Reducing the effect of nominal background samples on signal sample systematics using an adversarial neural network in the tW dilepton channel

Christian Kirfel

18th of September 2019

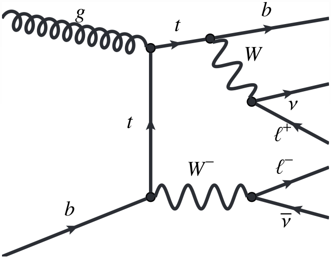
ATLAS EXPERIMENT

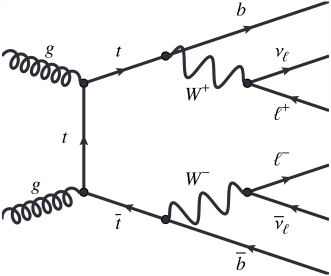
UNIVERSITÄT BONN

Outline

- tW and $t\bar{t}$ separation
- Artificial neural networks and adversarial neural networks as a possible solution
- Introduction to hyperparameters
- Preliminary training results for an adversarial neural network

tW and $t\bar{t}$ separation





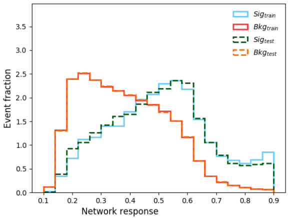
- Problem: Cross-sections of tW about 10 times smaller than $t\bar{t}$
- Interference in NLO order
- Instead of applying cut $\longrightarrow$ Neural networks

Setup of the classifier

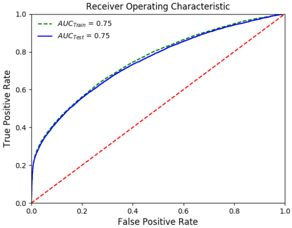
Hyper-parameter scan results

- Input: 14 variables motivated by a BDT variable scan.
- Hidden layers: 6 elu layers $\times$ 128 nodes each
- Output layer: 1 sigmoid node
- Optimisation: SGD, learning rate $= 0.06$, momentum $= 0.3$, no nesterov, no decay
- Duration: 600 epochs

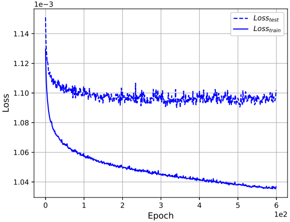
Simple network results



Separation



ROC curve

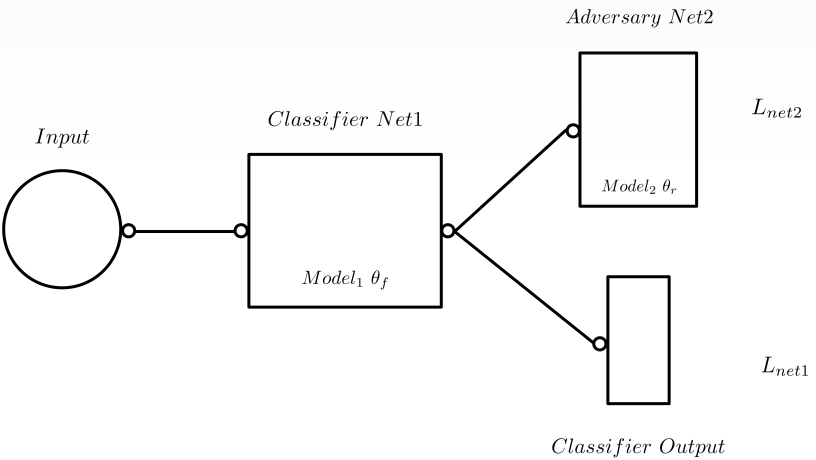


Losses

Adversarial Neural Networks

- Neural networks have no info on systematic uncertainties
- Introduction of a second, adversarial network classifying between nominal and systematic
- Combined loss function $\mathcal{L}_{adversarial}(\theta_f, \theta_t) = \mathcal{L}(\theta_f) - \lambda\mathcal{L}(\theta_f, \theta_r)$
- Network 1: signal/background separation
- Network 2: nominal/systematic separation
- Expectation: network 1 succeeds, network 2 fails

Setup of the ANN



ANN

	Discriminator	Adversary
$t\bar{t}$	0	1
tW DR	1	1
tW DS	1	0

ANN setup

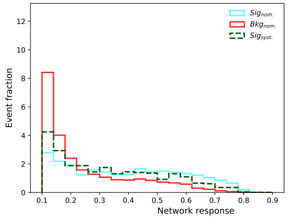
Discriminator setup

- Input: 14 variables motivated by a BDT variable scan.
- Hidden layers: 6 elu layers $\times$ 128 nodes each
- Output layer: 1 sigmoid node
- Optimisation: SGD, learning rate $= 0.01$, momentum $= 0.3$, no nesterov, no decay

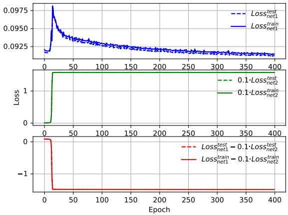
Adversary setup

- Input: 14 variables motivated by a BDT variable scan.
- Hidden layers: 6 elu layers $\times$ 128 nodes each
- Output layer: 1 sigmoid node
- Optimisation: SGD, learning rate $= 0.01$, momentum $= 0.3$, no nesterov, no decay

ANN results



Separation



Losses

- The separation is visibly bad.
- The agreement between nominal and systematics has barely improved
- Losses show bad behaviour

Improvement plans

Assumption

Labelling $t\bar{t}$ as a nominal sample results in a strong bias

Possible solution

- Randomly label $t\bar{t}$ events as either nominal or systematic
- Add additional weighting to the $t\bar{t}$ sample for the adversarial network only
- Exclude the $t\bar{t}$ sample for the adversarial training completely

Applied fixes

Re-labelling

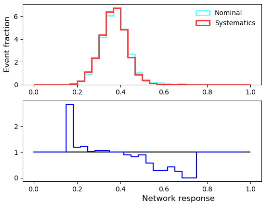
	Discriminator	Adversary
$t\bar{t}$	0	1/0 (50 % mix)
tW DR	1	1
tW DS	1	0

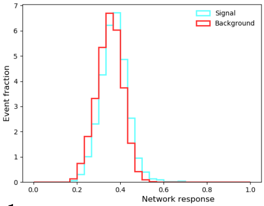
Weighting $t\bar{t}$ for the adversary

- Applying an additional weight to the $t\bar{t}$ events for the adversarial training only
- Varied the weights between 0.0 and 1.0

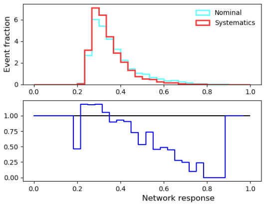
ttbar weights for the adversarial training

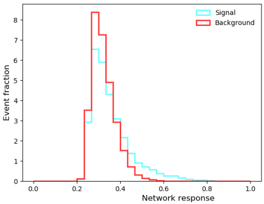
Weight = 0





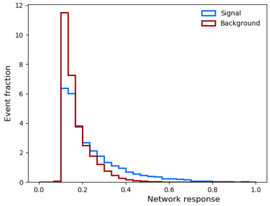
Weight = 1

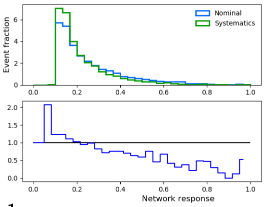




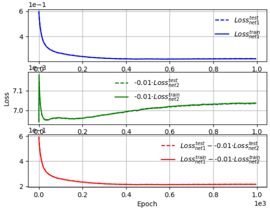
ttbar weights for the adversarial training

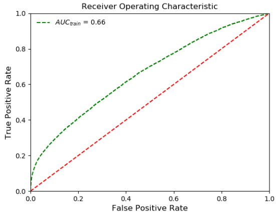
Weight $= 0$





Weight $= 1$





Conclusions

Improvements and insights

- sdfw

Weighting $t\bar{t}$ for the adversary

- Applying an additional weight to the $t\bar{t}$ events for the adversarial training only
- Varied the weights between 0.0 and 1.0