Search for Single Top Quark Production in the s-channel with the ATLAS Experiment using Machine-Learning Techniques

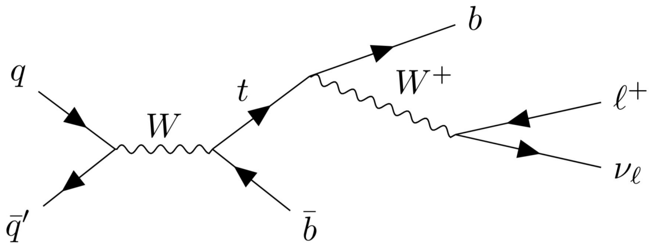
Niklas Düser

Supervisor: Andrea Knue

September 29, 2025

Single top quark production in the s-channel

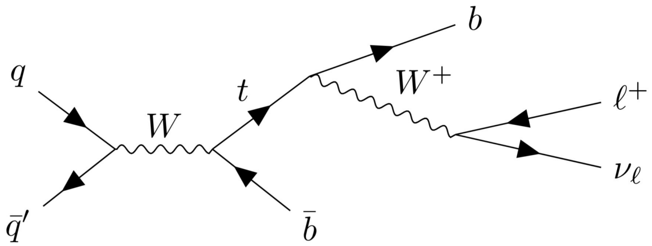
Electroweak s-channel production:



Single top quark production in the s-channel

Discovered at Tevatron with
6.3 standard deviations

Electroweak s-channel production:

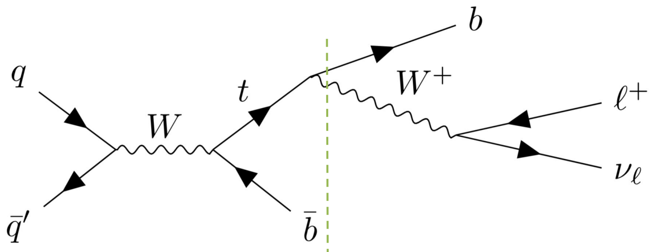


Single top quark production in the s-channel

Discovered at Tevatron with

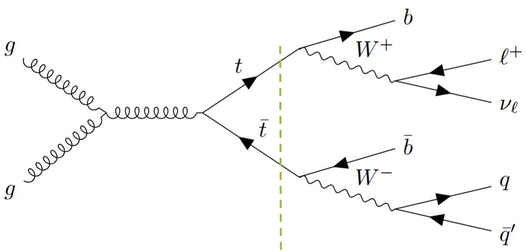
6.3 standard deviations

Electroweak s-channel
production:



$$\sigma = 10.32 \text{ pb}$$

Strong $t\bar{t}$ production:



$$\sigma = 832 \text{ pb}$$

Single top quark production in the s-channel

Discovered at Tevatron with

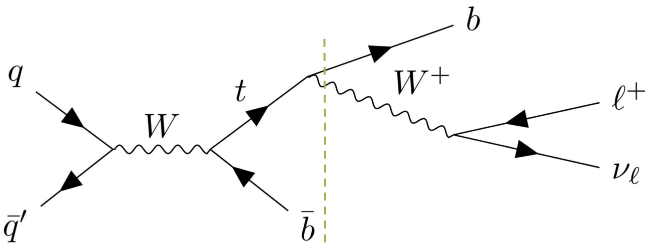
6.3 standard deviations

Best measurement at ATLAS:

$$\sigma = 8.2 \pm 0.6(\mathrm{stat.})\,^{+3.4}_{-2.8}\,(\mathrm{syst.})\,\mathrm{pb}$$

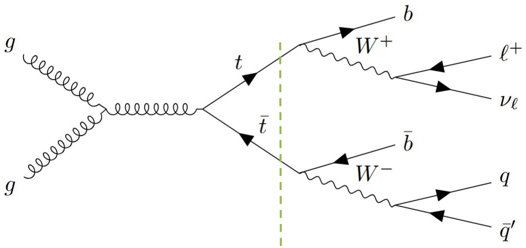
3.3 (3.9) standard deviations

Electroweak s-channel production:



$$\sigma = 10.32\ \mathrm{pb}$$

Strong $t\bar{t}$ production:



$$\sigma = 832\ \mathrm{pb}$$

Single top quark production in the s-channel

Discovered at Tevatron with
6.3 standard deviations

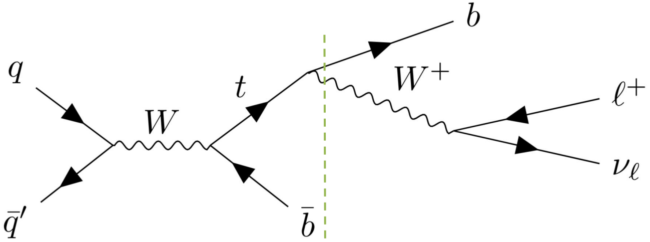
Best measurement at ATLAS:

$$\sigma = 8.2 \pm 0.6(\mathrm{stat.})\,^{+3.4}_{-2.8}\,(\mathrm{syst.})\ \mathrm{pb}$$

3.3 (3.9) standard deviations

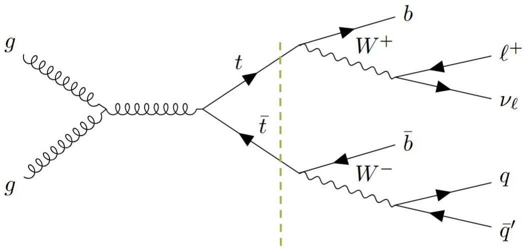
Source	$\Delta\sigma/\sigma$ [%]
$t\bar{t}$ normalisation	+24/−17
$t\bar{t}$ shape modelling	+18/−15
ISR/FSR	+13/−11
PS & had.	+12/−10
ME/PS matching	+10/−8
h_{damp}	< 1
s-channel modelling	+18/−8
PS & had.	+18/−8
ISR/FSR	+3/−1
Jet energy resolution	+18/−12
Jet energy scale	+18/−13
MC statistics	+13/−11
Flavour tagging	+12/−10
$W+$jets normalisation	+11/−8
Systematic uncertainties	+42/−34
Statistical uncertainty	±8
Total	+42/−35

Electroweak s-channel production:



$$\sigma = 10.32\ \mathrm{pb}$$

Strong $t\bar{t}$ production:



$$\sigma = 832\ \mathrm{pb}$$

Single top quark production in the s-channel

Discovered at Tevatron with

6.3 standard deviations

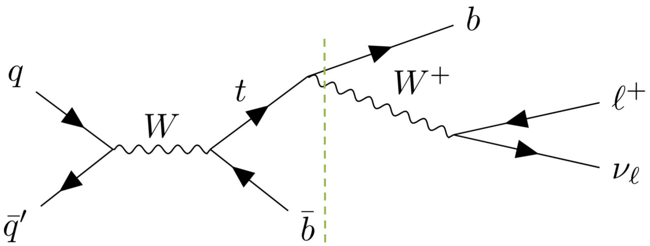
Best measurement at ATLAS:

$$\sigma = 8.2 \pm 0.6(\text{stat.})\,^{+3.4}_{-2.8}\,(\text{syst.})\text{ pb}$$

3.3 (3.9) standard deviations

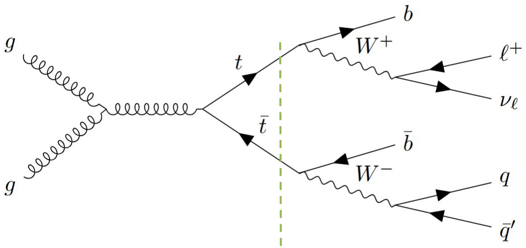
Source	$\Delta\sigma/\sigma$ [%]
$t\bar{t}$ normalisation	+24/−17
$t\bar{t}$ shape modelling	+18/−15
ISR/FSR	+13/−11
PS & had.	+12/−10
ME/PS matching	+10/−8
h_{damp}	< 1
s-channel modelling	+18/−8
PS & had.	+18/−8
ISR/FSR	+3/−1
Jet energy resolution	+18/−12
Jet energy scale	+18/−13
MC statistics	+13/−11
Flavour tagging	+12/−10
$W+$jets normalisation	+11/−8
Systematic uncertainties	+42/−34
Statistical uncertainty	±8
Total	+42/−35

Electroweak s-channel production:



$$\sigma = 10.32\text{ pb}$$

Strong $t\bar{t}$ production:



$$\sigma = 832\text{ pb}$$

Single top quark production in the s-channel

Discovered at Tevatron with 6.3 standard deviations

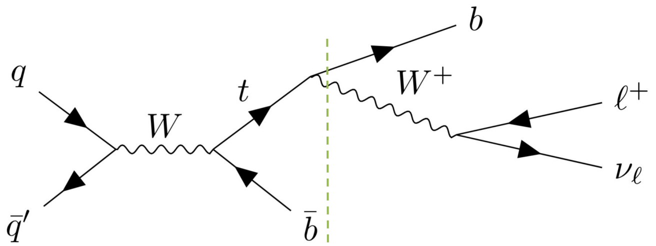
Best measurement at ATLAS:

$$\sigma = 8.2 \pm 0.6(\text{stat.})\,^{+3.4}_{-2.8}\,(\text{syst.})\text{ pb}$$

3.3 (3.9) standard deviations

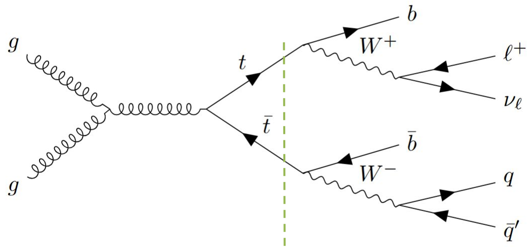
Source	$\Delta\sigma/\sigma$ [%]
$t\bar{t}$ normalisation	+24/−17
$t\bar{t}$ shape modelling	+18/−15
ISR/FSR	+13/−11
PS & had.	+12/−10
ME/PS matching	+10/−8
h_{damp}	< 1
s-channel modelling	+18/−8
PS & had.	+18/−8
ISR/FSR	+3/−1
Jet energy resolution	+18/−12
Jet energy scale	+18/−13
MC statistics	+13/−11
Flavour tagging	+12/−10
W+jets normalisation	+11/−8
Systematic uncertainties	+42/−34
Statistical uncertainty	±8
Total	+42/−35

Electroweak s-channel production:



$\sigma = 10.32$ pb

Strong $t\bar{t}$ production:



$\sigma = 832$ pb

Single top quark production in the s-channel

Discovered at Tevatron with

6.3 standard deviations

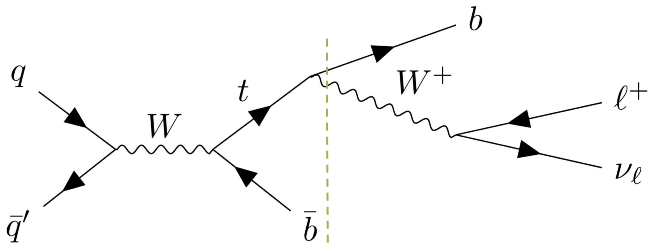
Best measurement at ATLAS:

$$\sigma = 8.2 \pm 0.6(\mathrm{stat.})\,{}^{+3.4}_{-2.8}\,(\mathrm{syst.})\ \mathrm{pb}$$

3.3 (3.9) standard deviations

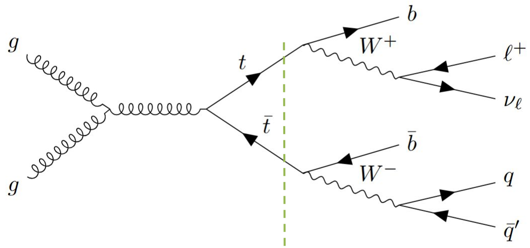
Source	$\Delta\sigma/\sigma$ [%]
$t\bar{t}$ normalisation	+24/−17
$t\bar{t}$ shape modelling	+18/−15
ISR/FSR	+13/−11
PS & had.	+12/−10
ME/PS matching	+10/−8
h_{damp}	< 1
s-channel modelling	+18/−8
PS & had.	+18/−8
ISR/FSR	+3/−1
Jet energy resolution	+18/−12
Jet energy scale	+18/−13
MC statistics	+13/−11
Flavour tagging	+12/−10
$W+$jets normalisation	+11/−8
Systematic uncertainties	+42/−34
Statistical uncertainty	±8
Total	+42/−35

Electroweak s-channel production:



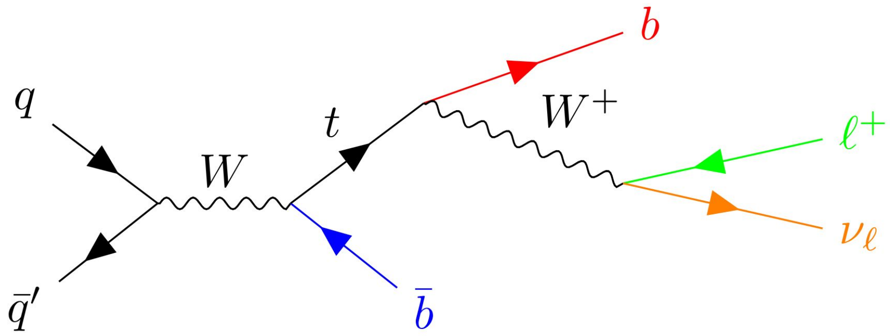
$$\sigma = 10.32\ \mathrm{pb}$$

Strong $t\bar{t}$ production:



$$\sigma = 832\ \mathrm{pb}$$

Signal Region definition

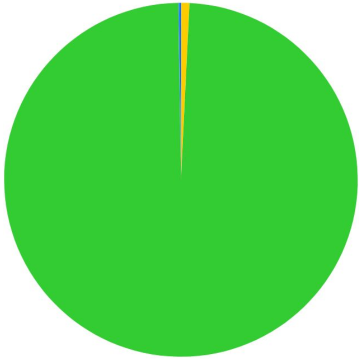


- exactly 1 lepton with p_{T}> 30 GeV
- missing transverse momentum > 35 GeV
- exactly 2 @77% b-tagged jets
- leading jet p_{T}> 40 GeV
- sub-leading jet p_{T}> 30 GeV

- transverse W mass condition
- further vetos

How does the SR change the data composition?

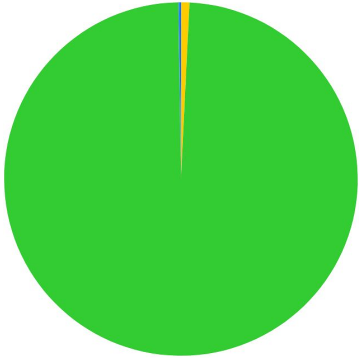
Pre SR, all samples:



99% V + jets

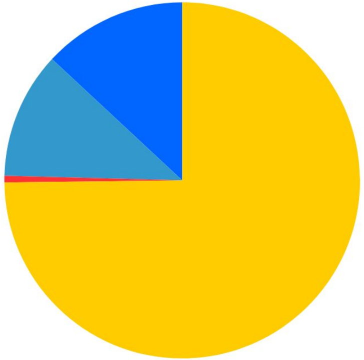
How does the SR change the data composition?

Pre SR, all samples:



99% $V + \text{jets}$

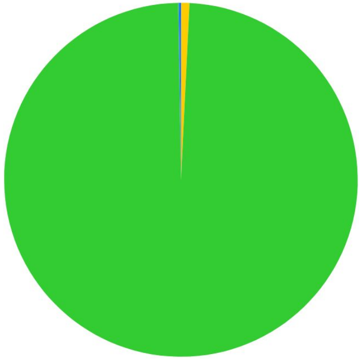
Pre SR, excluding V+jets:



74.8% $t\bar{t}$
0.6% s-channel

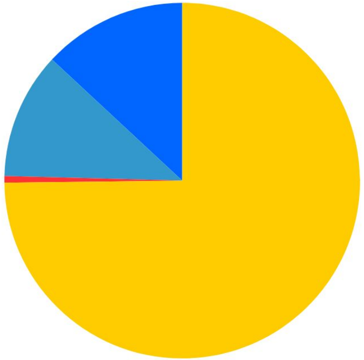
How does the SR change the data composition?

Pre SR, all samples:



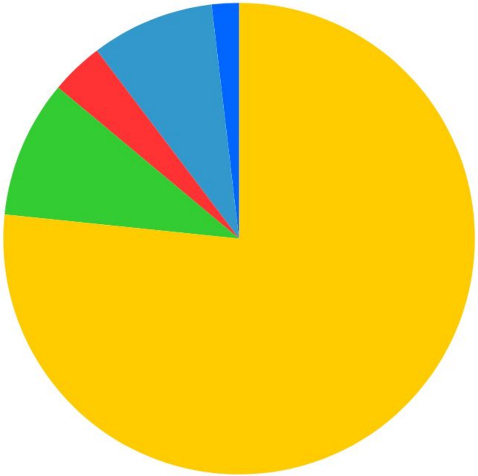
99% $V + \text{jets}$

Pre SR, excluding V+jets:



74.8% $t\bar{t}$
0.6% s-channel

SR:



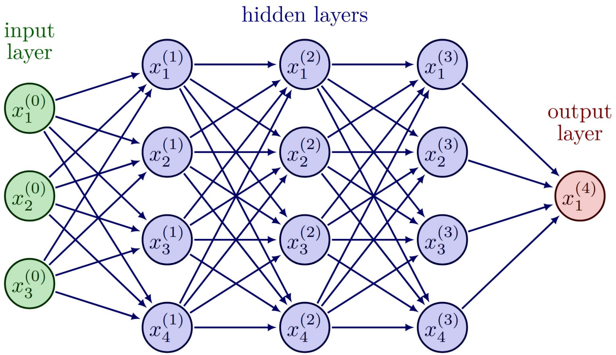
76.6% $t\bar{t}$ ⇒ 100476
3.6% s-channel ⇒ 4748

Machine Learning approach

- comparison of **Deep Neural Network** and **Graph Neural Network**
  1. training and evaluation on nominal data
  2. evaluation of systematic impact stability

- 50/50 even/odd data split
  - 80% training
  - 20% validation

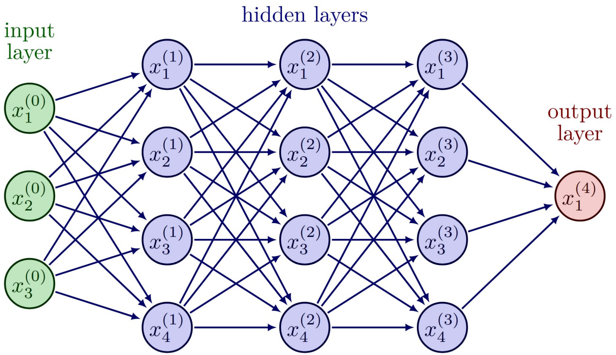
Deep Neural Network (DNN)



dimension: [59, 128, 256, 64, 32, 16, 1]

dropout: [0.1, 0.1, 0.1]

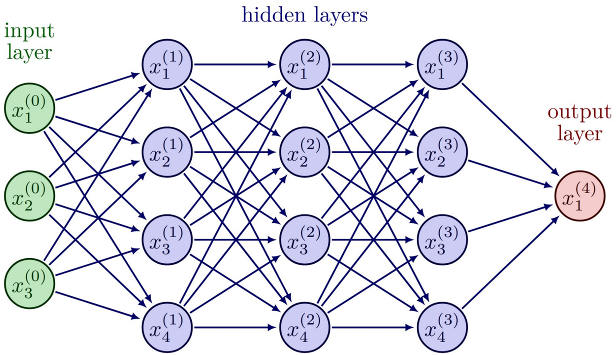
Deep Neural Network (DNN)



dimension: [59, 128, 256, 64, 32, 16, 1]
dropout: [0.1, 0.1, 0.1]

Feature	Description
$m_{W,\mathrm{prod}}$	Reconstructed mass of virtual production W-boson
m_t	Reconstructed mass of top quark
$m_{W,\mathrm{decay}}$	Reconstructed mass of decay W-boson
$m_{\mathrm{T},t}$	Reconstructed transverse mass of top quark
$m_{\mathrm{T},W,\mathrm{sec}}$	Reconstructed transverse mass of secondary W-boson
H_T	Summed total transverse energy
$H_{\mathrm{T,had}}$	Summed hadronic transverse energy
$H_{\mathrm{T,lep}}$	Summed leptonic transverse energy
Aplanarity	Deviation from planar event shape
Planarity	Degree of flatness in momentum flow
Sphericity	Isotropy of momentum distribution

Deep Neural Network (DNN)

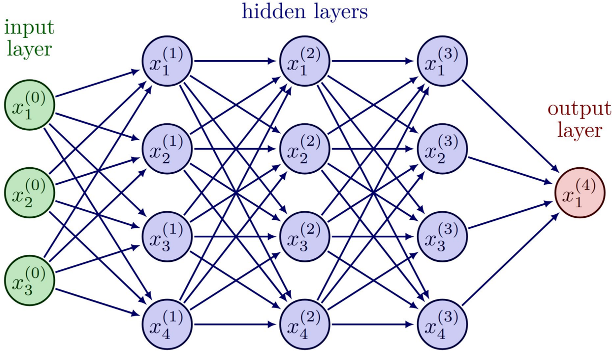


dimension: [59, 128, 256, 64, 32, 16, 1]

dropout: [0.1, 0.1, 0.1]

Feature	Description
$m_{W,\mathrm{prod}}$	Reconstructed mass of virtual production W-boson
m_t	Reconstructed mass of top quark
$m_{W,\mathrm{decay}}$	Reconstructed mass of decay W-boson
$m_{\mathrm{T},t}$	Reconstructed transverse mass of top quark
$m_{\mathrm{T},W,\mathrm{sec}}$	Reconstructed transverse mass of secondary W-boson
H_T	Summed total transverse energy
$H_{\mathrm{T,had}}$	Summed hadronic transverse energy
$H_{\mathrm{T,lep}}$	Summed leptonic transverse energy
Aplanarity	Deviation from planar event shape
Planarity	Degree of flatness in momentum flow
Sphericity	Isotropy of momentum distribution

Deep Neural Network (DNN)

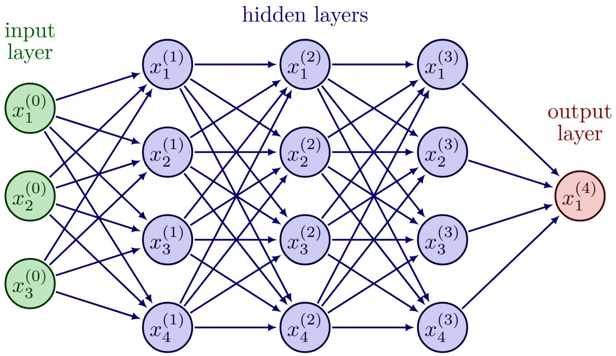


dimension: [59, 128, 256, 64, 32, 16, 1]

dropout: [0.1, 0.1, 0.1]

Feature	**Description**
$m_{W,\mathrm{prod}}$	Reconstructed mass of virtual production W-boson
m_t	Reconstructed mass of top quark
$m_{W,\mathrm{decay}}$	Reconstructed mass of decay W-boson
$m_{\mathrm{T},t}$	Reconstructed transverse mass of top quark
$m_{\mathrm{T},W,\mathrm{sec}}$	Reconstructed transverse mass of secondary W-boson
H_T	Summed total transverse energy
$H_{\mathrm{T,had}}$	Summed hadronic transverse energy
$H_{\mathrm{T,lep}}$	Summed leptonic transverse energy
Aplanarity	Deviation from planar event shape
Planarity	Degree of flatness in momentum flow
Sphericity	Isotropy of momentum distribution

Deep Neural Network (DNN)

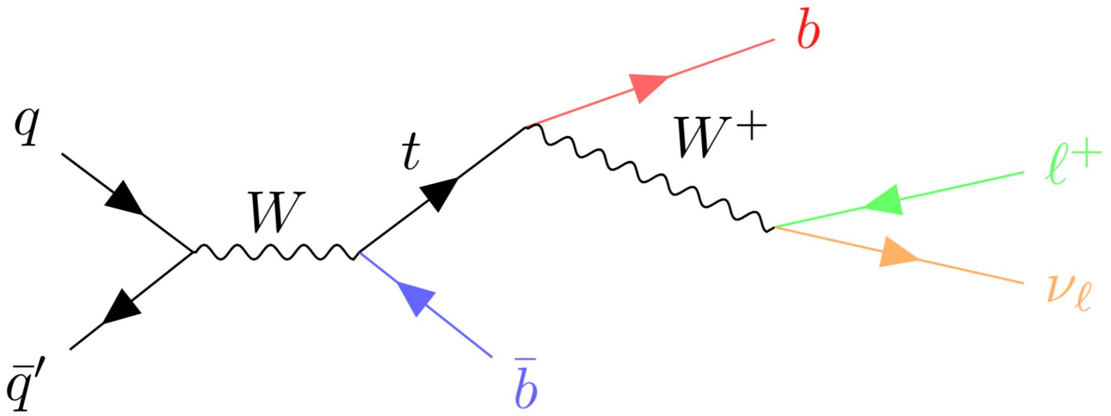


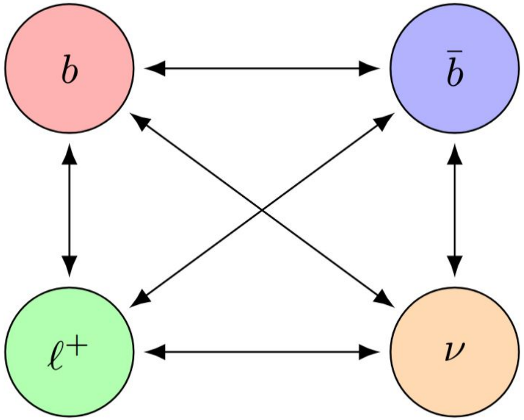
dimension: [59, 128, 256, 64, 32, 16, 1]

dropout: [0.1, 0.1, 0.1]

Feature	**Description**
$m_{W,\mathrm{prod}}$	Reconstructed mass of virtual production W-boson
m_t	Reconstructed mass of top quark
$m_{W,\mathrm{decay}}$	Reconstructed mass of decay W-boson
$m_{\mathrm{T},t}$	Reconstructed transverse mass of top quark
$m_{\mathrm{T},W,\mathrm{sec}}$	Reconstructed transverse mass of secondary W-boson
H_T	Summed total transverse energy
$H_{\mathrm{T,had}}$	Summed hadronic transverse energy
$H_{\mathrm{T,lep}}$	Summed leptonic transverse energy
Aplanarity	Deviation from planar event shape
Planarity	Degree of flatness in momentum flow
Sphericity	Isotropy of momentum distribution

Graph Neural Network





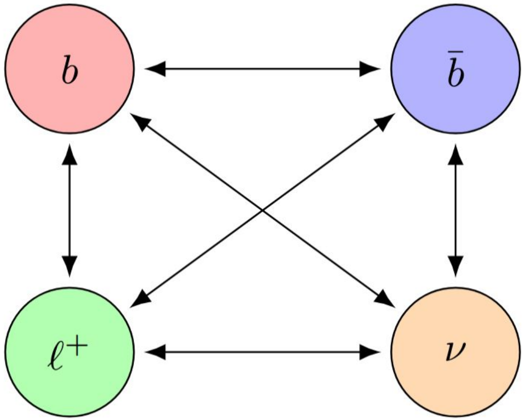
on each node:

[p_{T}, η, φ, 4 flags]

on each edge:

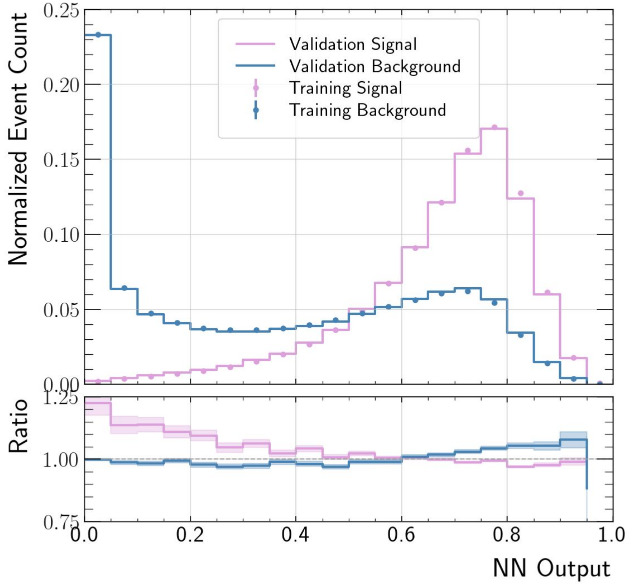
[$\Delta\eta$, $\Delta\varphi$, $\Delta\mathrm{R}$]

Graph Neural Network



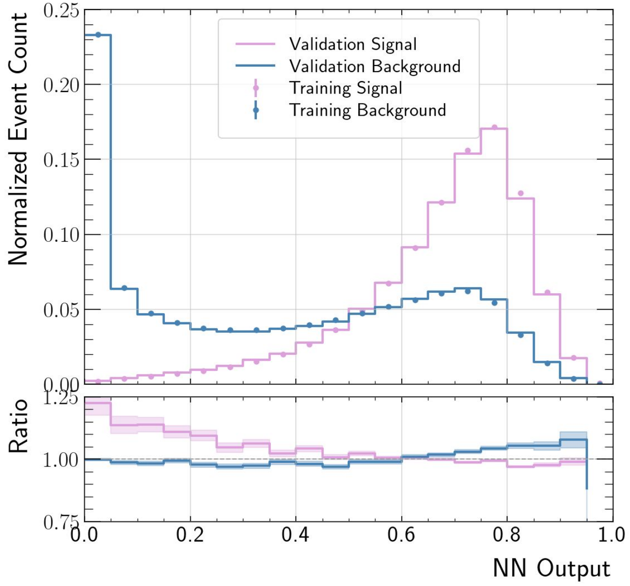
- convolution block
  - DNN applied on every edge
    - outputs aggregated at every node
    - used to update feature vector
- global mean
- final output DNN

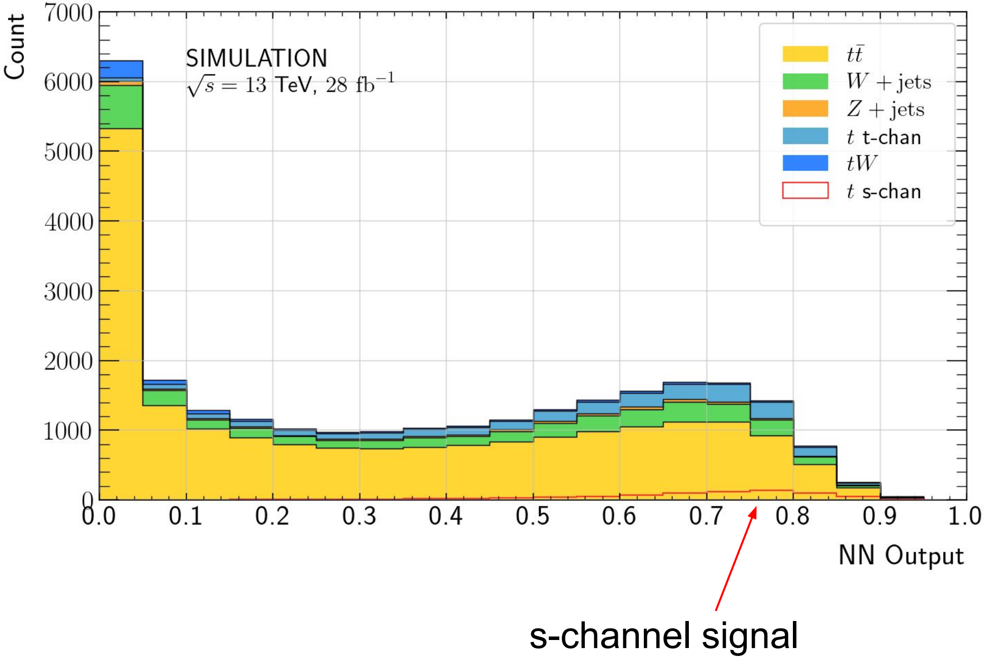
DNN Performance



No significant overtraining

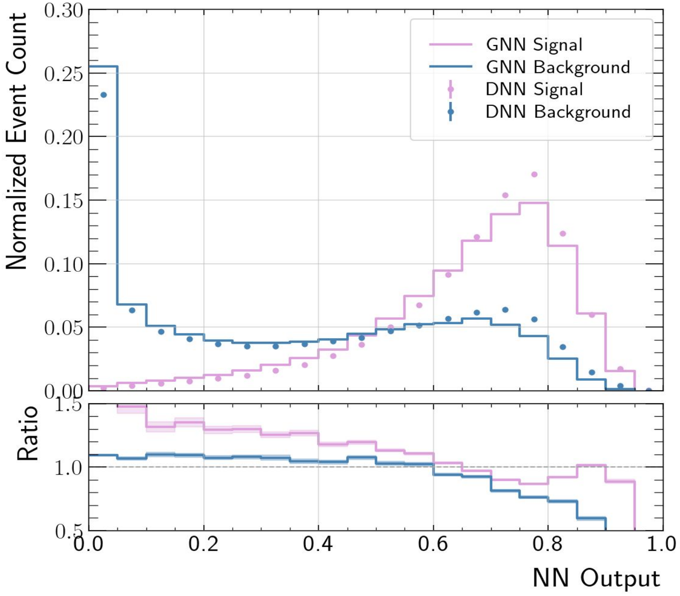
DNN Performance



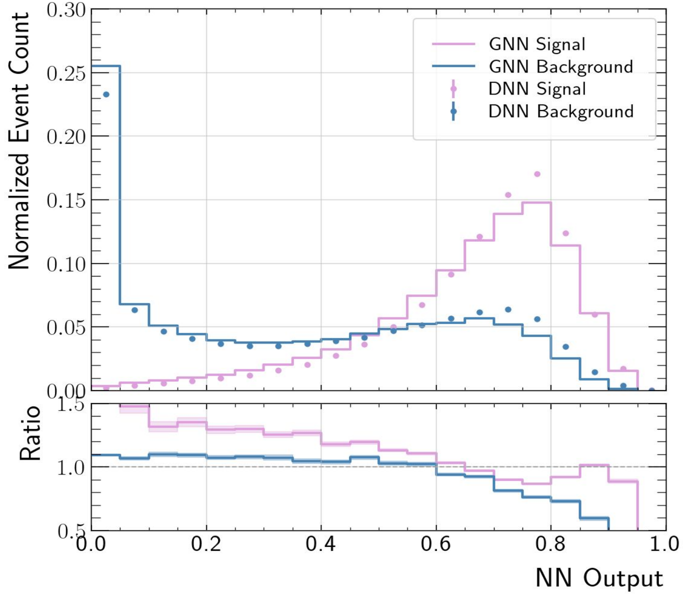


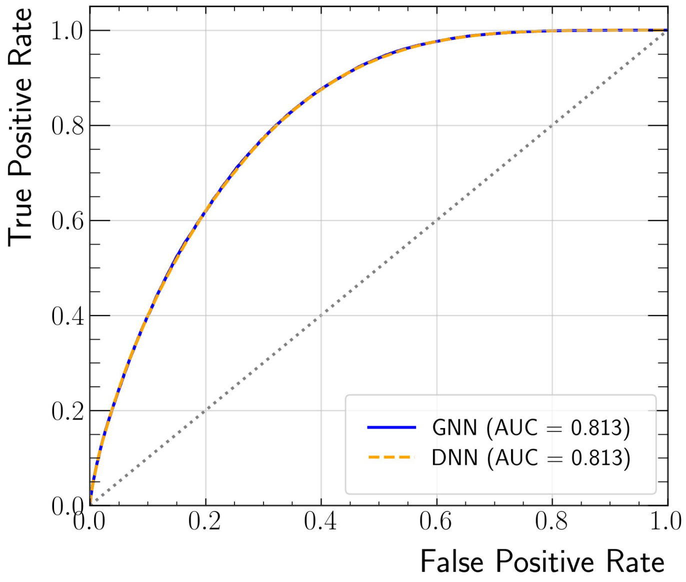
No significant overtraining

DNN/GNN Comparison



DNN/GNN Comparison





Same separation power at 0.315

Systematic uncertainties

- Every event in collider physics is a random occurrence
  - results in a poisson statistical uncertainty

- long list of potential systematic uncertainties
  - mainly detector effects or theory mismodelling
  - most significant in the current study:
    - Parton showering (PS) uncertainty
    - Matrix Element (ME) matching

Systematic errors

Systematic errors are evaluated based on their impact on the signal strength:

$$\mu = \frac{\sigma_{\mathrm{fit}}}{\sigma_{\mathrm{SM}}}$$

- individual fits with all but one systematic fixed
- all systematics free floating, evaluate the covariance of a systematic and μ

Systematic impact

Systematic	Past analysis $\frac{\Delta\mu}{\mu}$ [%]	DNN $\frac{\Delta\mu}{\mu}$ [%]	GNN $\frac{\Delta\mu}{\mu}$ [%]
s-channel: PS	$+18/-8$	$+16/-13$	$+17/-14$
s-channel: ME matching	-	$+5.5/-4.2$	$+4.7/-3.8$
$t\bar{t}$: h_{damp}	<1	$-2.1/+1.6$	$+0.2/-0.2$
$t\bar{t}$: PS	$+12/-10$	$+2.0/-1.5$	$+5.5/-4.4$
$t\bar{t}$: ME matching	$+10/-8$	$-0.5/+0.4$	$-9.4/+7.6$
luminosity	$+4/-3$	$-2.7/+2.0$	$-2.4/+1.9$

Systematic impact

Systematic	Past analysis $\frac{\Delta\mu}{\mu}$[%]	DNN $\frac{\Delta\mu}{\mu}$ [%]	GNN $\frac{\Delta\mu}{\mu}$ [%]
s-channel: PS	$+18/-8$	$+16/-13$	$+17/-14$
s-channel: ME matching	-	$+5.5/-4.2$	$+4.7/-3.8$
$t\bar{t}$: h_{damp}	<1	$-2.1/+1.6$	$+0.2/-0.2$
$t\bar{t}$: PS	$+12/-10$	$+2.0/-1.5$	$+5.5/-4.4$
$t\bar{t}$: ME matching	$+10/-8$	$-0.5/+0.4$	$-9.4/+7.6$
luminosity	$+4/-3$	$-2.7/+2.0$	$-2.4/+1.9$

Systematic impact

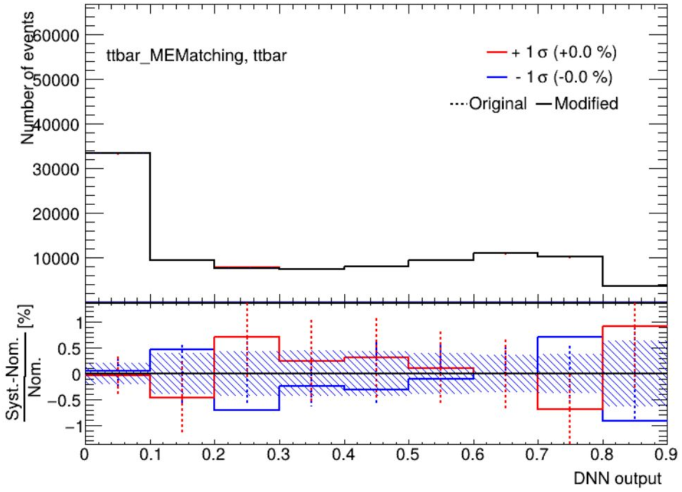
Systematic	Past analysis $\frac{\Delta\mu}{\mu}$[%]	DNN $\frac{\Delta\mu}{\mu}$ [%]	GNN $\frac{\Delta\mu}{\mu}$ [%]
s-channel: PS	$+18/-8$	$+16/{-13}$	$+17/{-14}$
s-channel: ME matching	-	$+5.5/{-4.2}$	$+4.7/{-3.8}$
$t\bar{t}$: h_{damp}	<1	$-2.1/{+1.6}$	$+0.2/{-0.2}$
$t\bar{t}$: PS	$+12/{-10}$	$+2.0/{-1.5}$	$+5.5/{-4.4}$
$t\bar{t}$: ME matching	$+10/{-8}$	$-0.5/{+0.4}$	$-9.4/{+7.6}$
luminosity	$+4/{-3}$	$-2.7/{+2.0}$	$-2.4/{+1.9}$

Systematic impact

Systematic	Past analysis $\frac{\Delta\mu}{\mu}$[%]	DNN $\frac{\Delta\mu}{\mu}$ [%]	GNN $\frac{\Delta\mu}{\mu}$ [%]
s-channel: PS	$+18/-8$	$+16/-13$	$+17/-14$
s-channel: ME matching	-	$+5.5/-4.2$	$+4.7/-3.8$
$t\bar{t}$: h_{damp}	<1	$-2.1/+1.6$	$+0.2/-0.2$
$t\bar{t}$: PS	$+12/-10$	$+2.0/-1.5$	$+5.5/-4.4$
$t\bar{t}$: ME matching	$+10/-8$	$-0.5/+0.4$	$-9.4/+7.6$
luminosity	$+4/-3$	$-2.7/+2.0$	$-2.4/+1.9$

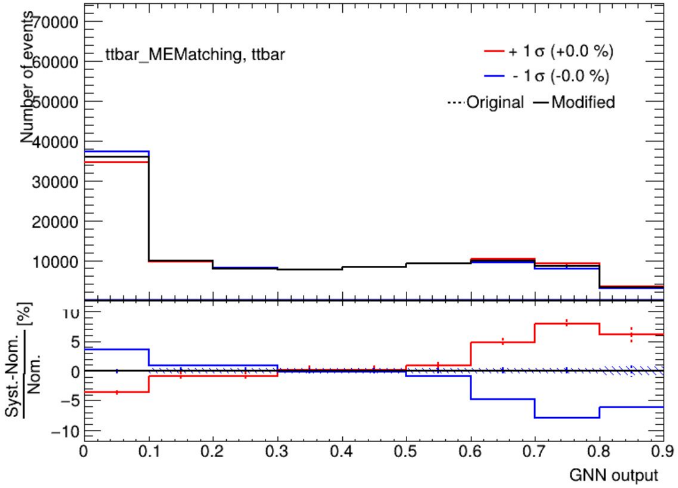
$t\bar{t}$ Matrix element matching

DNN



- almost within statistical error

GNN



- deviations of up to 8%

Conclusion

- data processing including SR implementation
- setup 2 machine learning models
- first modelling systematics evaluated

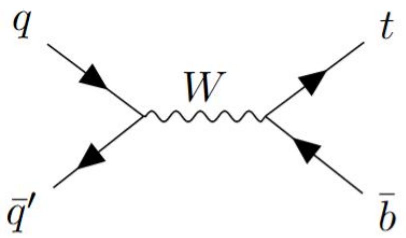
- While the GNN performed the same on nominal data
  - **DNN is clearly more stable towards systematic impacts**

Outlook:

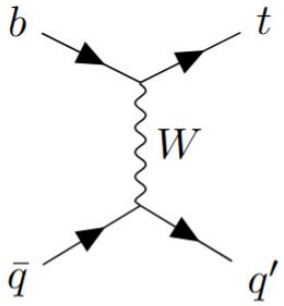
- more GNN testing
- DNN feature selection
- implementation of the remaining systematics
- implementation of control regions

Backup

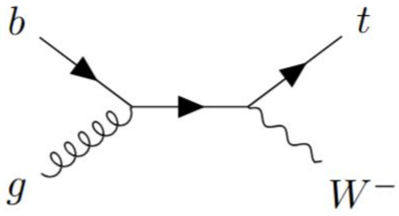
Backup, Single-top-quark produktion



s-channel



t-channel

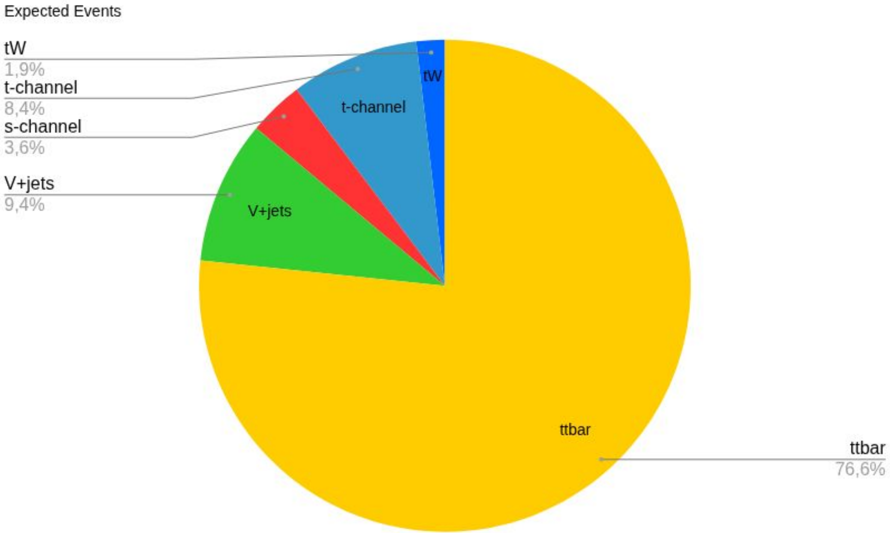


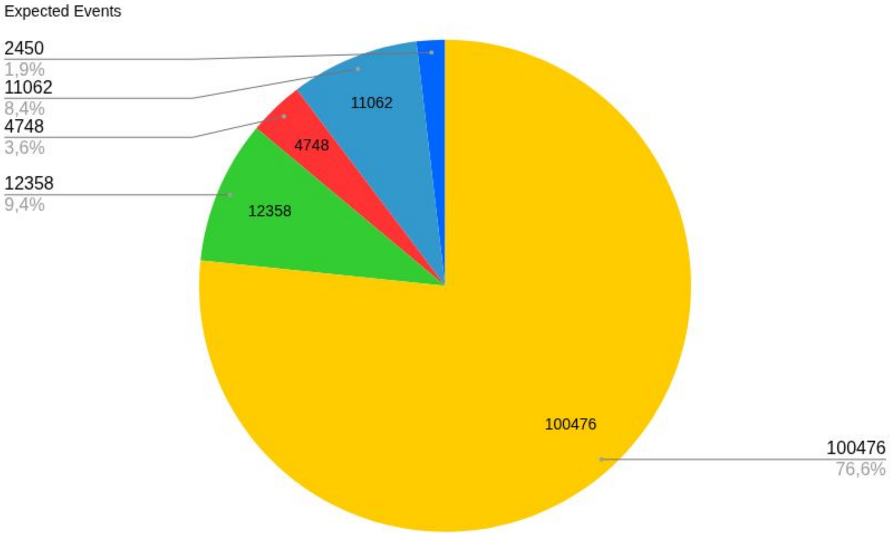
tW associated production

Signal Region definition

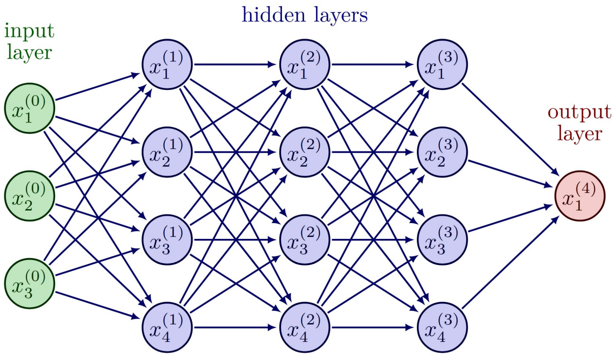
	SR
$\mathrm{n}(e, \mu)$	$== 1$ with $p_T >$ 30 GeV
n (jets)	$== 2$, in $\lvert\eta\rvert < 2.5$
n (b-tags)	$== 2$, @77%, in $\lvert\eta\rvert < 2.5$
Leading jet p_T	> 40 GeV, in $\lvert\eta\rvert < 2.5$
Subleading jet p_T	> 30 GeV, in $\lvert\eta\rvert < 2.5$
Missing transverse energy (E_T^{miss})	> 35 GeV
Transverse W boson mass	> 30 GeV
Additional jets (low p_T) veto	No jets with $p_T < 30$ GeV
Forward jets veto	No jets with $\lvert\eta\rvert > 2.5$
Additional leptons veto	No extra leptons with $p_T < 30$ GeV

Backup, SR Composition





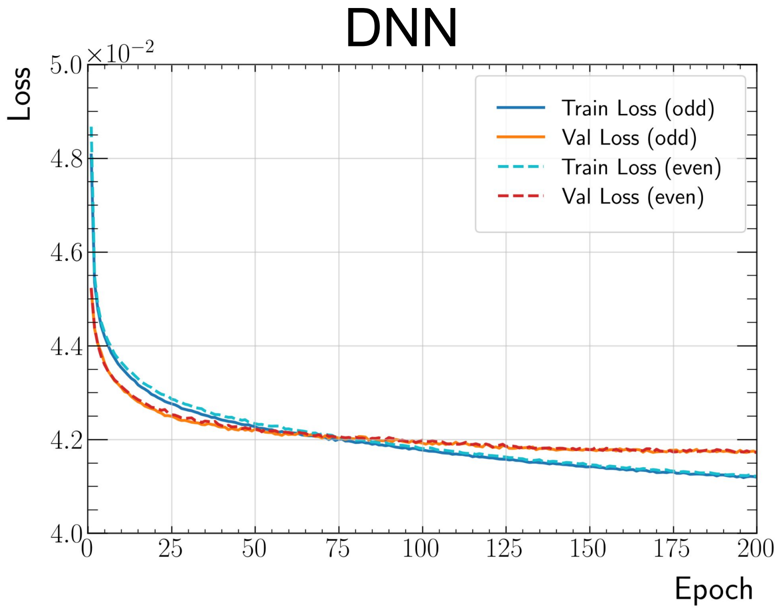
Deep Neural Network (DNN)

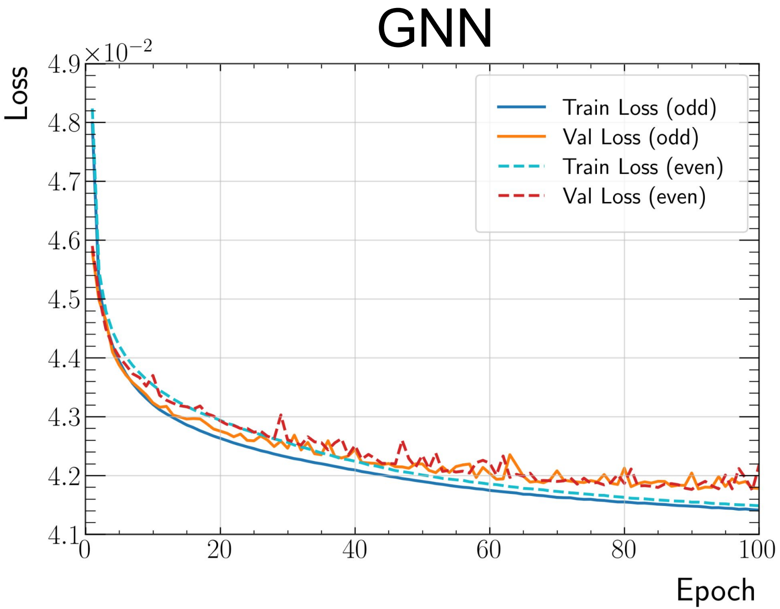


Hyperparameter	**Setting**
Batch size	2048
Epochs	200
Shape	[128, 256, 64, 32, 16]
Dropout	[0.1, 0.1, 0.1]
Loss function	Binary crossentropy
Optimizer	Adam
Learing rate	$7.1 \cdot 10^{-5}$

Feature	**Description**
$m_{W,\mathrm{prod}}$	Reconstructed mass of virtual production W-boson
m_t	Reconstructed mass of top quark
$m_{W,\mathrm{decay}}$	Reconstructed mass of decay W-boson
$m_{\mathrm{T},t}$	Reconstructed transverse mass of top quark
$m_{\mathrm{T},W,\mathrm{sec}}$	Reconstructed transverse mass of secondary W-boson
H_T	Summed total transverse energy
$H_{\mathrm{T,had}}$	Summed hadronic transverse energy
$H_{\mathrm{T,lep}}$	Summed leptonic transverse energy
Aplanarity	Deviation from planar event shape
Planarity	Degree of flatness in momentum flow
Sphericity	Isotropy of momentum distribution

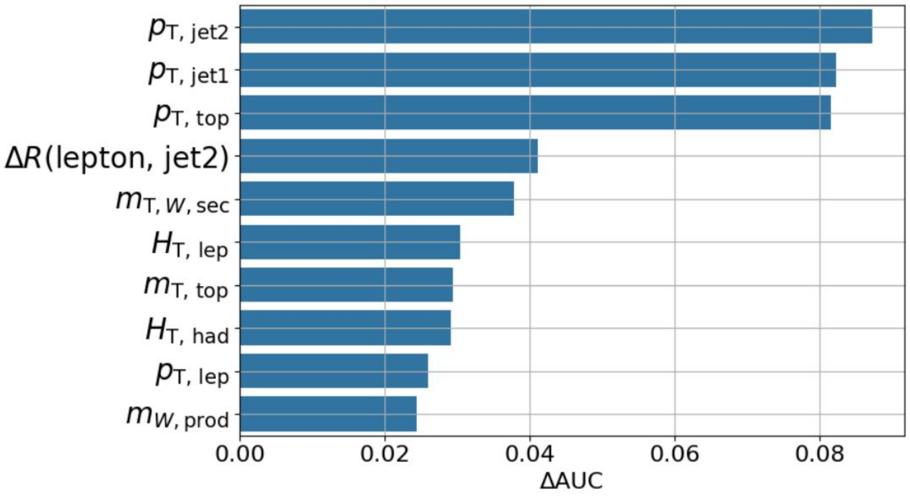
Training stability



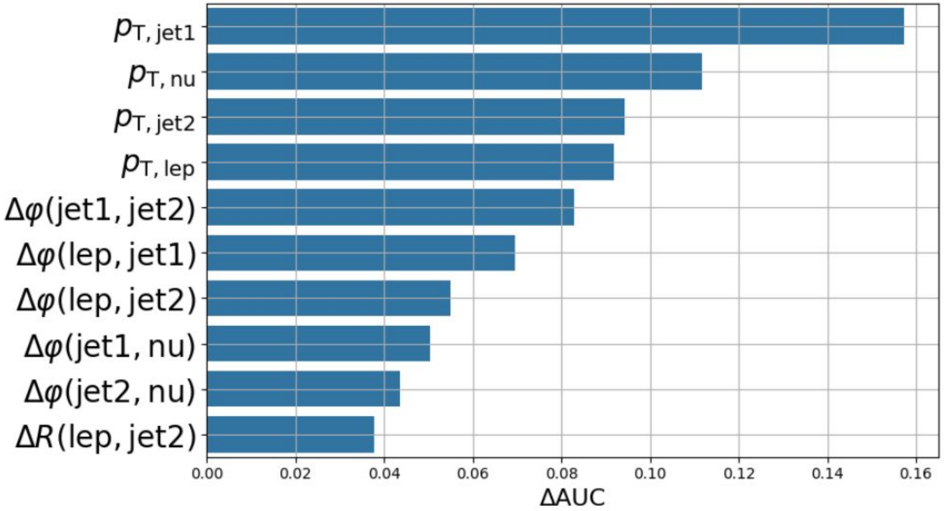


Feature Importance

DNN



GNN



- GNN shows larger reliance on individual features

Systematic correlation matrices

DNN

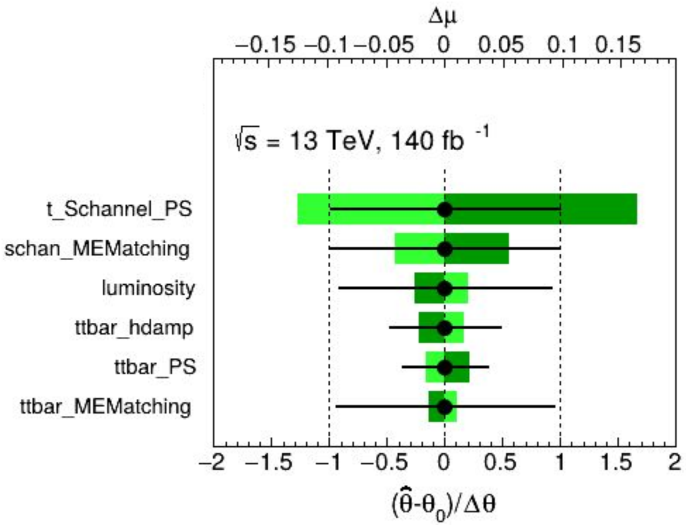
	μ	luminosity	t s-channel **ME**	t s-channel **PS**	$t\bar{t}$ **ME**	$t\bar{t}$ **PS**	$t\bar{t}$ h_{damp}
μ	100.0	-14.7	28.2	84.0	-3.7	28.4	-22.2
luminosity	-14.7	100.0	1.2	-0.0	-2.2	-11.6	70.8
t s-channel **ME**	28.2	1.2	100.0	0.3	-2.2	7.0	-5.5
t s-channel **PS**	84.0	-0.0	0.3	100.0	2.6	-2.0	1.0
$t\bar{t}$ **ME**	-3.7	-2.2	-2.2	2.6	100.0	-20.2	11.6
$t\bar{t}$ **PS**	28.4	-11.6	7.0	-2.0	-20.2	100.0	-71.4
$t\bar{t}$ h_{damp}	-22.2	70.8	-5.5	1.0	11.6	-71.4	100.0

GNN

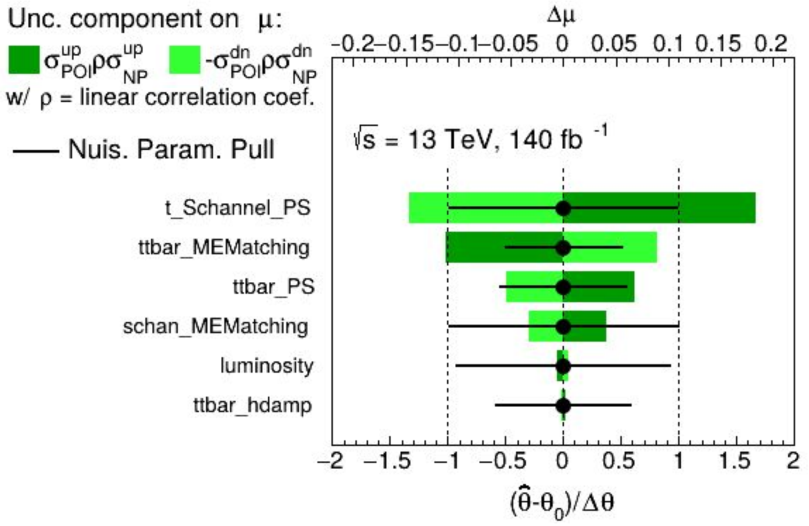
	μ	luminosity	t s-channel **ME**	t s-channel **PS**	$t\bar{t}$ **ME**	$t\bar{t}$ **PS**	$t\bar{t}$ h_{damp}
μ	100.0	-9.0	16.7	61.2	-74.0	56.2	1.8
luminosity	-9.0	100.0	0.7	-0.7	-0.9	-5.0	75.3
t s-channel **ME**	16.7	0.7	100.0	0.4	2.6	0.4	-3.1
t s-channel **PS**	61.2	-0.7	0.4	100.0	-3.4	-0.3	3.1
$t\bar{t}$ **ME**	-74.0	-0.9	2.6	-3.4	100.0	-85.0	3.7
$t\bar{t}$ **PS**	56.2	-5.0	0.4	-0.3	-85.0	100.0	-37.3
$t\bar{t}$ h_{damp}	1.8	75.3	-3.1	3.1	3.7	-37.3	100.0

Ranking Plots

DNN



GNN



Uncertainty on μ

DNN:

$$\mu_{\mathrm{DNN}} = 1^{+0.20}_{-0.15}$$

GNN:

$$\mu_{\mathrm{GNN}} = 1^{+0.32}_{-0.28}$$

Systematic errors

- Systematic errors are evaluated based on their impact on the signal strength: $\mu = \frac{\sigma_{\text{fit}}}{\sigma_{\text{SM}}}$

- either, individual fits with all but one systematic fixed
- or, all systematics free floating, evaluate the covariance of a systematic and μ

- impact: $I_{\theta_n \to \mu} = V_{\mu,n} = \rho_{\mu,n}\, \sigma_\mu\, \sigma_{\theta_n}$