Improved Constraints on Effective Top Quark Interactions using Edge Convolution Networks

Akanksha Bhardwaj

collaborators : O. Atkinson, AB, S. Brown, C. Englert, D. J. Miller, and P. Stylianou

Arxiv [2111.01838]

Anomaly 2021

Nov 11, 2021

Outline

- Introduction
- Effective interactions for top pair production with leptonic decays
- Graph representation of events
- Graph Neural Network with Edge Convolution
- GNN-improved Wilson coefficient constraints
- Results and Discussion

Introduction & Motivation

- Explore the potential of Graph Neural Networks (GNNs) to improve the performance of high-dimensional effective field theory parameter fits.

- Focusing on a SMEFT analysis of $pp \to t\bar{t}$ production, including top decays, where the linear effective field deformation is parametrised by thirteen independent Wilson coefficients.

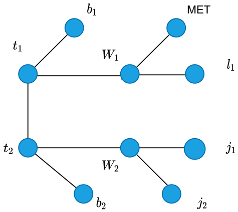
- The application of GNNs allows us to condense the multidimensional phase space information available for the discrimination of BSM effects from the SM expectation by considering all available final state correlations directly.

Effective interactions for top pair production with leptonic decays

- Any differential cross section can be written as

$$d\sigma = d\sigma_{\mathrm{SM}} + \frac{C_i}{\Lambda^2} d\sigma_i^{(1)} + \frac{C_i C_j}{\Lambda^4} d\sigma_{ij}^{(2)}, \tag{1}$$

  where the C_i are the Wilson Coefficients (WCs) and Λ is the generic scale of new physics (NP).
- We focus on EFT parameter constraints for the top sector, in particular, we focus on $pp \to t\bar{t}$ production with semi-leptonic top decays
  $pp \to t\bar{t} \to \ell b\bar{b}j + MET$.



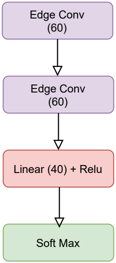


Figure: Representative diagram for the input graph and the network architecture.

- Nodes: Final state reconstructed particles.
- Edges: The final states are connected to the ones of the reconstructed objects from which they are derived.
- Node features: Each node is associated with a feature vector $[p_T, \eta, \phi, E, m, \text{PID}]$,

Graph Neural Network with Edge Convolution

During the message-passing phase a message $\vec{m}_{ij}^{(l)}$ is calculated between the two nodes by the following operation,

$$\vec{m}_{ij}^{(l)} = \vec{M}^{(l)}(\vec{x}_i^{(l)}, \vec{x}_j^{(l)}, \vec{e}_{ij}^{(l)})\,. \qquad (2)$$

Once the messages is calculated in a layer, each node feature is updated using an aggregation function

$$\vec{x}_i^{(l+1)} = \vec{A}(\vec{x}_i^{(l)}, \{\vec{m}_{ij}^{(l)} | \, j \in \mathcal{N}(i)\})\,, \qquad (3)$$

where $\mathcal{N}(i)$ are the nodes which are connected to ith node and $\vec{A}$ is the permutation invariant function. The edge convolution operation is defined with the following message-passing function

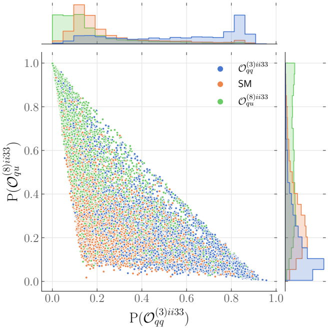
$$\vec{x}_i^{(l+1)} = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathrm{ReLU}\left((\vec{x}_j^{(l)} - \vec{x}_i^{(l)}) + (\vec{x}_i^{(l)})\right), \qquad (4)$$

GNN-improved Wilson coefficient constraints

- For illustration purposes, we first limit our study to a three-class classification problem.
- The network output in this example returns the probability of an event belonging to each of the three classes. An event is then assigned to the EFT/SM class with the greatest corresponding probability. We reflect this in our choice of operators for this section:

$$\begin{aligned} \mathcal{O}_{qu}^{(8)ii33} &= (\bar{q}_i \gamma_\mu T^A q_i)(\bar{u}_3 \gamma^\mu T^A u_3)\,, \\ \mathcal{O}_{qq}^{(3)ii33} &= (\bar{q}_i \gamma_\mu \tau^I q_i)(\bar{q}_3 \gamma^\mu \tau^I q_3)\,. \end{aligned} \tag{5}$$

Network performance and Output



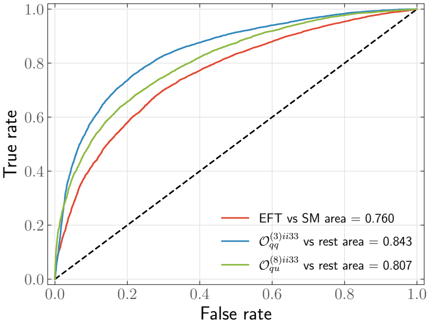
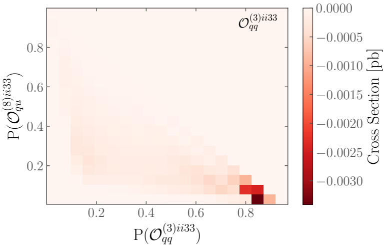
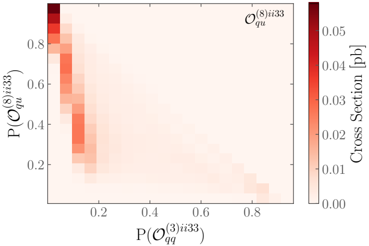


Figure: The probabilities calculated for each event to be a result of each SMEFT insertion is shown. On the right the Receiver Operator Characteristic (ROC) curves are shown. We calculate these in a one-vs-rest scheme for each operator.

Two-dimensional histograms for each contribution





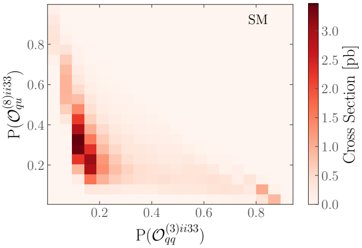


Figure: Example two-dimensional histograms for each contribution, normalised to the cross section rate.

Optimised χ^2 fit.

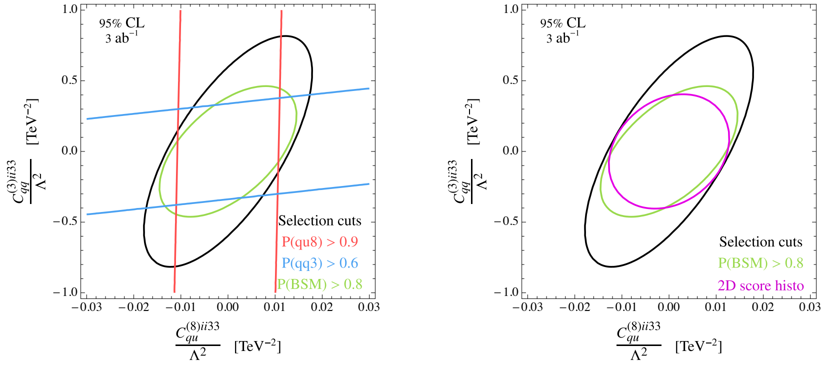
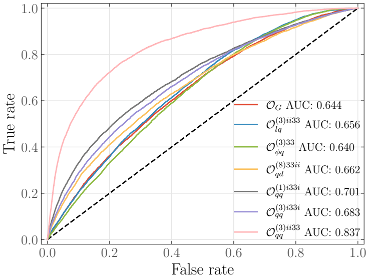


Figure: WC constraint contours at the 95% C.L. from χ^2 fitting; in black from the data of the baseline selection of which also passes the network requirements. The left plot shows the contours from cuts on the NN scores at the optimal value of these score cuts, with the analysis performed using $p_T(b_1)$ distributions. The right plot shows the BSM score cut as in the left plot, along with the contour from the 2D score histogram of (with no score cuts) analysis.

ROC for all operators



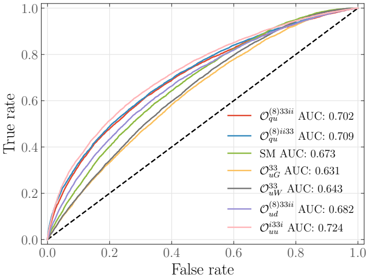


Figure: ROC curves for the scenario where multi-class classification is performed on thirteen SMEFT operators and the SM.

Improvement on the Wilson coefficient

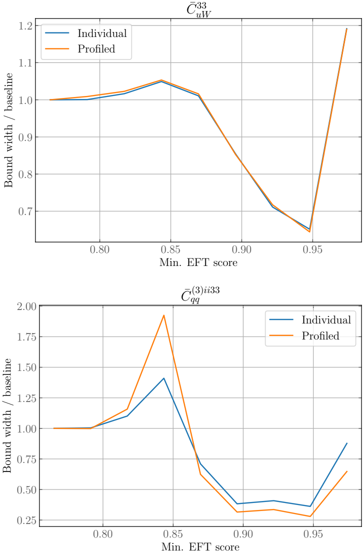


Figure: Representative relative improvement (decrease in the 2σ Wilson coefficient interval) over the individual (orange) and profiled (blue) operator constraints quoted in by imposing cuts on the ML score. Bounds were obtained at an integrated luminosity of 3/ab.

Results: Improved Individual and Profiled bounds on Wilson coefficient

	2.3 fb^{-1}		3 ab^{-1}	
	Individual	Profiled	Individual	Profiled
$\bar{C}_G$	0.07%	14.53%	0.07%	11.72%
$\bar{C}_{\varphi q}^{(3)33}$	33.74%	34.16%	33.73%	33.82%
$\bar{C}_{uG}^{33}$	28.29%	32.12%	28.28%	30.76%
$\bar{C}_{uW}^{33}$	34.86%	35.36%	34.85%	35.57%
$\bar{C}_{qq}^{(1)i33i}$	3.50%	3.52%	3.50%	3.23%
$\bar{C}_{qq}^{(3)i33i}$	4.35%	4.31%	4.35%	5.01%
$\bar{C}_{qq}^{(3)ii33}$	63.83%	–	63.83%	72.06%
$\bar{C}_{qu}^{(8)33ii}$	3.45%	3.45%	3.45%	3.39%
$\bar{C}_{qu}^{(8)ii33}$	3.74%	3.80%	3.74%	3.77%
$\bar{C}_{ud}^{(8)33ii}$	4.62%	4.63%	4.62%	4.64%
$\bar{C}_{uu}^{i33i}$	3.38%	3.41%	3.38%	3.83%
$\bar{C}_{lq}^{(3)ii33}$	–	–	10.57%	40.26%

Figure: Maximum improvements in 2σ bounds via a cut on the ML score.

Summary and Outlook

- We have focused employing on GNNs for EFT limit setting.
- GNNs are particularly motivated approaches for this purpose as they allow us to directly reflect the graph structure which is imposed by EFT interactions in the classification and eventual limit setting.
- We find that large improvements of the sensitivity become achievable when correlations are not yet fully exploited in the inclusive base selection.

Thank you