

New approaches for $hh \rightarrow b\bar{b}b\bar{b}$ as a probe of Higgs self-coupling

Jacob Amacker, William Balunas, Lydia Beresford, Daniela Bortoletto,
James Frost, Cigdem Issever, Jesse Liu, James McKee, Alessandro Michelis,
Santiago Paredes Saenz, Michael Spannowsky, and Bejan Stanislaus
santiago.paredes@cern.ch

Higgs 2021
October 2021



This Talk

- Introduction & Motivation
- Signal & Background Modelling
- Analysis Strategies
- Self-Coupling Constraints
- Conclusion

Based on [arXiv:2004.04240](https://arxiv.org/abs/2004.04240)



Cornell University

arXiv.org > hep-ph > arXiv:2004.04240

High Energy Physics – Phenomenology

[Submitted on 8 Apr 2020 (v1), last revised 12 Oct 2020 (this version, v3)]

Higgs self-coupling measurements using deep learning in the $b\bar{b}b\bar{b}$ final state

Jacob Amacker, William Balunas, Lydia Beresford, Daniela Bortoletto, James Frost, Cigdem Issever, Jesse Liu, James McKee, Alessandro Micheli, Santiago Paredes Saenz, Michael Spannowsky, Bejan Stanislaus

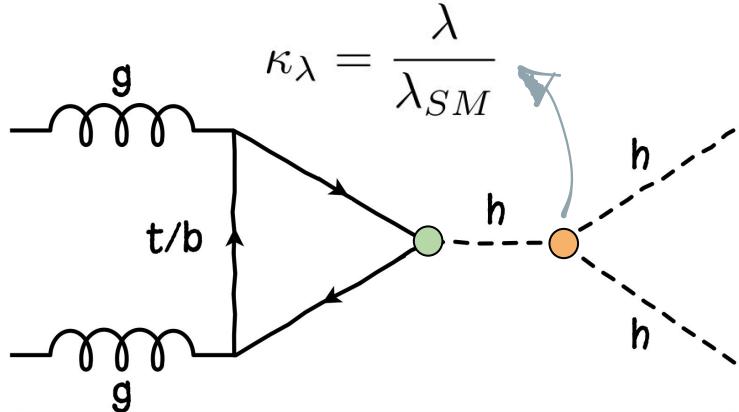
Measuring the Higgs trilinear self-coupling λ_{hhh} is experimentally demanding but fundamental for understanding the shape of the Higgs potential. We present a comprehensive analysis strategy for the HL-LHC using di-Higgs events in the four b -quark channel ($hh \rightarrow 4b$), extending current methods in several directions. We perform deep learning to suppress the formidable multijet background with dedicated optimisation for BSM λ_{hhh} scenarios. We compare the λ_{hhh} constraining power of events using different multiplicities of large radius jets with a two-prong structure that reconstruct boosted $h \rightarrow bb$ decays. We show that current uncertainties in the SM top Yukawa coupling y_t can modify λ_{hhh} constraints by $\sim 20\%$. For SM y_t , we find prospects of $-0.8 < \lambda_{hhh}/\lambda_{hhh}^{\text{SM}} < 6.6$ at 68% CL under simplified assumptions for 3000-fb^{-1} of HL-LHC data. Our results provide a careful assessment of di-Higgs identification and machine learning techniques for all-hadronic measurements of the Higgs self-coupling and sharpens the requirements for future improvement.

Comments: 36 pages, 15 figures + bibliography and appendices
Subjects: High Energy Physics – Phenomenology (hep-ph); High Energy Physics – Experiment (hep-ex)
Journal reference: JHEP 12 (2020) 115
DOI: [10.1007/JHEP12\(2020\)115](https://doi.org/10.1007/JHEP12(2020)115)
Report number: IPPV/20/11
Cite as: [arXiv:2004.04240 \[hep-ph\]](https://arxiv.org/abs/2004.04240)
(or [arXiv:2004.04240v3 \[hep-ph\]](https://arxiv.org/abs/2004.04240v3) for this version)

Introduction & Motivation



Why hh?



- **Key parameter** in the standard model
→ **Not only** for collider physics
- hh is the **most sensitive probe** of the Higgs self-coupling



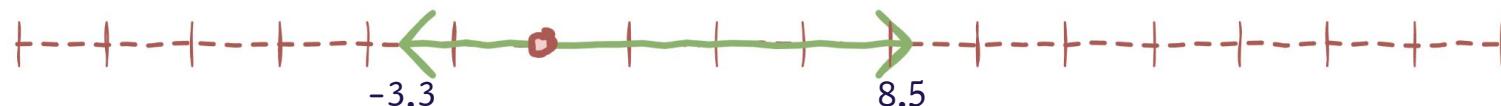
full Run II data - $b\bar{b}yy$ - 95% C.L. κ_λ constraints*



ATLAS-CONF-
-2021-016



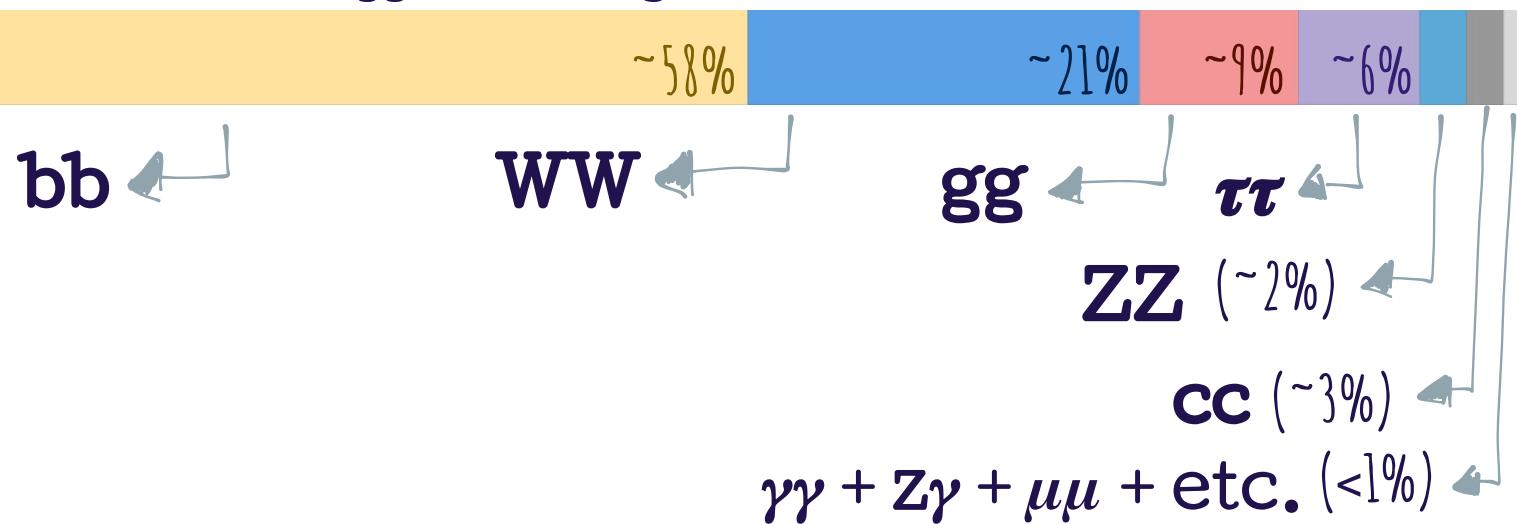
JHEP03(2021)
257



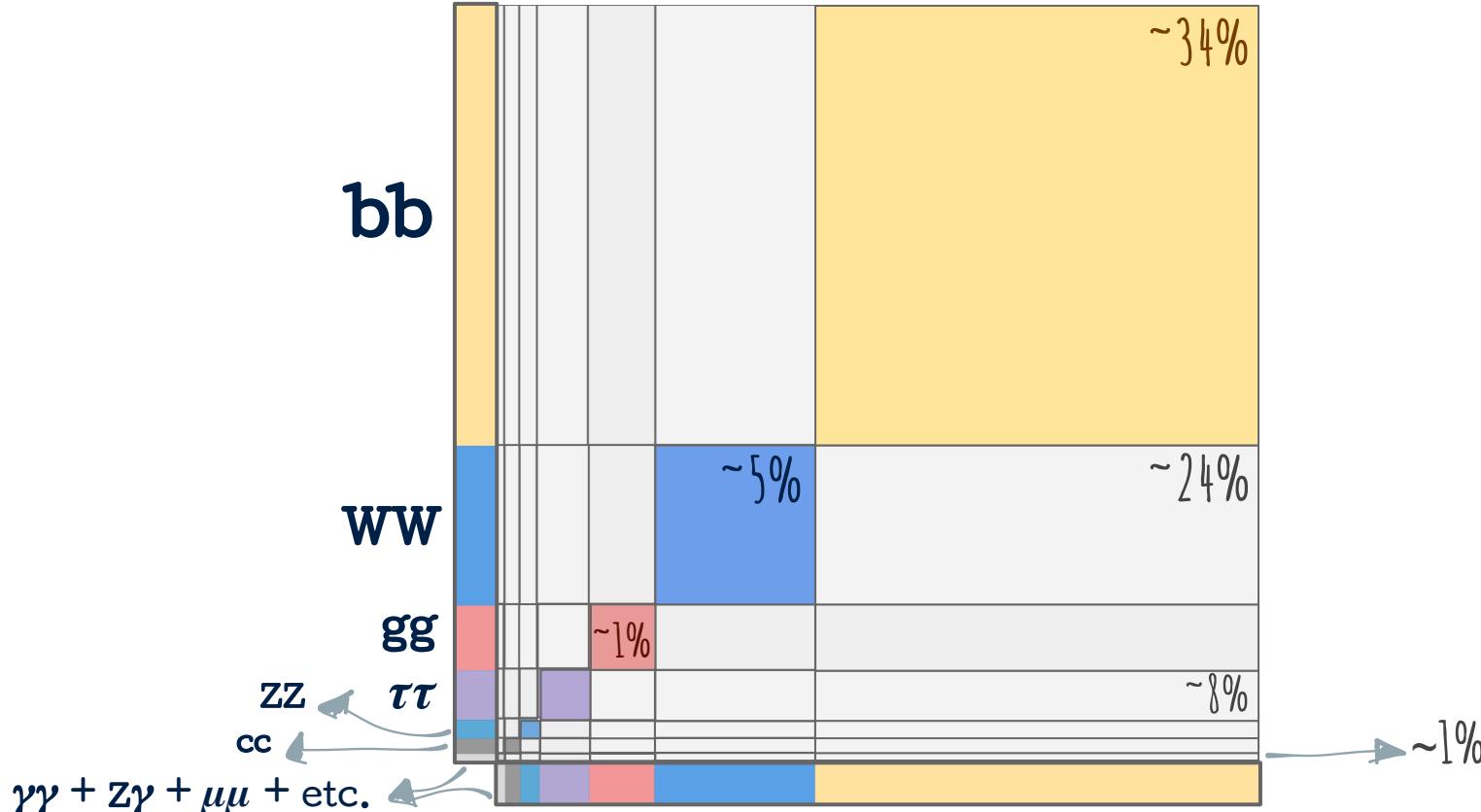
*Rough snapshot of our knowledge of κ_λ today, with run II data. Other channels being worked on. Probably already outdated since a few talks.

Why $hh \rightarrow 4b$?

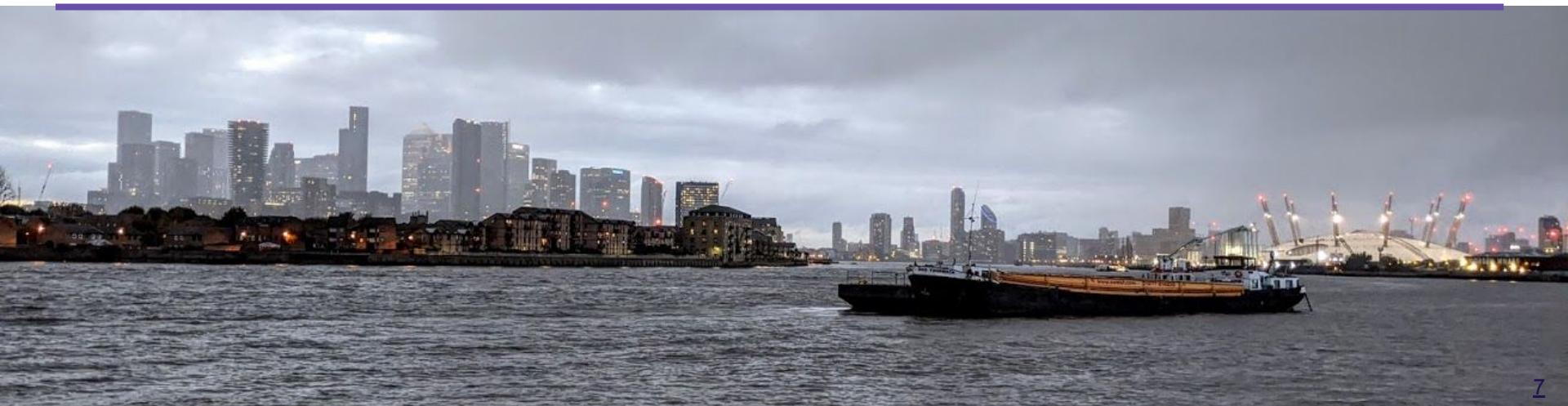
Higgs branching fraction



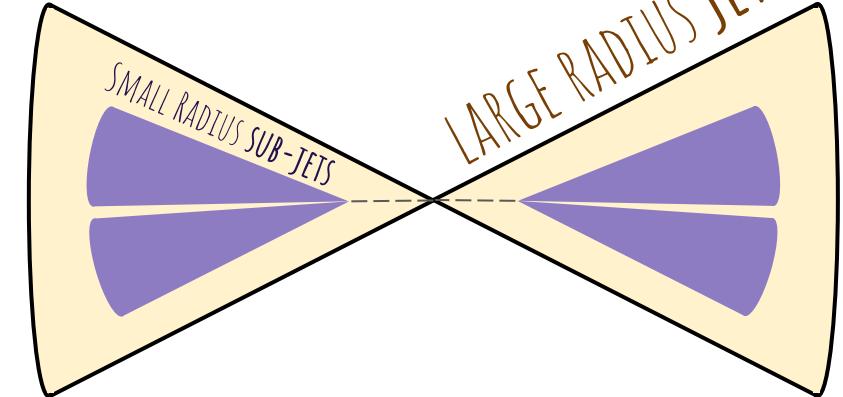
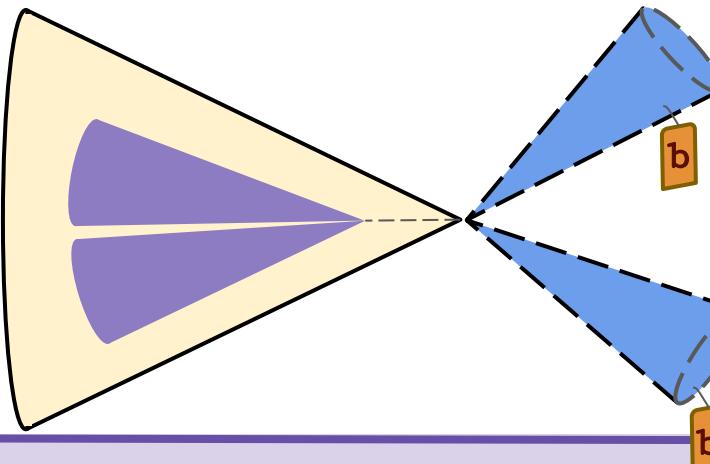
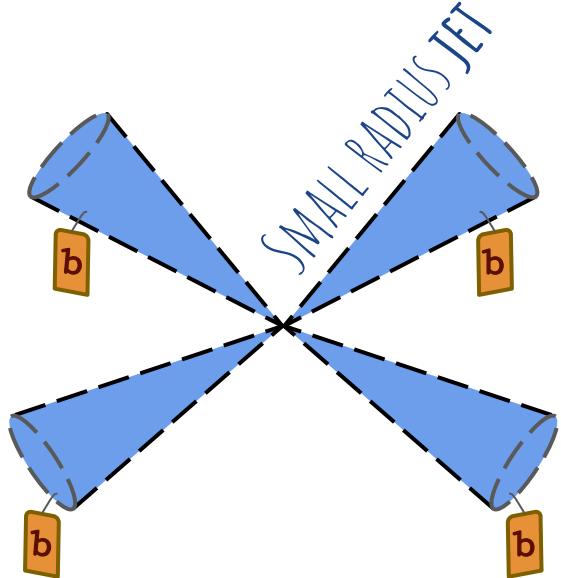
Why $hh \rightarrow 4b$?



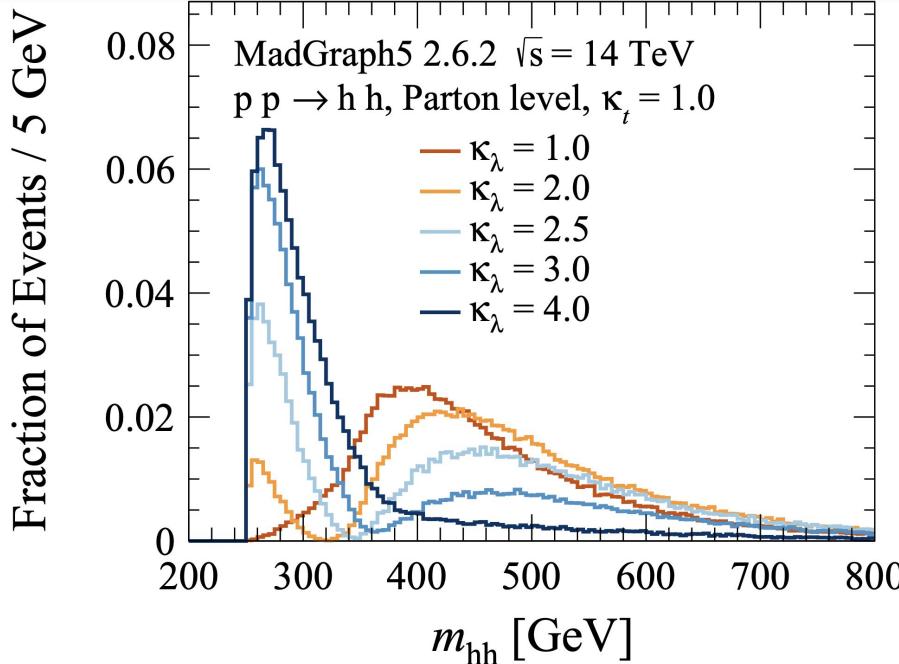
Signal & Background Modelling



Signal Topology



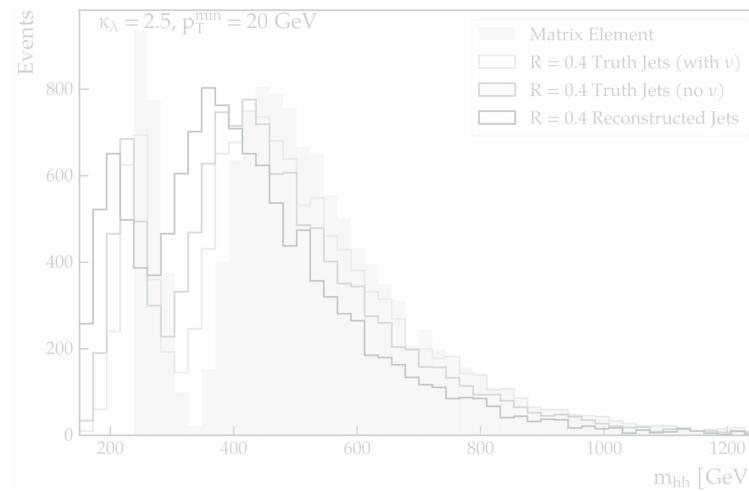
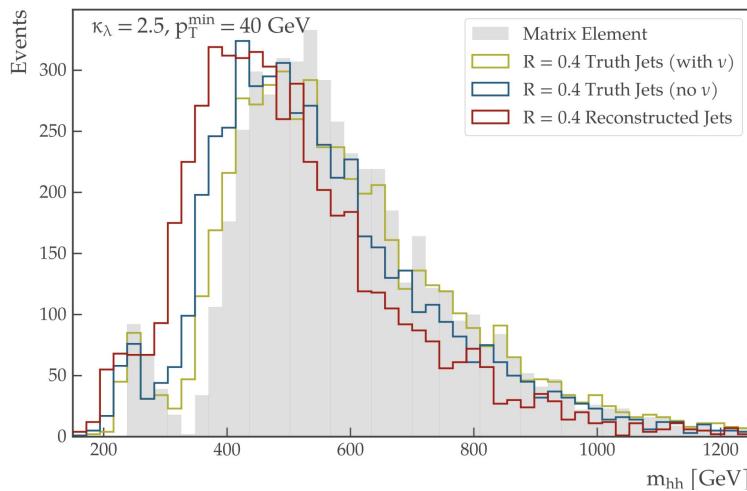
Signal Samples



- $gg \rightarrow hh$ production
 - ↪ **Inclusive h decay**
- Points with **varied** coupling to **top** quark and **self couplings**
- Extra $\kappa_t=1$ samples for **training**
 - ↪ **More** events per point
 - ↪ **Exclusive** decay $h \rightarrow b\bar{b}$

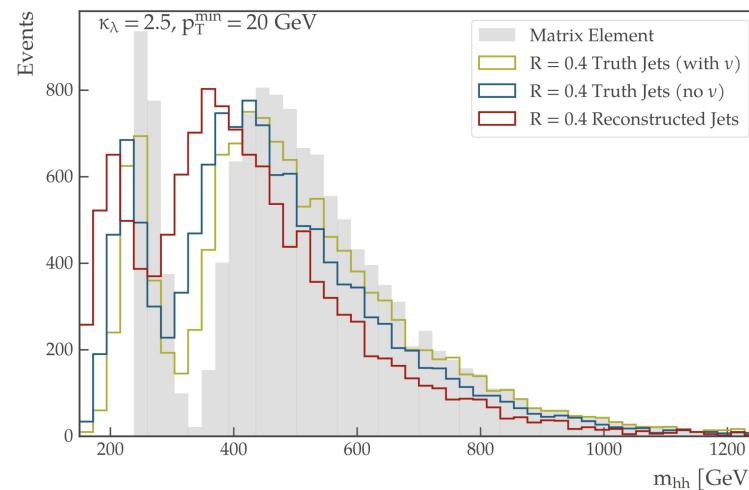
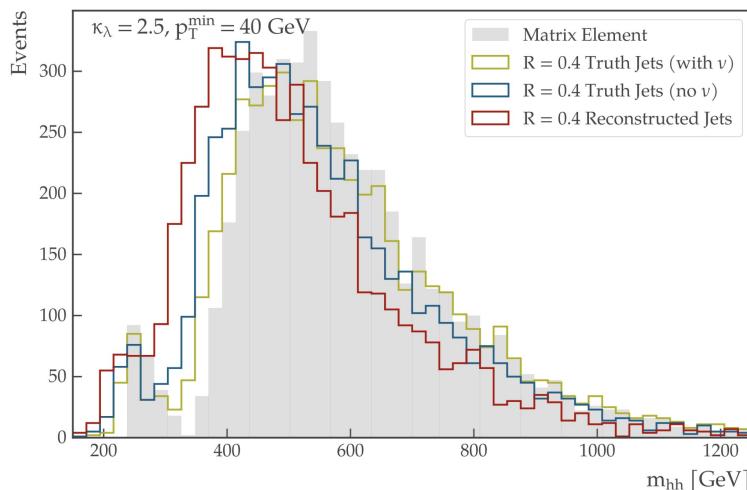
Parentheses - m_{hh} shape degradation

- m_{hh} spectrum, various jets
 - ↪ $p_T > 40 \text{ GeV} \rightarrow$ Same as analysis
 - ↪ $k_\lambda = 2.5 \rightarrow$ Max. interference
- **Double-peak is degraded**
- Same plot, except:
 - ↪ $p_T > 20 \text{ GeV}$
- **Recover double peak**



Parentheses - m_{hh} shape degradation

- m_{hh} spectrum, various jets
 - ↪ $p_T > 40 \text{ GeV} \rightarrow$ Same as analysis
 - ↪ $k_\lambda = 2.5 \rightarrow$ Max. interference
- **Double-peak is degraded**
- Same plot, except:
 - ↪ $p_T > 20 \text{ GeV}$
- **Recover double peak**



Background Samples

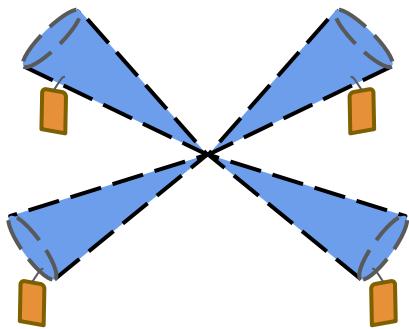
- Similar generation process to signals
- **Main** backgrounds:
 - ↪ **Multijet** → 4b and 2b-2j
 - ↪ **Top** quark **backgrounds** → $t\bar{t}$ (+ $b\bar{b}$) and $t\bar{t}h$
- Other backgrounds:
 - ↪ $b\bar{b}h$
 - ↪ ZZ
 - ↪ Zh
 - ↪ Wh

Analysis Strategies



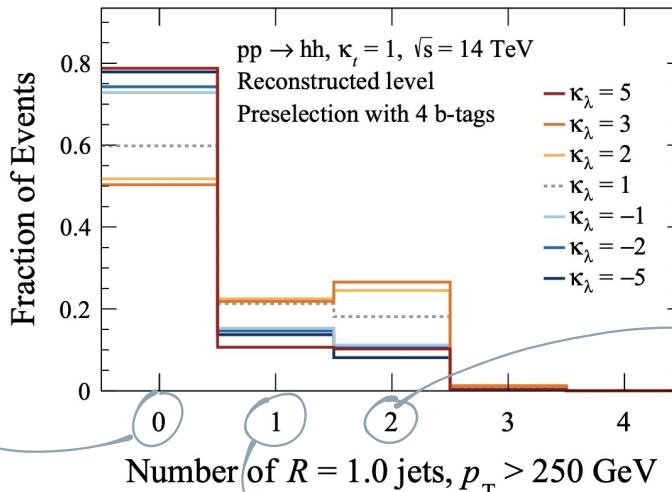
Channels

Resolved

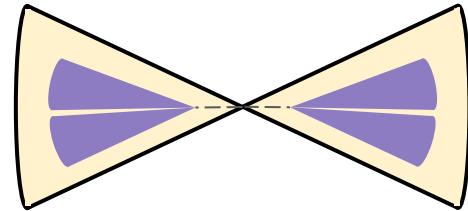


$$j_s \left| \begin{array}{l} R = 0.4 \\ p_T > 40 \text{ GeV} \\ |\eta| < 2.5 \end{array} \right.$$

$$\frac{j_T}{\left| \begin{array}{l} R = 0.2 \\ p_T > 20 \text{ GeV} \\ |\eta| < 2.5 \\ \Delta R(\triangleleft, \triangleright) < 1.0 \end{array} \right.}$$

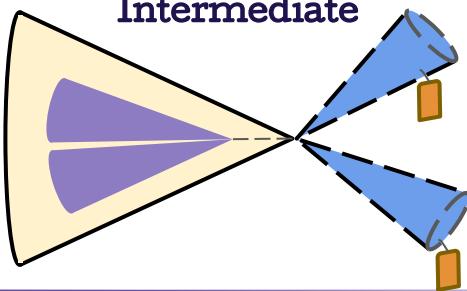


Boosted

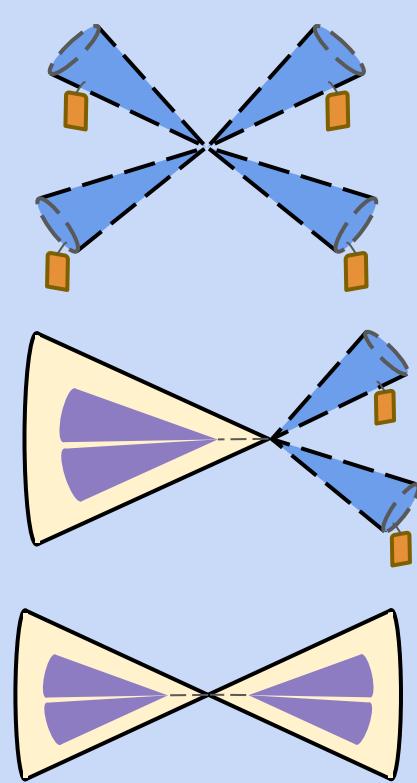


$$p_T > 250 \text{ GeV} \mid j_L \quad R = 1.0 \quad |\eta| < 2.0$$

Intermediate

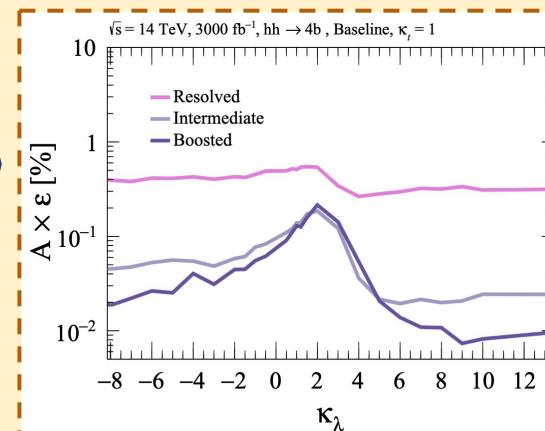


Analysis Strategy



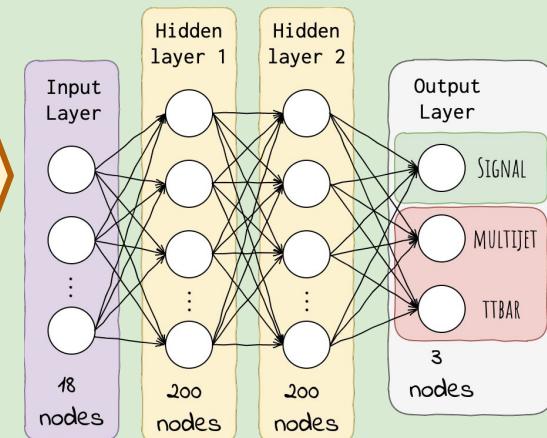
Baseline Analysis

- **Cut Based**
- **ATLAS/CMS-inspired**

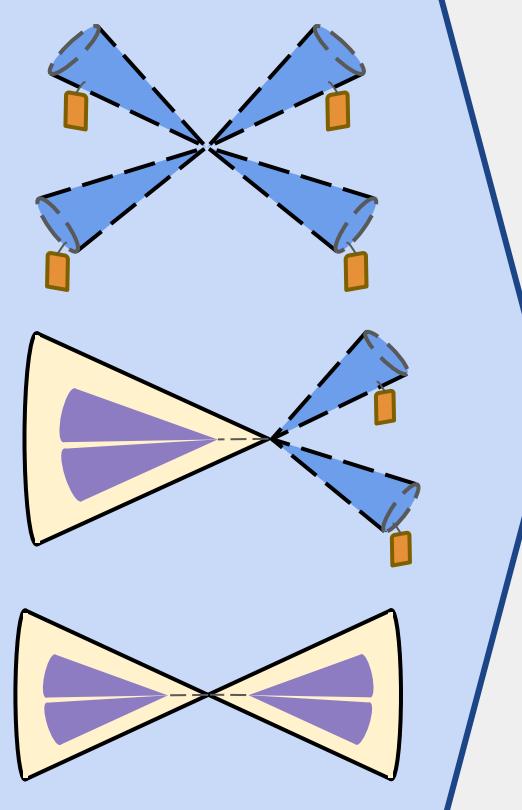


DNN Analysis

- Trained NN **classifier**
- **Cut on NN score**

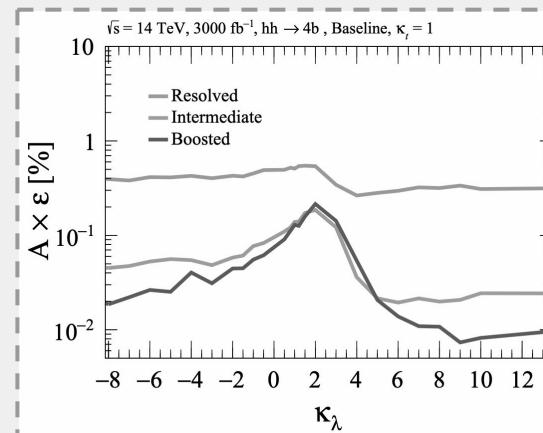


Analysis Strategy



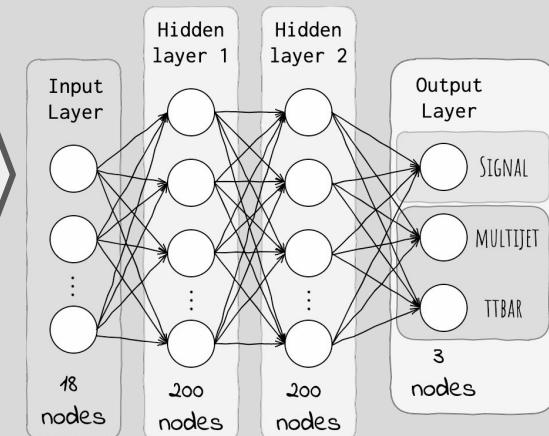
Baseline Analysis

- Cut Based
- ATLAS/CMS-inspired

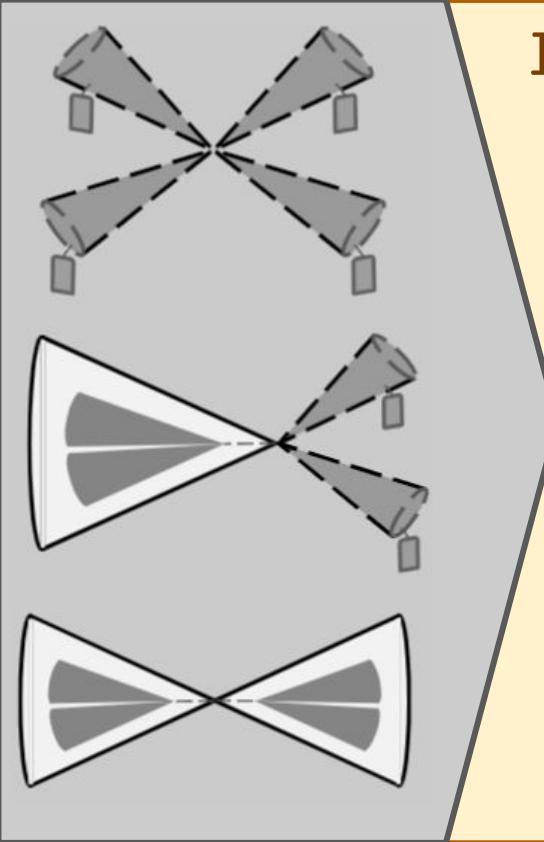


DNN Analysis

- Trained NN **classifier**
- Cut on **NN score**

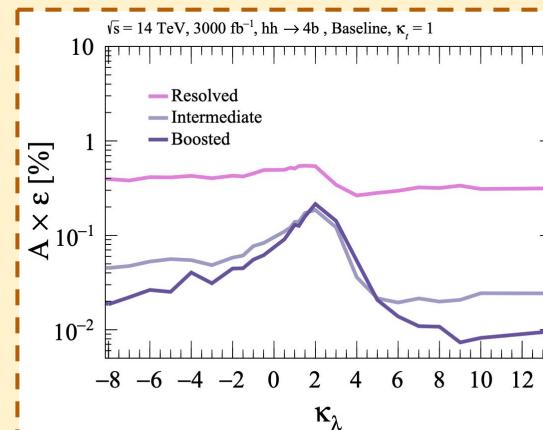


Analysis Strategy



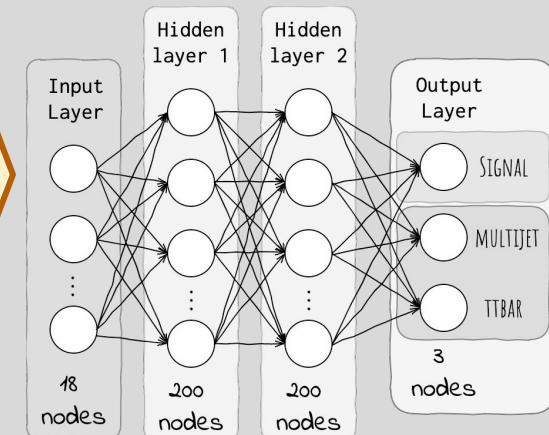
Baseline Analysis

- Cut Based
- ATLAS/CMS-inspired



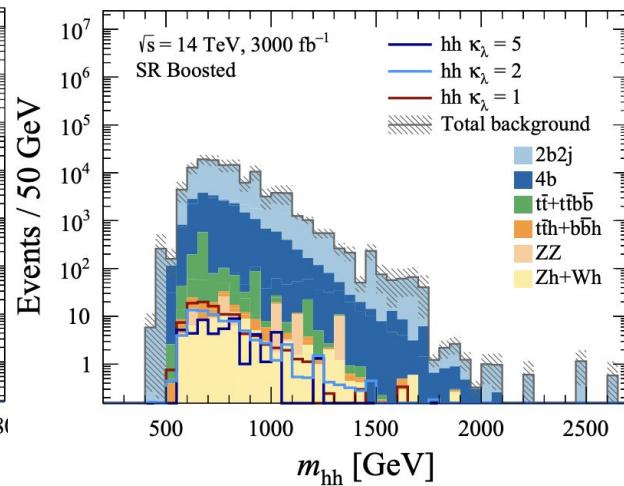
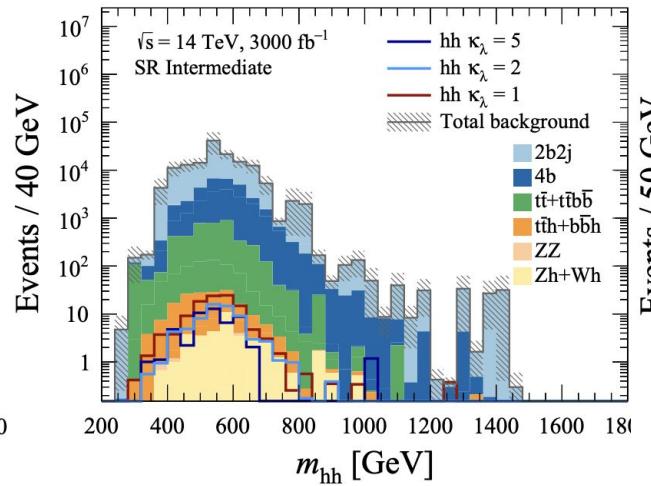
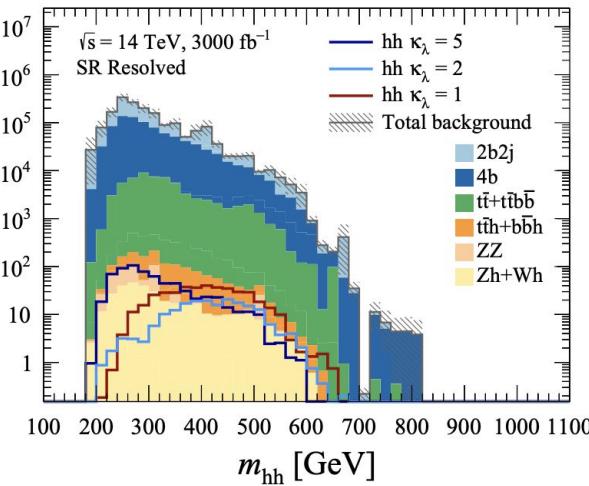
DNN Analysis

- Trained NN classifier
- Cut on NN score

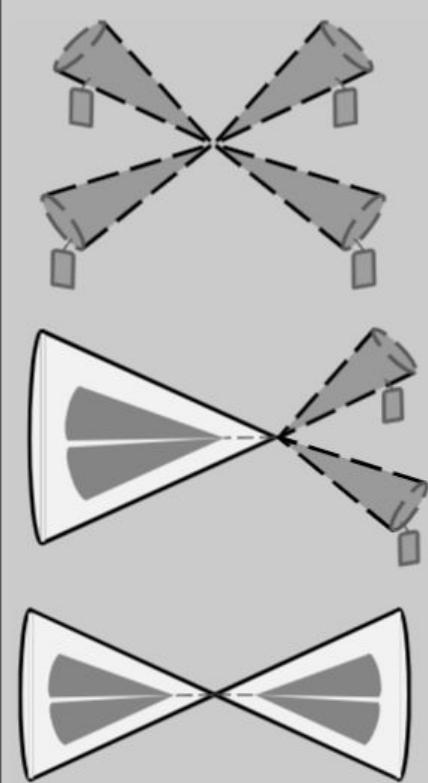


Baseline Analysis

- **Analysis-specific** cuts \Rightarrow define Signal Region (**SR**) in m_{hh}
 - $\hookrightarrow N(j_L \triangle) = 0$
 - $\hookrightarrow N(j_S \triangle) \geq 4$
 - \hookrightarrow Lepton, MET veto
 - \hookrightarrow 4 b-tags
 - $\hookrightarrow \Delta R(j_S^1 \triangle, j_S^2 \triangle)$ cut
- $N(j_L \triangle) = 1$
 - $\hookrightarrow N(j_S \triangle) \geq 2$
 - \hookrightarrow Lepton, MET veto
 - \hookrightarrow 4 b-tags
- $N(j_L \triangle) = 2$
 - $\hookrightarrow N(j_S \triangle) \geq 0$
 - \hookrightarrow Lepton, MET veto
 - \hookrightarrow 4 b-tags

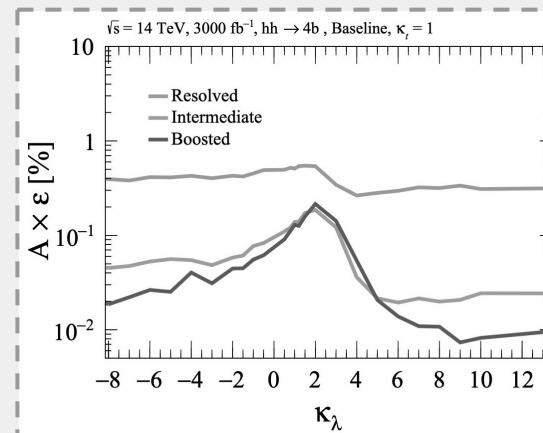


Analysis Strategy



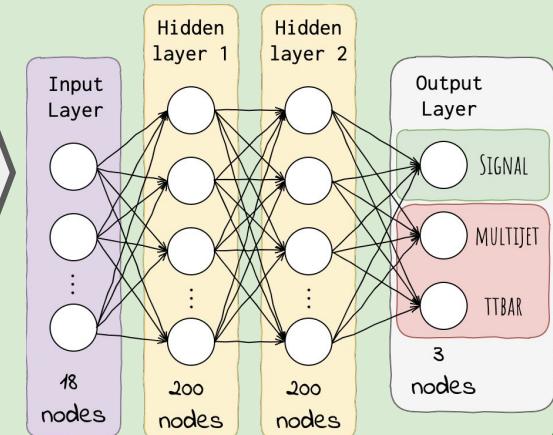
Baseline Analysis

- Cut Based
- ATLAS/CMS-inspired

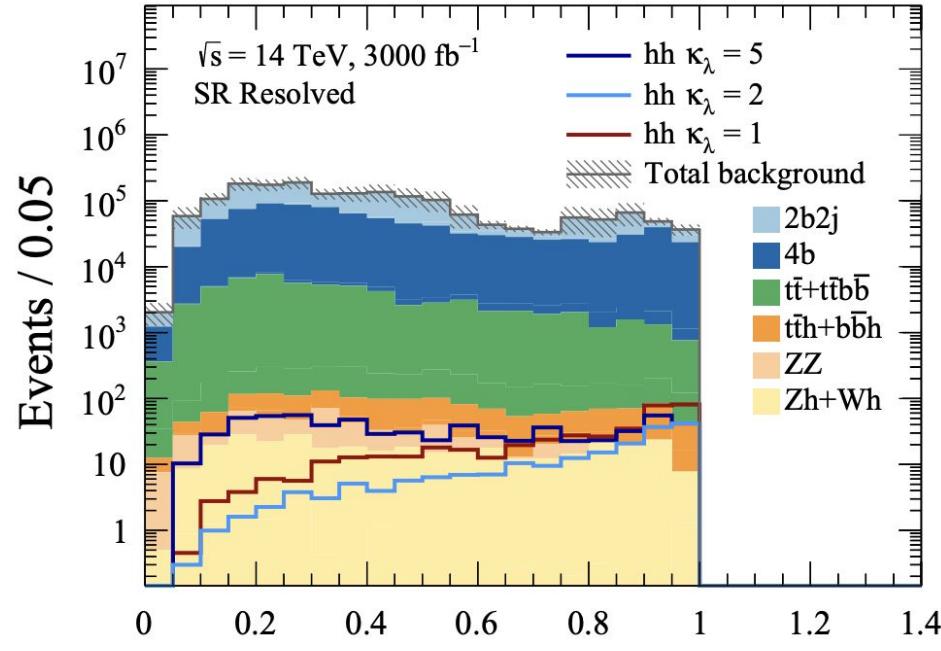


DNN Analysis

- Trained NN **classifier**
- **Cut on NN score**



DNN Analysis



$$\kappa_\lambda = \frac{\lambda}{\lambda_{SM}}$$

- **Multi-class** classifier
 - Signal vs multijet vs t̄t

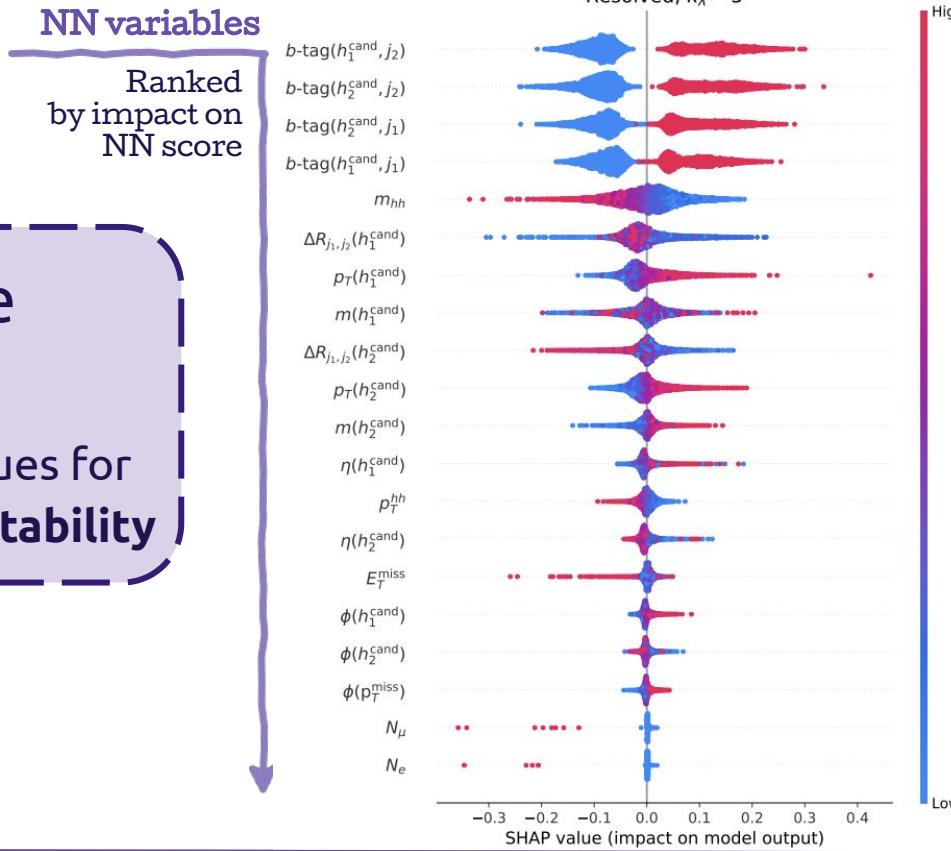
TRAINING VARIABLES

| | | |
|-------------------|--------------------------|------------------|
| ⇒ p_T^{HH} | ⇒ MET | ⇒ SUB-JET η |
| ⇒ M ^{HH} | ⇒ MET ϕ | ⇒ SUB-JET ϕ |
| ⇒ #MUONS | ⇒ SUB-JET MASS | ⇒ SUB-JETS ΔR |
| ⇒ #ELEC | ⇒ SUB-JET P _T | ⇒ SUB-JETS B-TAG |

- Cut ⇒ NN signal score > 0.75
- Trained with **multiple** κ_λ signals
 - **Use $\kappa_\lambda = 5$** network

Parentheses - What did our machine learn?

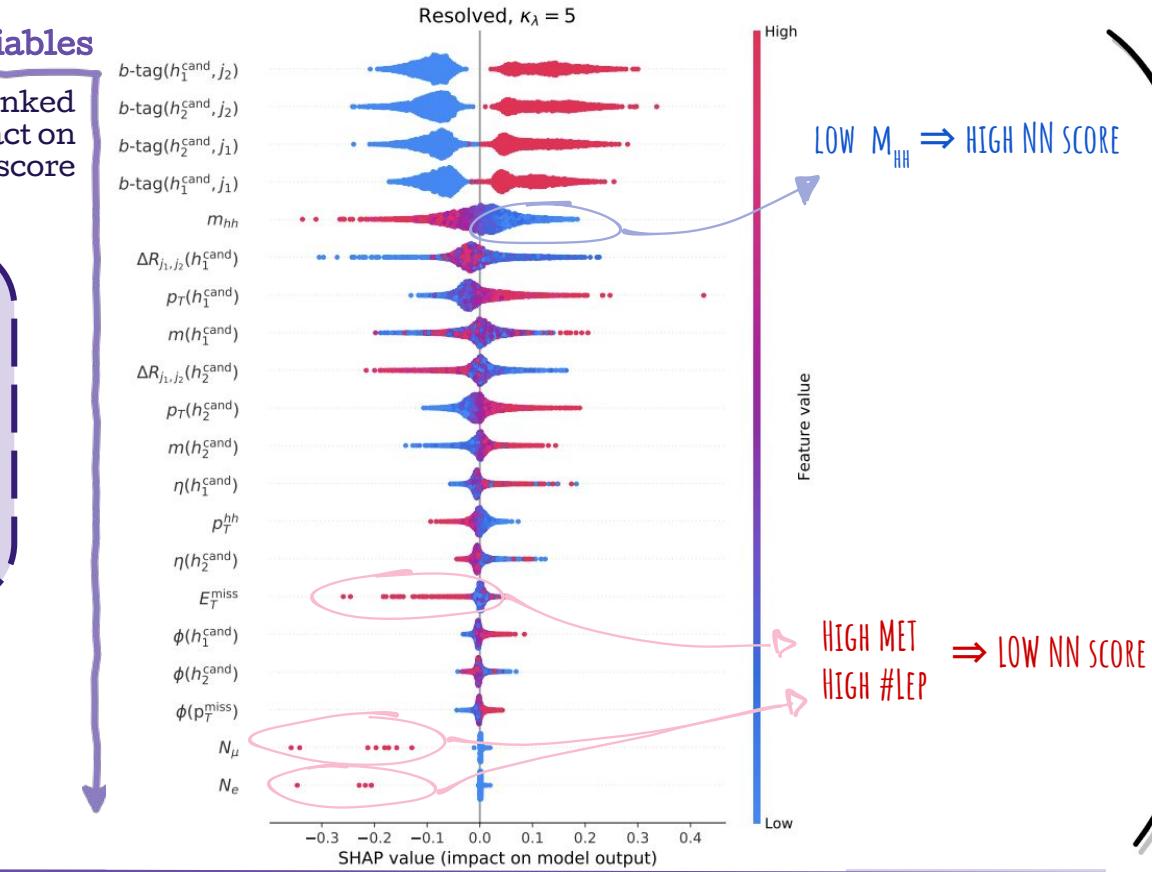
- SHAP value framework
 - ↪ Shapley values for ML interpretability



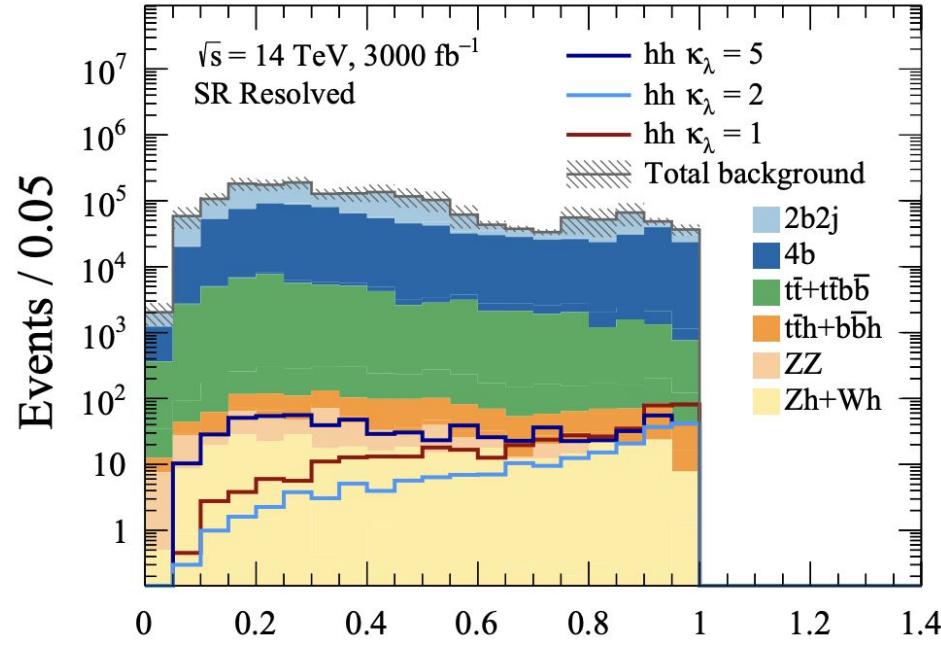
Parentheses - What did our machine learn?

- SHAP value framework
 - ↪ Shapley values for ML interpretability

NN variables
Ranked by impact on NN score



DNN Analysis



$$\kappa_\lambda = \frac{\lambda}{\lambda_{SM}}$$

NN signal score $\kappa_\lambda = 1$

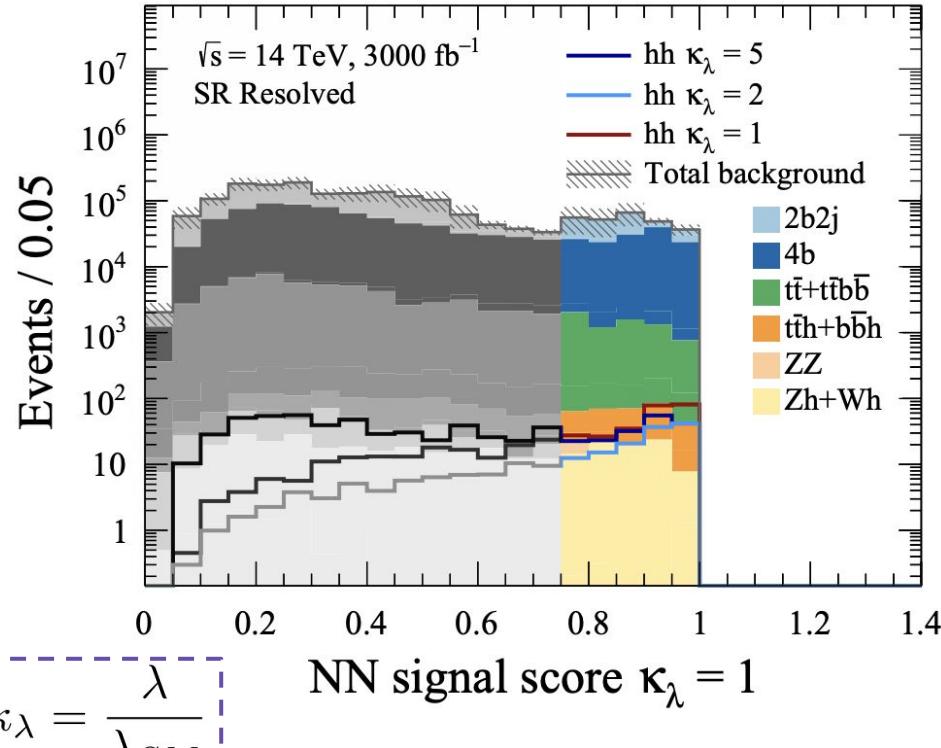
- **Multi-class** classifier
 - Signal vs multijet vs $t\bar{t}$

TRAINING VARIABLES

| | | |
|--------------|-----------------|-----------------------|
| ⇒ p_T^{HH} | ⇒ MET | ⇒ SUB-JET η |
| ⇒ M^{HH} | ⇒ MET ϕ | ⇒ SUB-JET ϕ |
| ⇒ #MUONS | ⇒ SUB-JET MASS | ⇒ SUB-JETS ΔR |
| ⇒ #ELEC | ⇒ SUB-JET P_T | ⇒ SUB-JETS B-TAG |

- Cut ⇒ NN signal score > 0.75
- Trained with **multiple** κ_λ signals
 - **Use $\kappa_\lambda = 5$ network**

DNN Analysis



- **Multi-class** classifier
 - Signal vs multijet vs $t\bar{t}$

TRAINING VARIABLES

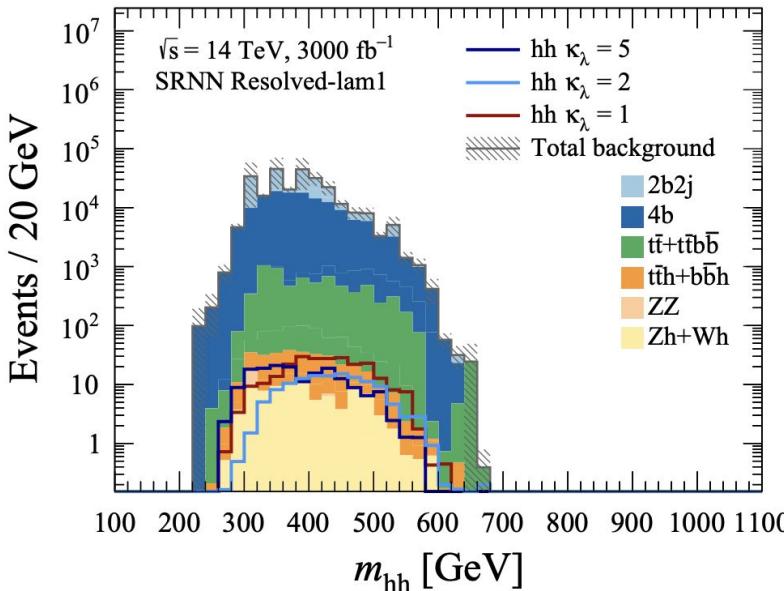
| | | |
|------------|---------------|---------------------|
| p_T^{HH} | MET | Sub-jet η |
| M^{HH} | MET ϕ | Sub-jet ϕ |
| #MUONS | Sub-jet mass | Sub-jets ΔR |
| #ELEC | Sub-jet P_T | Sub-jets B-TAG |

- Cut ⇒ NN signal score > 0.75
- Trained with **multiple κ_λ** signals
 - **Use $\kappa_\lambda = 5$** network

Parentheses - BSM k_λ training

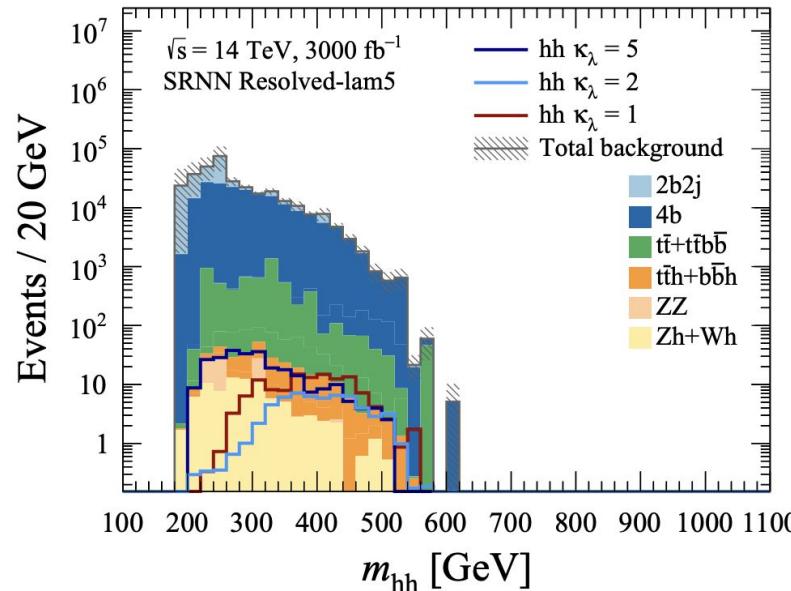
$k_\lambda = 1$ NN cut

- **Background** rejection ✓
- Signal **characterization** ✗

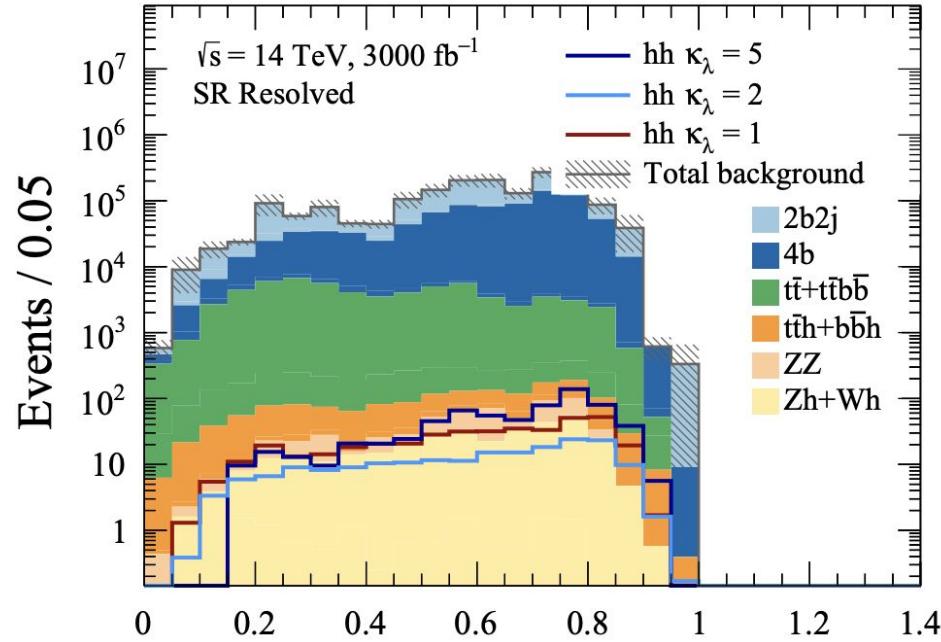


$k_\lambda = 5$ NN cut

- **Background** rejection ✓
- Signal **characterization** ✓



DNN Analysis



$$\kappa_\lambda = \frac{\lambda}{\lambda_{SM}}$$

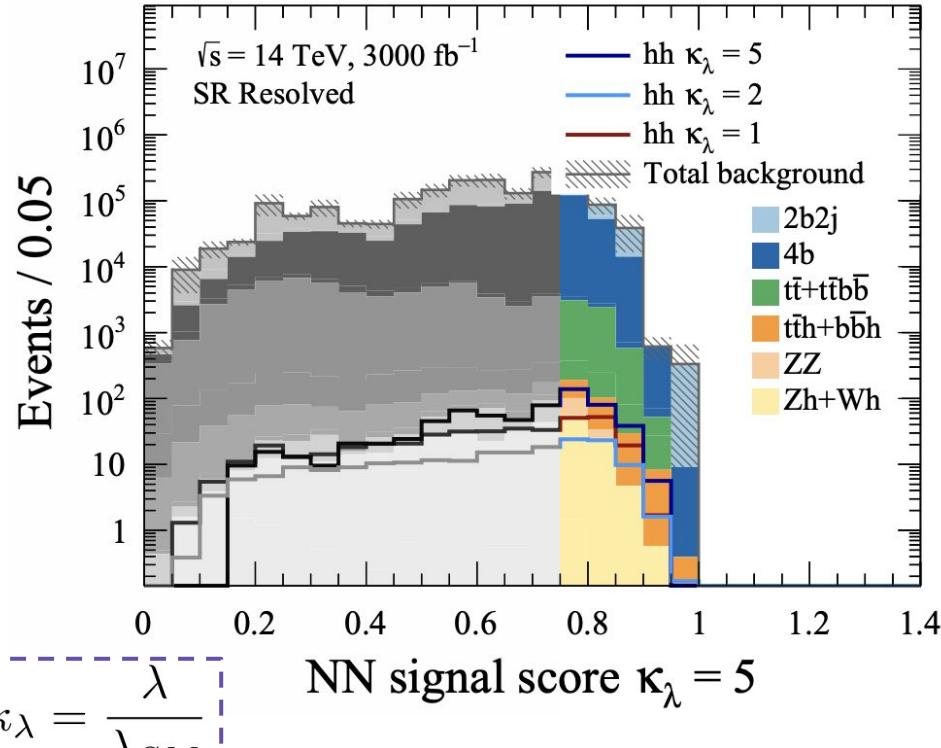
- **Multi-class** classifier
 - Signal vs multijet vs $t\bar{t}$

TRAINING VARIABLES

| | | |
|------------|---------------|---------------------|
| p_T^{HH} | MET | Sub-jet η |
| M^{HH} | MET ϕ | Sub-jet ϕ |
| #MUONS | Sub-jet mass | Sub-jets ΔR |
| #ELEC | Sub-jet p_T | Sub-jets B-TAG |

- Cut ⇒ NN signal score > 0.75
- Trained with **multiple** κ_λ signals
 - **Use $\kappa_\lambda = 5$** network

DNN Analysis



- **Multi-class** classifier
 - Signal vs multijet vs t̄t

TRAINING VARIABLES

| | | |
|-----------------------|---------------------------|---------------------------------|
| $\gg p_T^{HH}$ | $\gg \text{MET}$ | $\gg \text{SUB-JET } \eta$ |
| $\gg M^{HH}$ | $\gg \text{MET } \phi$ | $\gg \text{SUB-JET } \phi$ |
| $\gg \# \text{MUONS}$ | $\gg \text{SUB-JET MASS}$ | $\gg \text{SUB-JETS } \Delta R$ |
| $\gg \# \text{ELEC}$ | $\gg \text{SUB-JET } p_T$ | $\gg \text{SUB-JETS B-TAG}$ |

- Cut \Rightarrow NN signal score > 0.75
- Trained with **multiple κ_λ** signals
 - **Use $\kappa_\lambda=5$** network

Self-Coupling Constraints

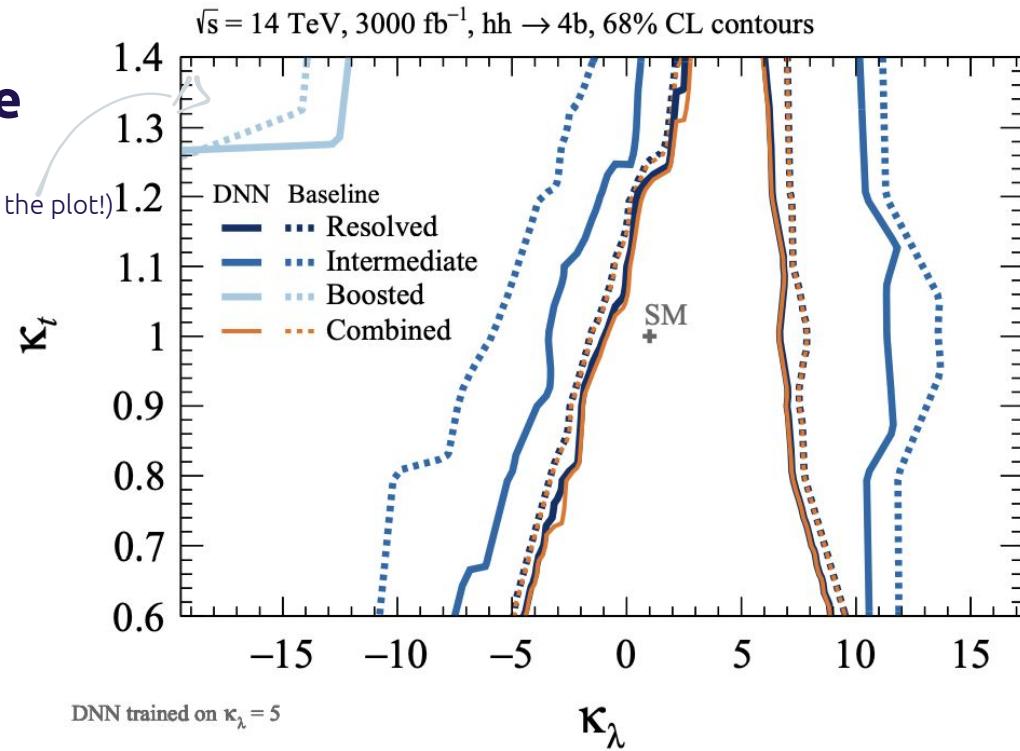


Constraints on k_λ - k_t Plane

- Resolved \rightarrow **most powerful**
 - ↪ Intermediate \rightarrow **non-negligible**
 - ↪ Boosted \rightarrow **negligible*** (but made it in the plot!)
- **Strong dependence** on k_t

Parentheses - Upper Bound for k_t

- Assuming $k_\lambda = 1 \rightarrow k_t < 1.22!$
 - ↪ (Not really a **safe** assumption)
- **Independent** of current methods
- Possible in **any hh process**
 - ↪ Likely better in **other decay modes**

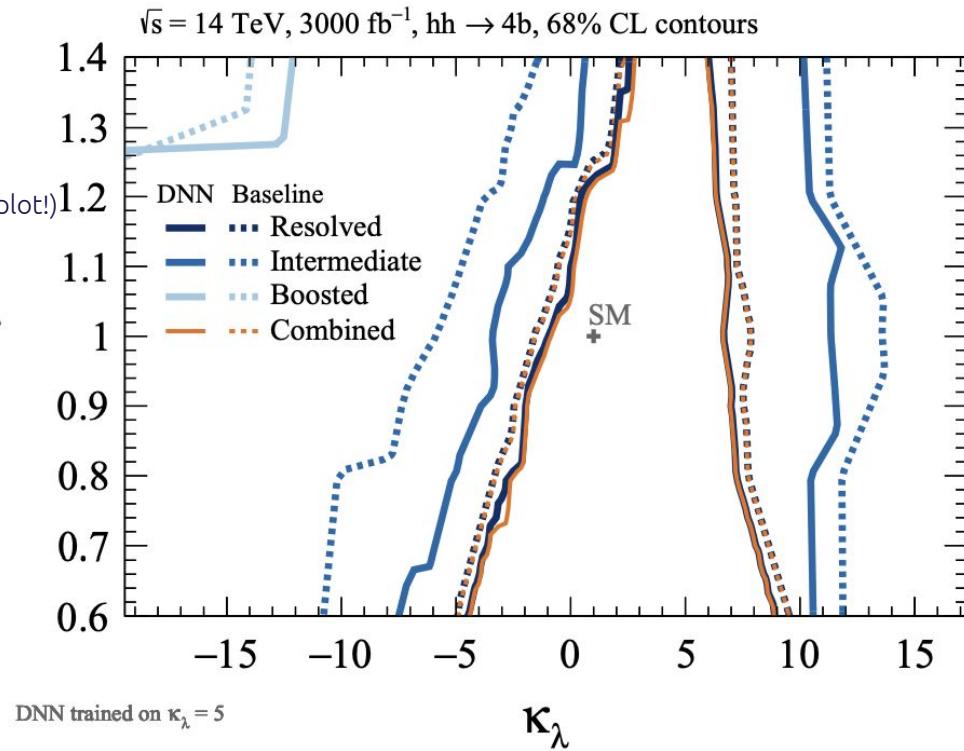


Constraints on k_λ - k_t Plane

- Resolved \rightarrow **most powerful**
 - ↪ Intermediate \rightarrow **non-negligible**
 - ↪ Boosted \rightarrow **negligible*** (but made it in the plot!)
- **Strong dependence** on k_t

Parentheses - Upper Bound for k_t

- Assuming $k_\lambda = 1 \rightarrow k_t < 1.22!$
 - ↪ (Not really a **safe** assumption)
- **Independent** of current methods
- Possible in **any hh process**
 - ↪ **Likely better** in **other decay modes**



*Note that this does not necessarily apply to analyses optimized for discovery of SM hh production - only those aiming to constrain k_λ .

Conclusion

Conclusions

- Showed a detailed **comparison** of λ_{hhh} **constraints** in $hh \rightarrow 4b$ resolved, intermediate and boosted channels, in the context of HL-LHC
 - ↪ **Resolved most constraining**, then intermediate and then boosted
- A basic **DNN analysis** provided **noticeable improvement** over the cut based baseline analysis
- Best constraints came from NN trained on BSM signal
 - ↪ $hh \rightarrow 4b$ analyses **optimized** for **discovery of SM hh** may be **suboptimal**

Conclusions

- **Experimental limitations**, triggering and jet reconstruction, **affect** the reconstruction of a **key discriminating variable** m_{hh}
- **Uncertainty** on \mathbf{k}_t has a strong impact on sensitivity to \mathbf{k}_λ
- **4b** is a **challenging** hh channel for λ_{hhh} constraints, but can provide important **independent information** for statistical **combinations**



Thanks! 😊



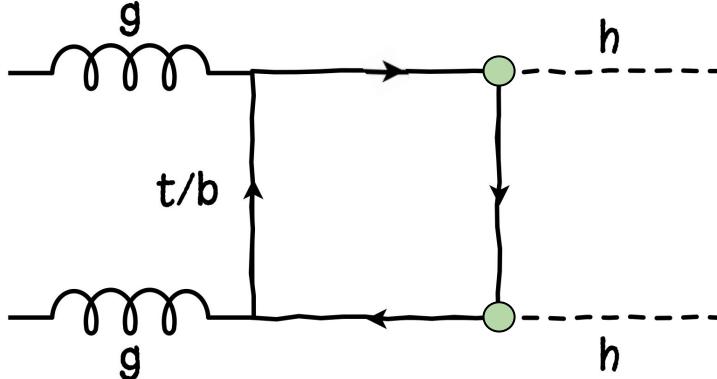
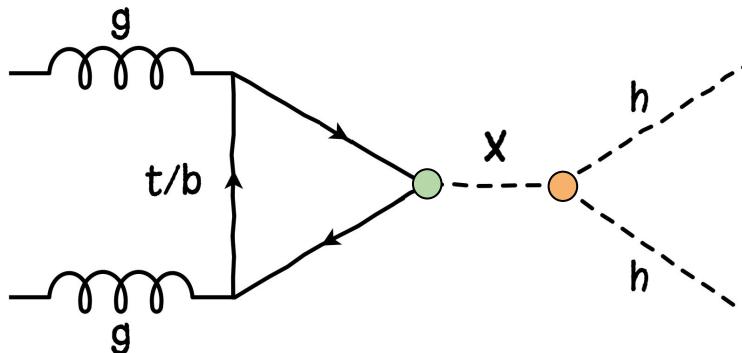
UNIVERSITÉ
LIBRE
DE BRUXELLES

Jacob Amacker, William Balunas, Lydia Beresford, Daniela Bortoletto,
James Frost, Cigdem Issever, Jesse Liu, James McKee, Alessandro Micheli,
Santiago Paredes Saenz, Michael Spannowsky, and Bejan Stanislaus
santiago.paredes@cern.ch

Higgs 2021
October 2021

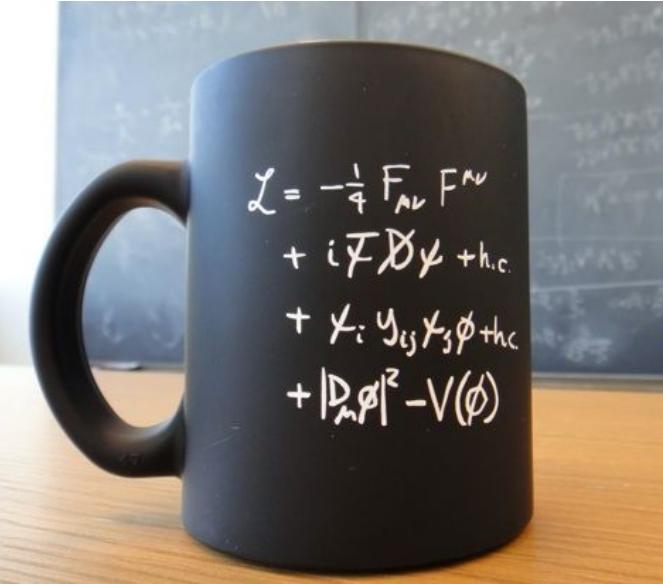
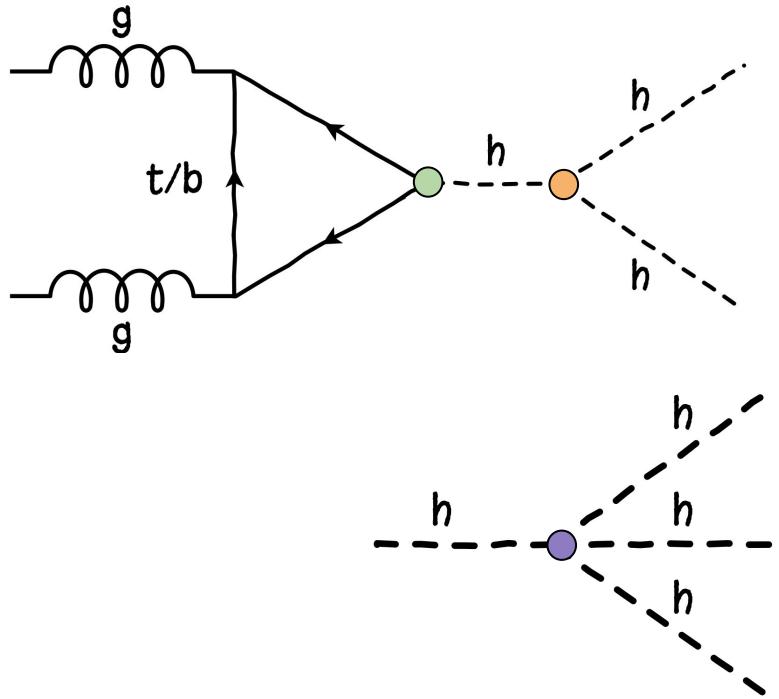


Why hh?



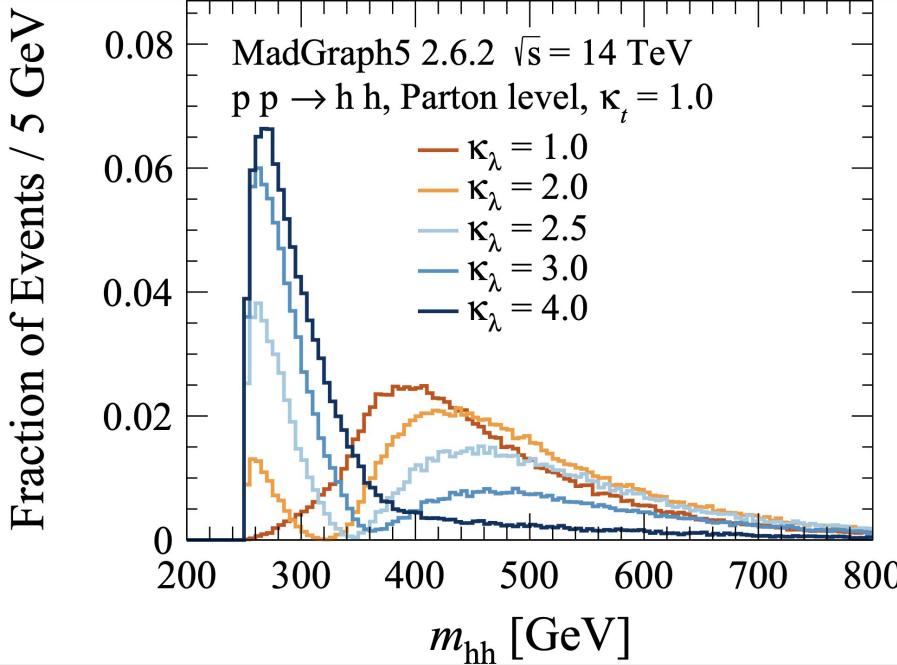
- Standard Model
 - ↪ Sensitive to the higgs self-coupling
 - ↪ Also to the $t\bar{t}h$ vertex
- Beyond the SM
 - ↪ New physics effects in loops
 - ↪ Heavy resonances (X) decaying to di-higgs

Why di-higgs?



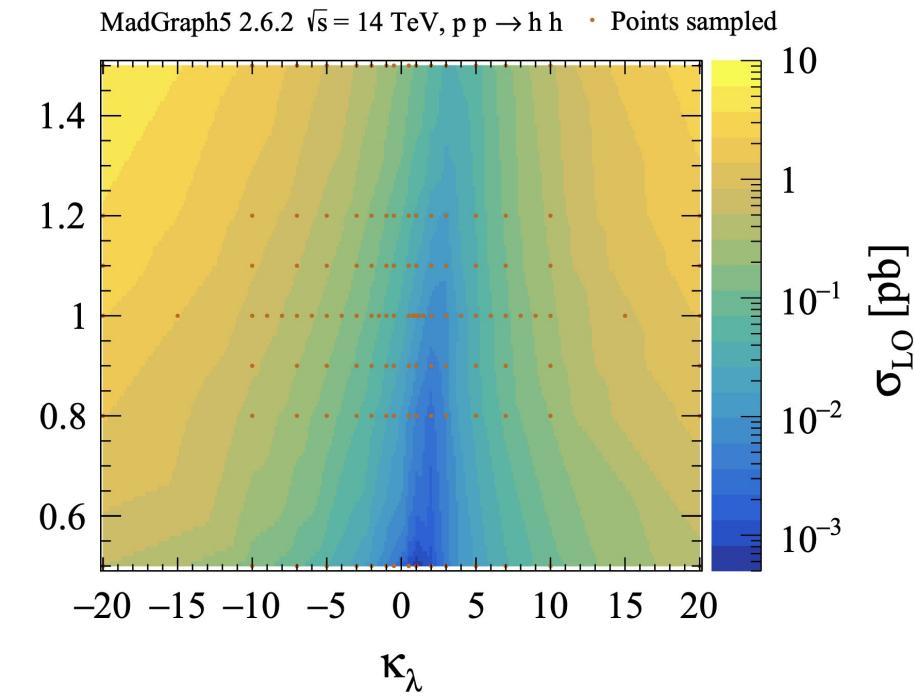
$$\lambda v^2 h^2 + \boxed{\lambda v h^3} + \boxed{\frac{\lambda}{4} h^4}$$

Signal Samples



- $gg \rightarrow hh$ production
 - ↪ **100k events** per point
 - ↪ MadGraph 2.6.2
 - ↪ **Inclusive h decay**
- Decay, parton shower, hadronization, and underlying event → Pythia 8.230
- **Varied** coupling to **top** quark and **self couplings**
 - ↪ All **BSM couplings** set to 0
- Extra $\kappa_t=1$ samples for **ML training**
 - ↪ **250k** events per point
 - ↪ Exclusive decay $h \rightarrow b\bar{b}$

Signal Samples



- $gg \rightarrow hh$ production
 - ↪ **100k events** per point
 - ↪ MadGraph 2.6.2
 - ↪ **Inclusive h decay**
- Decay, parton shower, hadronization, and underlying event
 - ↪ Pythia 8.230
- **Varied** coupling to **top** quark and **self couplings**
 - ↪ All **BSM couplings** set to 0
- Extra $k_t=1$ samples for **ML training**
 - ↪ **250k** events per point
 - ↪ Exclusive decay $h \rightarrow bb$

Event and Object Selection

| Observable | Preselection | | |
|---------------------|--|--|-------------|
| Large jet j_L | $R = 1.0, p_T > 250 \text{ GeV}, \eta < 2.0$ | | |
| Small jet j_S | $R = 0.4, p_T > 40 \text{ GeV}, \eta < 2.5$ | | |
| Track jet j_T | $R = 0.2, p_T > 20 \text{ GeV}, \eta < 2.5$ | | |
| $j_T \in j_L$ | $\Delta R(j_T, j_L) < 1.0$ | | |
| | Resolved | Intermediate | Boosted |
| $N(j_L)$ | $= 0$ | $= 1$ | $= 2$ |
| $N(j_S)$ | ≥ 4 | ≥ 2 | ≥ 0 |
| h_1^{cand} | $j_S^{(i)}$ pair | j_L | $j_L^{(1)}$ |
| h_2^{cand} | $j_S^{(i)}$ pair | $j_S^{(i)}$ pair, $\Delta R(j_S^{(i)}, j_L) > 1.2$ | $j_L^{(2)}$ |
| ΔR_{jj} | See Eqs. 3.2, 3.3 | — | — |

Signal region definitions

| Signal region | | | |
|--|---|--------------|----------|
| | Resolved | Intermediate | Boosted |
| $j_T \in h_1^{\text{cand}}$ | — | ≥ 2 | ≥ 2 |
| $j_T \in h_2^{\text{cand}}$ | — | — | ≥ 2 |
| <i>b</i> -tagging | Two <i>b</i> -tags for each h_i^{cand} | | |
| $ \Delta\eta(h_1, h_2) $ | < 1.5 | | |
| $E_{\text{T}}^{\text{miss}}$ | $< 150 \text{ GeV}$ | | |
| $p_{\text{T}}^{\ell}, \eta_{\ell} $ | $> 10 \text{ GeV}, < 2.5$ | | |
| N_{ℓ} | $= 0$ | | |
| $p_{\text{signal}}^{\text{DNN}}$ | > 0.75 (<i>neural network analysis</i> only) | | |
| Lower bin edges for m_{hh} binning [GeV] | | | |
| Resolved | [200, 250, 300, 350, 400, 500] | | |
| Intermediate | [200, 500, 600] | | |
| Boosted | [500, 800] | | |

Fixed $k_t=1$

BSM k_λ
yield

SM k_λ
yield

| Category | Systematic ζ_b |
|--------------|----------------------|
| Resolved | 0.3% |
| Intermediate | 1% |
| Boosted | 5% |

$$\chi^2 = \sum_i \left[\frac{(S - S_{\text{SM}})^2}{S + B + (\zeta_b B)^2 + (\zeta_s S)^2} \right]_i$$

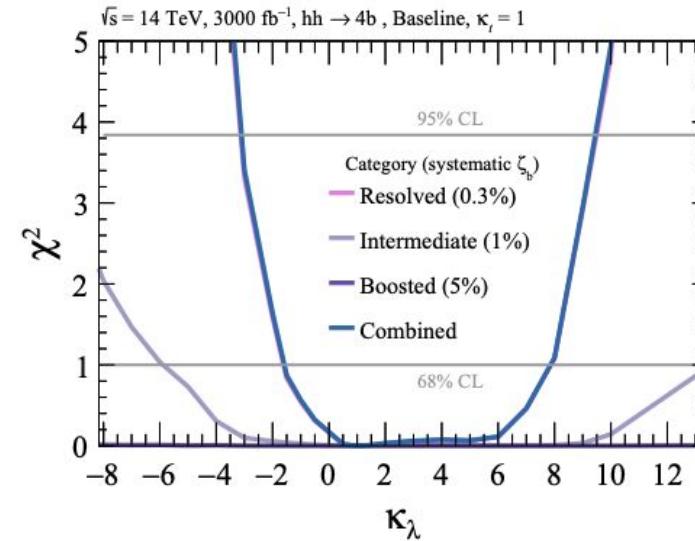
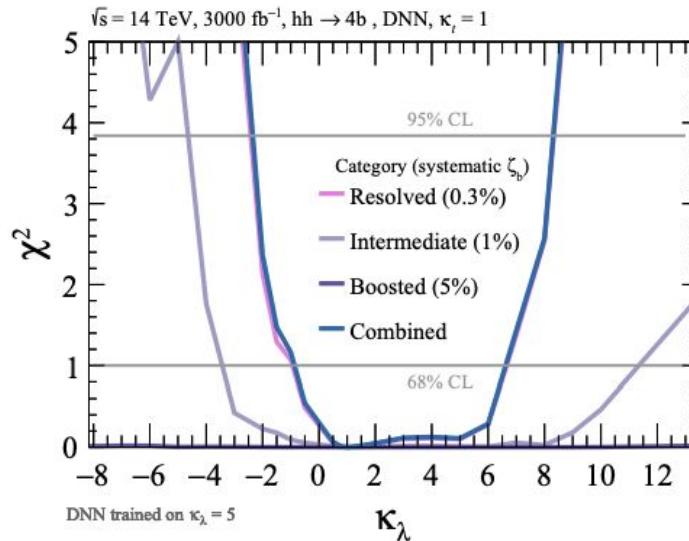
m_{hh}
bins

background
uncertainty

signal
uncertainty

Constraints on k_λ - Fixed $k_t=1$

- Resolved \rightarrow **most powerful**
 - ↪ Intermediate \rightarrow **non-negligible**
 - ↪ Boosted \rightarrow **negligible***
- **Basic DNN analysis improved sensitivity**



*Note that this does not necessarily apply to analyses optimized for discovery of SM hh production - only those aiming to constrain k_λ .

Parentheses - Impact of BKG Uncertainty

- Background uncertainty has **large impact** on sensitivity
 - ↪ Often a **large uncertainty** in $hh \rightarrow 4b$ searches

