Learning to pinpoint effective operators at the LHC: a study of the $t\bar{t}b\bar{b}$ signature

Jorgen D’Hondt\textsuperscript{(a)}, Alberto Mariotti\textsuperscript{(a,b)}, Ken Mimasu\textsuperscript{(c)}, Seth Moortgat\textsuperscript{(a)}, Cen Zhang \textsuperscript{(d)}

\textsuperscript{(a)} Vrije Universiteit Brussel – Interuniversity Institute for High Energies (IIHE)  
Pleinlaan 2, B-1050 Brussels, Belgium

\textsuperscript{(b)} Vrije Universiteit Brussel, Theoretische Natuurkunde  
Pleinlaan 2, B-1050 Brussels, Belgium

\textsuperscript{(c)} Université catholique de Louvain – Centre for Cosmology, Particle Physics and Phenomenology (CP3)  
Chemin du Cyclotron 2, B-1348 Louvain-La-Neuve, Belgium

\textsuperscript{(d)} Chinese Academy of Sciences, Institute of High Energy Physics  
Beijing 100049, China

\textbf{ArXiv:} 1807.02130  \textbf{JHEP11(2018)131}

\textit{EPS-HEP 2019 (Ghent) – 12\textsuperscript{th} July 2019}
1. SMEFT: $t\bar{t}b\bar{b}$ and its virtues
   a. Four-quark operators
   b. Complementarity to four top

2. Learning the effective operators
   1. Individual constraints
   2. Multiple operators

3. Conclusion
The Standard Model Effective Field Theory

Lack of direct evidence for BSM physics at the LHC → Standard Model Effective Field Theory (SMEFT):

- Model-independent interpretation
- New physics at high energy scales
- Heightened energy dependence and modified kinematics

Extend SM Lagrangian up to dim. 6:

(→ Leading B & L conserving contributions)

$$\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{C_i}{\Lambda^2} O_i^{(6)}$$

$$M^2 \equiv \Lambda^2 \gg p^2$$

$$\frac{g^2}{p^2 - M^2}$$

dim. 6

$$\begin{align*}
- \frac{g^2}{\Lambda^2} \left[ 1 + \frac{p^2}{\Lambda^2} + \frac{p^4}{\Lambda^4} + \cdots \right]
\end{align*}$$
$t\bar{t}b\bar{b}$ is sensitive to a set of four-quark dim. 6 operators.

MFV-inspired approach to separate 4-Heavy, 2-Heavy-2-Light and 4-Light operators

We focus on 4-Heavy operators

2H2L are constrained much more by $t\bar{t}$ and $b\bar{b}$ production via $q\bar{q}$ initial state

<table>
<thead>
<tr>
<th>Operator</th>
<th>$t\bar{t}b\bar{b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O^1_{QQ} = \frac{1}{2} (\bar{Q} \gamma_\mu Q) (\bar{Q} \gamma^\mu Q)$</td>
<td>✔</td>
</tr>
<tr>
<td>$O^8_{QQ} = \frac{1}{2} (\bar{Q} \gamma_\mu T^A Q) (\bar{Q} \gamma^\mu T^A Q)$</td>
<td>✔</td>
</tr>
<tr>
<td>$O^1_{tb} = (\bar{t} \gamma_\mu t) (\bar{b} \gamma_\mu b)$</td>
<td>✔</td>
</tr>
<tr>
<td>$O^8_{tb} = (\bar{t} \gamma_\mu A t) (\bar{b} \gamma_\mu A b)$</td>
<td>✔</td>
</tr>
<tr>
<td>$O^1_{tt} = (\bar{t} \gamma_\mu t) (\bar{t} \gamma_\mu t)$</td>
<td>✔</td>
</tr>
<tr>
<td>$O^1_{bb} = (\bar{b} \gamma_\mu b) (\bar{b} \gamma_\mu b)$</td>
<td>✔</td>
</tr>
<tr>
<td>$O^1_{Qt} = (\bar{Q} \gamma_\mu Q) (\bar{t} \gamma^\mu t)$</td>
<td>✔</td>
</tr>
<tr>
<td>$O^8_{Qt} = (\bar{Q} \gamma_\mu T^A Q) (\bar{t} \gamma^\mu T^A t)$</td>
<td>✔</td>
</tr>
<tr>
<td>$O^1_{Qb} = (\bar{Q} \gamma_\mu Q) (\bar{b} \gamma^\mu b)$</td>
<td>✔</td>
</tr>
<tr>
<td>$O^8_{Qb} = (\bar{Q} \gamma_\mu T^A Q) (\bar{b} \gamma^\mu T^A b)$</td>
<td>✔</td>
</tr>
<tr>
<td>$O^1_{QtQb} = (\bar{Q} t) \varepsilon (\bar{Q} b)$</td>
<td>✔</td>
</tr>
<tr>
<td>$O^8_{QtQb} = (\bar{Q} T^A t) \varepsilon (\bar{Q} T^A b)$</td>
<td>✔</td>
</tr>
</tbody>
</table>
Some operators can be constrained by four top as well

\[ O_{QQ}^1 = \frac{1}{2} (\bar{Q} \gamma_\mu Q) (\bar{Q} \gamma^\mu Q), \]
\[ O_{QQ}^8 = \frac{1}{2} (\bar{Q} \gamma_\mu T^A Q) (\bar{Q} \gamma^\mu T^A Q), \]
\[ O_{tb}^1 = (\bar{t} \gamma_\mu t) (\bar{b} \gamma_\mu b), \]
\[ O_{tb}^8 = (\bar{t} \gamma_\mu T^A t) (\bar{b} \gamma_\mu T^A b), \]
\[ O_{tt}^1 = (\bar{t} \gamma_\mu t) (\bar{t} \gamma_\mu t), \]
\[ O_{bb}^1 = (\bar{b} \gamma_\mu b) (\bar{b} \gamma_\mu b), \]
\[ O_{Qt}^1 = (\bar{Q} \gamma_\mu Q) (\bar{t} \gamma^\mu t), \]
\[ O_{Qt}^8 = (\bar{Q} \gamma_\mu T^A Q) (\bar{t} \gamma^\mu T^A t), \]
\[ O_{Qb}^1 = (\bar{Q} \gamma_\mu Q) (\bar{b} \gamma^\mu b), \]
\[ O_{Qb}^8 = (\bar{Q} \gamma_\mu T^A Q) (\bar{b} \gamma^\mu T^A b), \]
\[ O_{QtQb}^1 = (\bar{Q} t) \varepsilon (\bar{Q} b), \]
\[ O_{QtQb}^8 = (\bar{Q} T^A t) \varepsilon (\bar{Q} T^A b). \]

\[
C_{QQ}^{(+)} = \frac{1}{2} C_{QQ}^1 + \frac{1}{6} C_{QQ}^8
\]

Degeneracy in four-top, lifted for \( \bar{t} \bar{b} b \bar{b} \)!

**Pre-requisite:**

\( \bar{t} \bar{b} b \bar{b} \) has a sufficiently large production cross section (~3 pb) to exploit differential kinematical information with 300 fb-1 (after Run III)!

(*for comparison: \( \sigma_{tttt} \sim 9 \text{ fb} *)
The name of the game: Increasing sensitivity to SMEFT operators

\[ \sigma = \sigma_{SM} + \sum_i \frac{C_i}{\Lambda^2} \tilde{\sigma}_i + \sum_{i,j} \frac{C_i C_j}{\Lambda^4} \tilde{\delta}_{i,j} \]

1 operator:

\[ \sigma = \sigma_{SM} + p_1 \cdot \frac{C_i}{\Lambda^2} + p_2 \cdot \frac{C_i^2}{\Lambda^4} \]

Figure 1.11: Feynman diagrams describing a new mediator \( W \) with mass \( \Lambda \) that couples to the SM particles with a new physics coupling \( g \) (left) and the corresponding EFT vertex describing the point–like interaction (right).

The allowed operators need to obey the gauge invariance of the SM gauge group. There exists only one such operator of dimension five, which is a lepton–number-violating operator that could provide a mass term for the neutrinos [102]. At dimension six however, a whole new world of operators opens up which are suppressed by the new physics scale squared \( \Lambda \). Depending on the flavor assumptions, the number of dimension six operators can go as high as a few thousands if one assumes full flavor–non–universality. However, the minimal set of operators needed in a fully flavor–universal scenario is 59. There exists a freedom in choosing a particular basis of operators to fully describe the SMEFT at dimension six. A popular choice is the so–called Warsaw basis [101]. Higher order operators are suppressed by even higher orders of \( \Lambda \), and can often be neglected to first order. Nevertheless for some processes the higher order operators are relevant, or even the first relevant contributions to consider. Therefore one should always validate whether or not higher orders can safely be neglected. It is useful to note that all operators with odd-numbered dimensions can only generate baryon or lepton number violating processes. Any observable, such as a cross section, or a number of observed events in a given phase space region, can be expressed in terms of its SM value and its additional contributions due to the SMEFT effects. For contributions of dimension six operators, the functional form of an observable \( \bar{\sigma} \) can be expressed as in Eq. (1.28), where the indices \( i \) and \( j \) run over the number of operators that are considered for the given process and \( \tilde{\sigma}_i \) and \( \tilde{\delta}_{i,j} \) are coefficients to be determined. In this notation, \( \tilde{\sigma}_i \) signifies the strength of the interference of the SMEFT operators with the SM, whereas \( \tilde{\delta}_{i,j} \) represents the pure EFT contribution, including quadratic terms for a single operator and the interference effects amongst SMEFT operators.
The name of the game: Increasing sensitivity to SMEFT operators

\[ \sigma = \sigma_{SM} + \sum_i \frac{C_i}{\Lambda^2} \tilde{\sigma}_i + \sum_{i,j} \frac{C_i}{\Lambda^4} \tilde{\sigma}_{i,j} \]

1 operator:

\[ \sigma = \sigma_{SM} + p_1 \cdot \frac{C_i}{\Lambda^2} + p_2 \cdot \frac{C_i^2}{\Lambda^4} \]

Measurement → limits on Wilson Coeff.
Combine all available kinematics in a (shallow) neural network (NN) to select EFT enriched phase space.

Instead of a binary classifier (SM vs EFT), we exploit **multi-class** structure to also distinguish amongst **EFT operators** with left-handed top quark currents \((t_L)\) and with right-handed top quark currents \((t_R)\)! 

18 kinematic input observables

<table>
<thead>
<tr>
<th>(\Delta R)</th>
<th>(m_{\text{inv}})</th>
<th>(p_T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta R(\ell_1,\ell_2))</td>
<td>(m_{\text{inv}}(\ell_1,\ell_2))</td>
<td>(p_T(\ell_1))</td>
</tr>
<tr>
<td>(\Delta R(b_1,b_2))</td>
<td>(m_{\text{inv}}(b_1,b_2))</td>
<td>(p_T(b_2))</td>
</tr>
<tr>
<td>(\Delta R(b_1,\ell_2))</td>
<td>(m_{\text{inv}}(b_1,\ell_2))</td>
<td>(p_T(b_1))</td>
</tr>
<tr>
<td>(\Delta R(b_2,\ell_1))</td>
<td>(m_{\text{inv}}(b_2,\ell_1))</td>
<td>(p_T(b_2))</td>
</tr>
<tr>
<td>(\Delta R(\text{add}_1,\text{add}_2))</td>
<td>(m_{\text{inv}}(\text{add}_1,\text{add}_2))</td>
<td>(p_T(\text{add}_1))</td>
</tr>
<tr>
<td>(\text{add}_1,\text{add}_2)</td>
<td>(m_{\text{inv}}(b_1,b_2,\text{add}_1,\text{add}_2))</td>
<td>(p_T(\text{add}_2))</td>
</tr>
<tr>
<td>(\ell_1,\ell_2,b_1,b_2,\text{add}_1,\text{add}_2)</td>
<td>(m_{\text{inv}}(\ell_1,\ell_2,b_1,b_2,\text{add}_1,\text{add}_2))</td>
<td>(p_T(\text{add}_2))</td>
</tr>
</tbody>
</table>

50 nodes (10% dropout) 3 output classes

\[
\sum_{i=1}^{3} P_i = 1
\]
Learning the effective operators combining NN outputs

<table>
<thead>
<tr>
<th>Desired Discrimination</th>
<th>Combined NN Output used for limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>only $t_L$ operator</td>
<td>$\frac{P(t_L)}{P(t_L)+P(SM)}$</td>
</tr>
<tr>
<td>only $t_R$ operator</td>
<td>$\frac{P(t_R)}{P(t_R)+P(SM)}$</td>
</tr>
<tr>
<td>including both $t_L$ and $t_R$ operators</td>
<td>$P(t_L) + P(t_R)$</td>
</tr>
<tr>
<td>$t_L$ vs $t_R$</td>
<td>$\frac{P(t_L)}{P(t_L)+P(t_R)}$</td>
</tr>
</tbody>
</table>

One operator at a time: dedicated SM vs $t_L/t_R$ outputs

Multiple operators: SM vs EFT and $t_L$ vs $t_R$ outputs
Summary of the obtained (projected) 95% CL constraints on all relevant operators (one-by-one).

Limits on individual operators
Sensitivity study

Factor ~2 improvement from fiducial phase space definition to EFT-enriched NN selection!
## Contributions from multiple operators one LH and one RH top-quark operator

<table>
<thead>
<tr>
<th>Operator</th>
<th>tibb</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_{QQ}^{1} = \frac{1}{2} (\bar{Q} \gamma_{\mu} Q) (\bar{Q} \gamma_{\mu} Q)$,</td>
<td>✓</td>
</tr>
<tr>
<td>$O_{QQ}^{8} = \frac{1}{2} (\bar{Q} \gamma_{\mu} T^A Q) (\bar{Q} \gamma_{\mu} T^A Q)$,</td>
<td>✓</td>
</tr>
<tr>
<td>$O_{tb}^{1} = (\bar{t} \gamma_{\mu} t) (\bar{b} \gamma_{\mu} b)$,</td>
<td>✓</td>
</tr>
<tr>
<td>$O_{tb}^{8} = (\bar{t} \gamma_{\mu} T^A t) (\bar{b} \gamma_{\mu} T^A b)$,</td>
<td>✓</td>
</tr>
<tr>
<td>$O_{tt}^{1} = (\bar{t} \gamma_{\mu} t) (\bar{t} \gamma_{\mu} t)$,</td>
<td>✓</td>
</tr>
<tr>
<td>$O_{tb}^{1} = (\bar{b} \gamma_{\mu} b) (\bar{b} \gamma_{\mu} b)$,</td>
<td>✓</td>
</tr>
<tr>
<td>$O_{Qt}^{1} = (\bar{Q} \gamma_{\mu} Q) (\bar{t} \gamma_{\mu} t)$,</td>
<td>✓</td>
</tr>
<tr>
<td>$O_{Qt}^{8} = (\bar{Q} \gamma_{\mu} T^A Q) (\bar{t} \gamma_{\mu} T^A t)$,</td>
<td>✓</td>
</tr>
<tr>
<td>$O_{Qb}^{1} = (\bar{Q} \gamma_{\mu} Q) (\bar{b} \gamma_{\mu} b)$,</td>
<td>✓</td>
</tr>
<tr>
<td>$O_{Qb}^{8} = (\bar{Q} \gamma_{\mu} T^A Q) (\bar{b} \gamma_{\mu} T^A b)$,</td>
<td>✓</td>
</tr>
<tr>
<td>$O_{QtQb}^{1} = (\bar{Q} t) \varepsilon (\bar{Q} b)$,</td>
<td>✓</td>
</tr>
<tr>
<td>$O_{QtQb}^{8} = (\bar{Q} T^A t) \varepsilon (\bar{Q} T^A b)$.</td>
<td>✓</td>
</tr>
</tbody>
</table>

**Case study:** operators with right-handed top currents ($t_R$) or left-handed top currents ($t_L$)
Contributions from multiple operators
one LH and one RH top-quark operator

2-dim phase space of NN outputs
x-axis: SM vs EFT ($t_L$ and $t_R$)
y-axis: $t_L$ vs $t_R$
Contributions from multiple operators
SM-hypothesis $\rightarrow$ limits

Case-study: Consider two non-zero Wilson coefficients: $C_{tb}^1$ and $C_{Qb}^1$

$\rightarrow$ Assume an observation of the SM: $(C_{tb}^1, C_{Qb}^1) = (0,0)$
Contributions from multiple operators
SM-hypothesis → limits

Case-study: Consider two non-zero Wilson coefficients: $C_{tb}^1$ and $C_{Qb}^1$

→ Assume an observation of the SM: $(C_{tb}^1, C_{Qb}^1) = (0,0)$
Contributions from multiple operators
SM-hypothesis $\rightarrow$ limits

Case-study: Consider two non-zero Wilson coefficients: $C_{tb}^1$ and $C_{Qb}^1$

$\Rightarrow$ Assume an observation of the SM: $(C_{tb}^1, C_{Qb}^1) = (0,0)$
Contributions from multiple operators
SM-hypothesis → limits

Case-study: Consider two non-zero Wilson coefficients: $C^1_{tb}$ and $C^1_{Qb}$

→ Assume an observation of the SM: $(C^1_{tb}, C^1_{Qb}) = (0,0)$
Contributions from multiple operators
SM-hypothesis → limits

Case-study: Consider two non-zero Wilson coefficients: $C_{tb}^1$ and $C_{Qb}^1$

→ Assume an observation of the SM: $(C_{tb}^1, C_{Qb}^1) = (0,0)$
Contributions from multiple operators
SMEFT-hypothesis → observation?

Case-study: Consider two non-zero Wilson coefficients: $C_{tb}^1$ and $C_{Qb}^1$

→ Assume an observation of EFT signal: $(C_{tb}^1, C_{Qb}^1) = (5,3)$
Contributions from multiple operators
SMEFT-hypothesis $\rightarrow$ observation?

Case-study: Consider two non-zero Wilson coefficients: $C_{tb}^1$ and $C_{Qb}^1$

$\Rightarrow$ Assume an observation of EFT signal: $(C_{tb}^1, C_{Qb}^1) = (5, 3)$
Contributions from multiple operators
SMEFT-hypothesis → observation?

Case-study: Consider two non-zero Wilson coefficients: $C_{tb}^1$ and $C_{Qb}^1$

→ Assume an observation of EFT signal: $(C_{tb}^1, C_{Qb}^1) = (5,3)$
t\bar{t}b\bar{b} is an indispensable component in a global fit of the top-quark interactions in the SMEFT at the LHC!

- Large enough cross section to exploit differential information
- First direct constraints on a specific set of operators

Multi-class machine learning algorithms are a suitable tool for interpreting LHC data in this framework!
- Intrinsically large SMEFT parameter space
- High-multiplicity final states with inter-correlated information

Probing multiple SMEFT couplings simultaneously allow to pinpoint (or constrain) more efficiently the origin (absence) of a possible excess!
Backup
Introduction: $t\bar{t}b\bar{b}$ production

A. $t\bar{t}b\bar{b}$ is important background for $t\bar{t}H$ ($H \rightarrow b\bar{b}$). Recent discovery of this Higgs production mode


B. $t\bar{t}b\bar{b}$ ($t\bar{t}b\bar{b}$/ttjj) has therefore been measured by CMS and ATLAS (7, 8 & 13 TeV)


C. Difficult modeling (different mass scales, collinear splitting,...) $\rightarrow$ large effort from theory community


D. ??? $\rightarrow$ Indispensable component in global fit of top-quark interactions!
Model building and generator software details

**Dimension-6 four-fermion EFT operators**
*Feynrules model provided by LHC TOP WG*

**UFO output:** “Dim6Top”

**LO matrix element calculation**
*MadGraph*

**Parton showering**
*Pythia 8*

**Detector simulation and event reconstruction**
*Delphes*

Visible phase space at particle level

Phase space after event reconstruction and selection
EFT validity

\[ \frac{C_i}{\Lambda^2} E^2 \equiv C_i E^2 < C_i M_{\text{cut}}^2 \lesssim (4\pi)^2 \]

Fix \( \Lambda = 1 \) TeV and express limits in [TeV^{-2}]

All energy scales associated to the final state are imposed to be below \( M_{\text{cut}} \).
\( \rightarrow H_T \) (scalar sum of all visible final state objects) is a good example.

\[ M_{\text{cut}} = 2 \text{ TeV} \]
Strategy

Learning effective operators: Combine kinematic information of the \(ttbb\) final state into machine learning tools
→ Select EFT enriched phase space
→ Distinguish amongst EFT operators!

Cross section measurement in the fiducial detector volume
→ CMS \(ttbb/\text{ttjj} \geq 13\) TeV


Selection of kinematic phase space to enrich in EFT contributions (using \(m_{4b}\))
→ reconstructed phase space needed!
Cross section in the fiducial detector volume


Integrated luminosity = 2.3 fb$^{-1}$
Visible phase space definition:
$\sigma_{t\bar{t}b\bar{b},CMS} = 88 \pm 12(\text{stat.}) \pm 29(\text{syst.})$ fb

Projections for 300 fb$^{-1}$:
- scaled stat. unc. and fixed syst. unc. of 10%
- measured xsec = prediction of MadGraph
Cross section in the fiducial detector volume

CSM Collaboration, Measurements of $t\bar{t}$ cross sections in association with $b$ jets and inclusive jets and their ratio using dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV, 

Integrated luminosity = 2.3 fb$^{-1}$
Visible phase space definition:
$\sigma_{t\bar{t}b\bar{b},CMS} = 88 \pm 12(\text{stat.}) \pm 29(\text{syst.})$ fb

Projections for 300 fb$^{-1}$: scaled stat. unc. and fixed syst. unc. of 10%
measured xsec = prediction of MadGraph

$$\sigma_{fit} = \sigma_{SM} \left( 1 + p_1 \cdot C_i + p_2 \cdot C_i^2 \right)$$

95% CL

$$\Delta \chi^2(C_i|p_1,p_2) = \chi^2(C_i|p_1,p_2) - \chi^2_{min} = \frac{(\sigma_{fit}(C_i|p_1,p_2) - \sigma_{obs})^2}{\delta \sigma^2} - \chi^2_{min}$$
Cross section in the fiducial detector volume

\[ \sigma_{fit} = \sigma_{SM} \left( 1 + p_1 \cdot C_i + p_2 \cdot C_i^2 \right) \]

Color singlet operators have small interference but larger squared order contributions

Color octet operators have larger interference (SM ~ gluon induced) but suppressed squared order contributions (color factor 2/9)
Tailoring the kinematical phase space

Step 1: move to the reconstructed phase space: **Dileptonic decays of the top quarks**

Step 2: identify quantities that are sensitive to the EFT operators ($\Delta R, M_{\text{inv}}, p_T, \eta$) \(\rightarrow M_{4b}\)

Step 3: Make a selection on this quantity and derive the effective cross section dependence

![Diagram showing the decay process of top quarks into dileptons and other particles.](diagram.png)
Question: What cut to choose on M4b?
Answer: The one that optimizes the sensitivity!
   → increase relative population of EFT contributions
   → without blowing up statistical uncertainty on the SM measurement

\[
M_{4b} > 1.1 \text{ TeV} \quad (> 2 \text{ TeV!!})
\]
Learning the effective operators

*multiple operators*

The NN has indeed learned to distinguish amongst $t_L$ and $t_R$ operators!
Tailoring the kinematical phase space

Prospects for 300 fb$^{-1}$ after event reconstruction/selection and $M_{4b} > 1.1$ TeV

$$\sigma [fb] = 0.2977 \left( 1 + 0.0052 C + 0.0332 C^2 \right)$$

![Graph showing sample points and fitted cross section](image1)

$$\sigma [fb] = 0.2901 \left( 1 + 0.0023 C + 0.0078 C^2 \right)$$

![Graph showing sample points and fitted cross section](image2)

$C^{1}_{Qb} \in [-3, +3]$ TeV$^{-2}$

$xsec: C^{1}_{Qb} \in [-6, +6]$ TeV$^{-2}$

$C^{8}_{Qb} \in [-6.5, +7]$ TeV$^{-2}$

$xsec: C^{8}_{Qb} \in [-15, +10]$ TeV$^{-2}$

$\Rightarrow$ Improvement with a factor \(~2\)!
Learning the effective operators

one operator at a time

Once again the cut value is chosen to optimize the sensitivity

\[
\sigma \text{[fb]} = 0.2064 \left( 1 - 0.0016 C + 0.0671 C^2 \right)
\]

\[
\sigma \text{[fb]} = 0.2026 \left( 1 + 0.0166 C + 0.0153 C^2 \right)
\]

\( C^1_{Qb} \in [-2.1, +2.3] \text{ TeV}^{-2} \)

\( M_{4b} : C^1_{Qb} \in [-3, +3] \text{ TeV}^{-2} \)

\( C^8_{Qb} \in [-5, +4.3] \text{ TeV}^{-2} \)

\( M_{4b} : C^8_{Qb} \in [-6.5, +7] \text{ TeV}^{-2} \)

→ significant further improvement!
Learning the effective operators
\textit{one operator at a time}

Question: What cut to choose on the NN output?
Answer: The one that optimizes the sensitivity!
\begin{itemize}
  \item increase relative population of EFT contributions
  \item without blowing up statistical uncertainty on the SM measurement
\end{itemize}

NN output > 0.83
Backup: Neural Network training

- 18 inputs + RELU + 1 hidden layer (50 neurons) + RELU + Dropout (10%) + 3 outputs + SOFTMAX (sum=1)
- Mini-batches of size 128, training for 100 epochs
- Loss function: Categorical cross entropy
- Optimizer: Stochastic gradient descent
  - Initial learning rate = 0.005
  - Decay = $10^{-6}$
  - Nestrov momentum = 0.8

### Variables used in the network

<table>
<thead>
<tr>
<th>$\Delta R$</th>
<th>$m_{inv}$</th>
<th>$p_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta R(\ell_1, \ell_2)$</td>
<td>$m_{inv}(\ell_1, \ell_2)$</td>
<td>$p_T(\ell_1)$</td>
</tr>
<tr>
<td>$\Delta R(b_1, b_2)$</td>
<td>$m_{inv}(b_1, b_2)$</td>
<td>$p_T(b_2)$</td>
</tr>
<tr>
<td>$\Delta R(b_1, \ell_2)$</td>
<td>$m_{inv}(b_1, \ell_2)$</td>
<td>$p_T(b_1)$</td>
</tr>
<tr>
<td>$\Delta R(b_2, \ell_1)$</td>
<td>$m_{inv}(b_2, \ell_1)$</td>
<td>$p_T(b_2)$</td>
</tr>
<tr>
<td>$\Delta R(add_1, add_2)$</td>
<td>$m_{inv}(add_1, add_2)$</td>
<td>$p_T(add_1)$</td>
</tr>
<tr>
<td>$m_{inv}(b_1, b_2, add_1, add_2)$</td>
<td>$p_T(add_2)$</td>
<td></td>
</tr>
<tr>
<td>$m_{inv}(\ell_1, \ell_2, b_1, b_2, add_1, add_2)$</td>
<td>$p_T(add_2)$</td>
<td></td>
</tr>
</tbody>
</table>
Outlook

• Fully marginalized limits when more precise measurements become available

• Method is generic and can be applied to other topologies / final states!

• Increased complexity of the network (Deep learning) or more advanced machine learning techniques may result in better sensitivity.

• Question for the future: How much can we push these algorithms to distinguish different EFT operators.
  o We demonstrated a distinction between $t_L$ and $t_R$ operators
  o Distinguish color singlet operators from color octet ones would be possible if one includes interference effects during the training phase!
    (becomes dependent on the value of the Wilson coefficient
    $\rightarrow$ Parametrized learning approach?)
  o Can you (ideally) distinguish each individual operator or are some of them indistinguishable?
Chapter 6. Probing new physics in the SMEFT

In this chapter are compared to other existing bounds from four–top–quark measurements in Tab. 6.1. The first column shows individual 95% CL intervals (for projections to 300 fb⁻¹) from four–top–quark production [101] assuming a value of \( M_{\text{cut}} \) at 4 TeV. These constraints assume an upper limit of the four–top–quark signal strength, \( \mu < 1.87 \), obtainable at the LHC with 300 fb⁻¹ at 13 TeV [271]. The second column quotes results from the four–top–quark measurement by CMS in the single lepton and opposite–sign dilepton final–states [100], using 35.8 fb⁻¹ at 13 TeV. No upper threshold on the allowed energy scales has been applied in this analysis. The third column shows the bounds from a global fit of multiple SMEFT operators to top–quark related measurements at center–of–mass energies of 8 and 13 TeV [98]. These bounds result mostly from the CMS measurement of four–top–quark production in the same–sign dilepton and multilepton final–states [66] at 13 TeV using 35.9 fb⁻¹. The final column represents the best sensitivities obtained from the study presented in this chapter for comparison.

Table 6.1: Comparison between the sensitivity of \( t\bar{t}b\bar{b} \) and \( tttb \) production to the mutual SMEFT operators. The first column shows individual 95% CL intervals (for projections to 300 fb⁻¹) from four–top–quark production [101] at 13 TeV, assuming a value of \( M_{\text{cut}} \) at 4 TeV. The second column quotes 95% CL intervals from the four–top–quark measurement by CMS in the single lepton and opposite–sign dilepton final–states [100], using 35.8 fb⁻¹ at 13 TeV. The third column shows the bounds from a global fit of multiple SMEFT operators to top–quark related measurements at center–of–mass energies of 8 and 13 TeV [66, 98]. The last column compares these intervals to the best constraints from \( t\bar{t}b\bar{b} \) production (assuming 300 fb⁻¹ at 13 TeV) obtained in this work.

4-top (300 fb⁻¹) \((M_{\text{cut}} = 4 \text{ TeV})\)

4-top (35.8 fb⁻¹) \((\text{no } M_{\text{cut}})\)

global fit \((\text{no } M_{\text{cut}})\)

\(t\bar{t}b\bar{b} \) (300 fb⁻¹) \((M_{\text{cut}} = 2 \text{ TeV})\)

<table>
<thead>
<tr>
<th>(C_{1QQ}^1)</th>
<th>([-2.8, 2.5])</th>
<th>([-2.2, 2.0])</th>
<th>([-5.4, 5.2])</th>
<th>([-2.1, 2.3])</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C_{8QQ}^8)</td>
<td>([-8.4, 7.4])</td>
<td>n.a.</td>
<td>([-21, 16])</td>
<td>([-4.5, 3.1])</td>
</tr>
<tr>
<td>(C_{Qt}^1)</td>
<td>([-2.2, 2.3])</td>
<td>([-3.5, 3.5])</td>
<td>([-4.9, 4.9])</td>
<td>([-2.1, 2.3])</td>
</tr>
<tr>
<td>(C_{Qt}^8)</td>
<td>([-5.1, 4.1])</td>
<td>([-7.9, 6.6])</td>
<td>([-11, 8.7])</td>
<td>([-3.9, 3.8])</td>
</tr>
</tbody>
</table>

\(\mu_{4t} < 1.87\) \(\mu_{4t} < 5.22\)