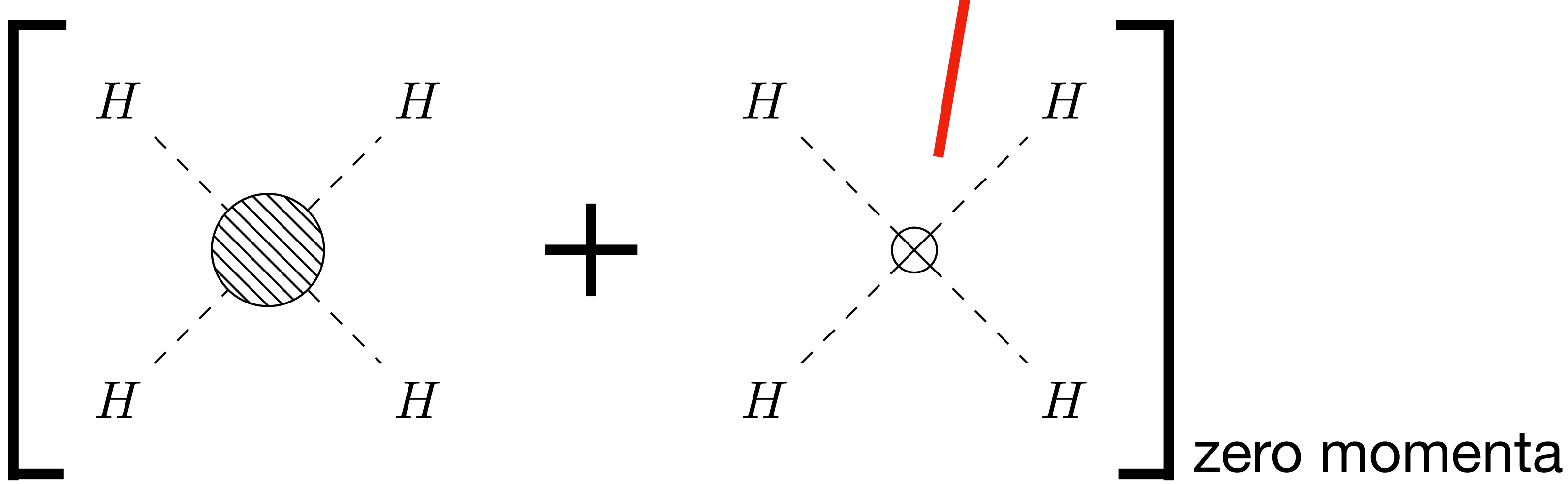


- 2HDM renormalisation constants calculated with on-shell conditions
- divergence cancellation checked analytically

$$\Gamma^{\mathrm{CT}}_{3H} = -\frac{3}{v}\left[\delta M_H^2 + \frac{\delta T_H}{v} + M_H^2\left(\delta Z_e + \frac{3}{2}\delta Z_H - \frac{\delta M_W^2}{2M_W^2} - \frac{\delta s_W}{s_W}\right)\right]$$

$$\Gamma^{\mathrm{CT}}_{4H} = -\frac{3}{v^2}\left[\delta M_H^2 + \frac{\delta T_H}{v} + 2M_H^2\left(\delta Z_e + \delta Z_H - \frac{\delta M_W^2}{2M_W^2} - \frac{\delta s_W}{s_W}\right)\right]$$

$\hat{\Gamma}^{(1)}_{4H} =$



Edge Convolution

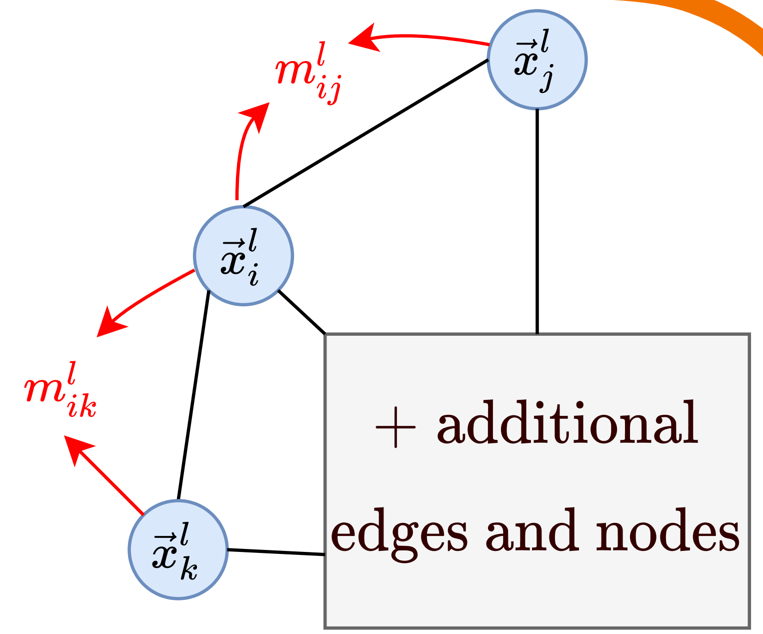
Input features: $\vec{x}_i^{(0)}$ → update iteratively with **Edge Convolution** operation:

Edge Convolution operation

'Message' calculation:

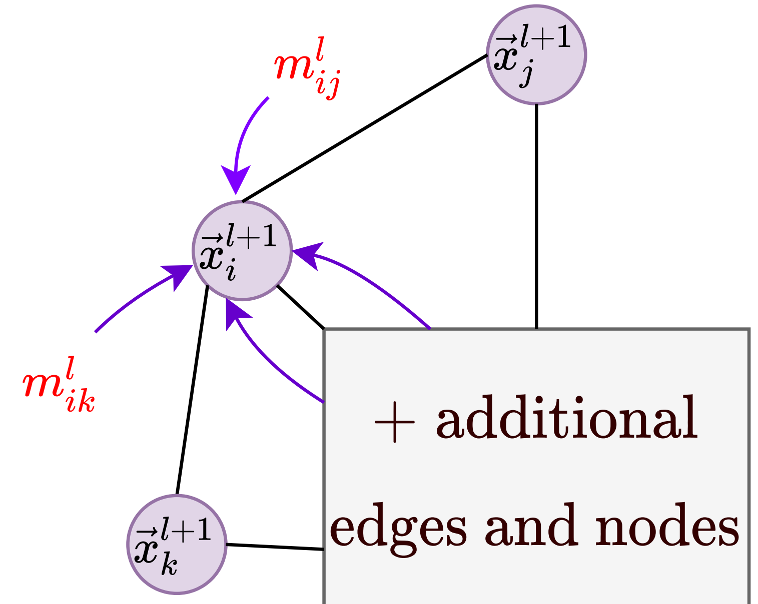
$$m_{ij}^{(l)} = \text{ReLU}\left(\Theta(\vec{x}_j^{(l)} - \vec{x}_i^{(l)}) + \Phi(\vec{x}_i^{(l)})\right)$$

linear layers



Aggregation: update node features

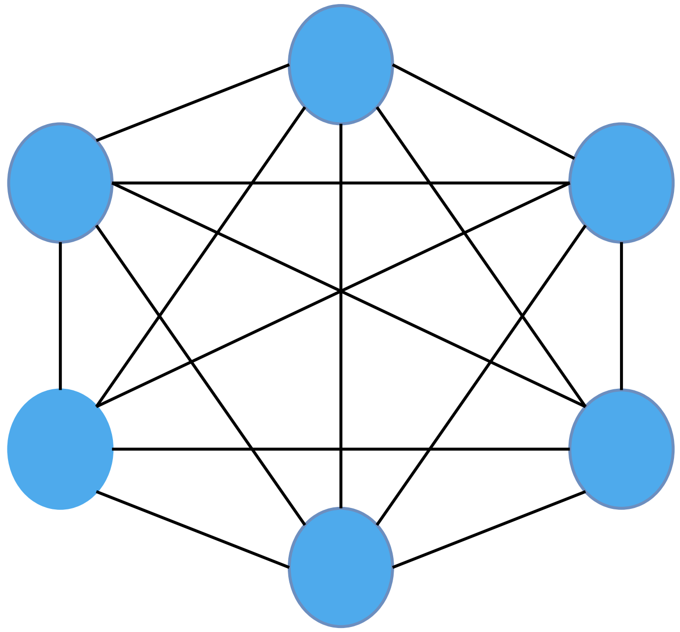
$$\vec{x}_i^{(l+1)} = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} m_{ij}^{(l)}$$



Graph Embedding

1.

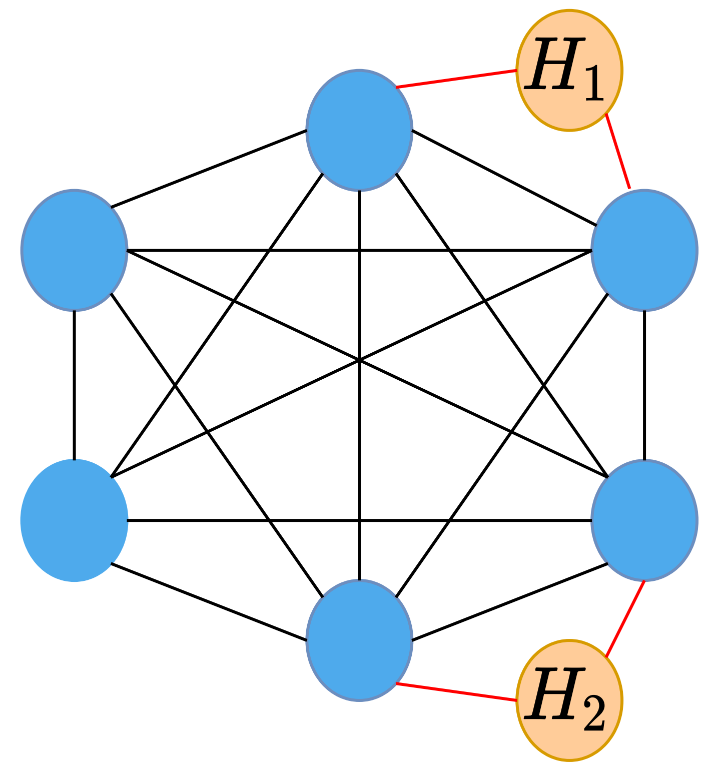
- Fully-connected nodes for b and τ final states
- **Input features**: $[p_T, \eta, \phi, E, m,$ PDGID$]$
- Additional node for Missing Transverse Momentum (MTM) in showered & reconstructed events



FC: Fully-Connected

2.

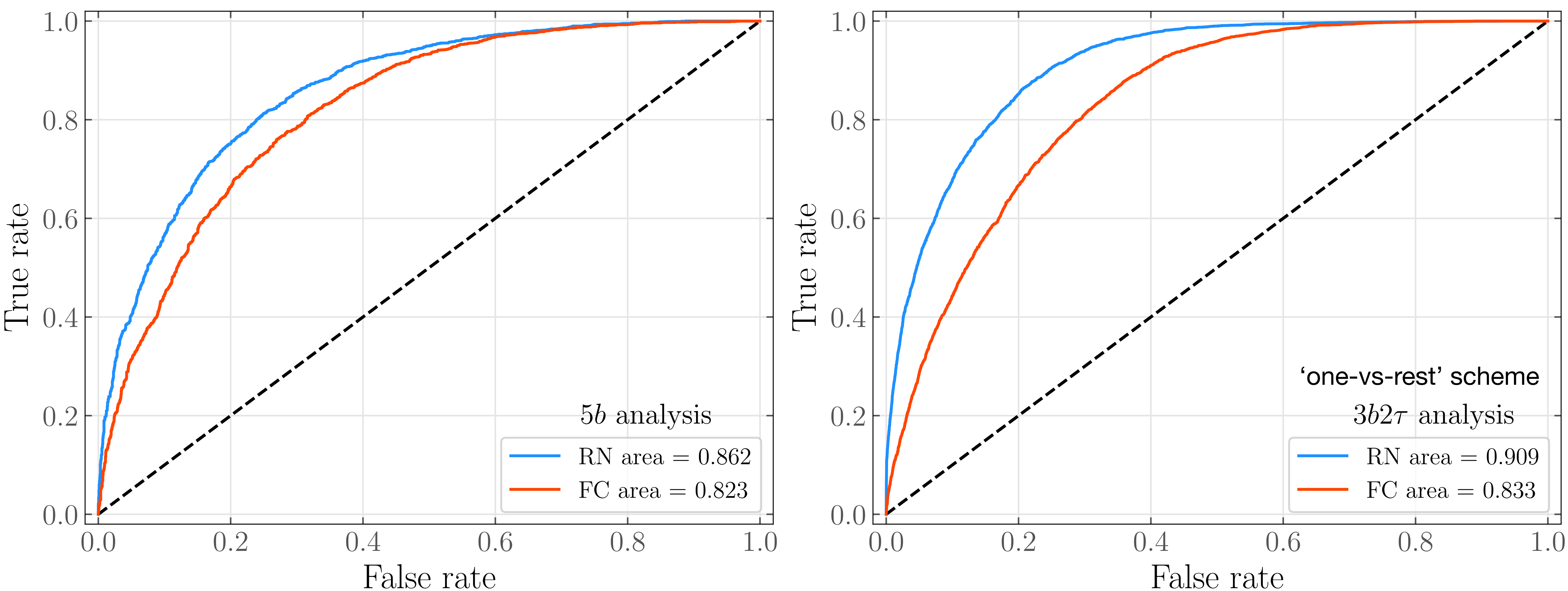
- Consider combinations of b-quarks and τ with reconstructed four-momentum $(p_i + p_j)$
- If $m_{ij} \in [100,150]$ (GeV) add extra node H_i



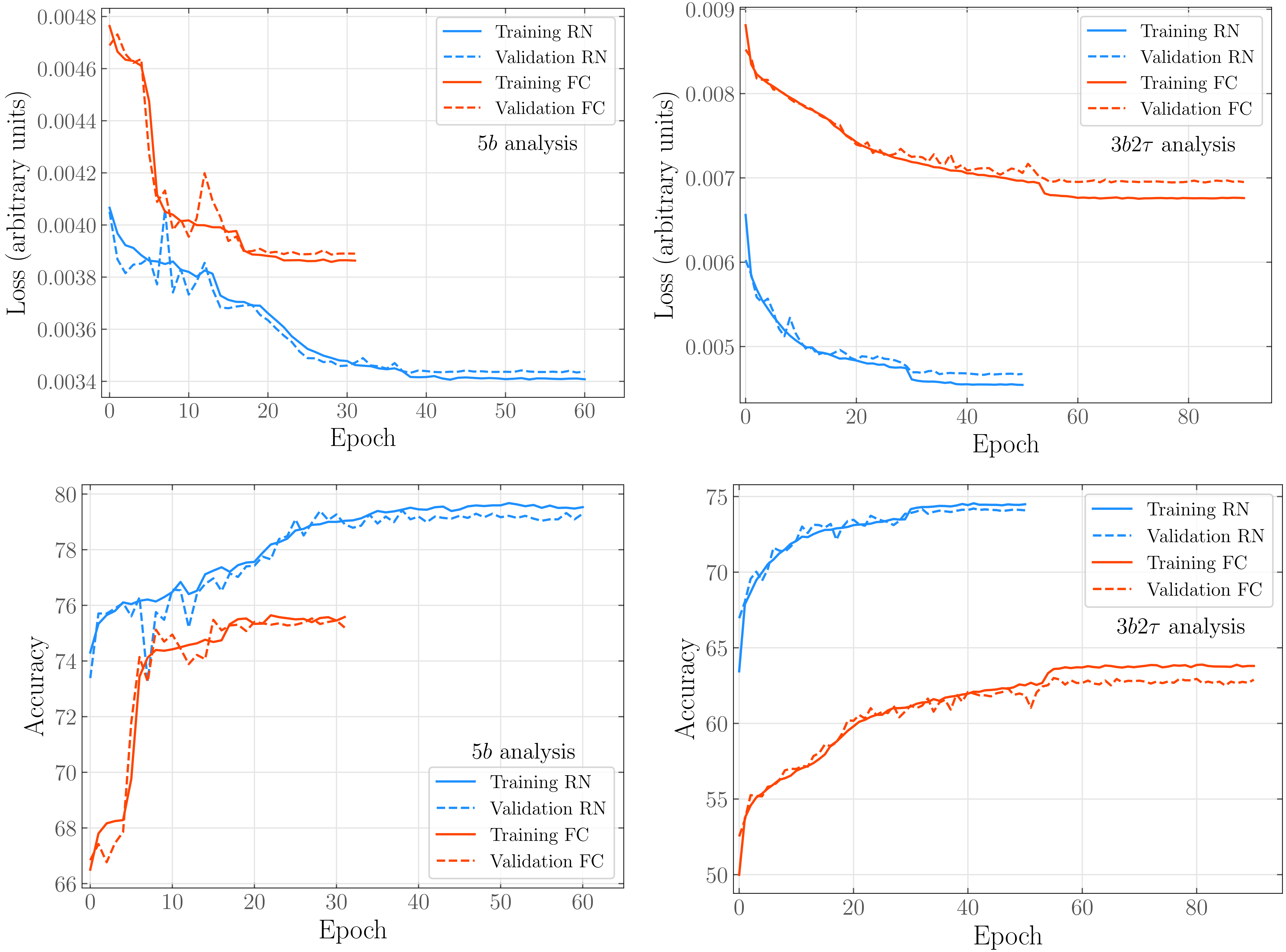
RN: Reconstructed Nodes

Embedding performance

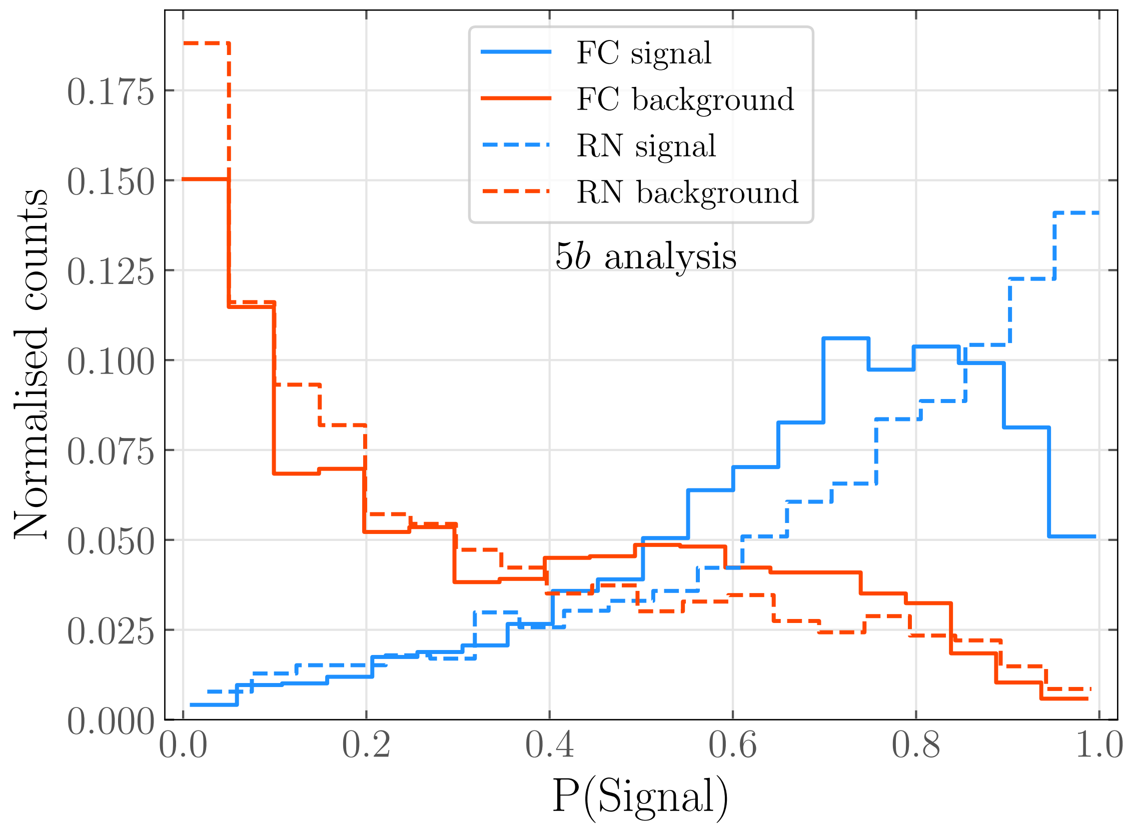
- GNN trained on $(\kappa_3, \kappa_4) = (1,1)$ sample
- Evaluate performance with Receiver Operating Characteristic (ROC) curves

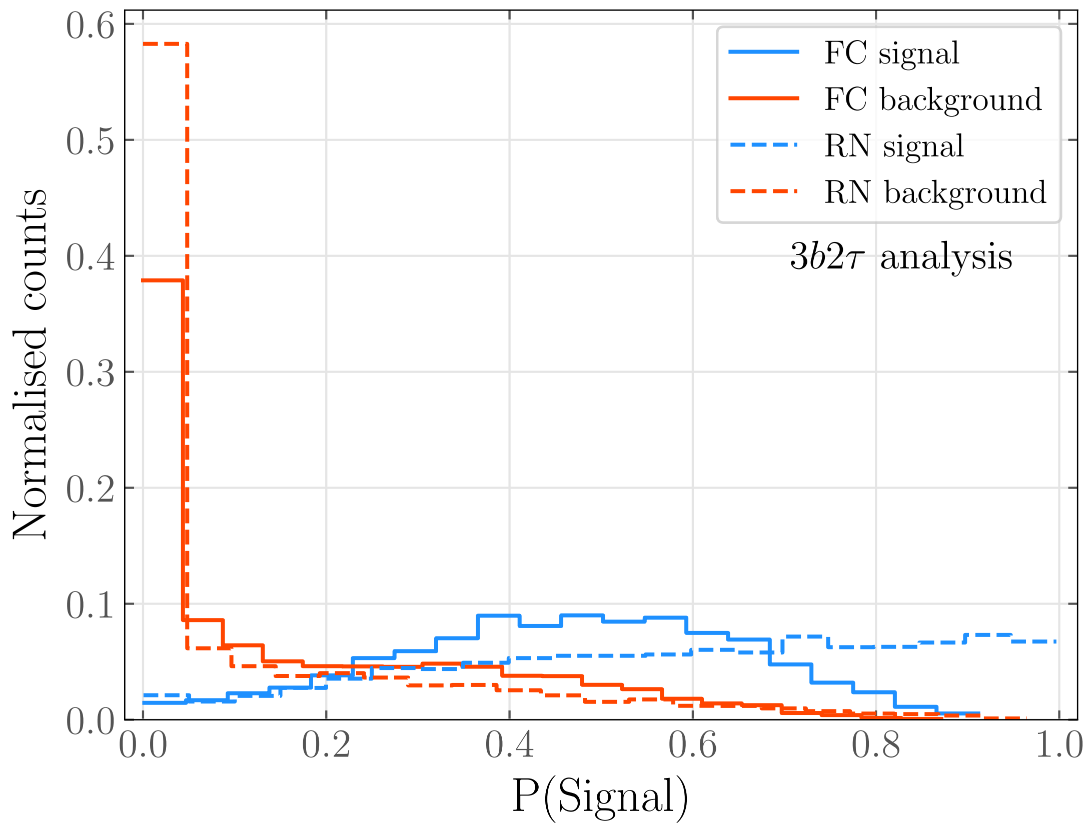


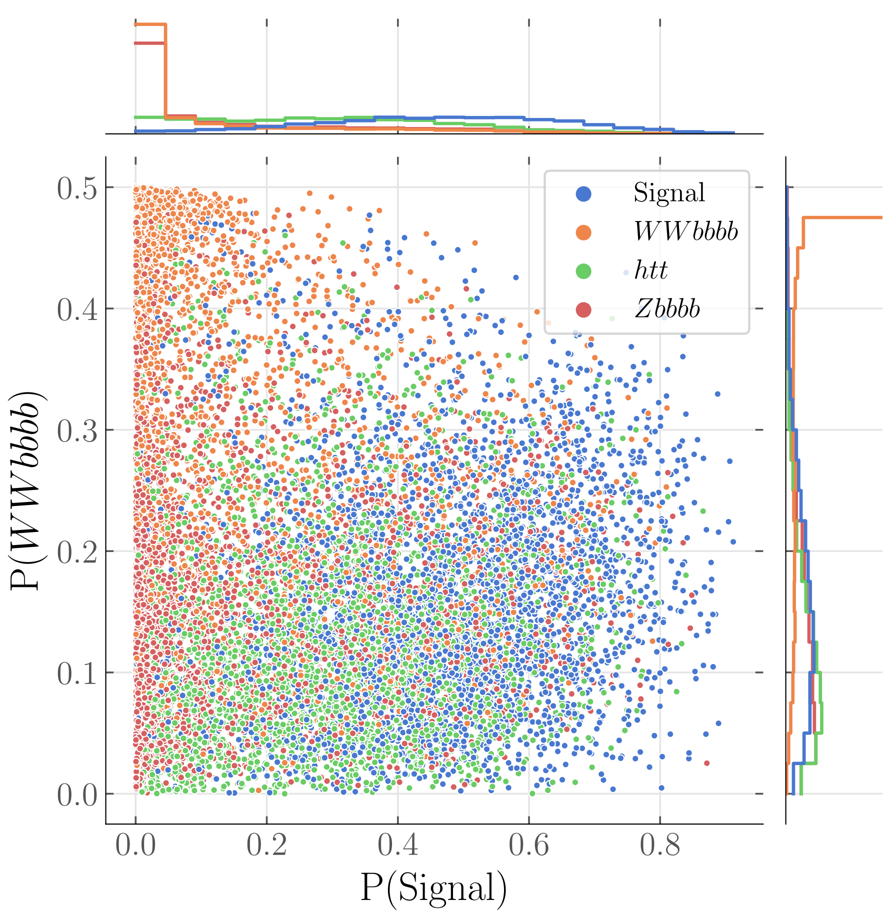
Training loss and accuracy

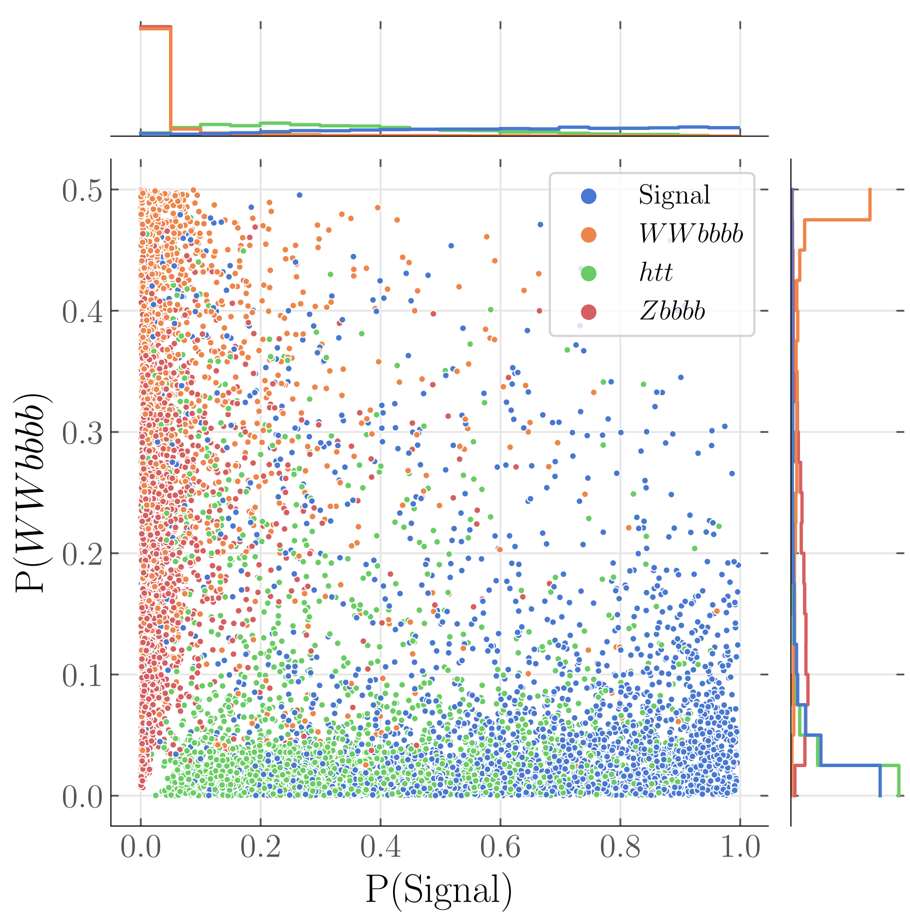


Score distributions





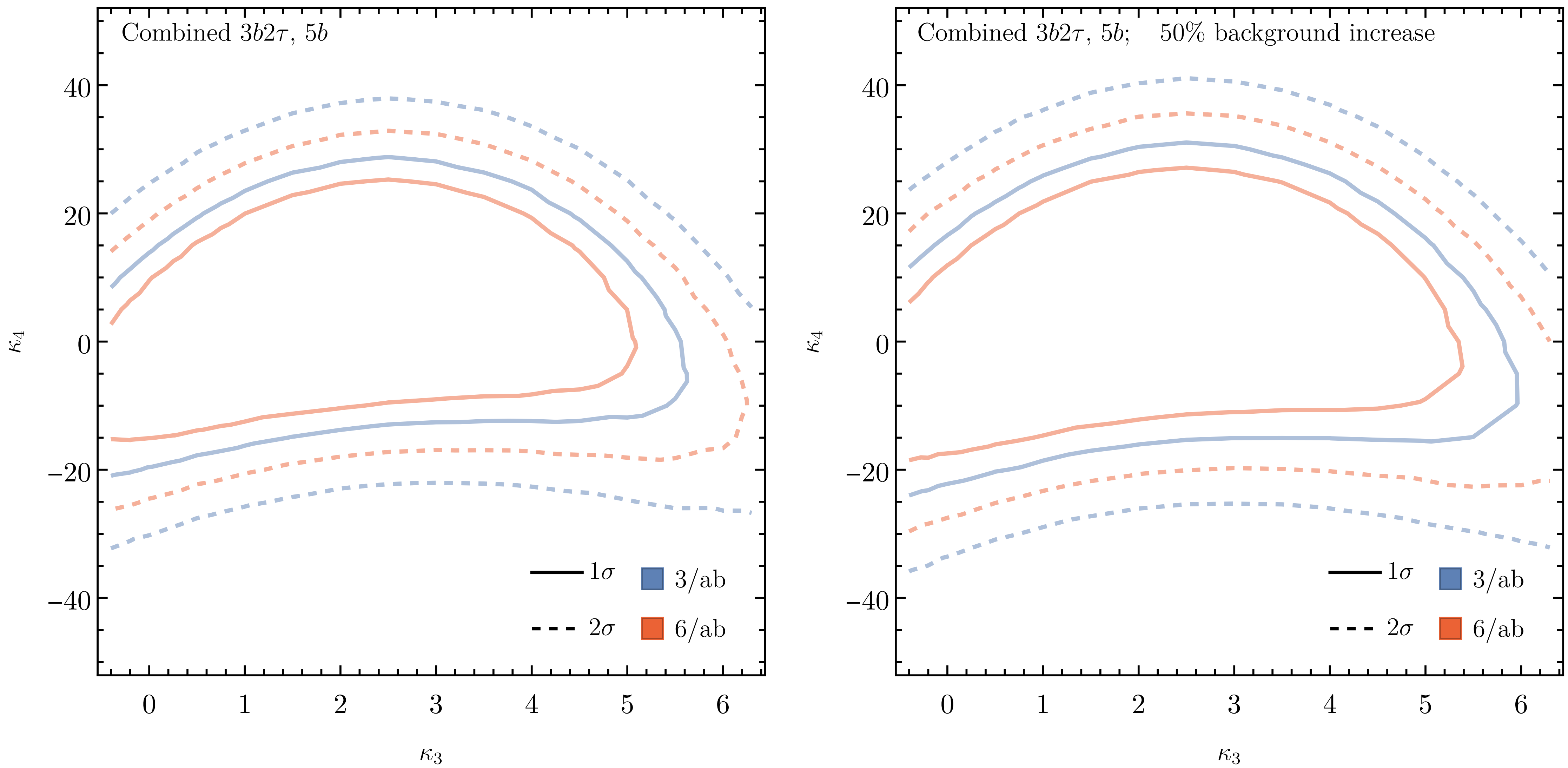




Combined Results

- **Assumption:** No correlations → combine significances

$$Z_{\text{comb}} = \sqrt{Z_{3b2\tau}^2 + Z_{5b}^2}$$



Combination of further channels and improvements of **tagging/ reconstruction** methods could enhance results further

Integrated Gradients

→ **Integrated Gradients**: [Sundararajan, Taly, Yan 1703.01365]

- axiomatic method
- uses Neural Network gradients → **fast!**
- requires a differentiable model

→ **suitable for Neural Networks!**

- Definition:

input   baseline

$$I_i(x) = (x_i - x'_i) \int_0^1 d\alpha \frac{\partial F(x' + \alpha(x - x'))}{\partial x_i}$$

Attribution scores
→ importance of feature

Gradient of Neural Network F

- Easy to implement for Graph Neural Networks as well

- Does **not** take into account graph structures → work in progress in Deep Learning community
- Viable to understand important features → expect mass of reconstructed Higgs to be important

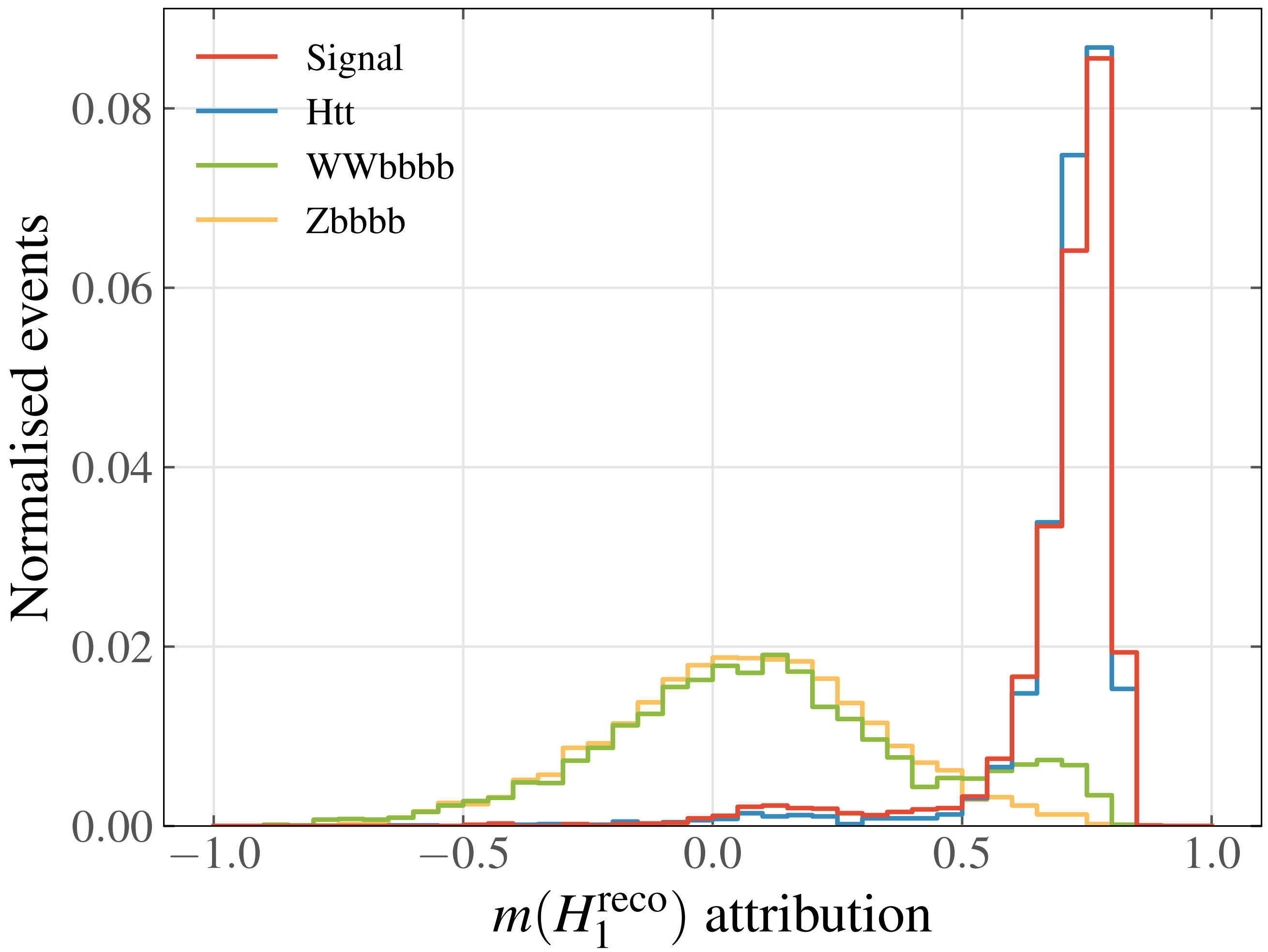
NN Interpretation axioms

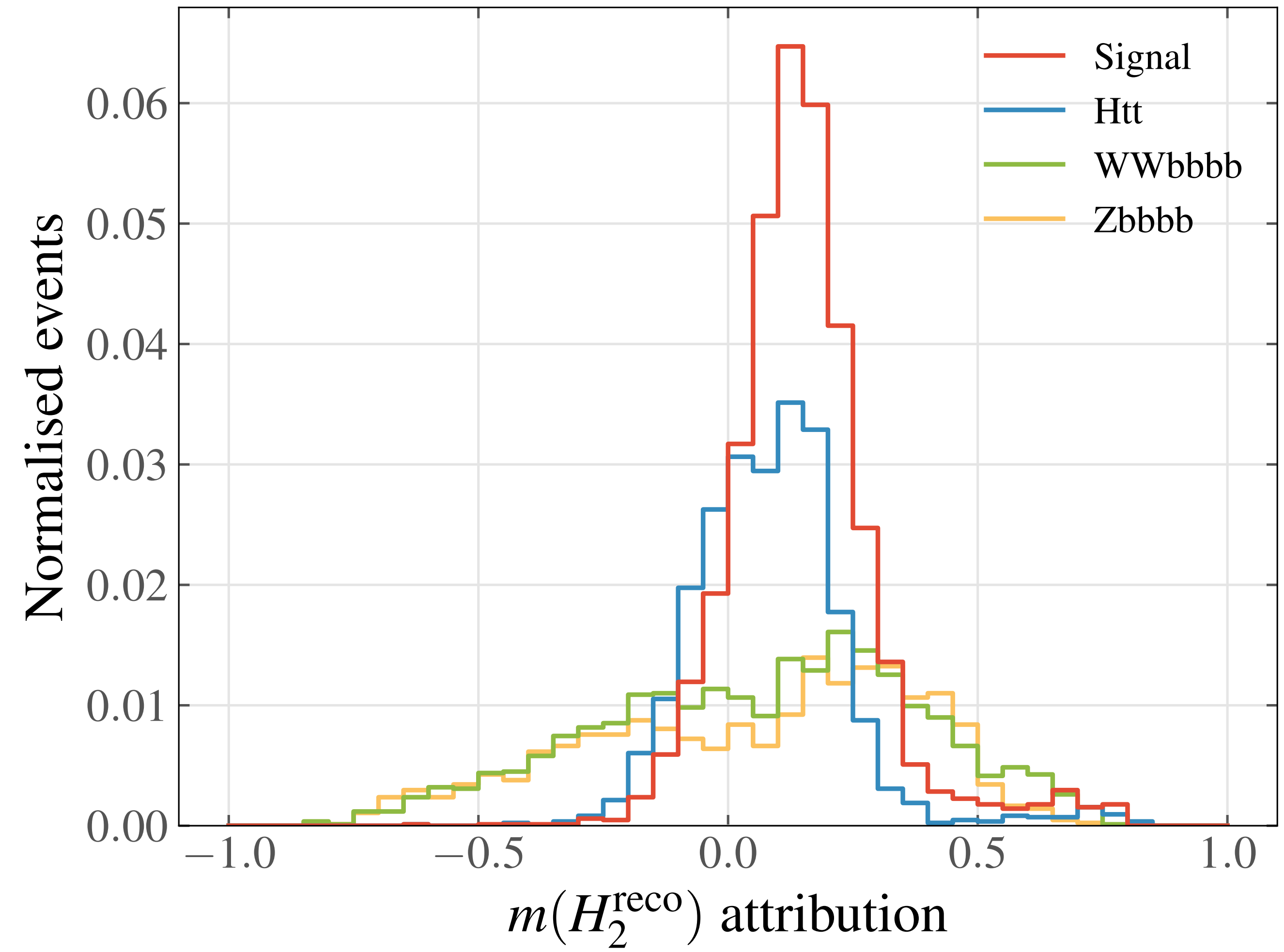
- **<u>Integrated Gradients Axioms:</u>**
  - **<u>Completeness</u>**: sum of attributions equal to difference of network output for input and baseline values
  - **<u>Sensitivity</u>**: when baseline and input have different values and different NN outputs, attributions should also be different
  - **<u>Dummy</u>**: A zero input should yield no attribution
  - **<u>Implementation Invariance</u>**: If two methods are equivalent (i.e. yield same scores for all inputs despite being different) then attributions should be identical
  - **<u>Linearity</u>**: Attributions should be linear for linear combinations of networks $aF_1 + bF_2$
  - **<u>Symmetry</u>**: For a network symmetric for two variables $F(x, y) = F(y, x)$, the attributions should be the same

Reconstructed Higgs Mass

- **Interpretation as expected:**
  If a Higgs close to 125 GeV can be found $\implies$ signal

- Complete understanding would require to study correlations between observables → <u>**future work**</u>

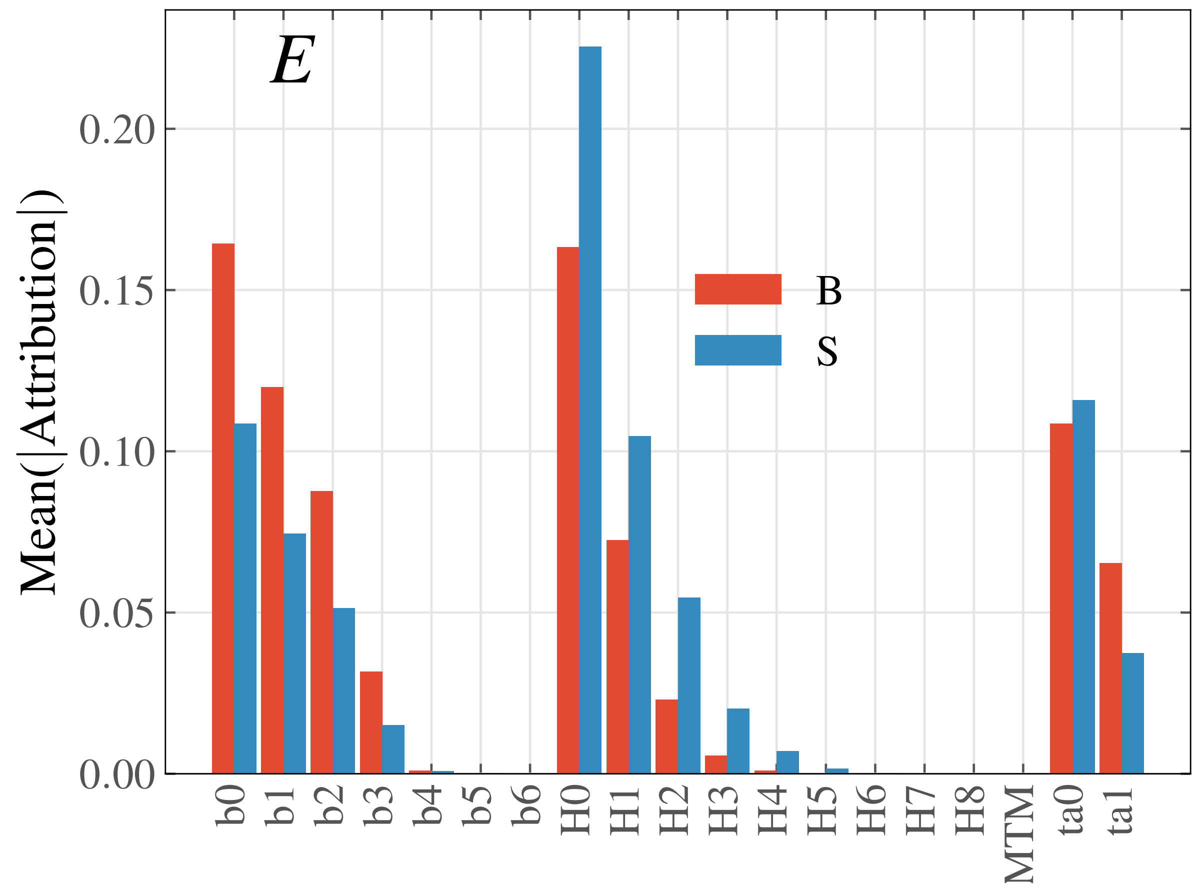


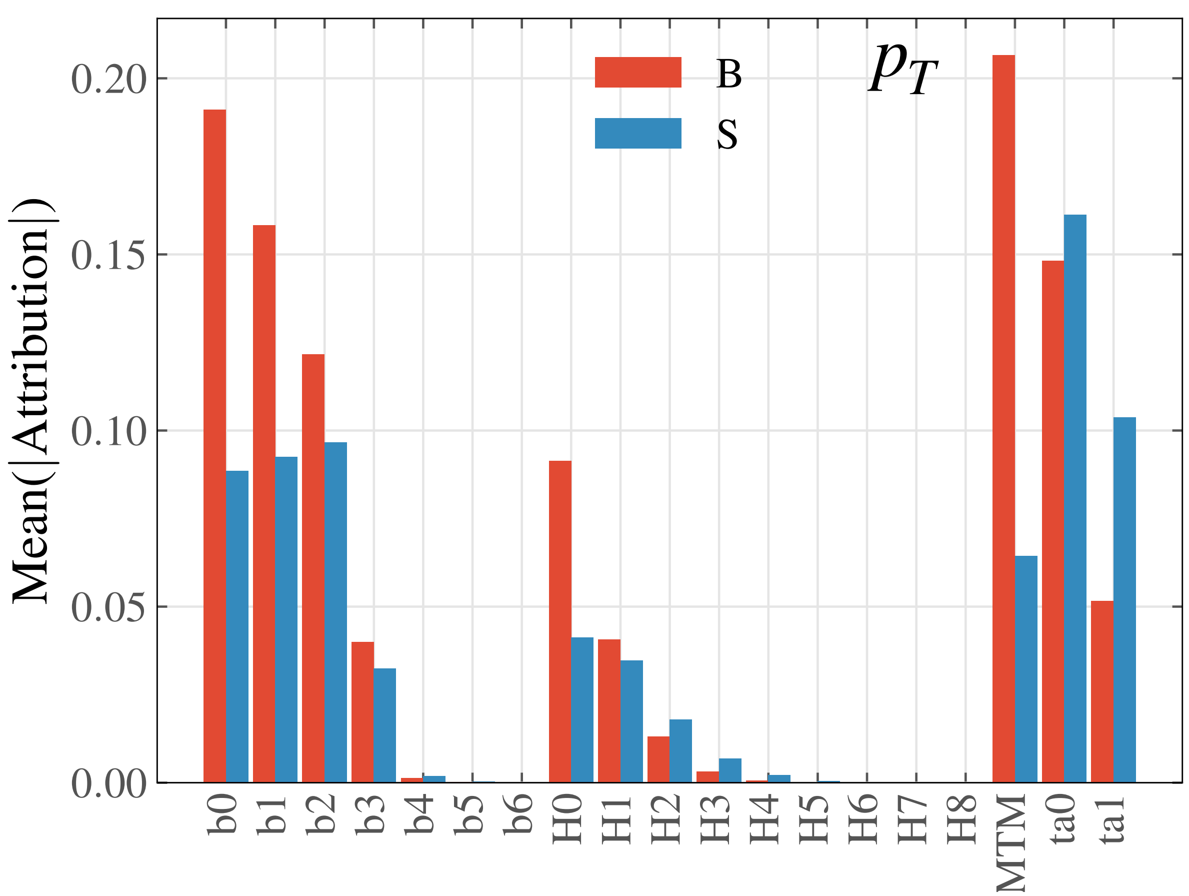


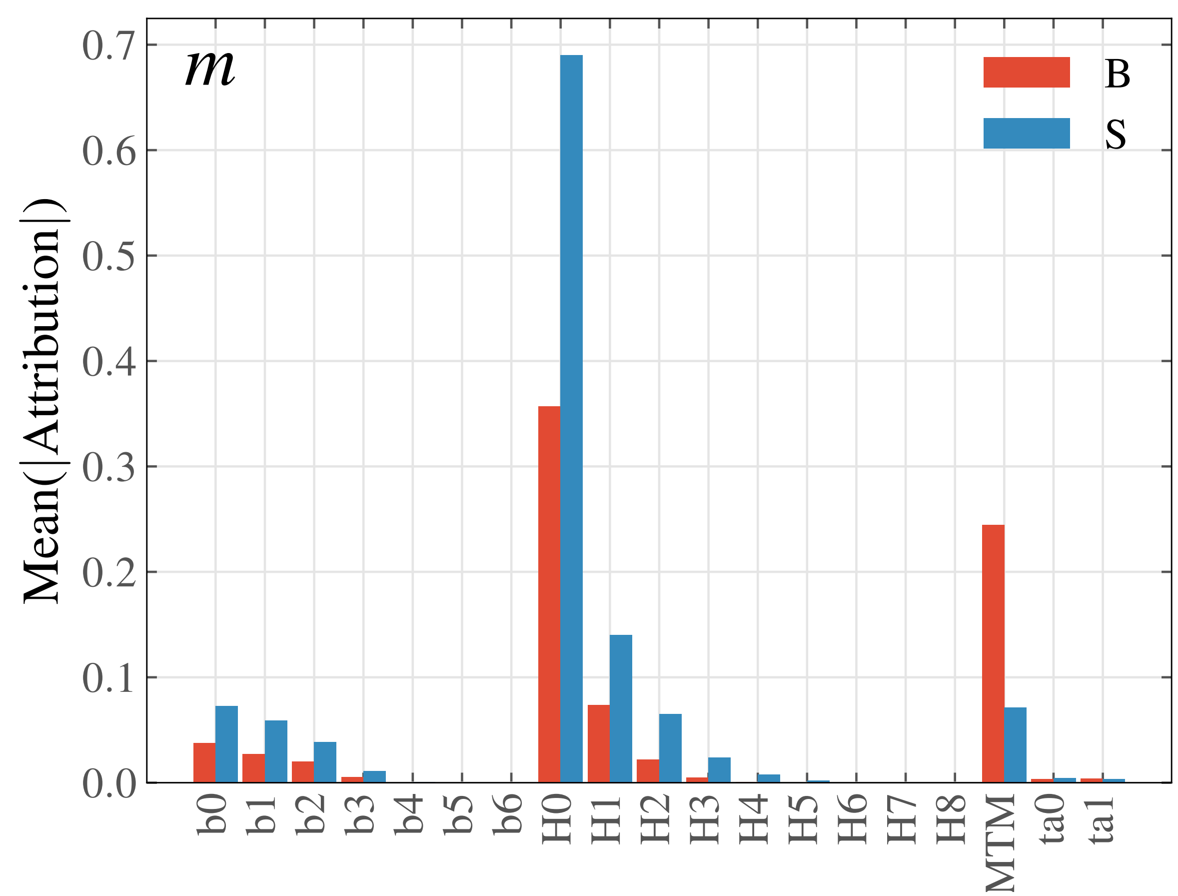
Attribution vs. nodes

- E and p_T from leading order particles is more important

- m is more important for the reconstructed Higgs closest to the SM mass value







Lepton collider cross sections

- Inclusive $\ell\ell \to HHH + X$ analysis with $H \to b\bar{b}$
- Cross sections small below 1 TeV
- **Note:** $\mu^+\mu^-$ vs. e^+e^- collider at 10 TeV has difference of less than $5\,\%$ on cross sections