

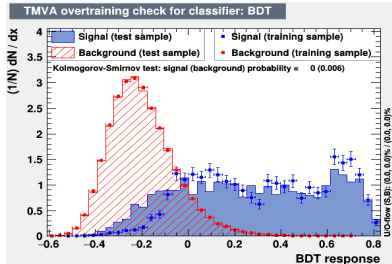
Search for dark matter production in association with a single top quark or a top quark pair in the dilepton final state at $\sqrt{s} = 13$ TeV

Pablo Martínez Ruíz del Árbol, Jónatan Piedra Gomez, **Cédric Prieëls**
February 9th 2022

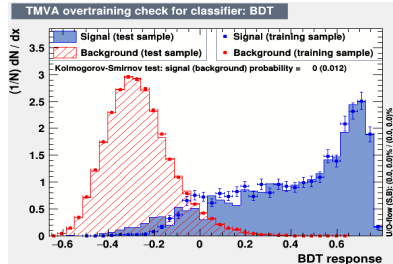
Thesis Endorsement - Answer to questions
Instituto de Física de Cantabria

Worst cases

Scalar 100 GeV



Pseudo 100 GeV

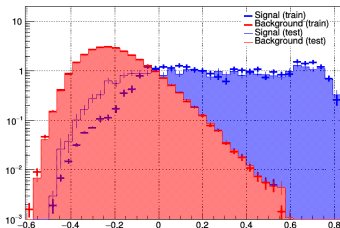


"You showed that signal in the low BDT region has large data/simulation disagreement. This should be investigated. In addition, it would be good to see the overtraining plots in log-scale so that we ensure there is no large further disagreement in the tails of the background in the high BDT region."

We first of all produced the exact same plots in log scale, as required.

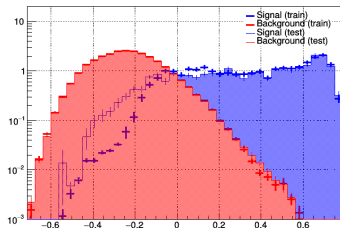
Scalar 100 GeV

Signal (train)



Pseudo 100 GeV

Signal (train)



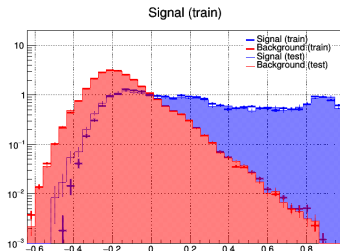
These plots comfort us in the fact that no sign of test/train discrepancy is observed for the background samples, even in the tail.

Low BDT region large discrepancies

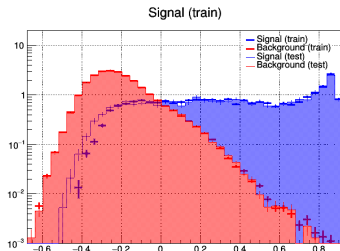
For the training process, both the t/\bar{t} and $t\bar{t}$ +DM are mixed together, with a weight corresponding to their respective cross-sections. One hypothesis for this effect might be the fact that the weight associated to the t/\bar{t} +DM is so low, that the test sample, which has less statistics, lose its contribution.

To test this hypothesis, we reproduced the same plot while assigning the same weight to both the signal samples, obtaining a much better agreement.

Scalar 100 GeV



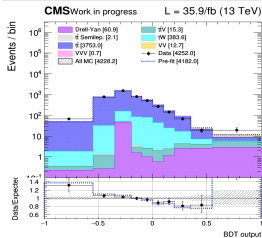
Pseudo 100 GeV



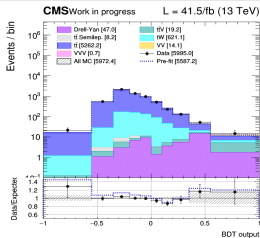
These discrepancies should not have any impact on the final results as they impact a region with very low amounts of signal.

t/\bar{t} +DM region (scalar 100 GeV)

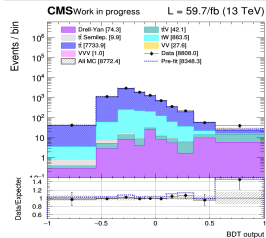
2016



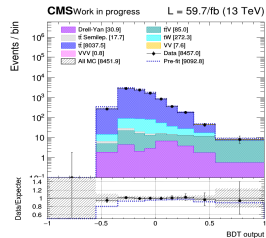
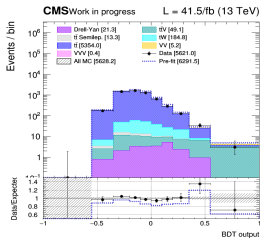
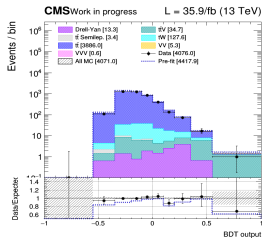
2017



2018



$t\bar{t}$ +DM region

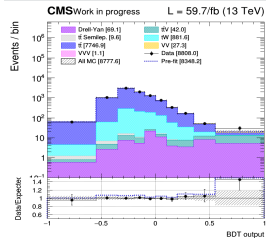
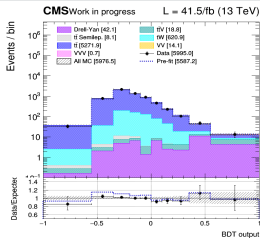
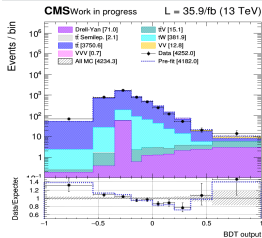


t/\bar{t} +DM region (pseudoscalar 100 GeV)

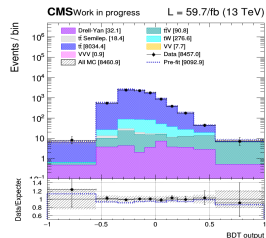
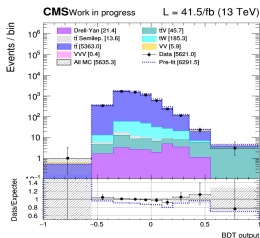
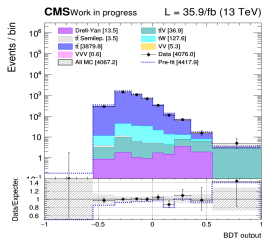
2016

2017

2018



$t\bar{t}$ +DM region

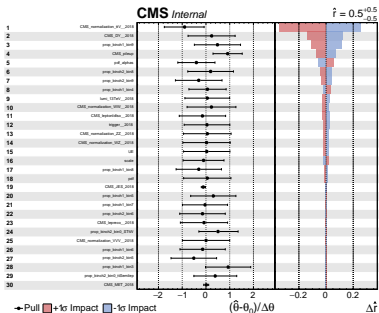


All the post-fit plots can be found in the v4 of the AN-22-014.

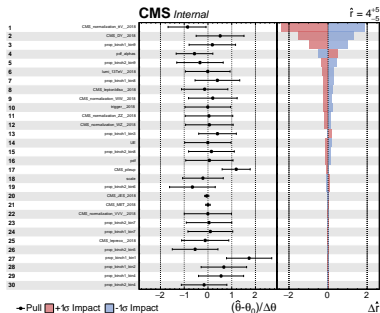
" In addition, please provide information on the fit itself: such as pulls with the data."

2018, scalar

100 GeV



500 GeV



All the impact plots with real data can also be found in the newest version of the AN. In general no large pulls, the most significant effect is the constraint of the JES/MET.

Pulls and impact plots

These strong constraints seem to come from the low BDT output region of the plots fed into the algorithm, where a lot of \bar{t} can be found. This region might actually play the role of a $t\bar{t}$ control region, therefore constraining a lot these two systematic.

We therefore tried removing the BDT output < 0 region, getting results less constrained.

2018, scalar

100 GeV

500 GeV

