Hunting coloured scalars with machine learning

Manuel Kunkel
in collaboration with Thomas Flacke, Jeong Han Kim, Jun Seung Pi, Werner Porod

supported by

Studienstiftung des deutschen Volkes

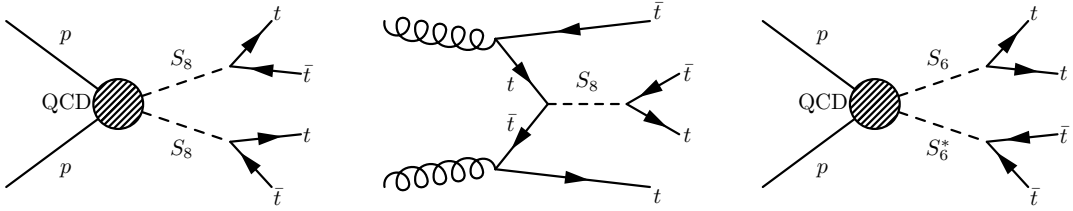
GRK 2994

DFG

September 25, 2025
DESY Theory Workshop 2025

Model

Simplified model: Extend the SM by S_8 or S_6 with couplings to top quarks:



(motivated by Composite Higgs models)

Preselection cuts

Design a NN-based search for

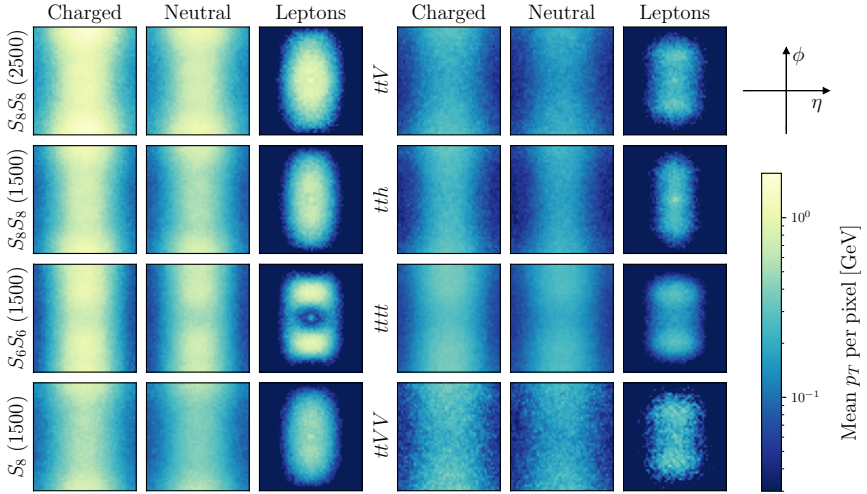
$$t\bar{t}t\bar{t} \to (W^+b)(W^-\bar{b})(W^+b)(W^-\bar{b}) \to \ell^\pm\ell^\pm + 4b + 4j + \not{p}_T,$$

Preselection:

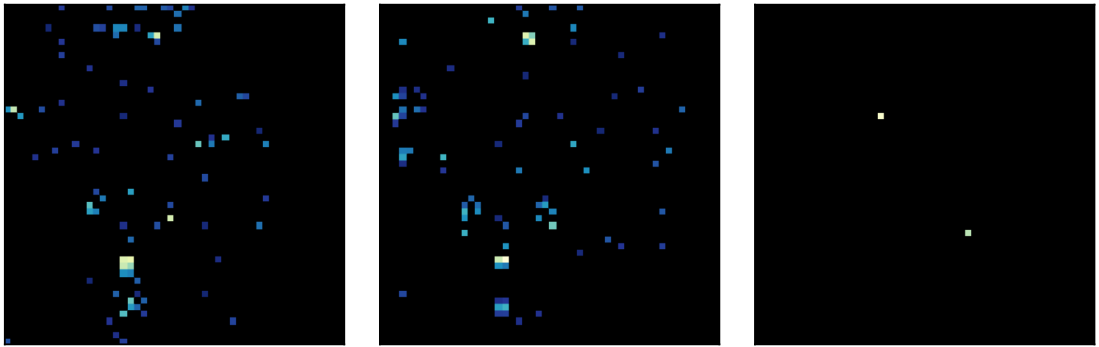
- 2 same-sign leptons
- ≥ 3 light jets
- ≥ 3 b-tagged jets
- and some kinematic cuts

$\Rightarrow$ Backgrounds: $tttt$, ttV, $ttVV$, tth

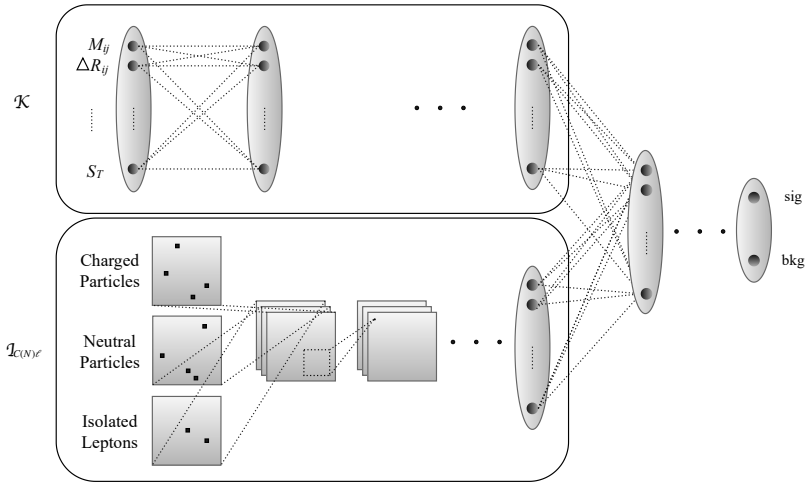
Jet images


Charged
Neutral
Leptons
Charged
Neutral
Leptons
S_8S_8 (2500)
S_8S_8 (1500)
S_6S_6 (1500)
S_8 (1500)
$t\bar{t}V$
$t\bar{t}h$
$t\bar{t}t\bar{t}$
$t\bar{t}VV$
ϕ
η
10^0
10^{-1}
Mean p_T per pixel [GeV]

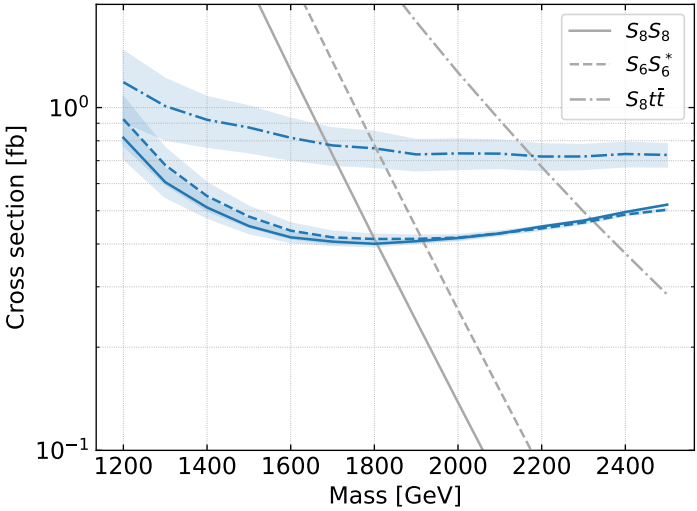
Jet images



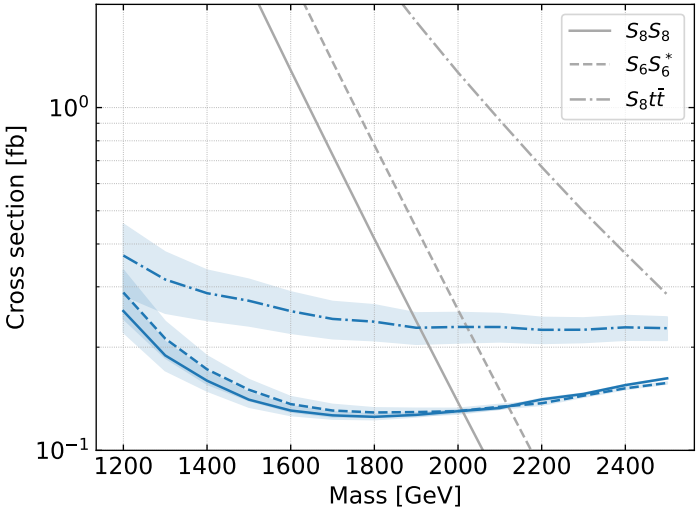
Neural network architecture



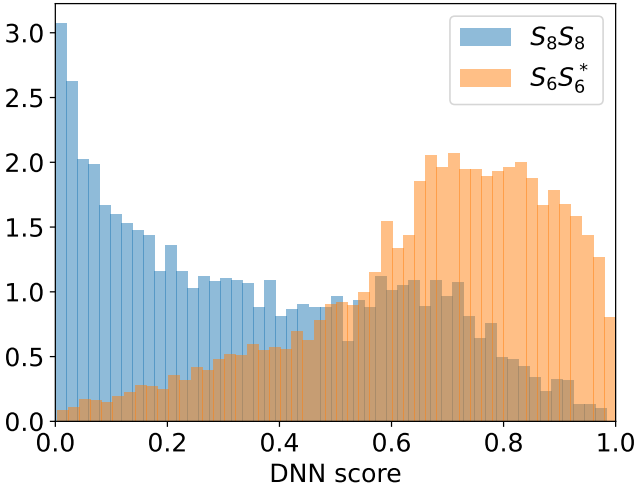
Expected discovery reach at 3 ab^{-1}



Expected exclusion limit at 3 ab^{-1}

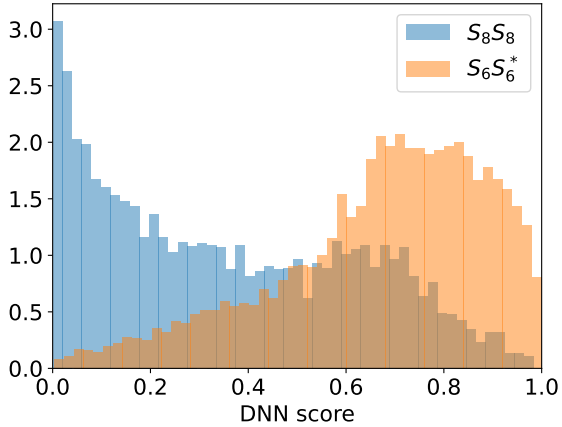


Identifying the signal process: octet or sextet?



Identifying the signal process: octet or sextet?

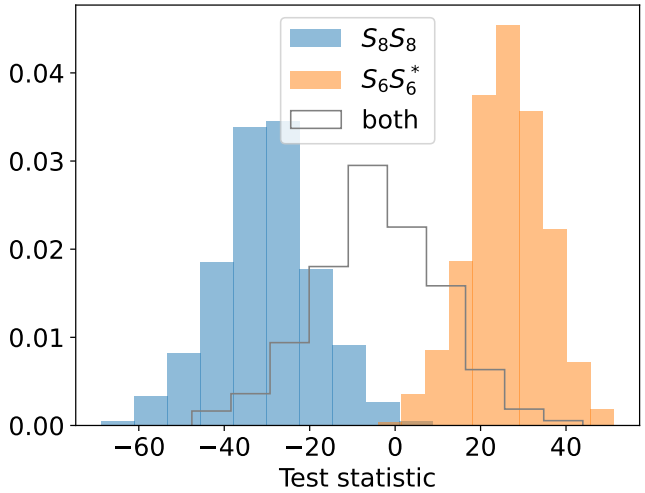
Given DNN scores $\{x_i\}$, which distribution did they come from?



$$\mathcal{L}_j(\{x_i\}) = \prod_{i=1}^{N} f_j(x_i)$$

$$t = -2 \ln \frac{\mathcal{L}_1(\{x_i\})}{\mathcal{L}_2(\{x_i\})}$$

Identifying the signal process: octet or sextet?



Summary

With computer vision tools we can

- search for processes with busy final states,
- discover coloured scalars in four top quark events,
- identify the underlying resonance.
- More details: `arXiv:2506.04318`

Backup

Technical details

- Simplified models implemented in FeynRules
- Generate events with MadGraph5 at LO or NLO
- PDF set NNPDF 3.0
- Showering with Pythia8
- Detector simulation with Delphes
- Machine learning with PyTorch

Pile-up (different signal: 2304.09195)