FACULTY OF SCIENCES

DEPARTMENT: PHYSICS & ASTRONOMY
RESEARCH GROUP: EXPERIMENTAL PARTICLE PHYSICS

Disentangling triple and four top quark production at the LHC

Karam Kaspar

19 June 2024

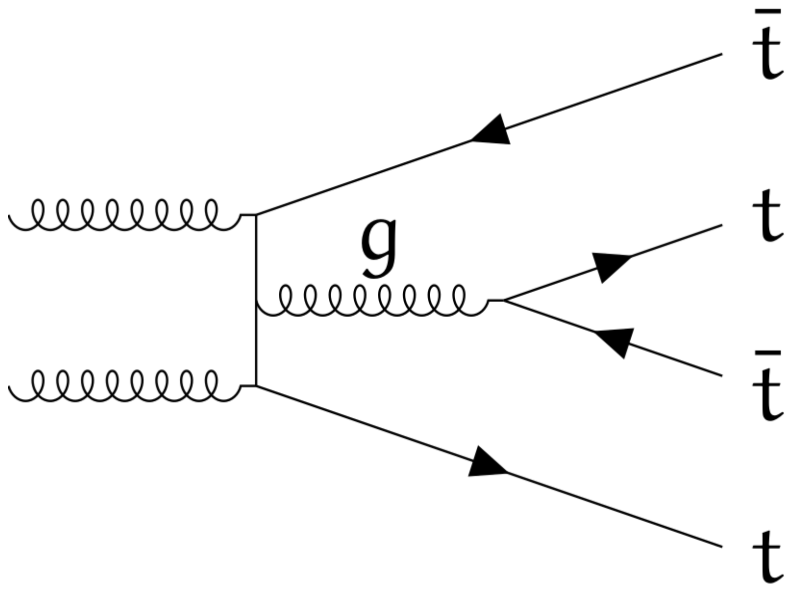
GHENT UNIVERSITY

Multi-top production

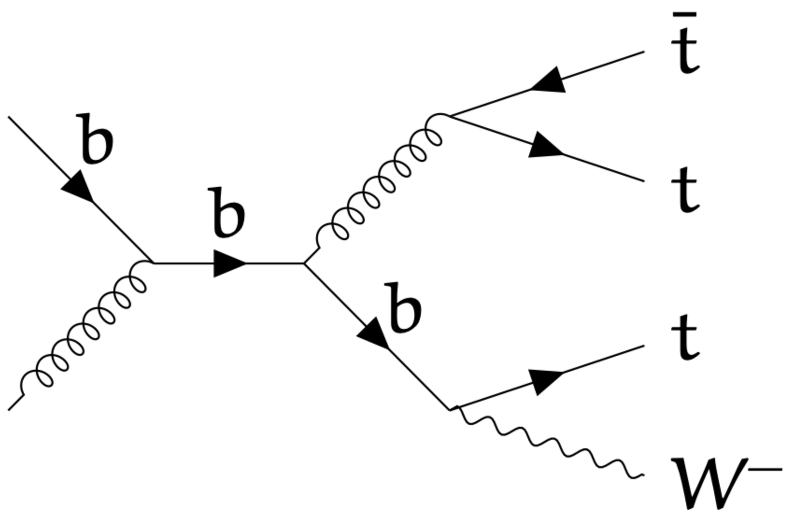
- Top quark discovery in 1995 at Fermilab.
- Heaviest elementary SM particle with the largest Yukawa coupling.
- pp-collisions produce tttt, tttW & tttq events at the LHC.

Process	xsection [fb]	Final configuration
tttt	13.4	(4W, 4b)
tttW	1.3	(4W, 3b)
tttq	0.7	(3W, 3b, q)

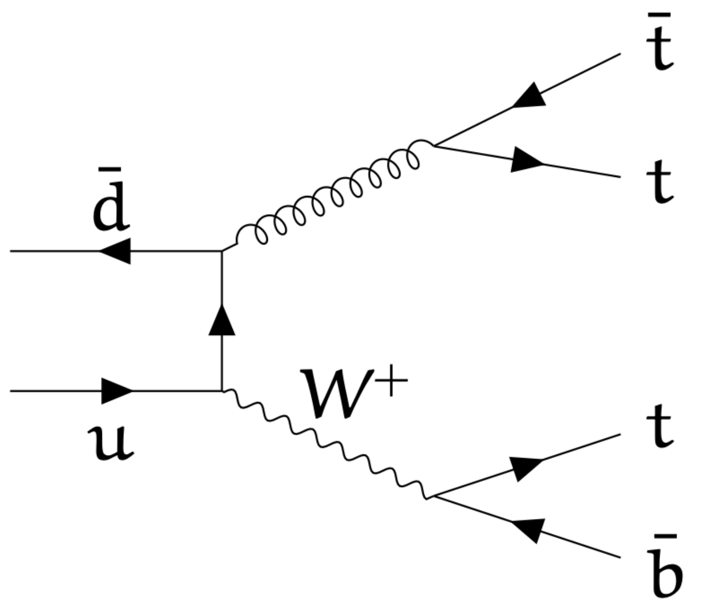
tttt



tttW

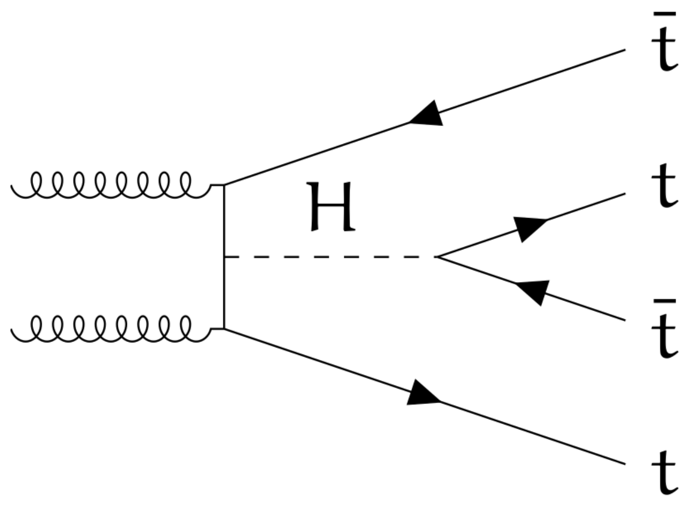


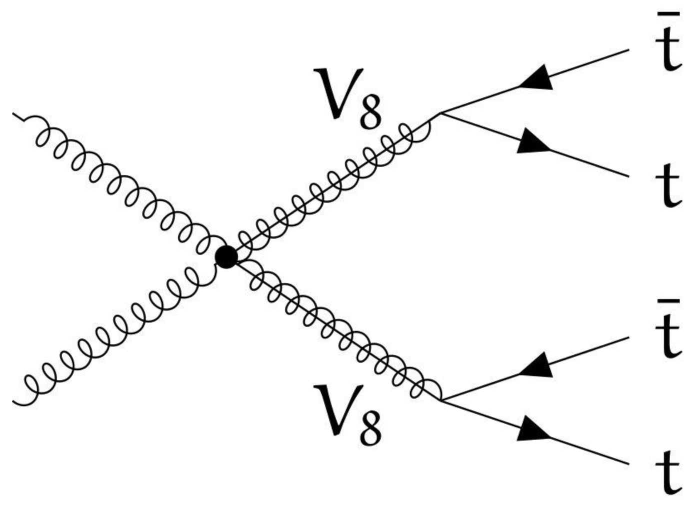
tttq



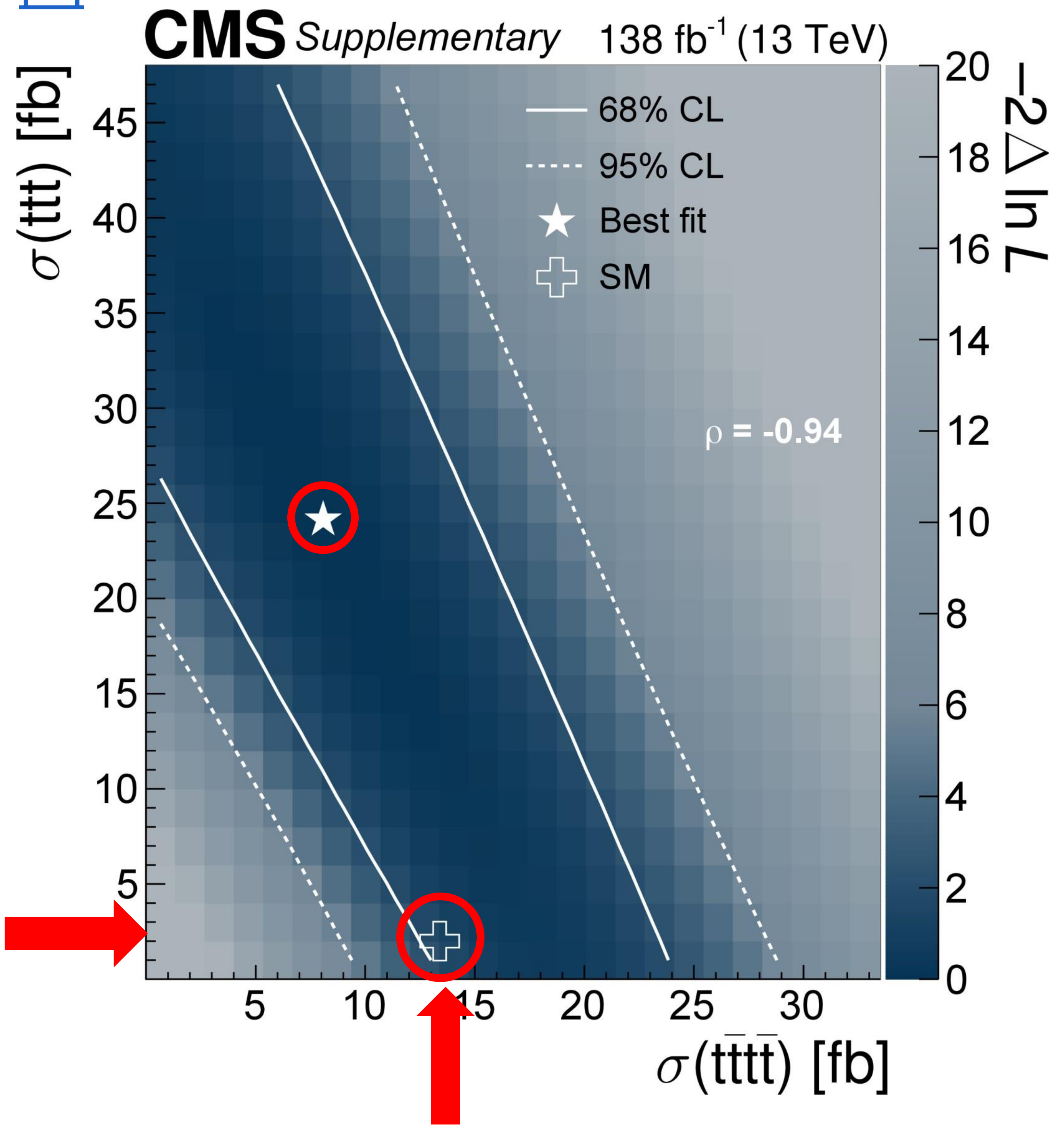
Goal & Motivations

- **Goal:**
  - Distinction between tttt and ttt signals.
- **Motivation:**
  - ttt is important background for tttt with similar final states.
- **Importance of tttt:**
  - tttt is heaviest observed SM process.
  - Determination of top Yukawa coupling.
  - Enhancement via BSM physics.
- **Importance of ttt:**
  - Occurrence of FVNI & FCNI.



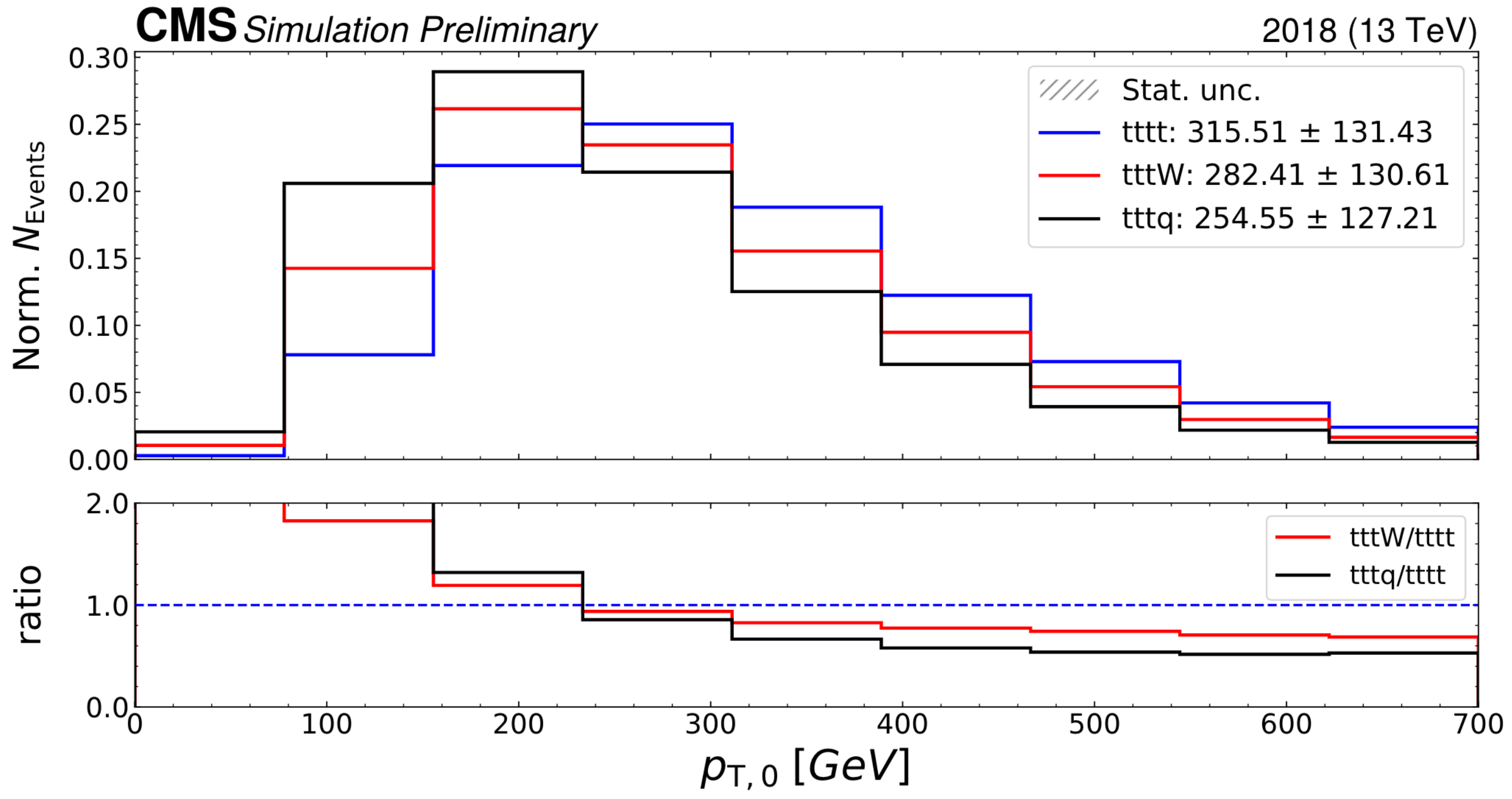


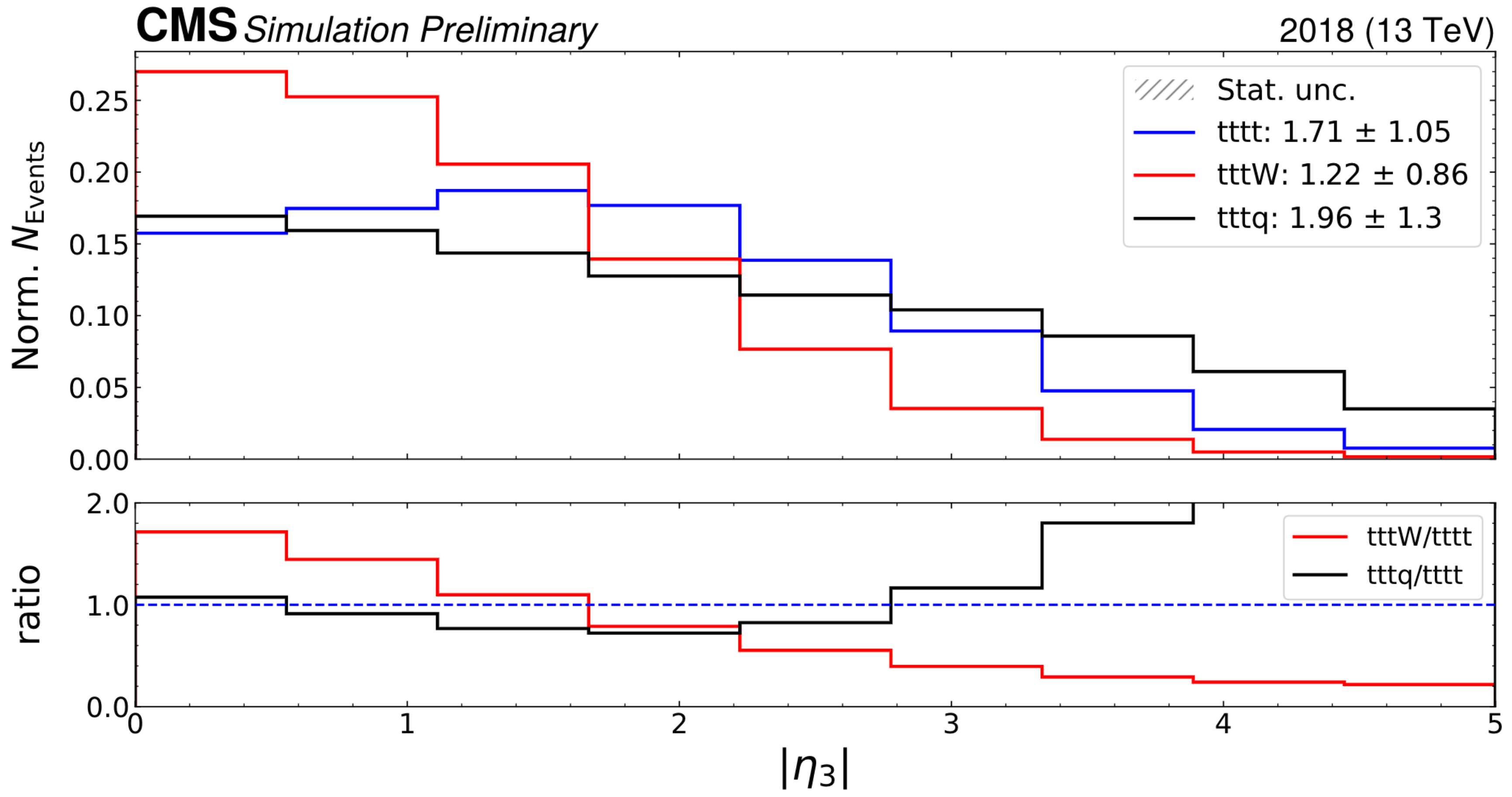
[1]



Generator level

- Investigation of generator particle characteristics
  → Distinctive patterns observed





GHENT UNIVERSITY

tttt

tttW

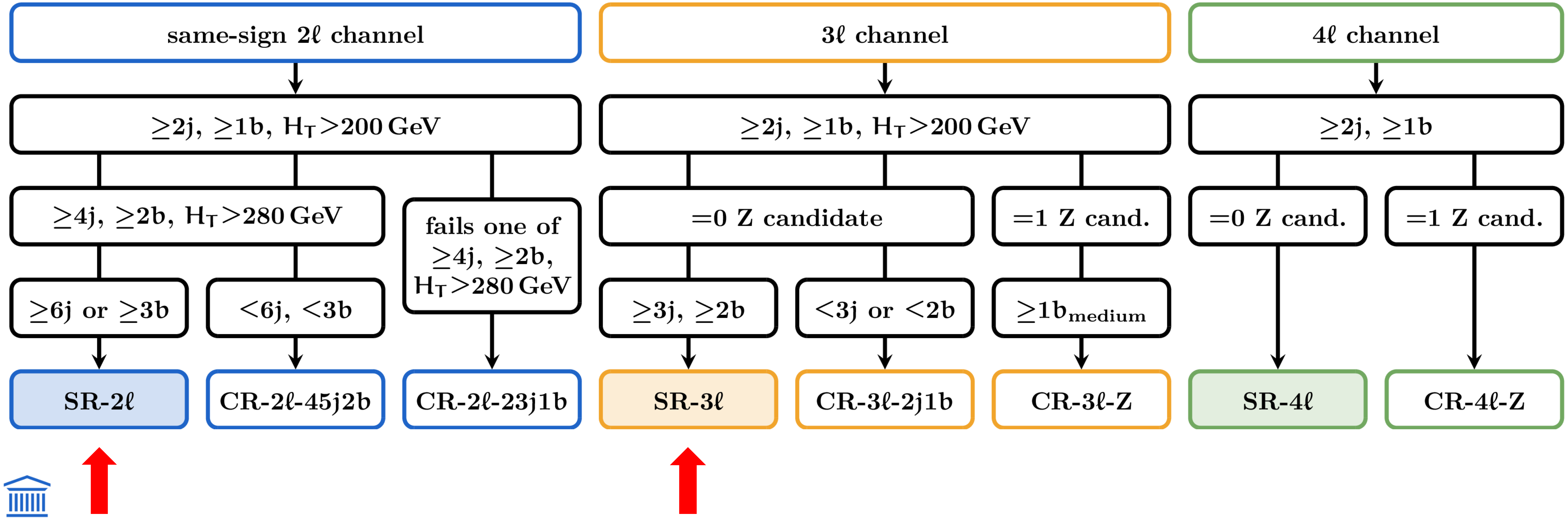
tttq

tttW/tttt

tttq/tttt

Event selection

- Goal: Filtering high-quality reconstructed objects.
- Exact implementation of tttt selection from observation paper.
- Same-sign di-lepton (SR-2l) & triple lepton (SR-3l) channels.



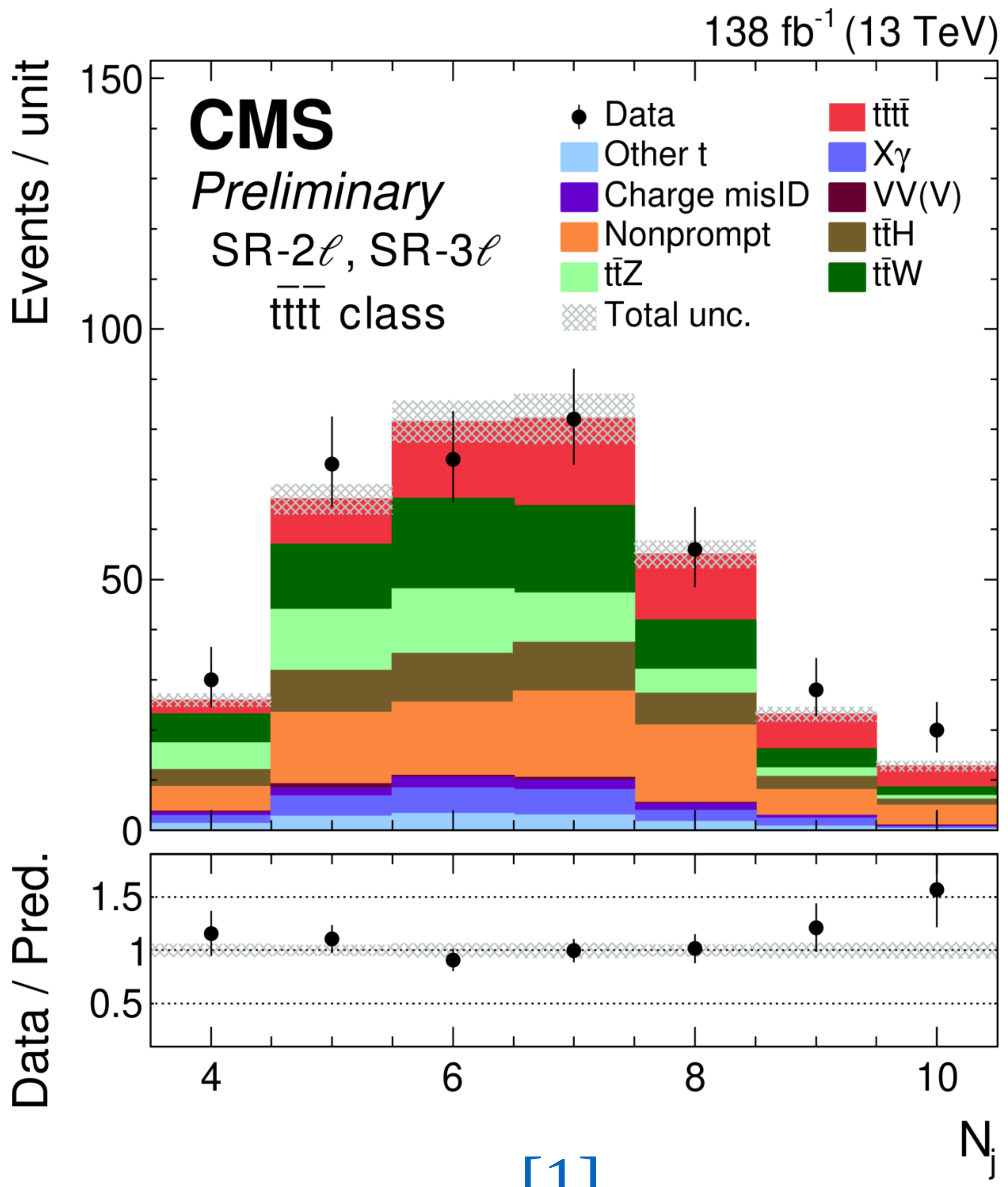
Reconstruction level variables

- Goal: Find observables for the distinction of tttt, tttW and tttq events.
- Explored about 100 observables each in SR-2l and SR-3l.

N_{j}	H_{T}	L_{T}
N_b^{loose}	N_b^{medium}	N_b^{tight}
$\mathrm{p}_{T,\ \mathrm{MET}}$	$p_{\mathrm{T}}(\ell_{0,1,2})$	$\lvert\eta(\ell_{0,1,2})\rvert$
$\mathrm{p}_T(\mathrm{j}_{0,1,2,3,4})$	$\lvert\eta(\mathrm{j}_{0,1,2,3,4})\rvert$	$\mathrm{DF}(\mathrm{j}_{0,1,2,3,4})$
$\mathrm{p}_T(\mathrm{b}_{0,1,2,3,4})$	$\lvert\eta(\mathrm{b}_{0,1,2,3,4})\rvert$	$\max\ \mathrm{DF}(\mathrm{j}_{0,1,2,3,4})$
$H_{\mathrm{T,\ f}}$	$\mathrm{p}_T(\mathrm{fj}_{0,1})$	$\lvert\eta(\mathrm{fj}_{0,1})\rvert$
min, max $\Delta R(\ell\ell)$	min, max $\lvert\Delta\eta(\ell\ell)\rvert$	min, max $\lvert\Delta\phi(\ell\ell)\rvert$
$\Delta R(\ell,\ \mathrm{j})$	$\lvert\Delta\eta(\ell,\ \mathrm{j})\rvert$	$\lvert\Delta\phi(\ell,\ \mathrm{j})\rvert$
$\Delta R(\mathrm{j},\ \mathrm{b})$	$\lvert\Delta\eta(\mathrm{j},\ \mathrm{b})\rvert$	$\lvert\Delta\phi(\mathrm{j},\ \mathrm{b})\rvert$
$\Delta R(\mathrm{j},\ \mathrm{fj})$ $\Delta R(\mathrm{nb},\ \mathrm{b})$	$\lvert\Delta\eta(\mathrm{j},\ \mathrm{fj})\rvert$ $\lvert\Delta\eta(\mathrm{nb},\ \mathrm{b})\rvert$	$\lvert\Delta\phi(\mathrm{j},\ \mathrm{fj})\rvert$ $\lvert\Delta\phi(\mathrm{nb},\ \mathrm{b})\rvert$
$M_{\mathrm{inv},\ \ell}$	$\max\ M_{\mathrm{inv,\ j\ j}}$	$\max\ p_{\mathrm{T,\ j\ j}}$
$\Sigma_\ell q_\ell$	$\max\ M_{\mathrm{inv,\ j\ fj}}$	$\max\ p_{\mathrm{T,\ j\ fj}}$
m_{T_0}	m_{W_0}	$m_T(\ell_{0,1,2})$

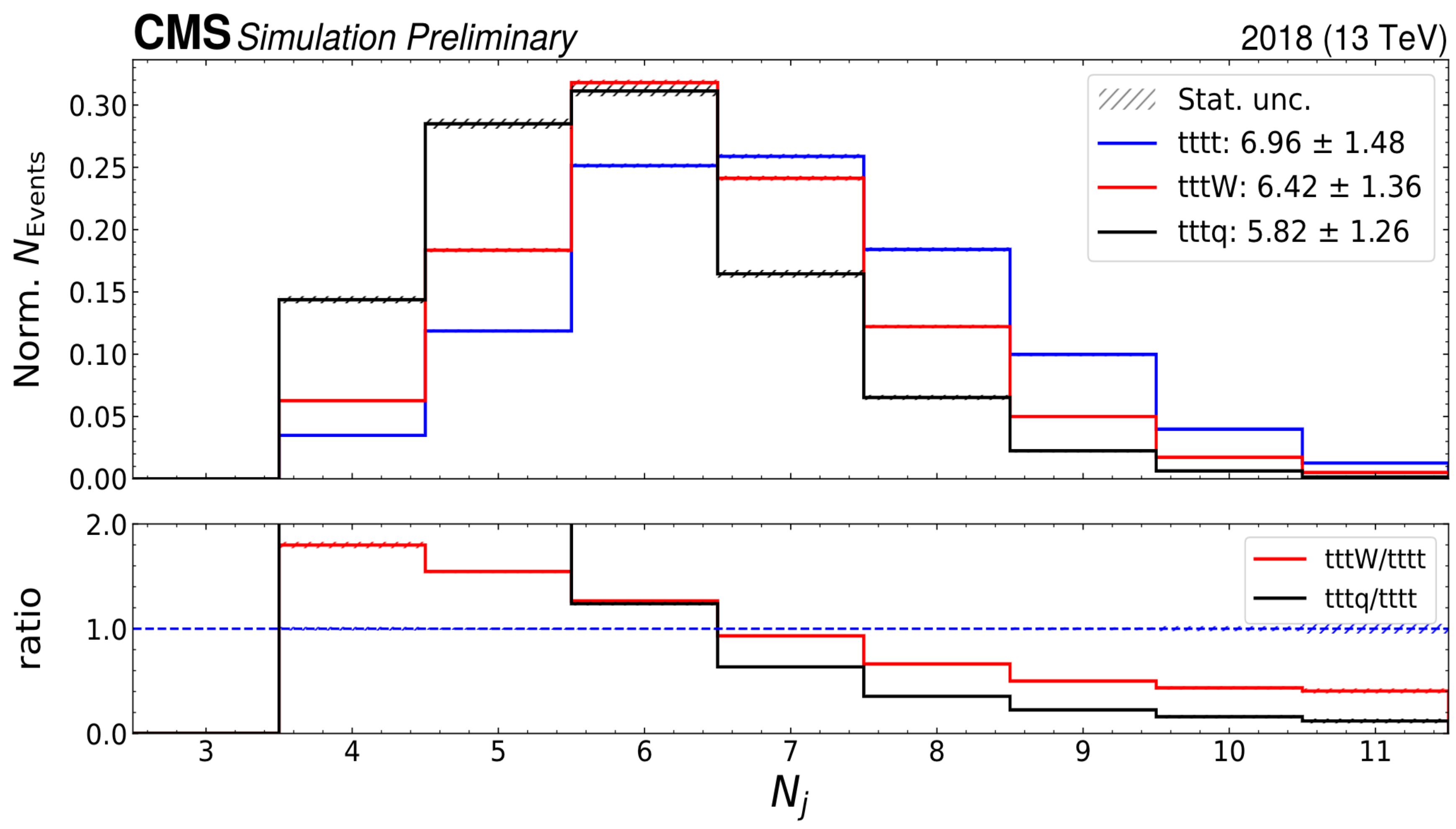
GHENT UNIVERSITY

Most differentiating variables



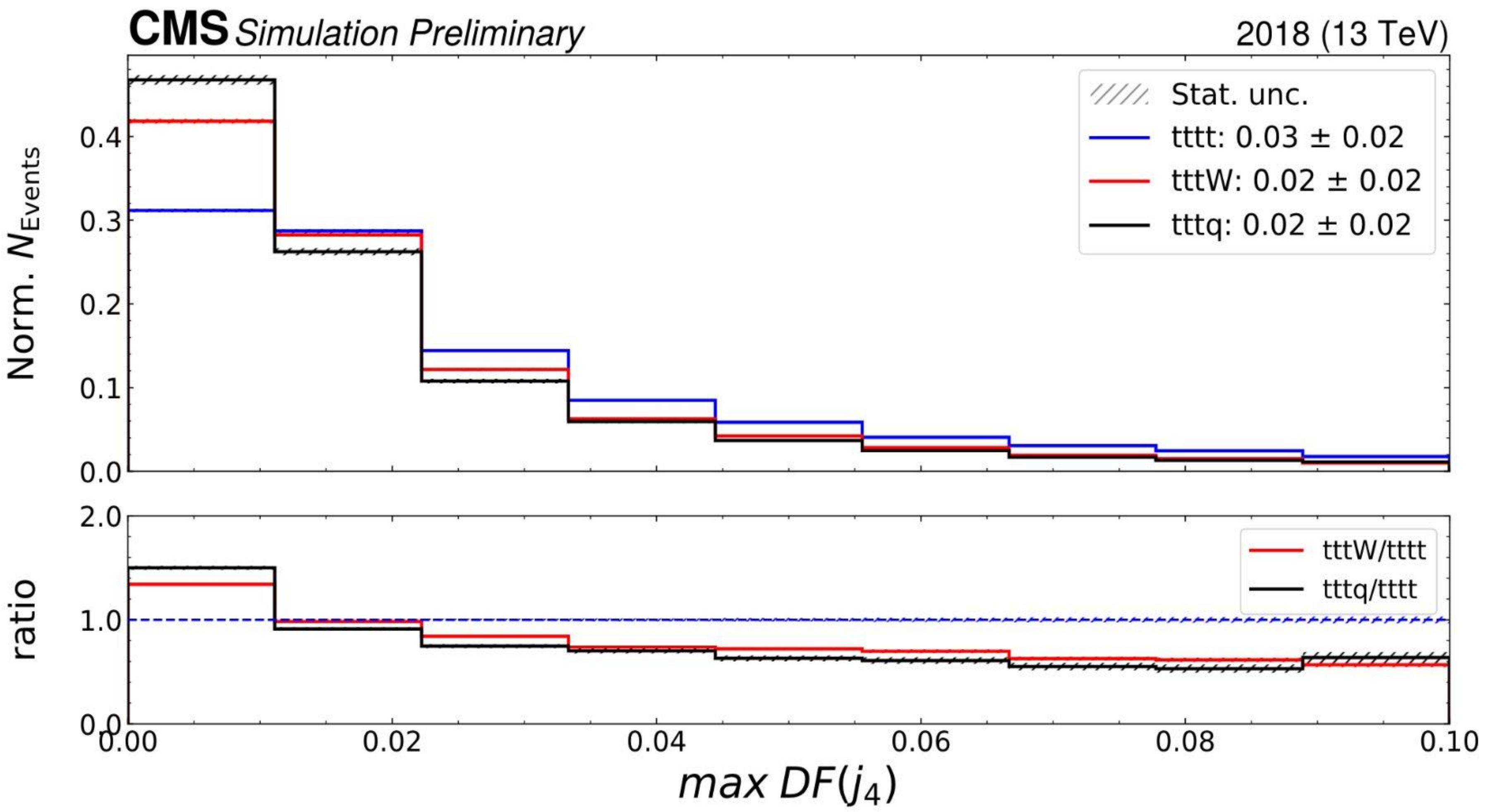
[1]

SR-2l

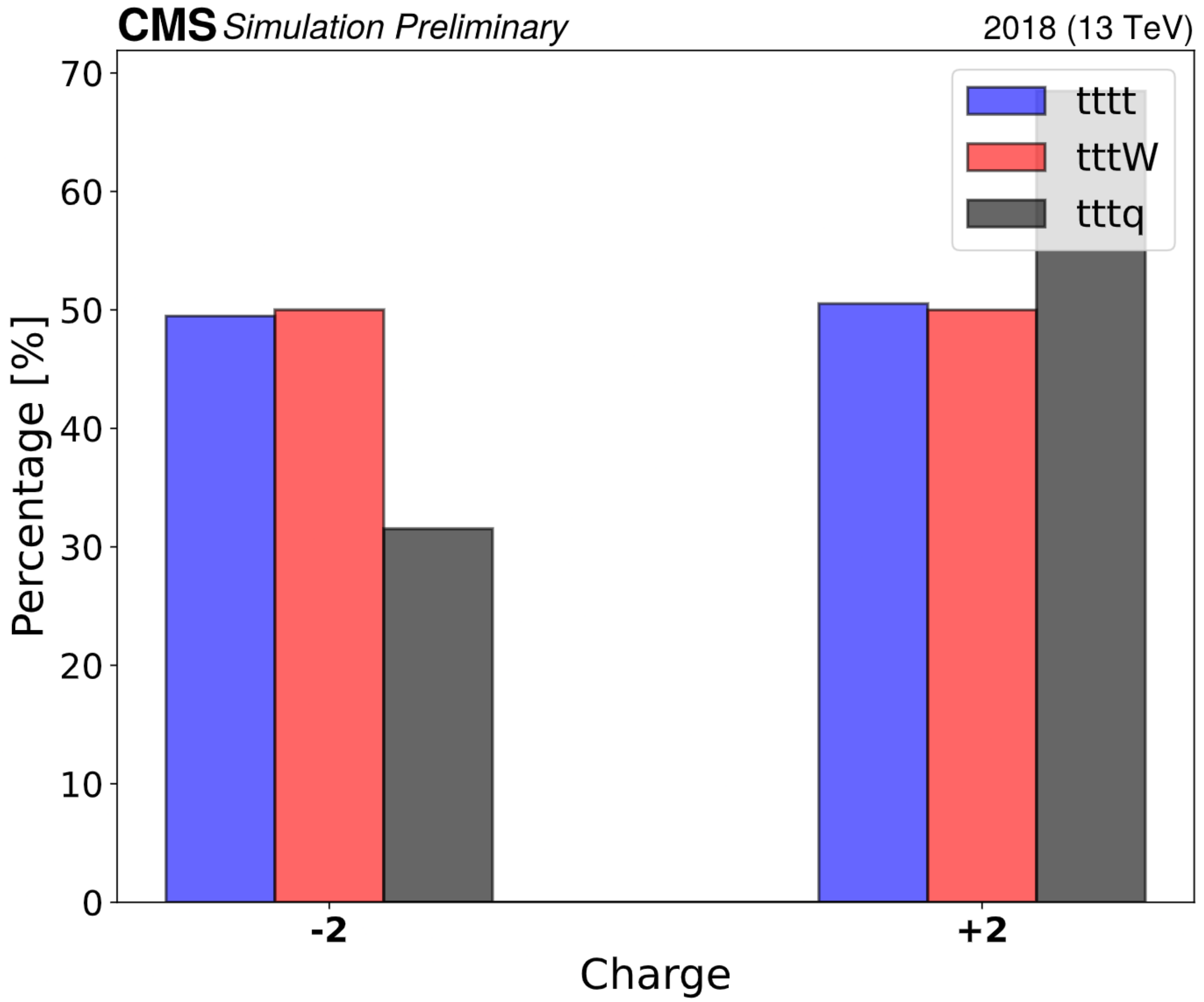


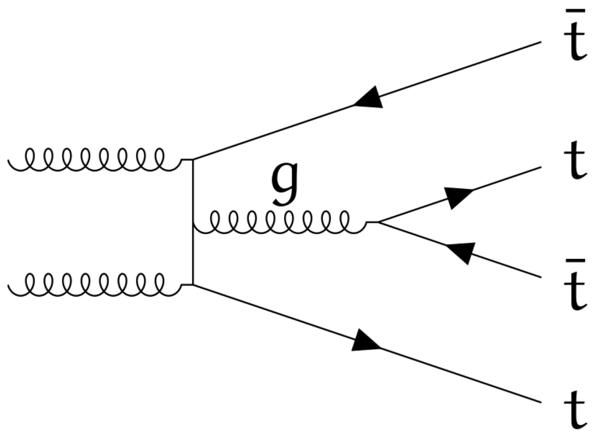
Most differentiating variables

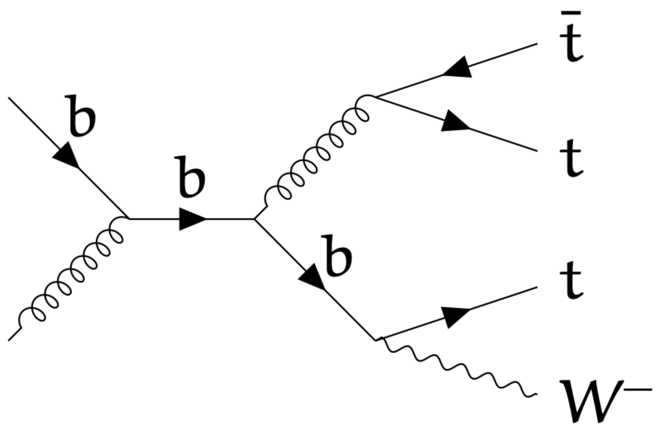
SR-2l

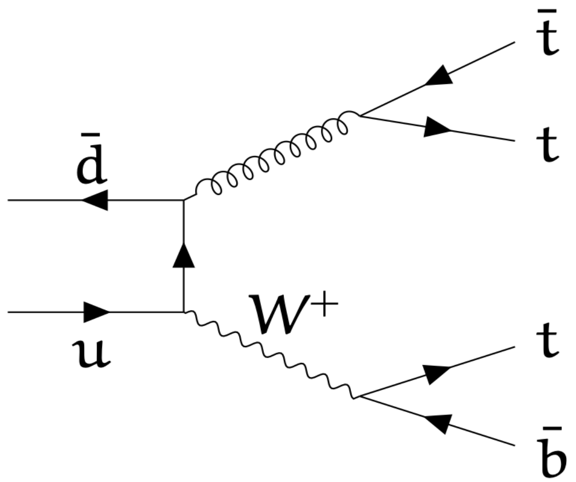


SR-2l







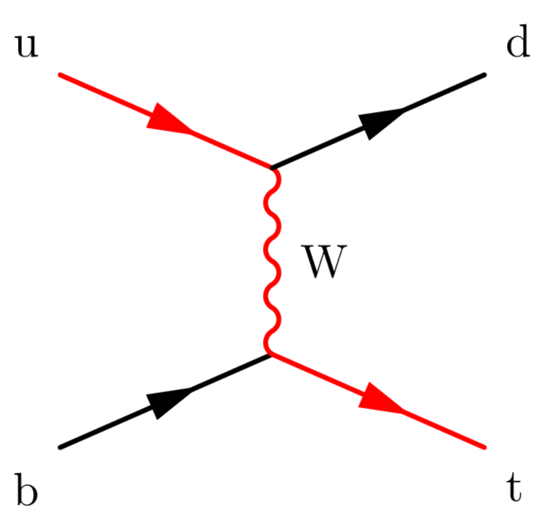


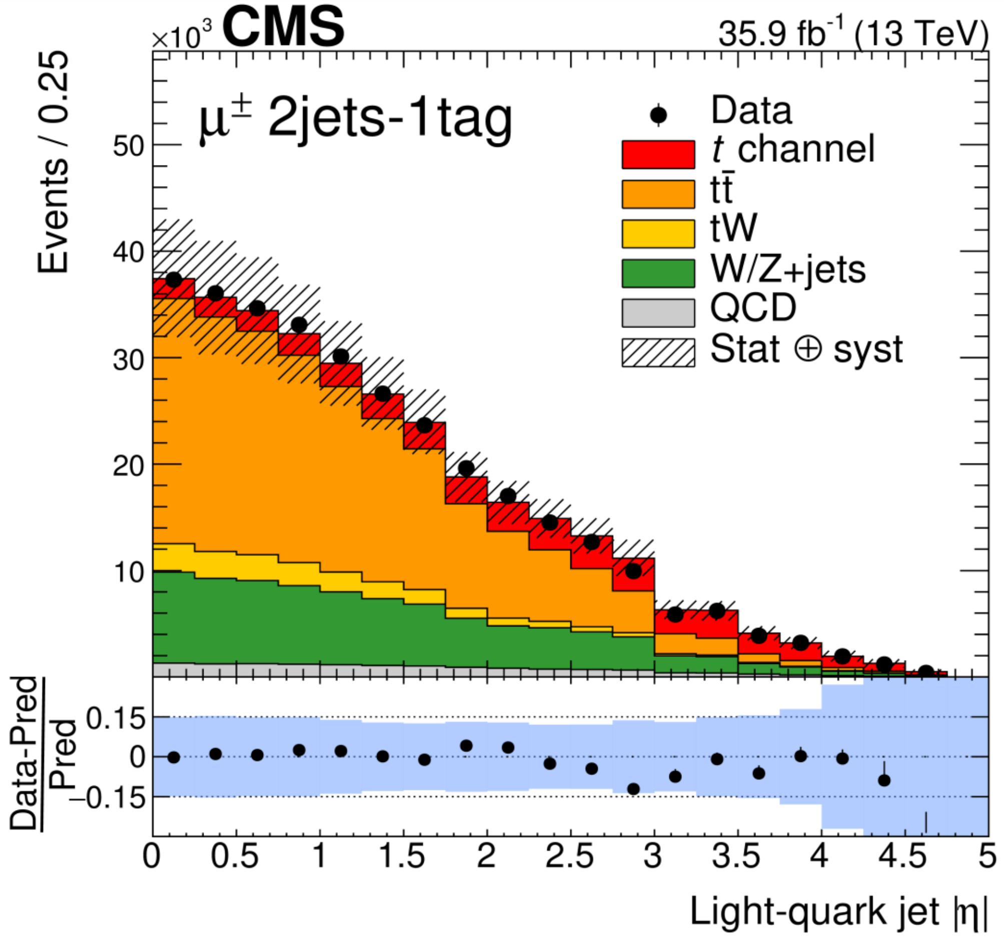
Forward jets

tq vs tt analysis

- Jet recoil due to a top quark jet.

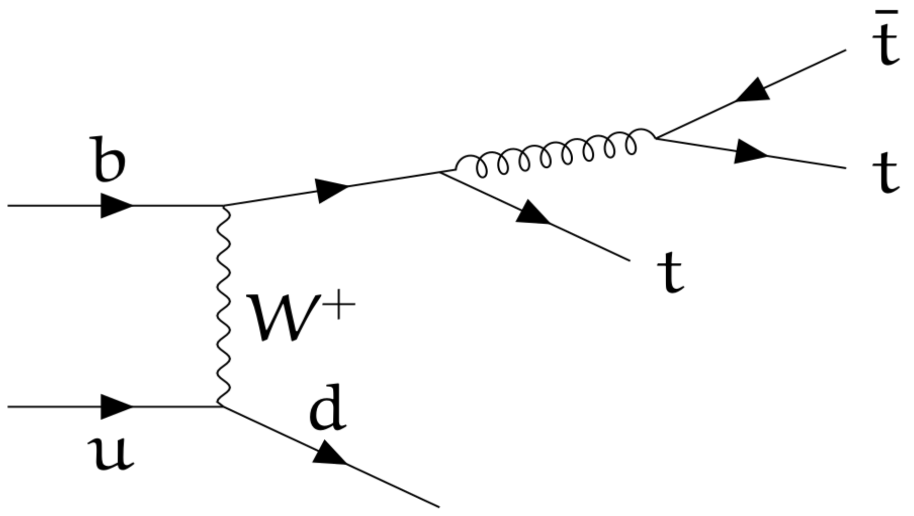
[2]



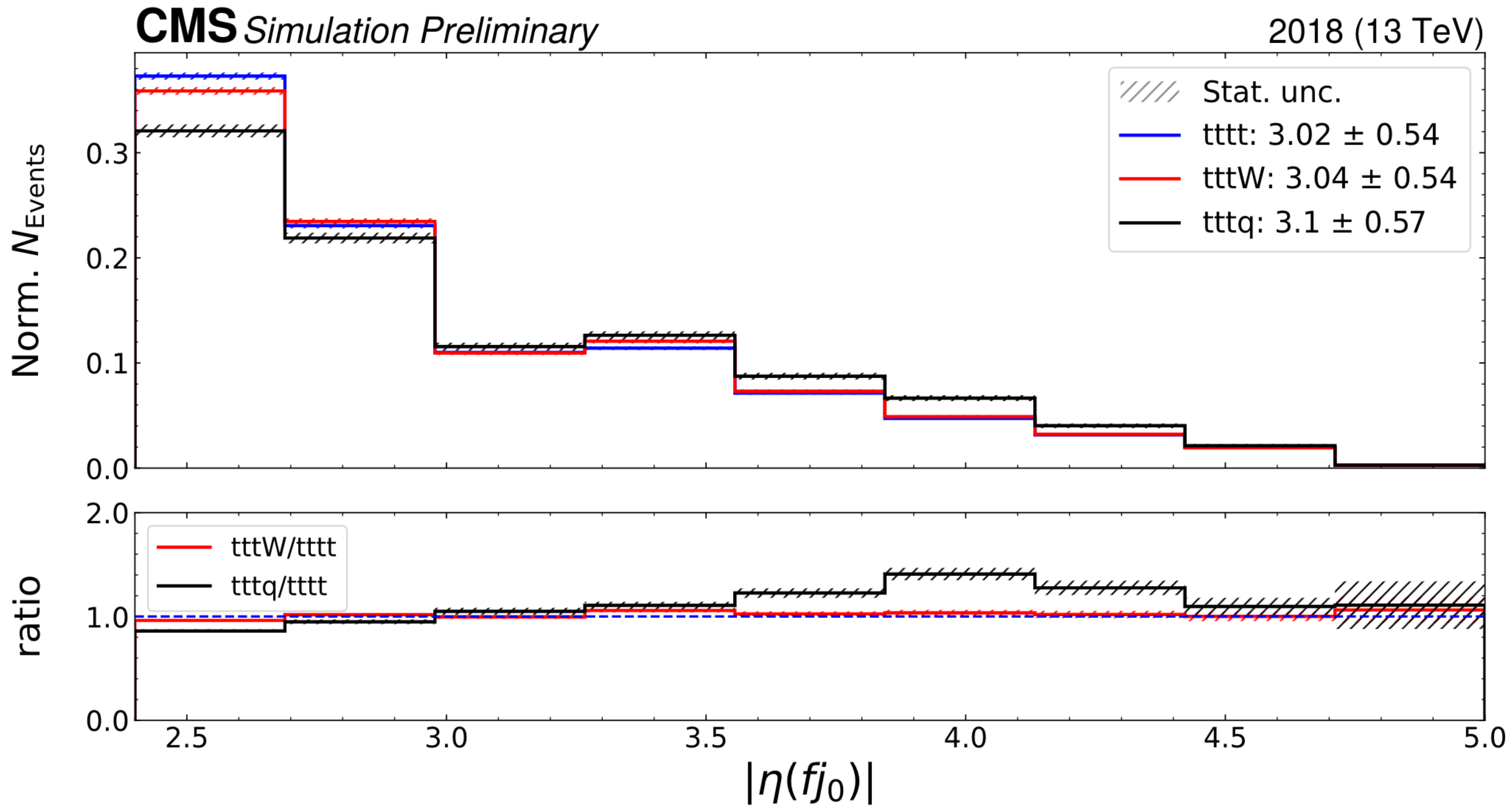


ttt vs tttt analysis

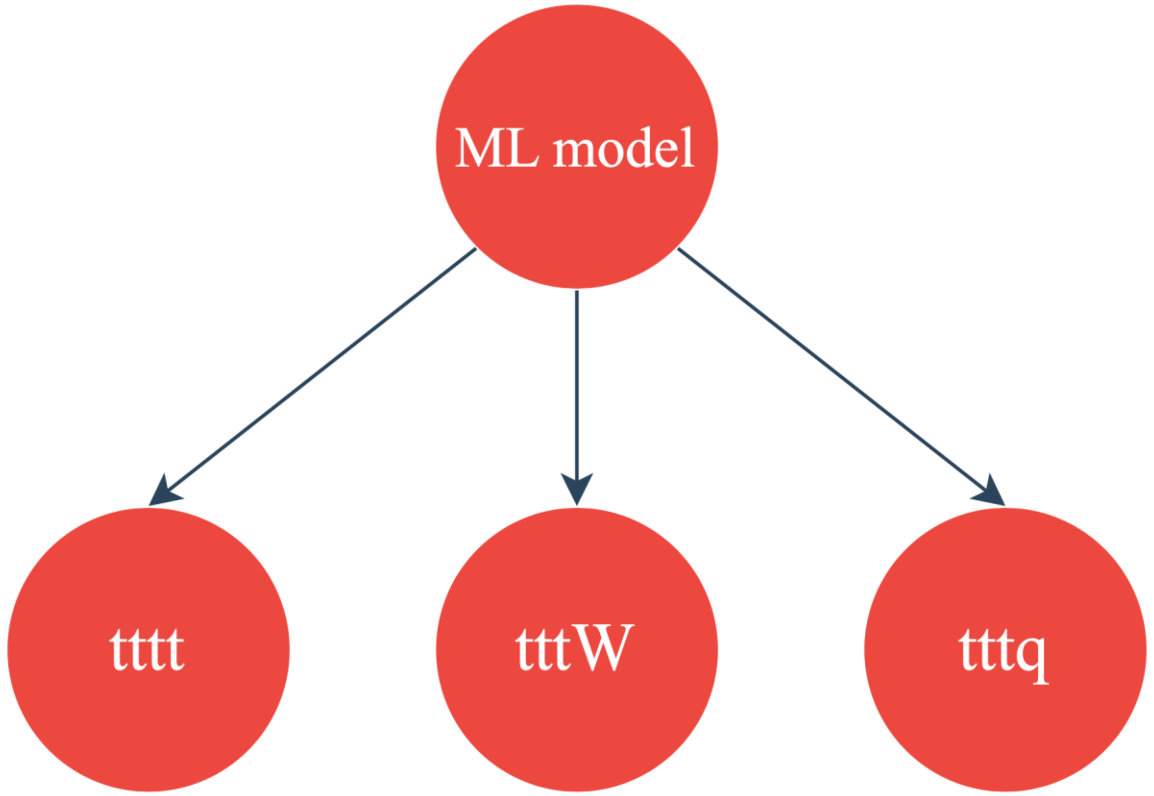
- Similar diagrams present in tttq production
  But: More jets are present.
- Indeed: Subtle effects are visible!



SR-2l



Machine Learning



- Construction of binary and triple class ML models for SR-2l and SR-3l.
- Extreme Gradient Boosting Decision Trees (XGBoost).
- ML procedure:
  - Class imbalance → Introduction of weights .
  - Hypertuning of model parameters with full data.
  - Data splitting into:
    - Training set: Classification known **(80%)**.
    - Testing sets: Classification unknown **(20%)**.
  - Final model creation with hypertuned parameters and training set.
  - Model application on test set.

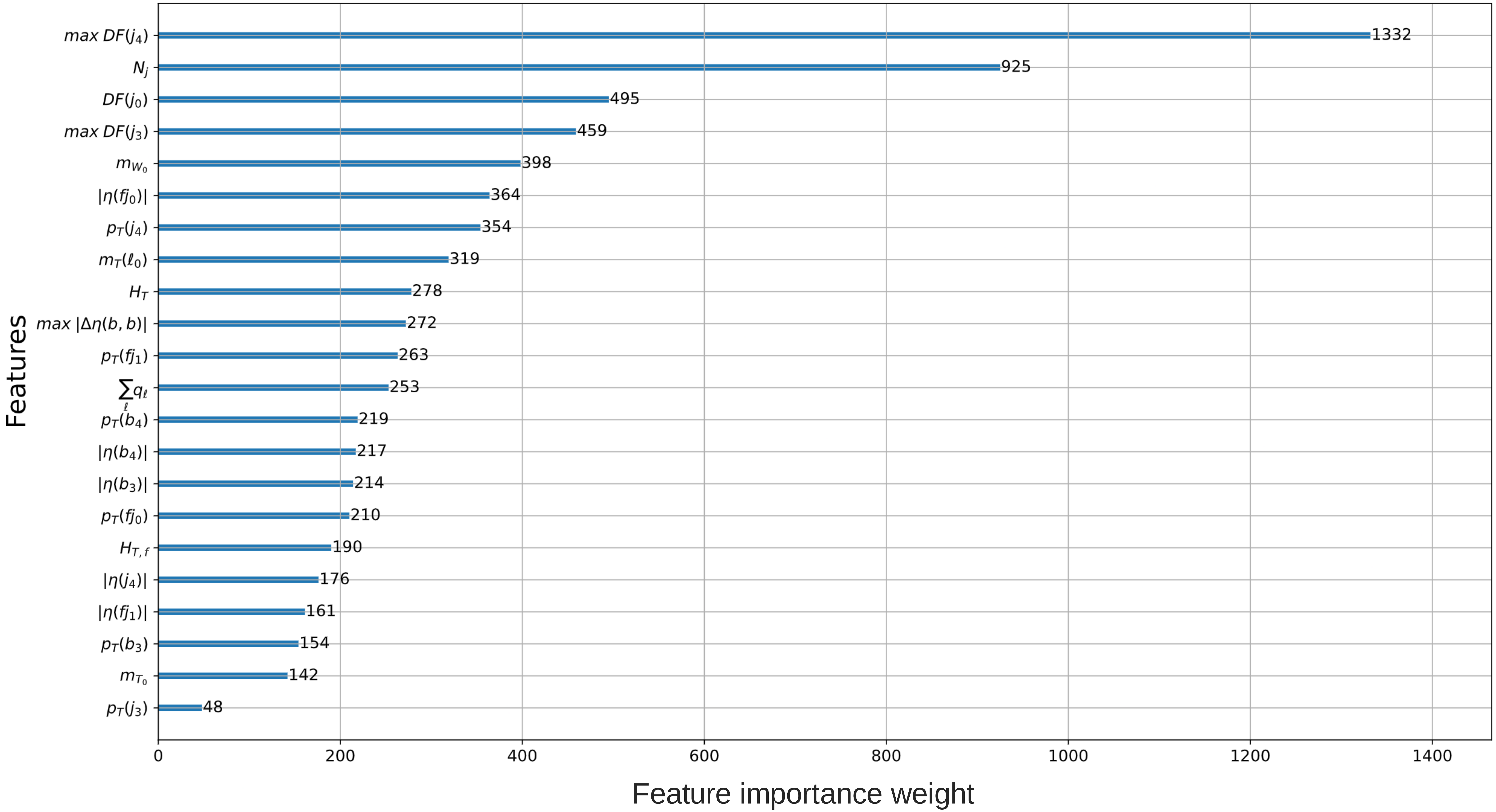
Feature importance

- Feature sets obtained for the four ML models.
- Consideration of four importance types.
- Forward jet variables are also utilized.

Binary		**Triple**	
SR-2ℓ	**SR-3ℓ**	**SR-2ℓ**	**SR-3ℓ**
N_j	N_j	N_j	N_j
H_{T}	H_{T}	H_{T}	H_{T}
$\sum_\ell q_\ell$	$\sum_\ell q_\ell$	$\sum_\ell q_\ell$	$\sum_\ell q_\ell$
$\lvert\eta(\mathrm{fj}_0)\rvert$	$\lvert\eta(\mathrm{fj}_0)\rvert$	$\lvert\eta(\mathrm{fj}_0)\rvert$	$p_{\mathrm{T}}(\mathrm{fj}_0)$
m_{T_0}	$\min\ \Delta R(b,b)$	m_{T_0}	m_{T_0}
m_{W_0}	$\max\ \lvert\Delta\eta(b,b)\rvert$	m_{W_0}	m_{W_0}
$\mathrm{DF}(j_0)$	$p_{\mathrm{T}}(j_2)$	$m_{\mathrm{T}}(\ell_0)$	$\lvert\Delta\eta(nb_0,b_0)\rvert$
$p_{\mathrm{T}}(j_4)$	$\lvert\Delta\phi(nb_0,b_0)\rvert$	$p_{\mathrm{T}}(\mathrm{fj}_1)$	$\max\ \mathrm{DF}(j_3)$
$\max\ \lvert\Delta\eta(b,b)\rvert$	$\min\ \Delta R(j,j)$	L_{T}	$\lvert\eta(j_4)\rvert$
...	...	...	...

GHENT UNIVERSITY

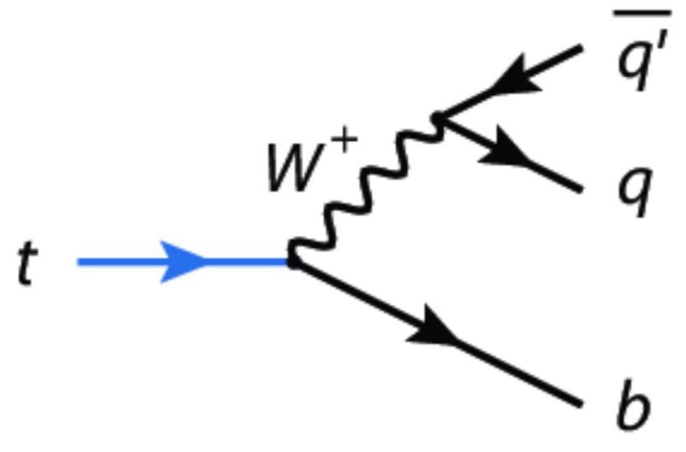
Binary SR-2l



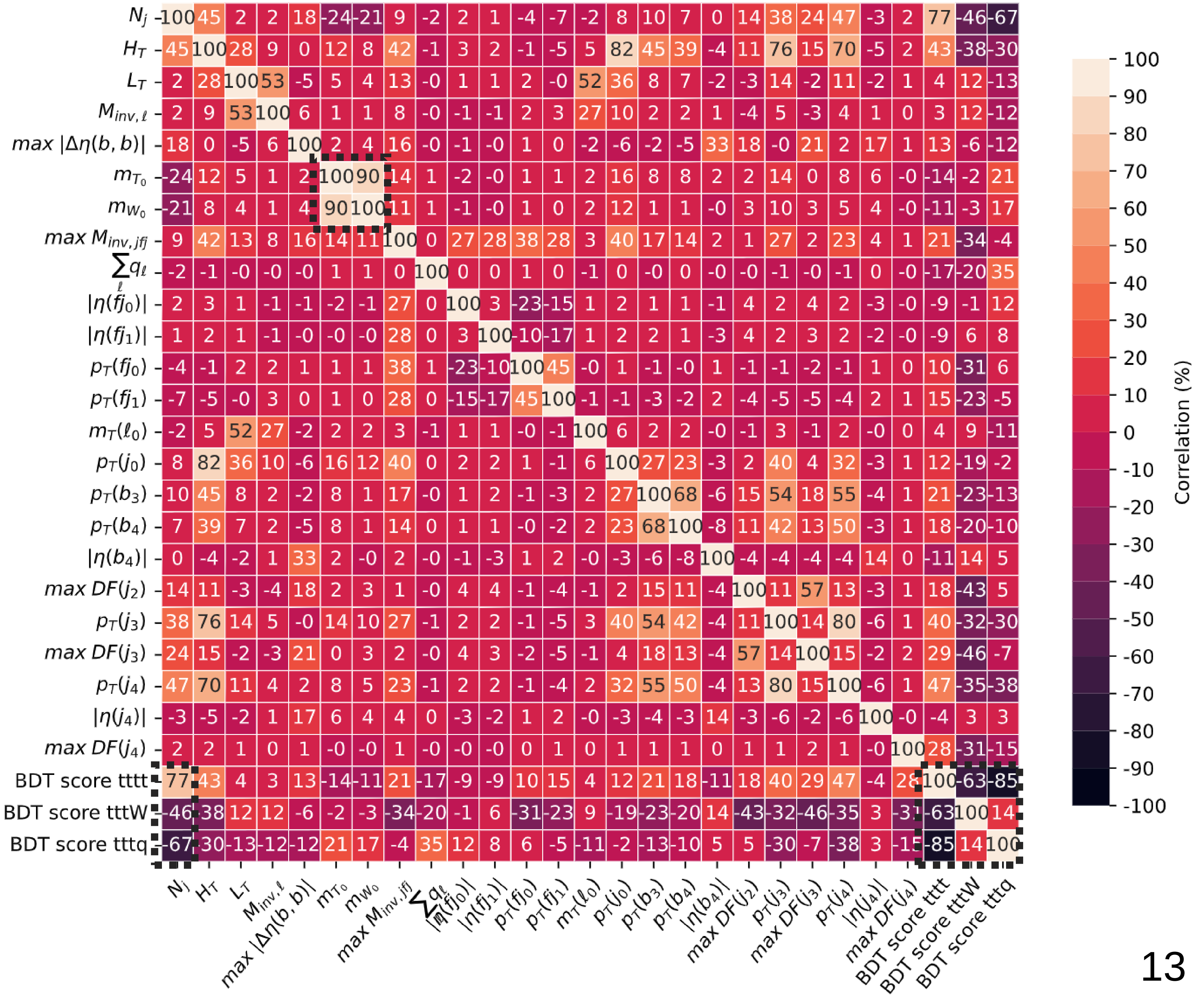
max DF(j4)
Nj
DF(j0)
max DF(j3)
mW0
|η(fj0)|
pT(j4)
mT(ℓ0)
HT
max |Δη(b, b)|
pT(fj1)
∑ qℓ
pT(b4)
|η(b4)|
|η(b3)|
pT(fj0)
HT,f
|η(j4)|
|η(fj1)|
pT(b3)
mT0
pT(j3)
1332
925
495
459
398
364
354
319
278
272
263
253
219
217
214
210
190
176
161
154
142
48
0
200
400
600
800
1000
1200
1400
Features
Feature importance weight

Correlation matrix

- Correlation between features is possible (e.g. m_{T0} & m_{W0}).
- Correlation between features & BDT scores resembles patterns observed on reconstruction level.
- Correlation between BDT scores is also visible.



Triple SR-2l

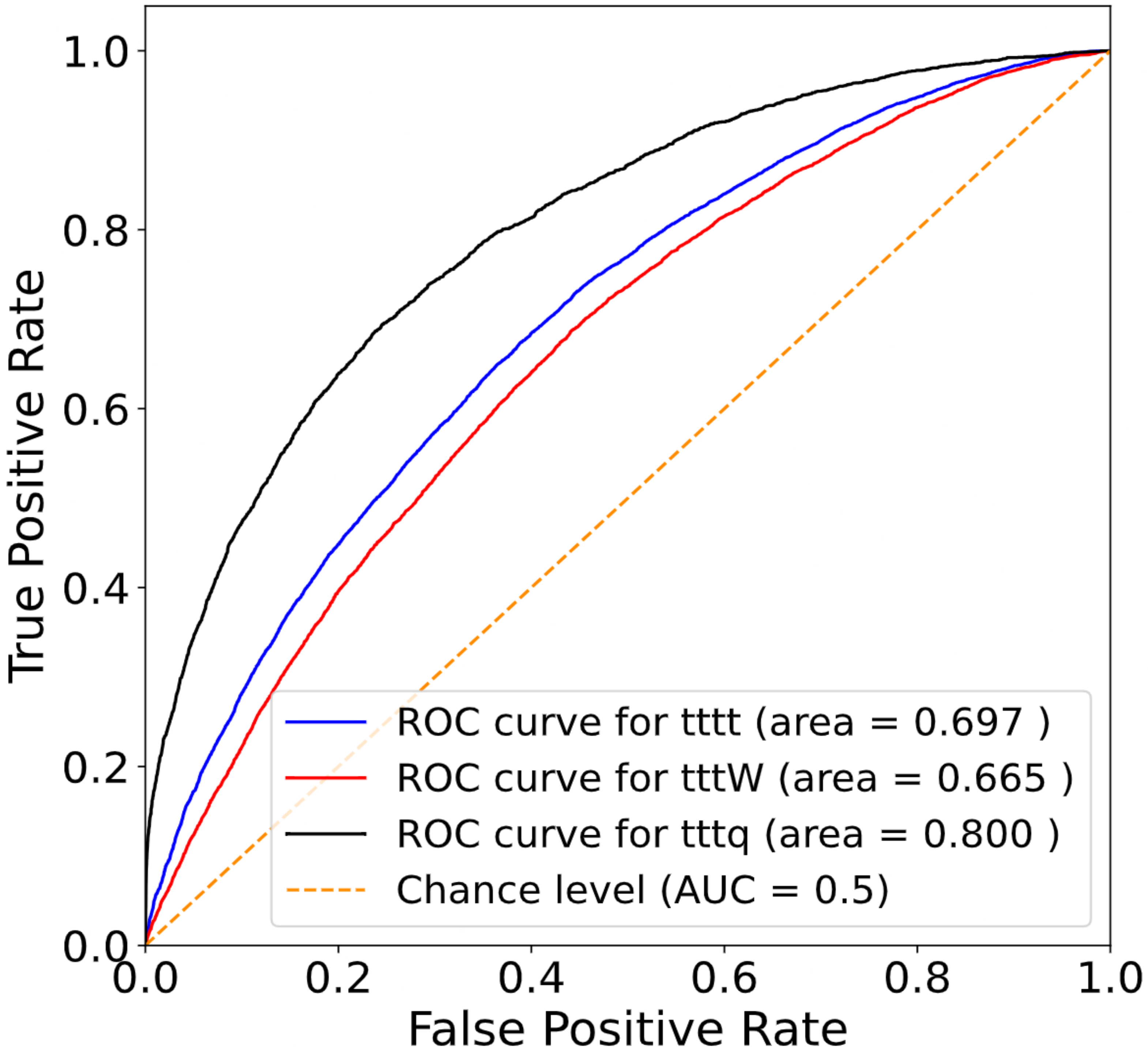


GHENT UNIVERSITY

ROC curves

- Binary ROC AUC values:
  - SR-2l: 71.4%
  - SR-3l: 68.7%
- Best identification for tttq events, followed by tttt.
- tttW highly confused with other processes.

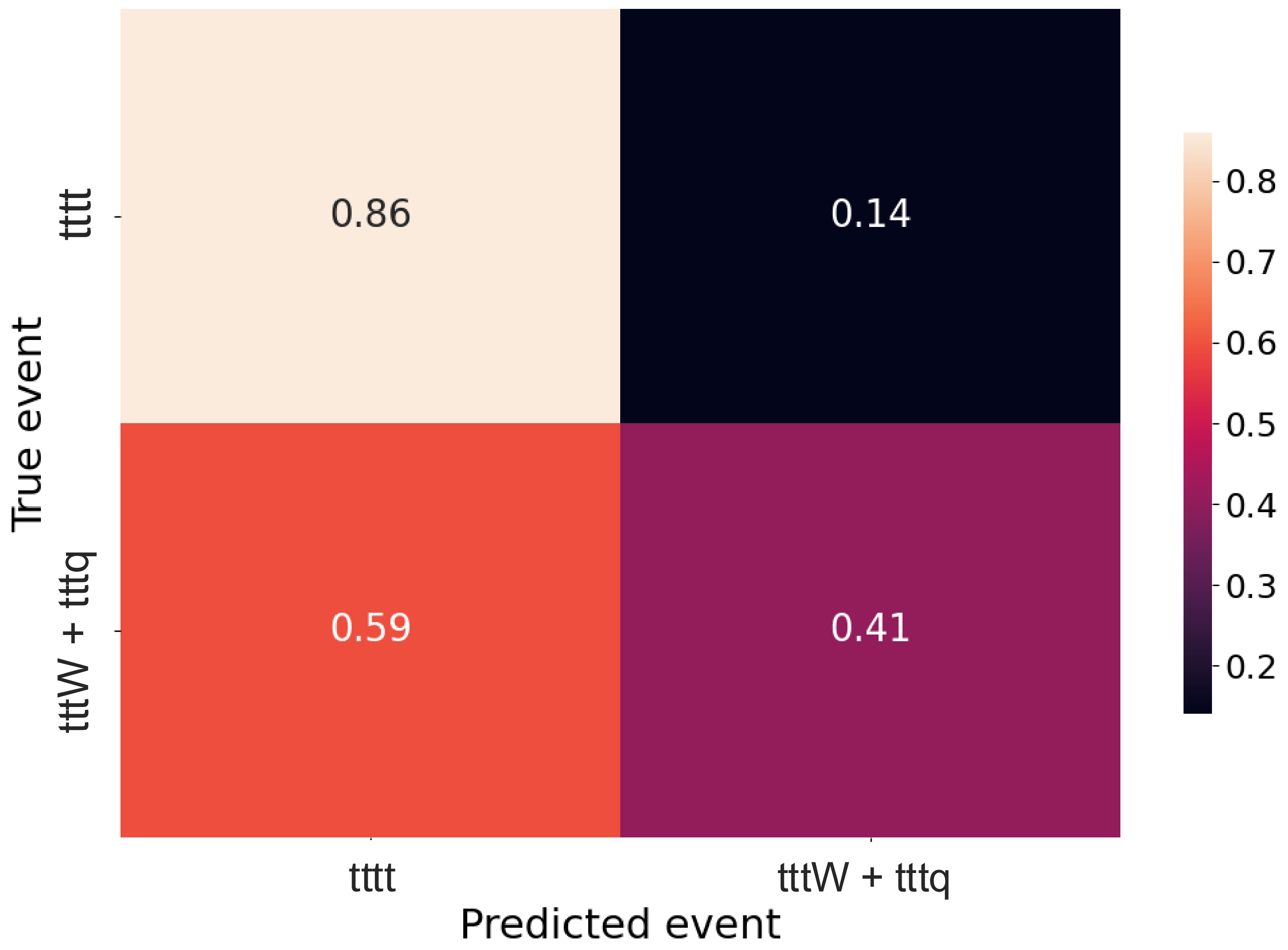
Triple SR-2l



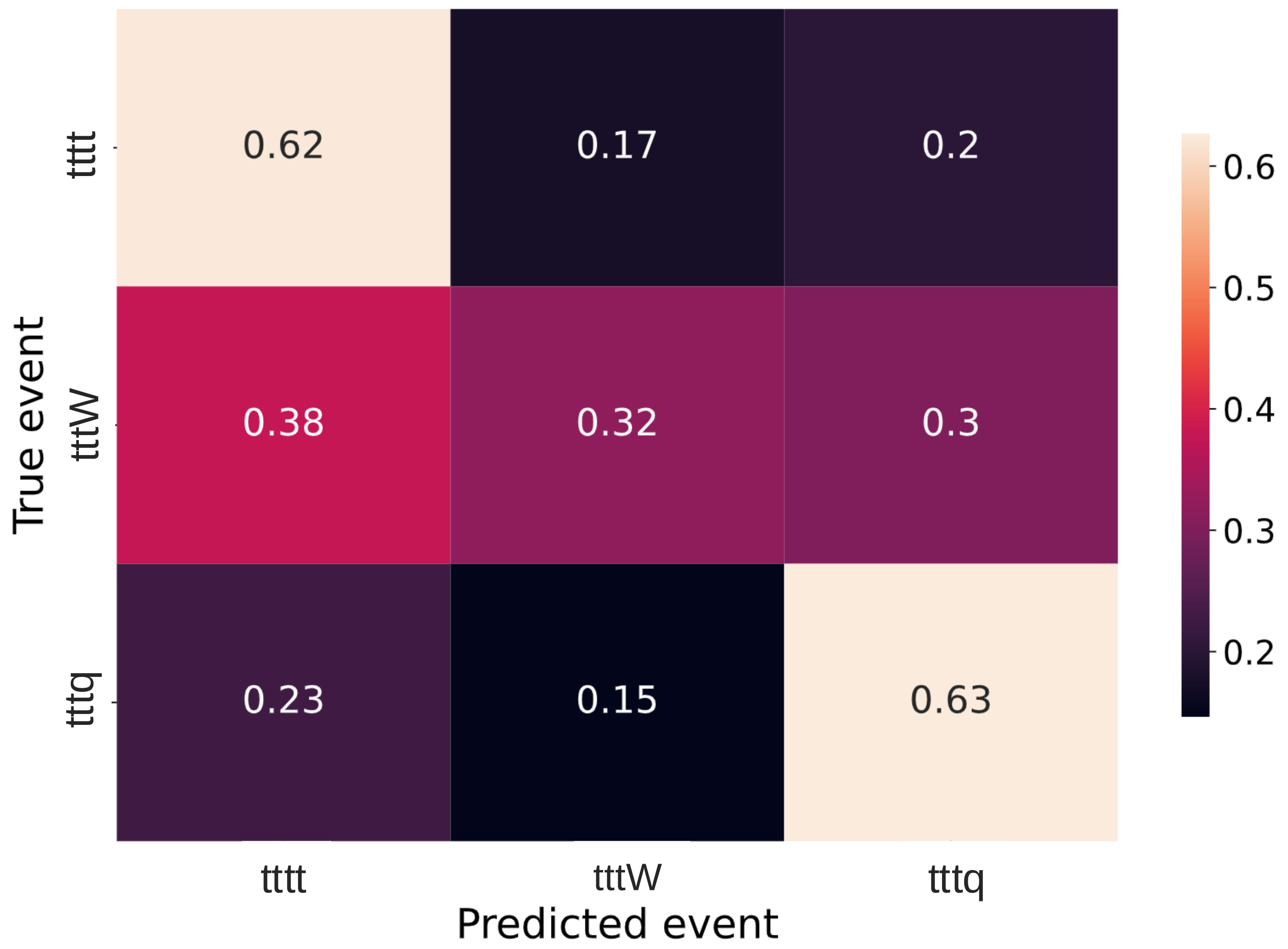
GHENT UNIVERSITY

Confusion matrices

Binary SR-2l



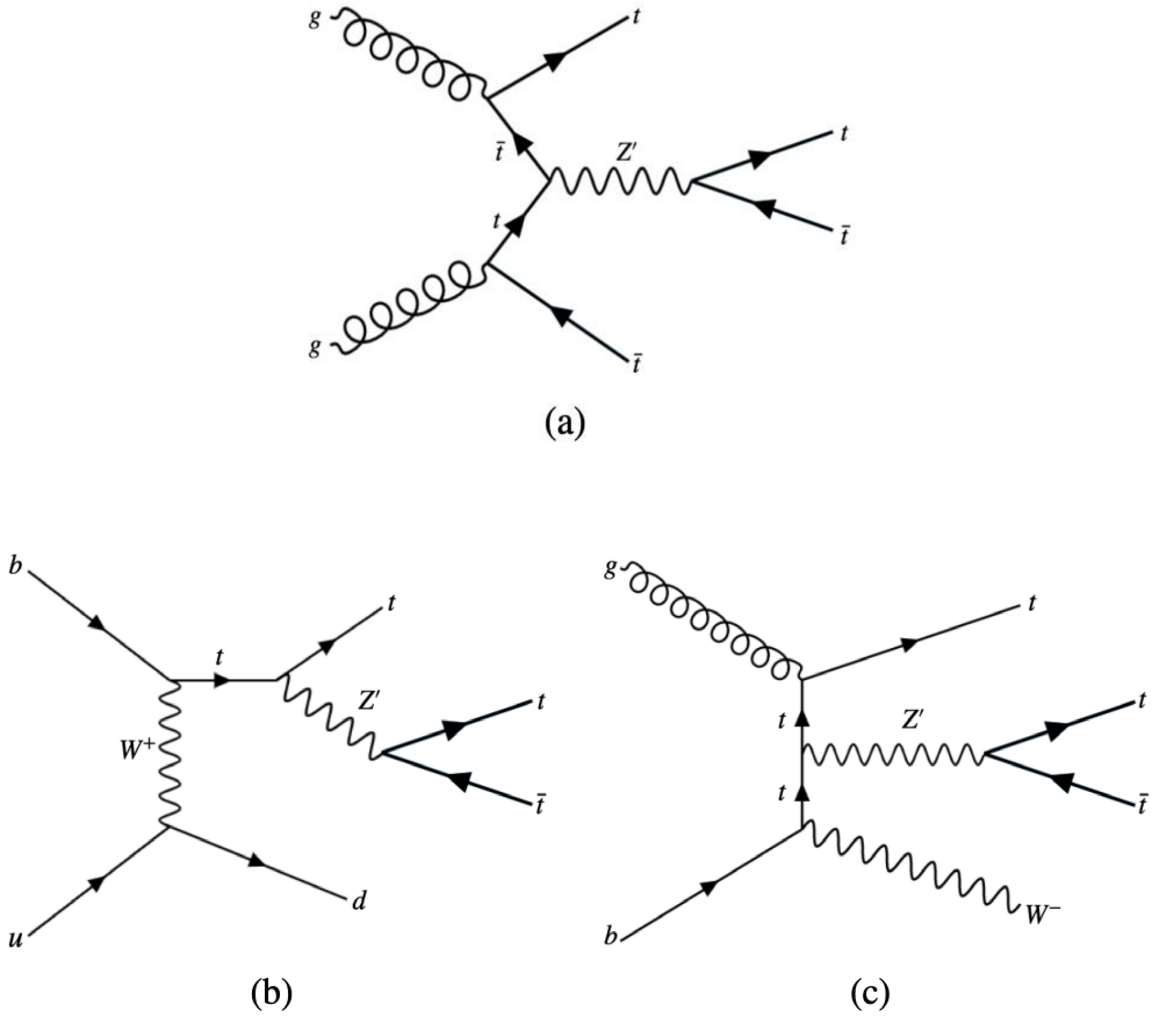
Triple SR-2l



GHENT UNIVERSITY

Top-philic heavy resonances

- Consideration of additional spin-0 and spin-1 states, either color-singlets or color-octets.
- New particles also present in 2HDM.
- Large top quark coupling, other couplings are negligible.
- Ongoing active searches.
- In this thesis: Scalar singlet boson with m = 600 GeV.

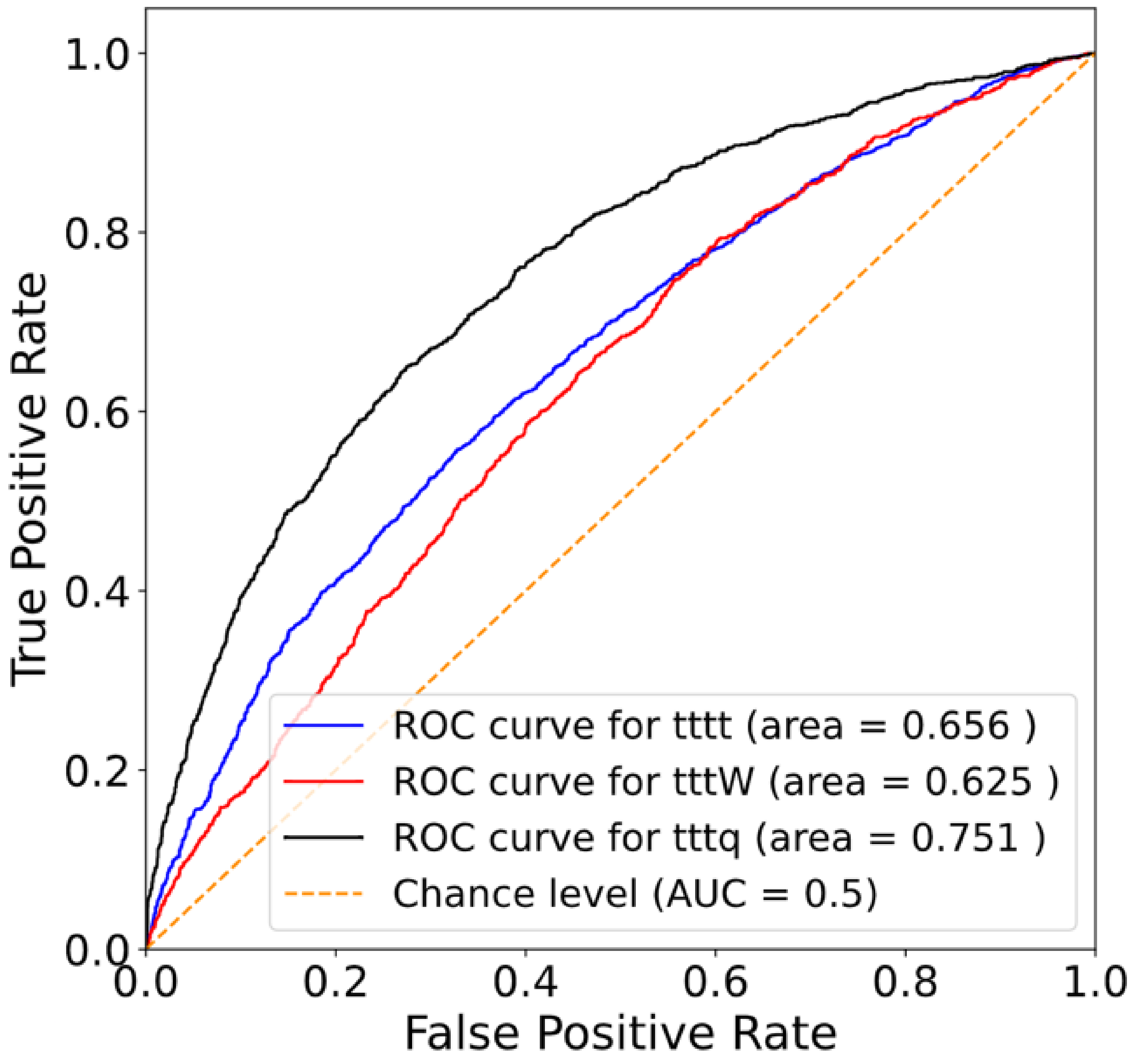


Application on BSM data

- Binary ROC AUC values:
  - SR-2l: 67.6%
  - SR-3l: 71.2%

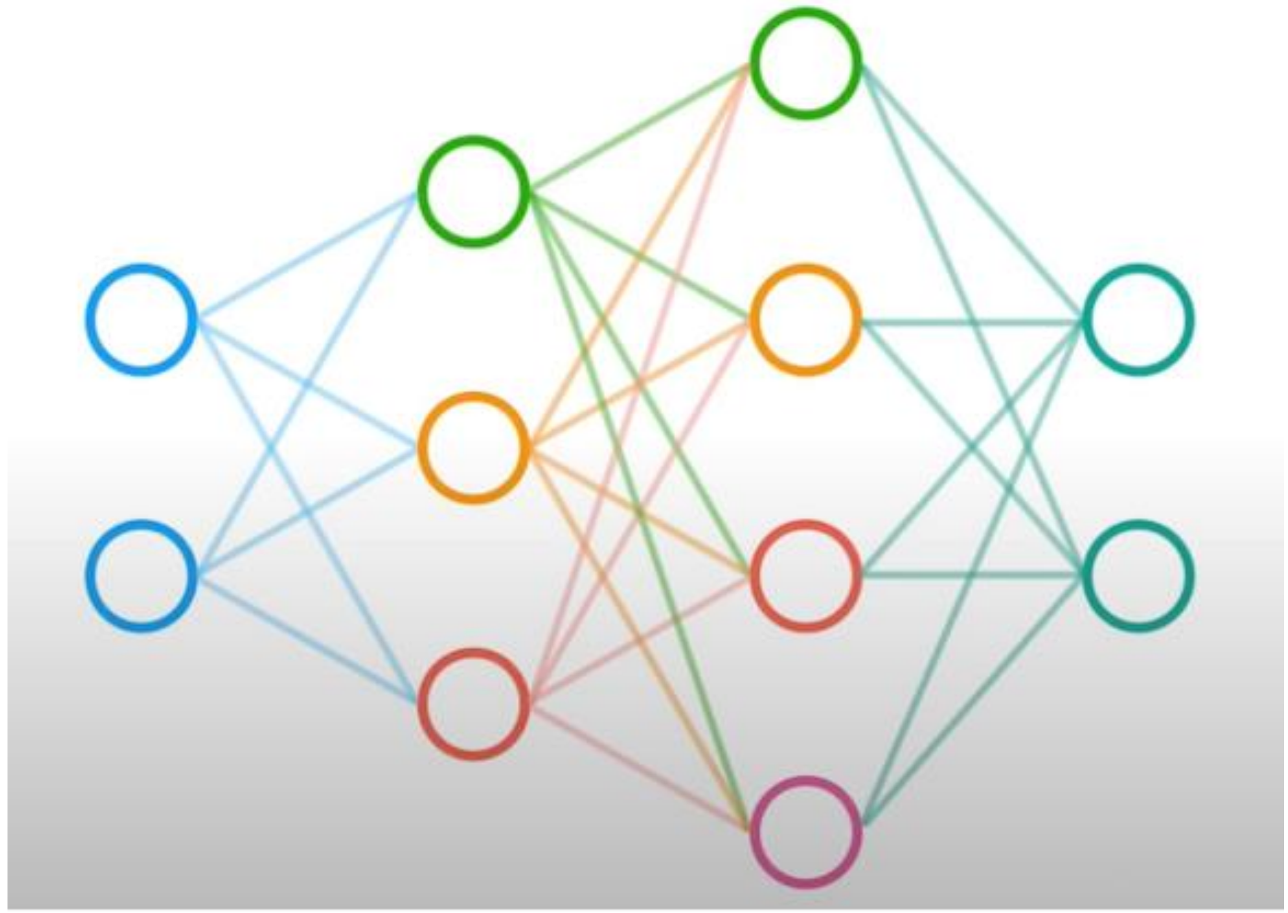
- Best identification for tttq events, followed by tttt.
- tttW highly confused with other processes.

Triple SR-2l



GHENT UNIVERSITY

Outlook



- Fine-tuning of selection criteria for ttt production.
- Introduction of background processes.
- Advanced W boson and top quark reconstruction.
- Variables related to spin correlations.
- Implementation of graph neural networks.
- Application of PCA & LDA.

GHENT UNIVERSITY

Any questions?

GHENT UNIVERSITY

FACULTY OF SCIENCES

Karam Kaspar

Master student

EXPERIMENTAL PARTICLE PHYSICS

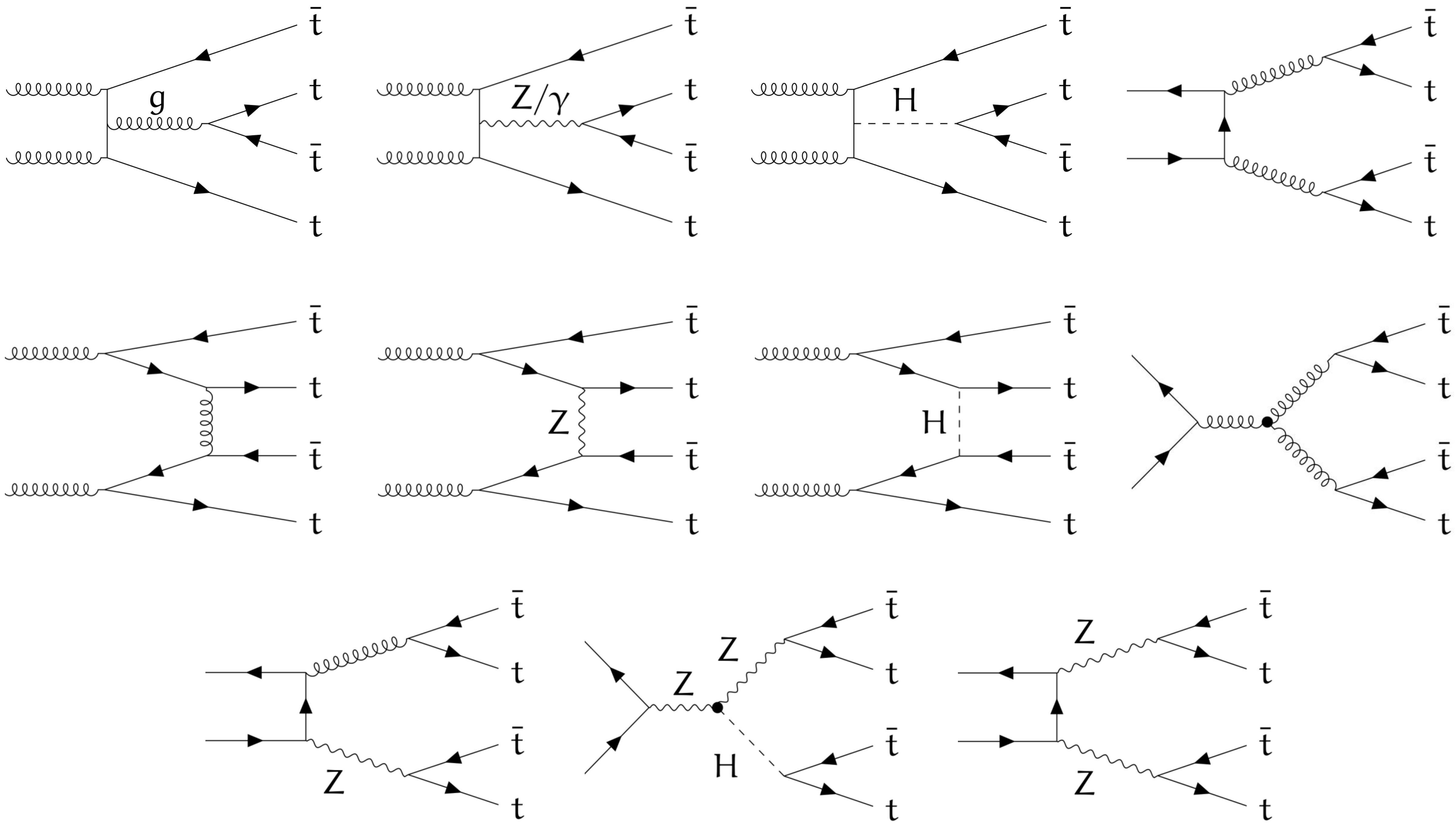
www.ugent.be

E-mail: karam.kaspar@ugent.be

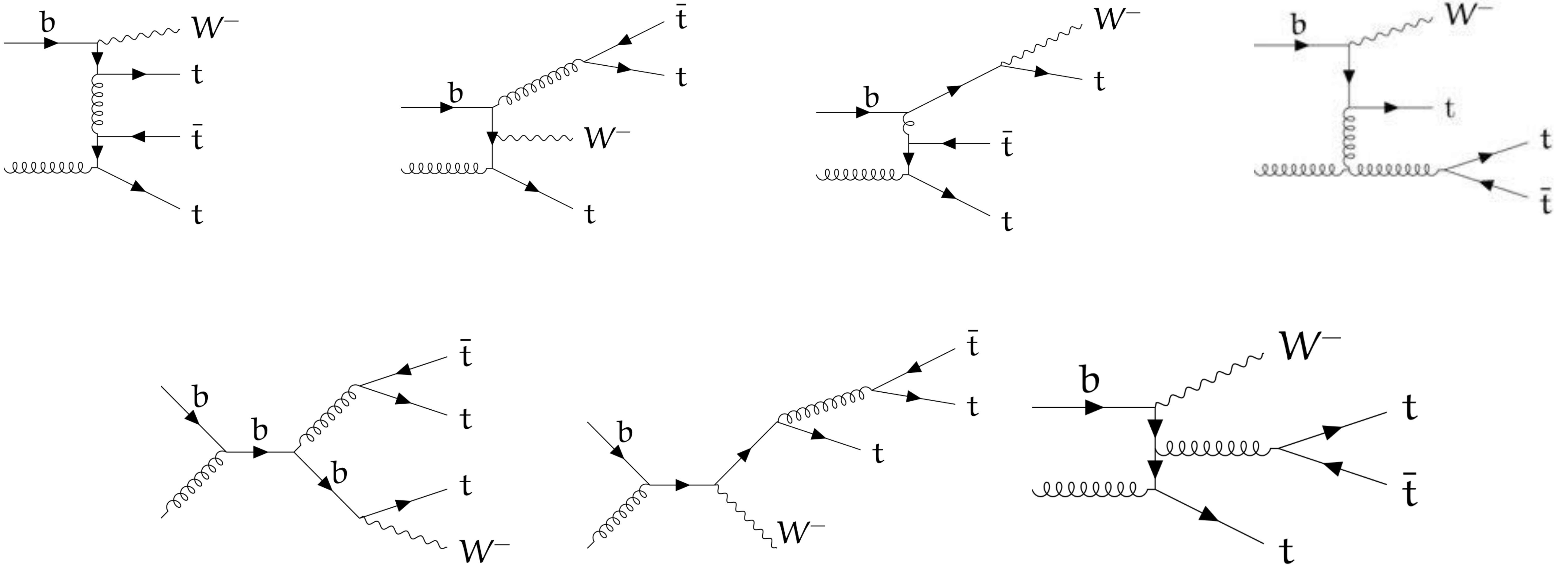
GHENT UNIVERSITY

Appendix

Feynman diagrams: tttt

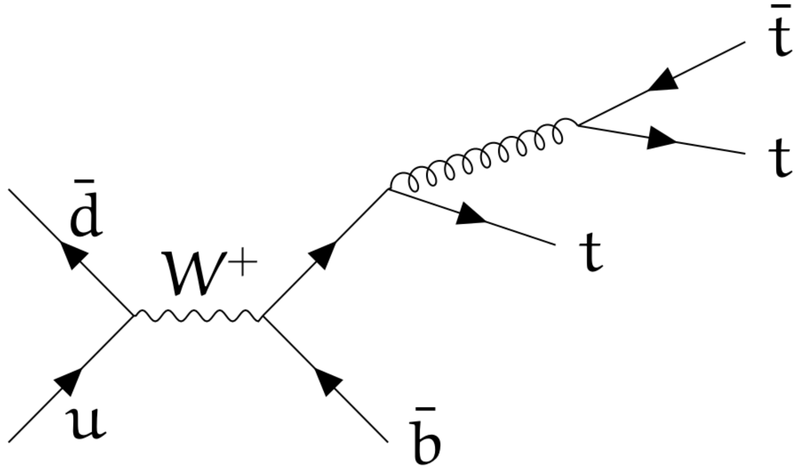


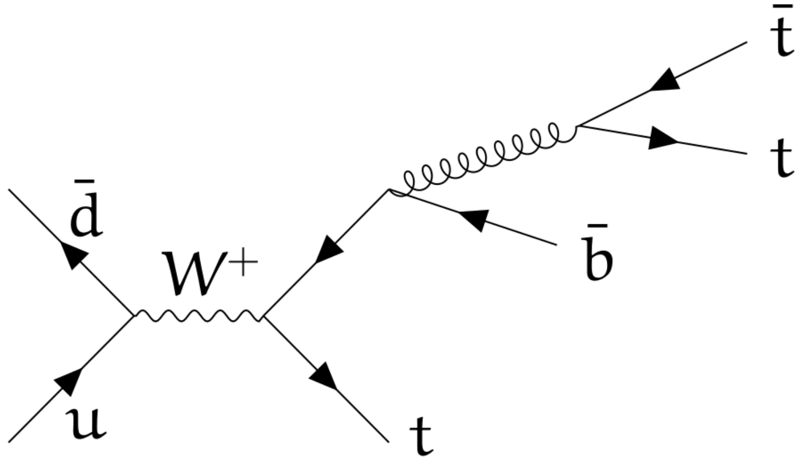
Feynman diagrams: tttW

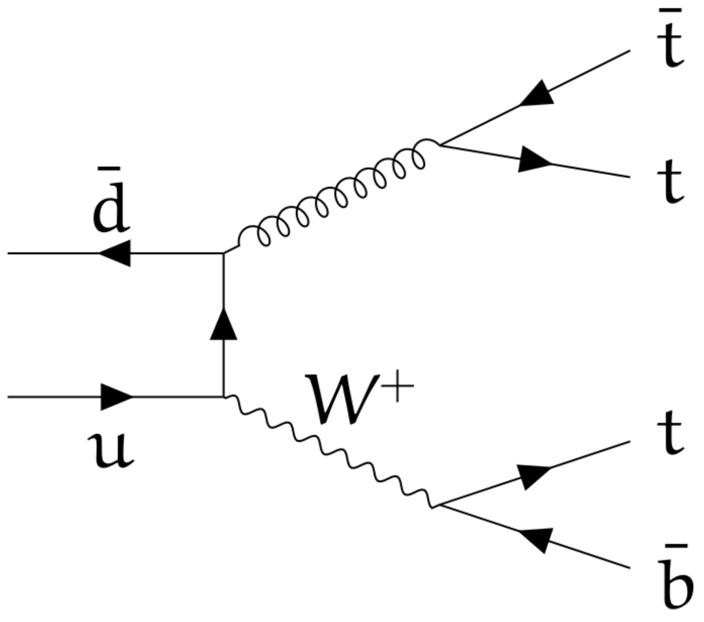


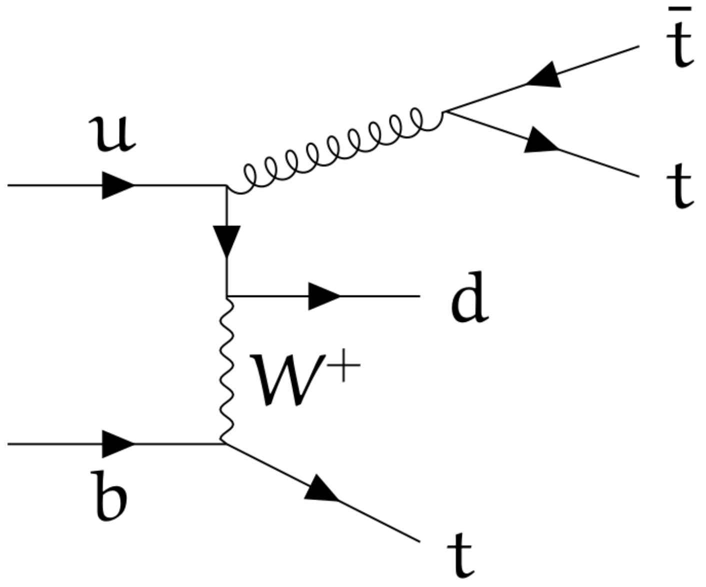
GHENT UNIVERSITY

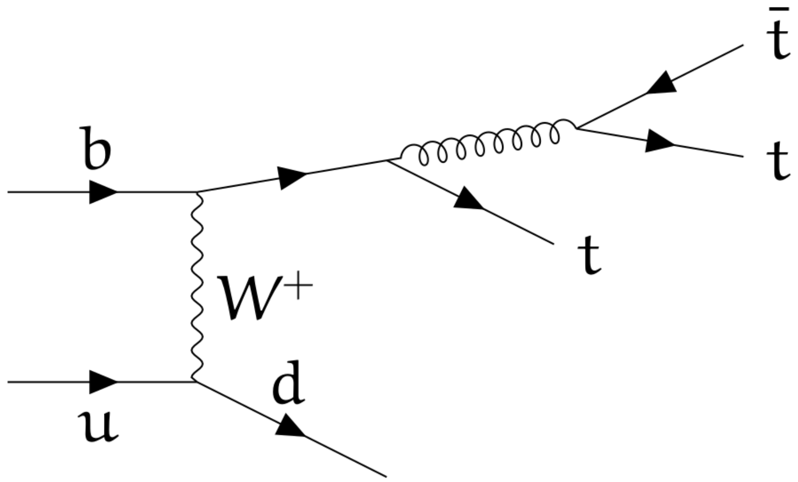
Feynman diagrams: tttq

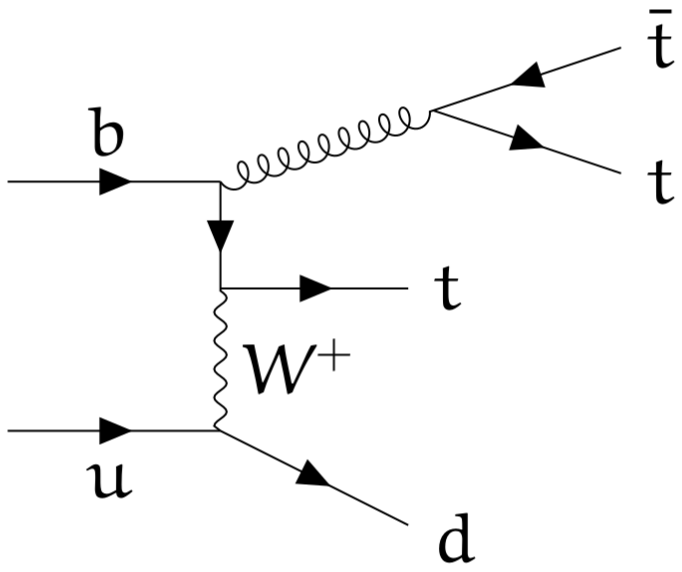


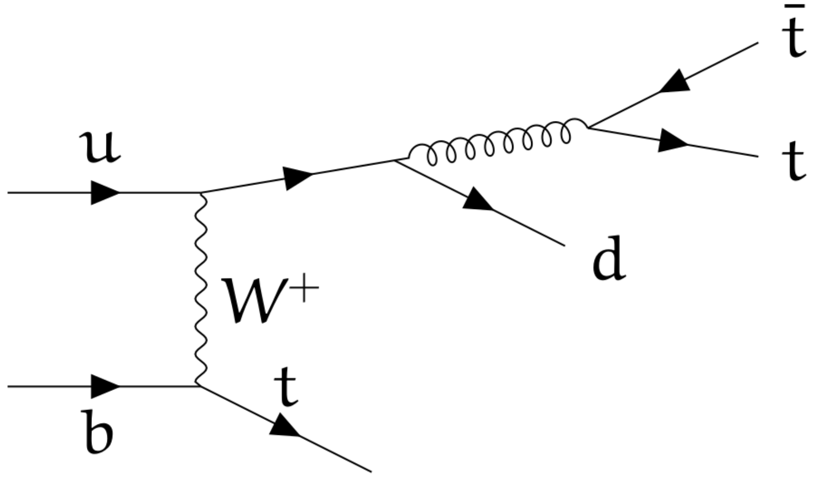






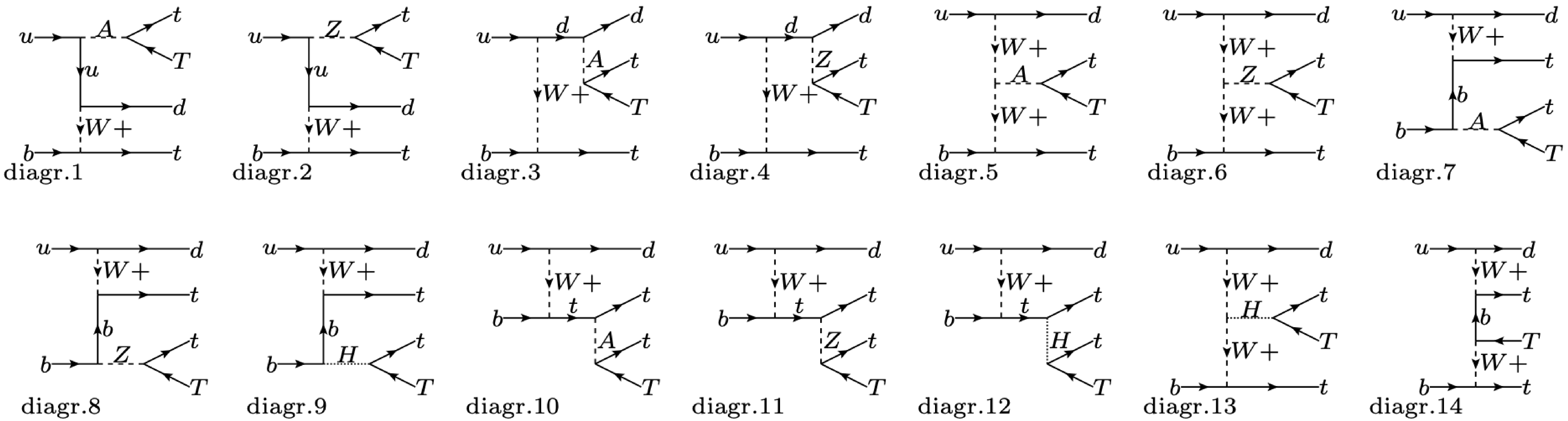






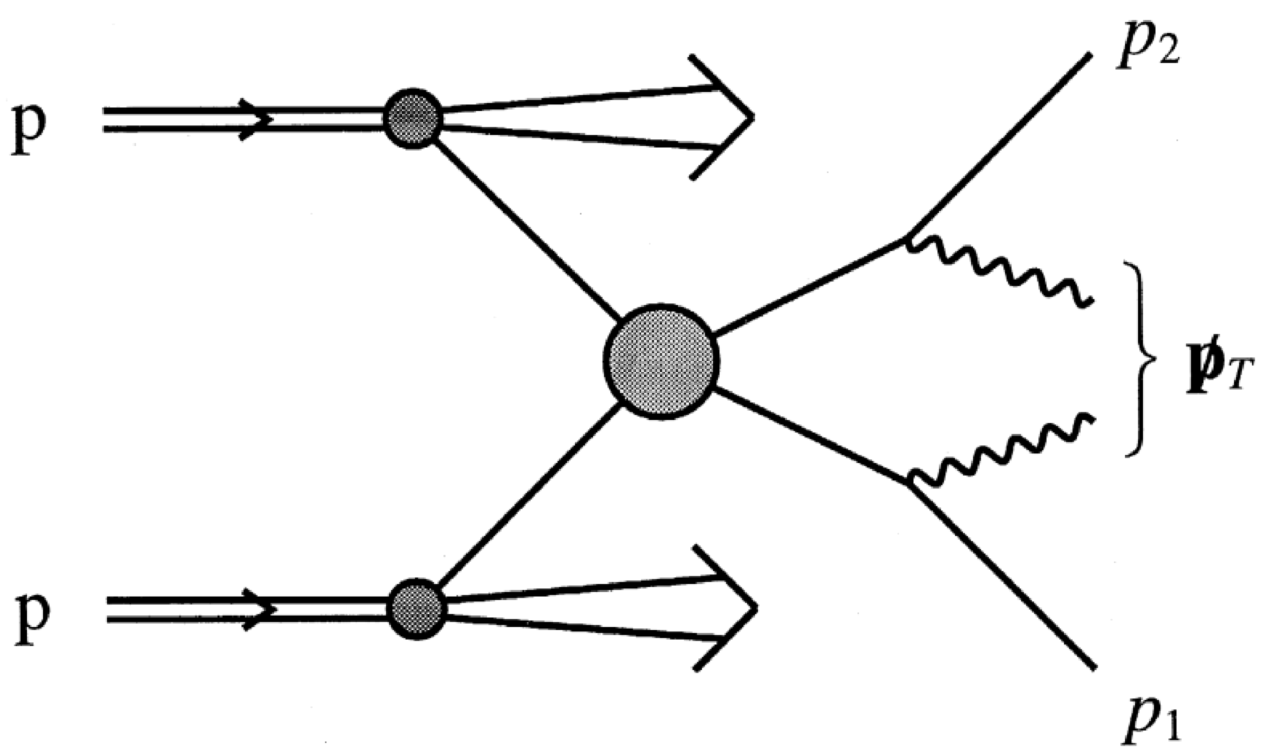
GHENT
UNIVERSITY

Feynman diagrams: tttq



GHENT UNIVERSITY

Advanced reconstruction



- Transverse mass variable:

$$m_{\mathrm{T}}^2\left(\vec{p}_{\mathrm{T}}(\ell), \vec{p}_{\mathrm{T}}^{\mathrm{miss}}\right) = 2\vec{p}_{\mathrm{T}}(\ell) p_{\mathrm{T}}^{\mathrm{miss}}\left[1 - \cos(\Delta\phi)\right] \leq m_W^2$$

- But p_T^{miss} contains all neutrino contributions.
- Consider:

$$pp \to X + \tilde{l}^+ \tilde{l}^- \to X + l^+ l^- \tilde{\chi}_1^0 \tilde{\chi}_1^0.$$

$$\rightarrow \quad m_{\tilde{l}}^2 = m_l^2 + m_{\tilde{\chi}}^2 + 2\left(E_{Tl} E_{T\tilde{\chi}} \cosh(\Delta\eta) - \boldsymbol{p}_{Tl} \cdot \boldsymbol{p}_{T\tilde{\chi}}\right)$$

- Additional variable defined as:

$$m_{\tilde{l}}^2 \geq M_{T2}^2$$
$$\equiv \min_{\not{\boldsymbol{p}}_1 + \not{\boldsymbol{p}}_2 = \not{\boldsymbol{p}}_T}\left[\max\{m_T^2(\boldsymbol{p}_{Tl^-}, \not{\boldsymbol{p}}_1), m_T^2(\boldsymbol{p}_{Tl^+}, \not{\boldsymbol{p}}_2)\}\right]$$

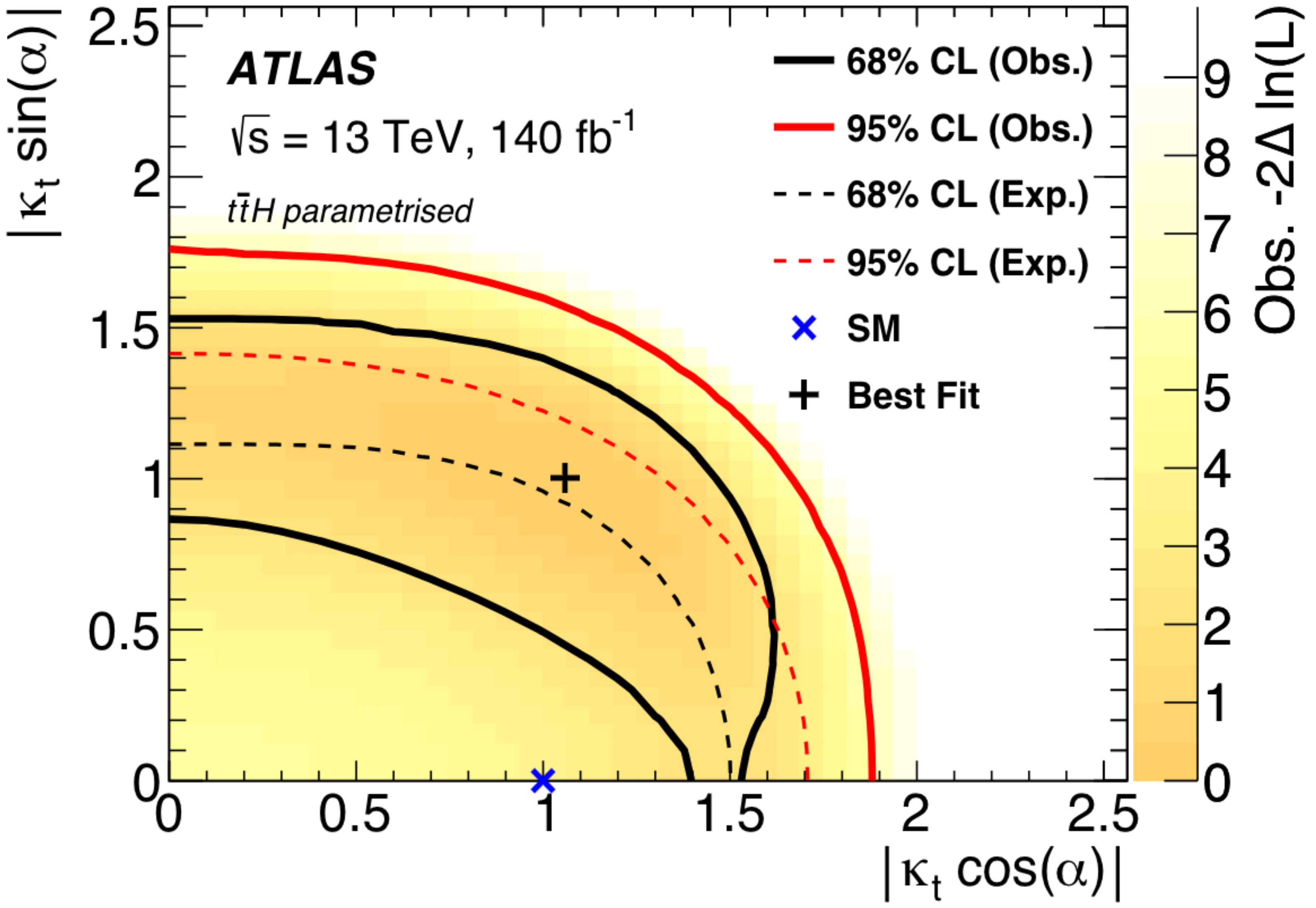
- Application to W boson and top quark decays.

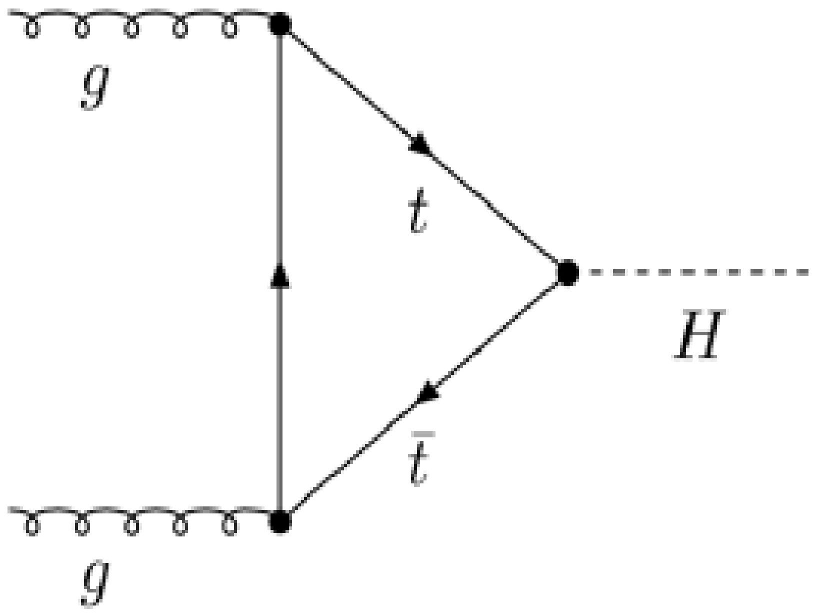
Top quark Yukawa coupling

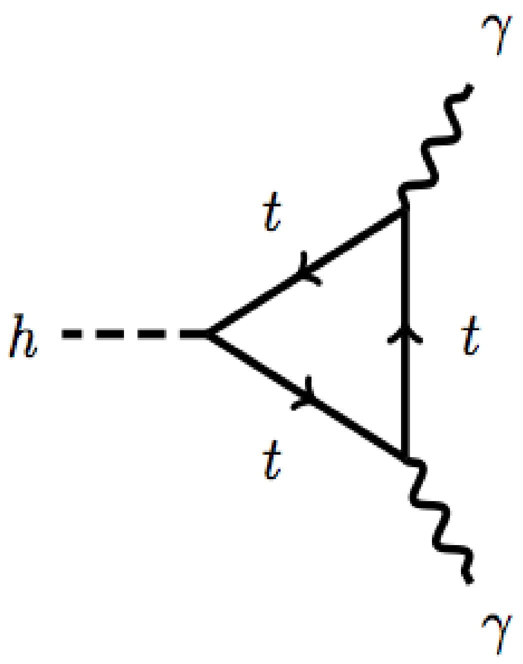
- y_t coupling measurement via several processes

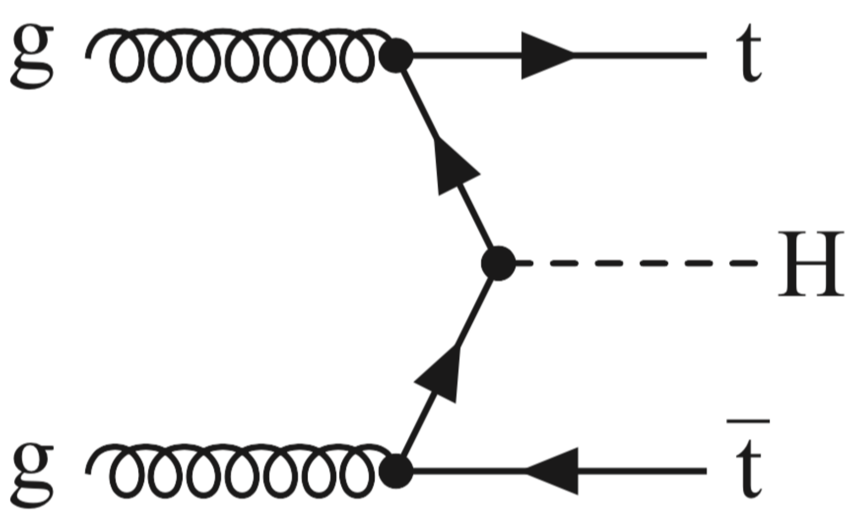
  But: Contamination with the Higgs width

- tttt analysis is useful:
  - tttt cross section ~ y_t^4
  - No dependence of decay width





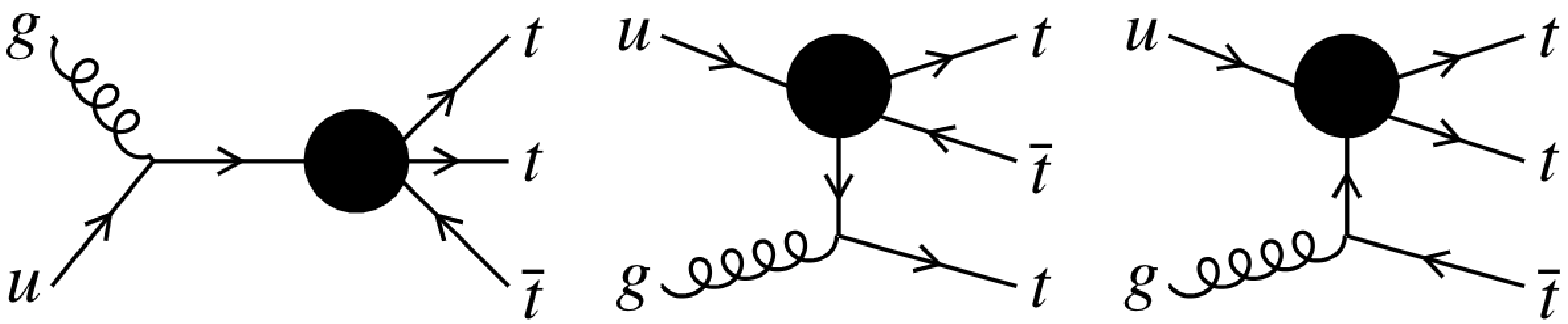




GHENT UNIVERSITY

FVNI & FCNI in ttt production

- ttt contains FVNI and FCNI.
- But: FVNI forbidden at tree level and suppressed in loop corrections by GIM mechanism.
- BSM theories might enhance FVNIs (e.g. minimal supersymmetry).
- Absence of t in proton → FVNI.
- Operator must connect to u-, t-quark on one side and to a top pair on the other side → FCNI.



GHENT UNIVERSITY

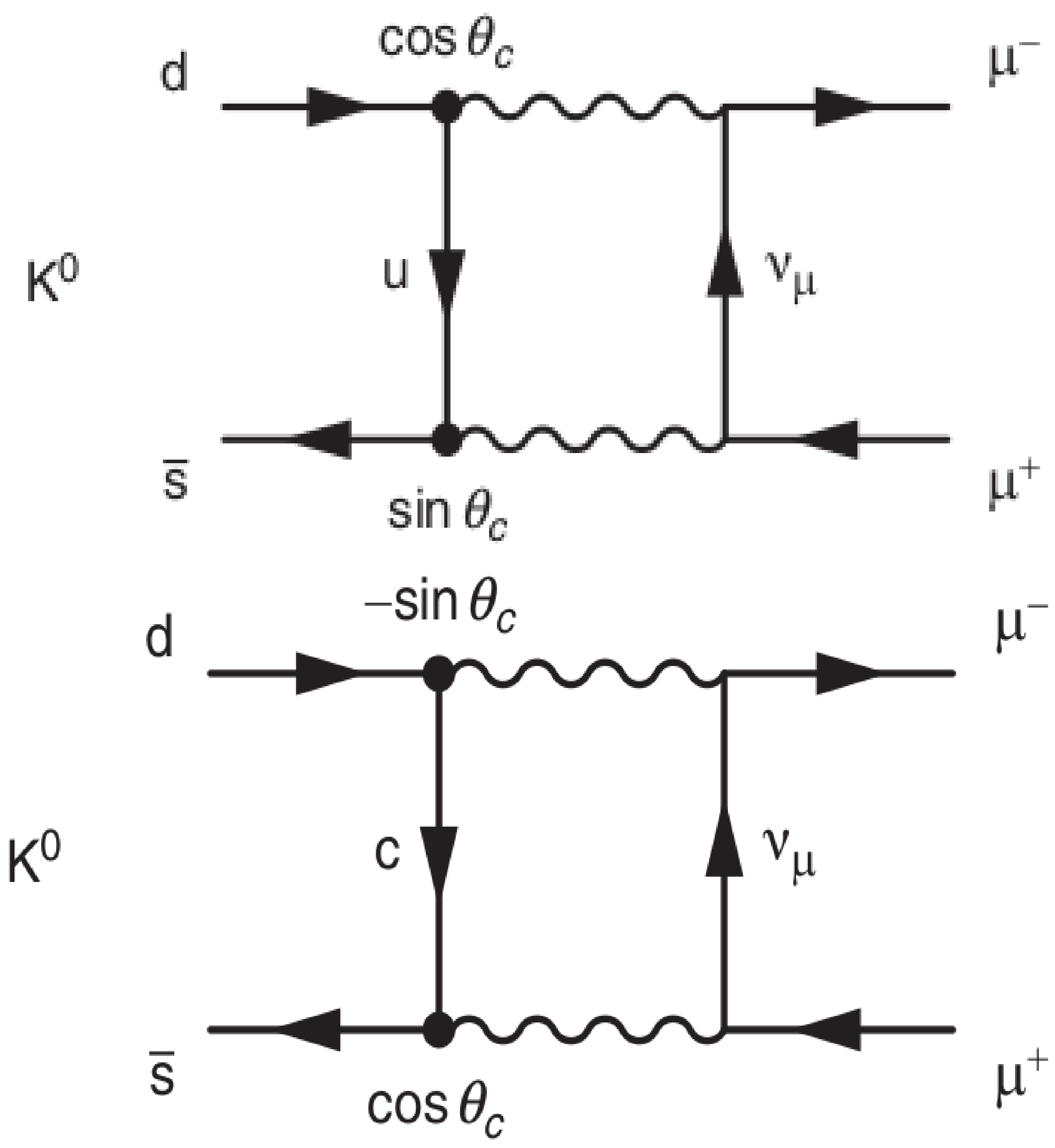
GIM mechanism

- Glashow–Iliopoulos–Maiani mechanism
- Cabbibo mechanism allows $K_L \rightarrow \mu^+ \mu^-$ with FVNI
- $BR_{measured}$ $(K_L \rightarrow \mu+ \mu-) = (6.84 \pm 0.11).10^{-9}$
- Proposal of charm quark:

$$\mathcal{M}_u \propto g_W^4 \cos\theta_c \sin\theta_c \quad \text{and} \quad \mathcal{M}_c \propto -g_W^4 \cos\theta_c \sin\theta_c$$

$$|\mathcal{M}|^2 = |\mathcal{M}_u + \mathcal{M}_c|^2 \approx 0$$

→ Suppression of FVNI interactions



GHENT UNIVERSITY

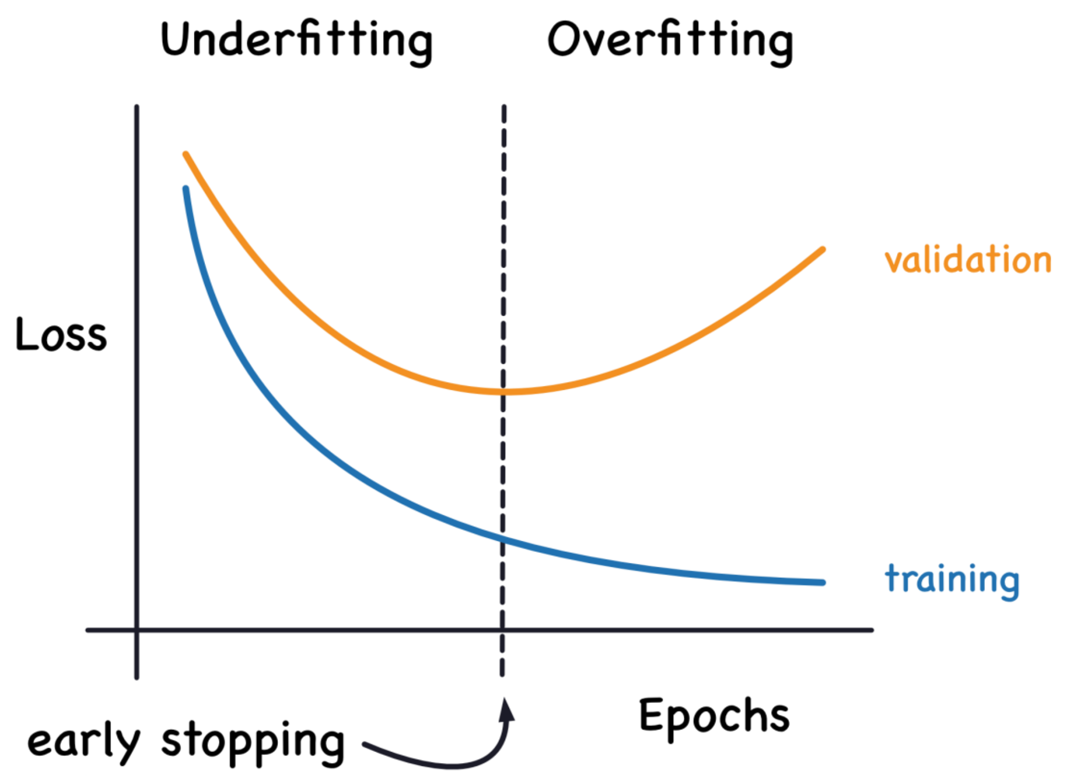
Feature sets

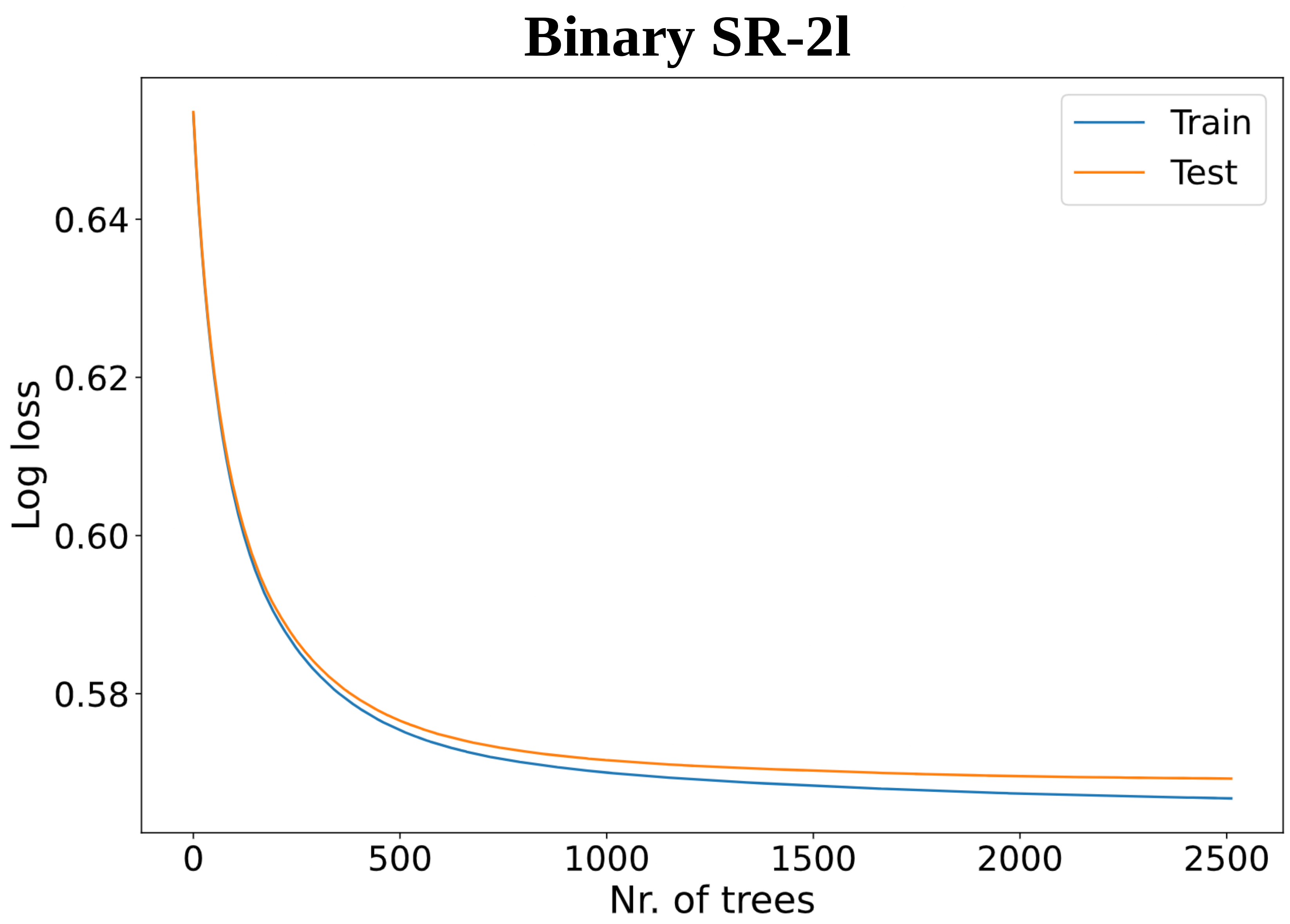
Binary		Triple	
SR-2ℓ	**SR-3ℓ**	**SR-2ℓ**	**SR-3ℓ**
N_j	N_j	N_j	N_j
H_T	H_T	H_T	H_T
max $\lvert\Delta\eta(b,b)\rvert$	min $\Delta R(j,j)$	L_T	$\lvert\eta(j_4)\rvert$
m_{T_0}	min $\Delta R(b,b)$	$M_{\mathrm{inv},\ell}$	$M_{\mathrm{inv},\ell}$
m_{W_0}	max $\lvert\Delta\eta(b,b)\rvert$	max $\lvert\Delta\eta(b,b)\rvert$	min $\Delta R(j,j)$
$\sum_\ell q_\ell$	$M_{\mathrm{inv,\,j\,fj}}$	m_{T_0}	m_{T_0}
$H_{T,f}$	$\sum_\ell q_\ell$	m_{W_0}	m_{W_0}
$\lvert\eta(\mathrm{fj}_0)\rvert$	$H_{T,f}$	max $M_{\mathrm{inv},j\,\mathrm{fj}}$	max $M_{\mathrm{inv},j,j}$
$\lvert\eta(\mathrm{fj}_1)\rvert$	$\lvert\eta(\mathrm{fj}_0)\rvert$	$\sum_\ell q_\ell$	$\sum_\ell q_\ell$
$p_\mathrm{T}(\mathrm{fj}_0)$	$\lvert\eta(\mathrm{fj}_1)\rvert$	$\lvert\eta(\mathrm{fj}_0)\rvert$	$H_{T,f}$
$p_\mathrm{T}(\mathrm{fj}_1)$	$p_\mathrm{T}(\mathrm{fj}_0)$	$\lvert\eta(\mathrm{fj}_1)\rvert$	$\lvert\eta(j_4)\rvert$
$\lvert\eta(j_4)\rvert$	$\Delta R(nb_0,b_0)$	$p_\mathrm{T}(\mathrm{fj}_0)$	$p_\mathrm{T}(\mathrm{fj}_0)$
$p_\mathrm{T}(j_4)$	$\lvert\Delta\phi(nb_0,b_0)\rvert$	$p_\mathrm{T}(\mathrm{fj}_1)$	max $\mathrm{DF}(j_3)$
$m_\mathrm{T}(\ell_0)$	$\lvert\eta(b_3)\rvert$	max $\mathrm{DF}(j_4)$	$\Delta R(nb_0,b_0)$
$\mathrm{DF}(j_0)$	$p_\mathrm{T}(j_2)$	$m_\mathrm{T}(\ell_0)$	$\lvert\Delta\eta(nb_0,b_0)\rvert$
max $\mathrm{DF}(j_4)$	$p_\mathrm{T}(j_3)$	$p_\mathrm{T}(j_0)$	$p_\mathrm{T}(b_3)$
$p_\mathrm{T}(b_3)$	$\lvert\eta j_3\rvert$	$p_\mathrm{T}(j_4)$	max $\mathrm{DF}(j_4)$
$\lvert\eta(b_3)\rvert$	$\mathrm{DF}(j_3)$	$\lvert\eta(j_4)\rvert$	$\mathrm{DF}(j_1)$
$p_\mathrm{T}(b_4)$	max $\mathrm{DF}(j_3)$	max $\mathrm{DF}(j_3)$	$\mathrm{DF}(j_4)$
$\lvert\eta(b_4)\rvert$	$p_\mathrm{T}(j_4)$	$\lvert\eta(b_4)\rvert$	$\mathrm{DF}(j_2)$
$p_\mathrm{T}(j_3)$	$\lvert\eta(j_4)\rvert$	$p_\mathrm{T}(b_3)$	max $\mathrm{DF}(j_2)$
max $\mathrm{DF}(j_3)$	max $\mathrm{DF}(j_4)$	$p_\mathrm{T}(b_4)$	$p_\mathrm{T}(j_3)$
		max $\mathrm{DF}(j_2)$	$\mathrm{DF}(j_3)$
		$\lvert p_T(j_3)\rvert$	

GHENT UNIVERSITY

Loss function

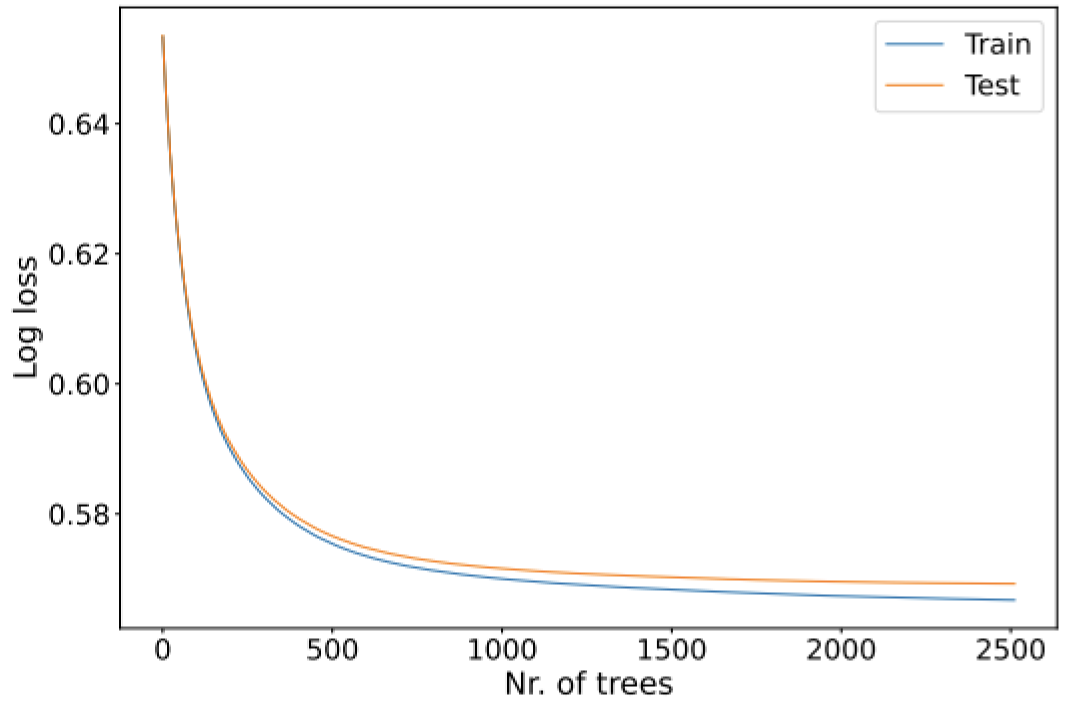
- The loss function quantifies the error between predicted and actual values
- Loss function minimization using training set
- Underfitting or overfitting → early stopping implementation



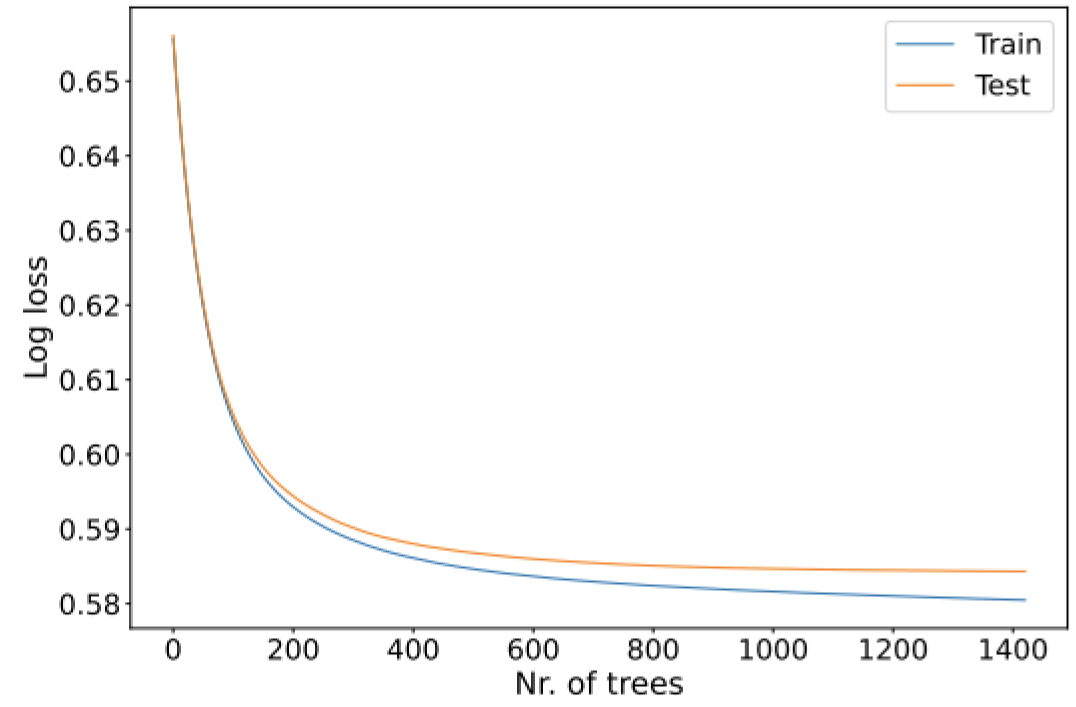


GHENT UNIVERSITY

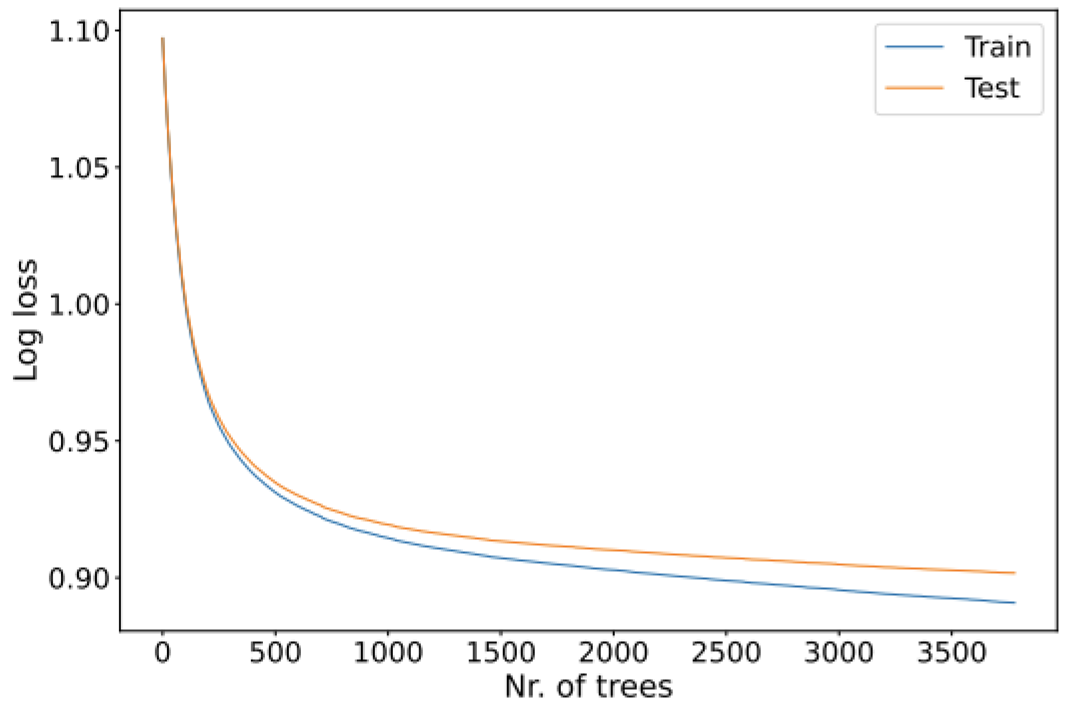
Loss functions



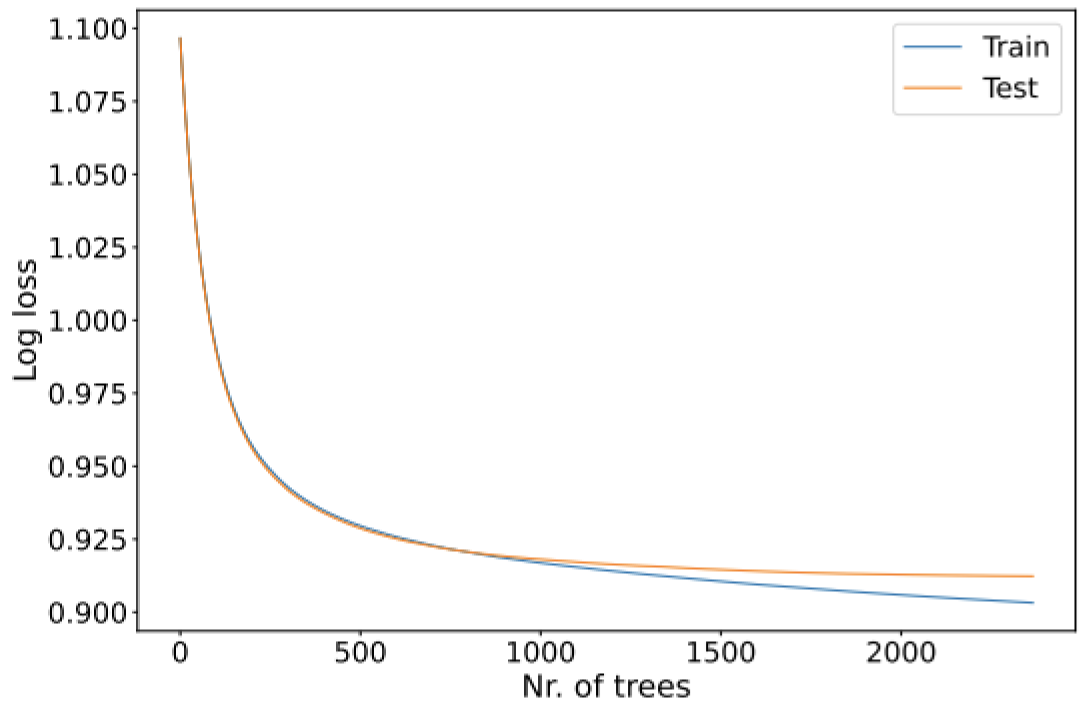
(a) Binary SR-2ℓ model.



(b) Binary SR-3ℓ model.



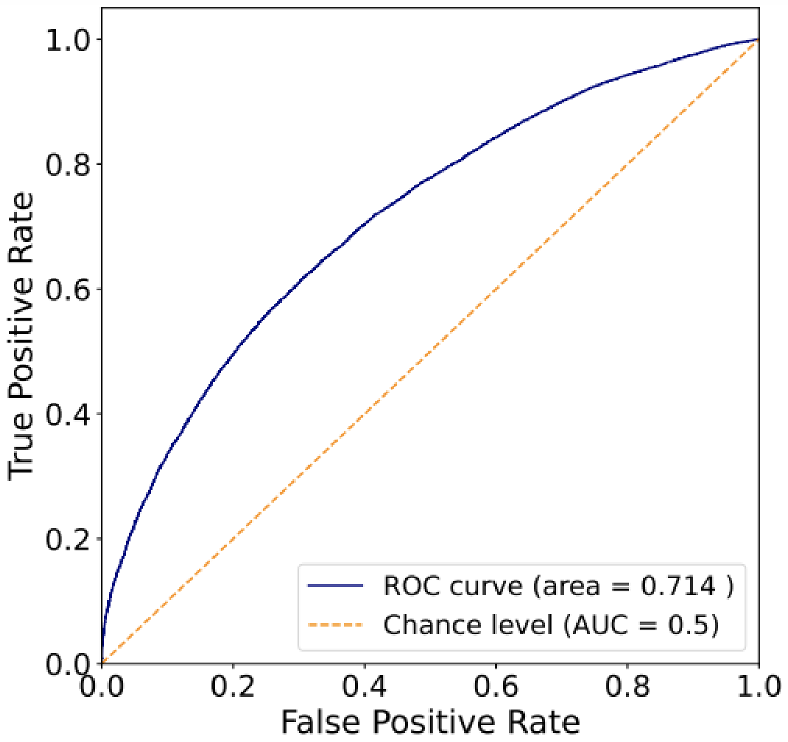
(c) Triple SR-2ℓ model.



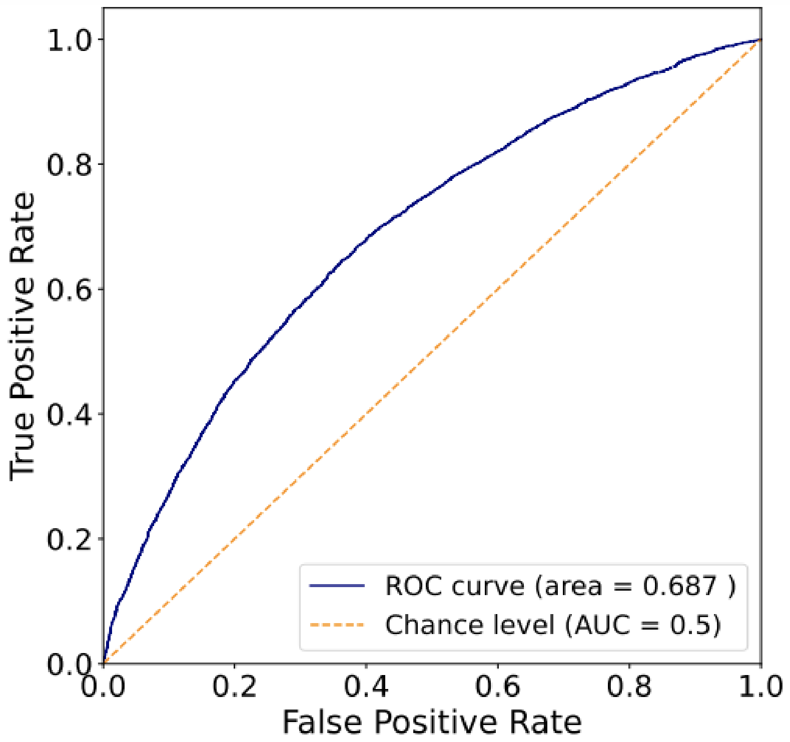
(d) Triple SR-3ℓ model.

GHENT
UNIVERSITY

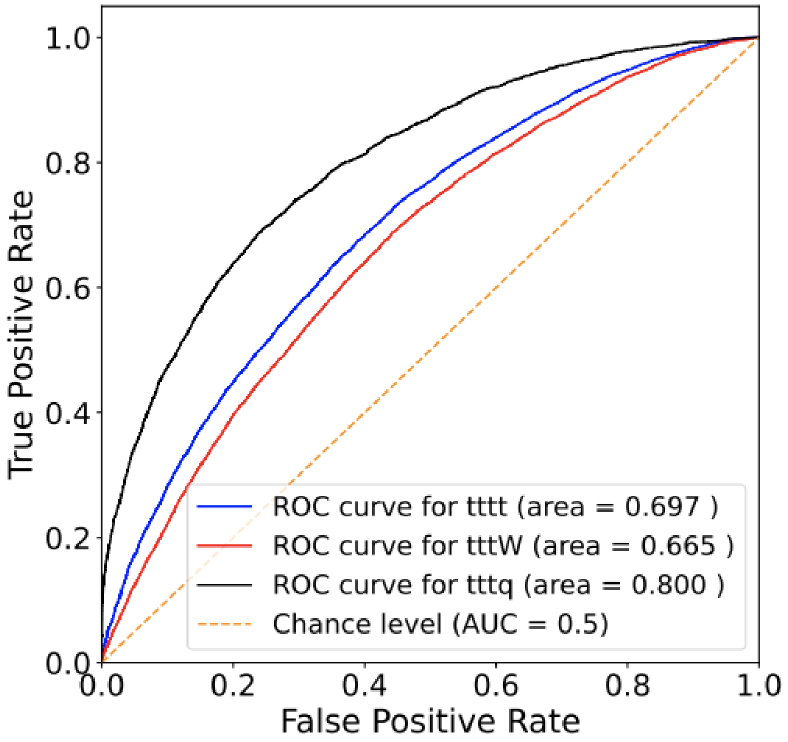
ROC curves



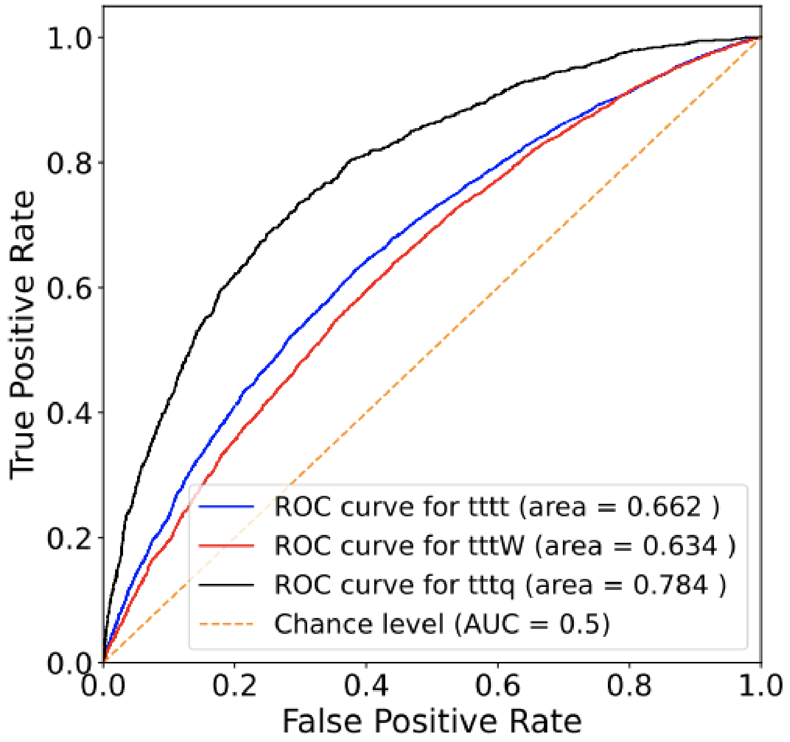
(a) Binary SR-2ℓ model.



(b) Binary SR-3ℓ model.



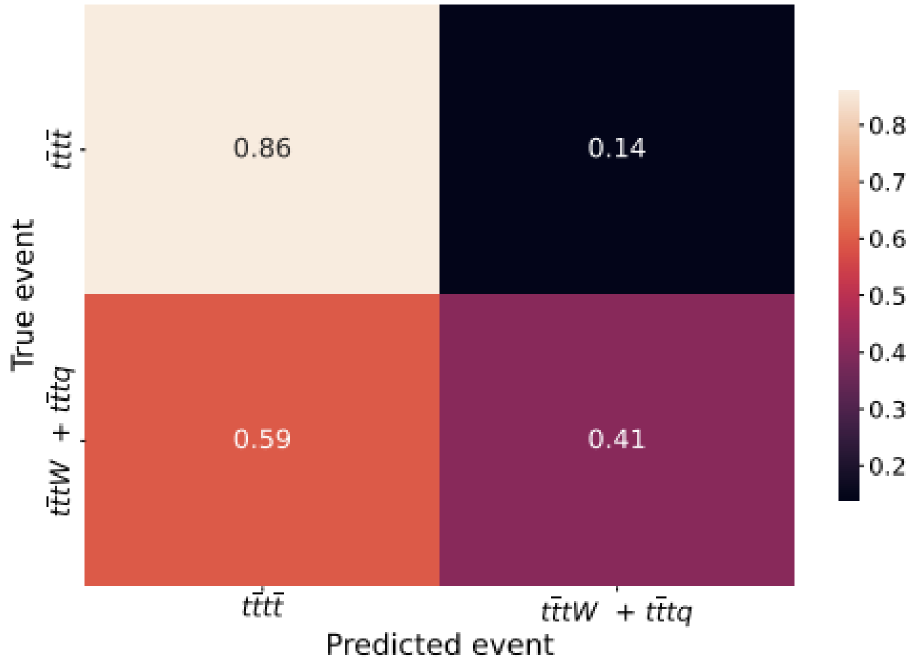
(c) Triple SR-2ℓ model.



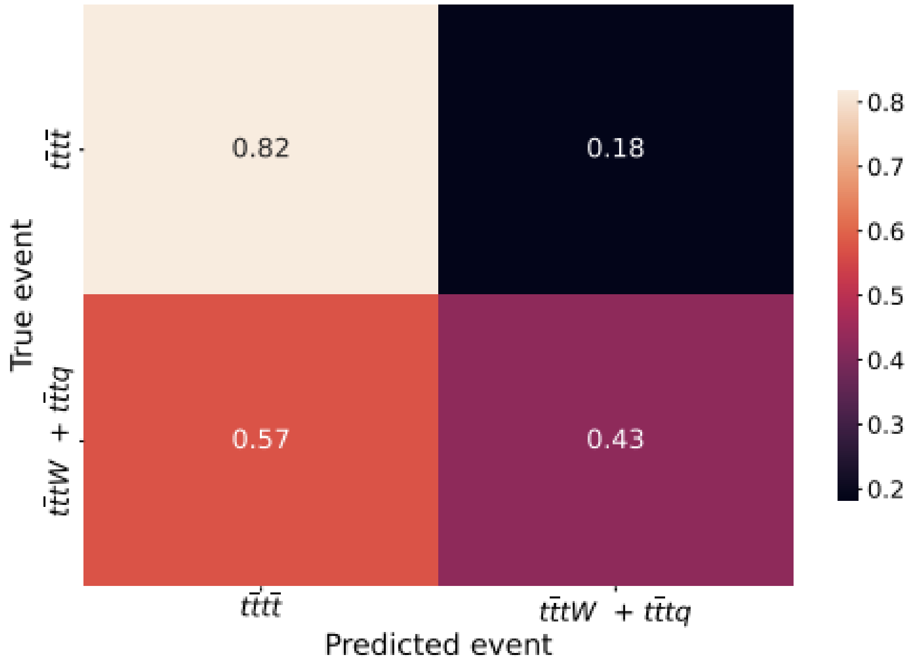
(d) Triple SR-3ℓ model.

GHENT UNIVERSITY

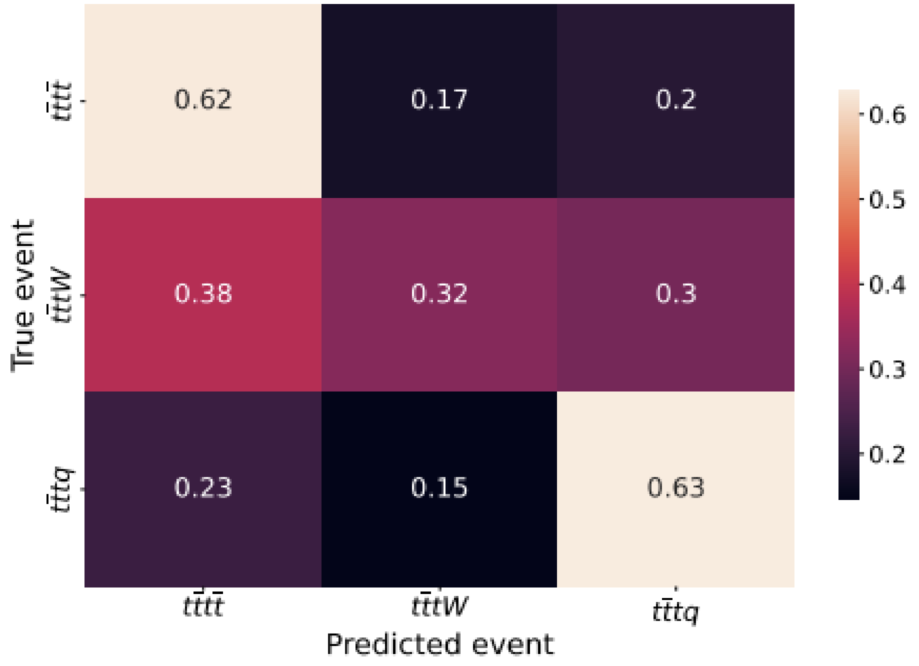
Confusion matrices



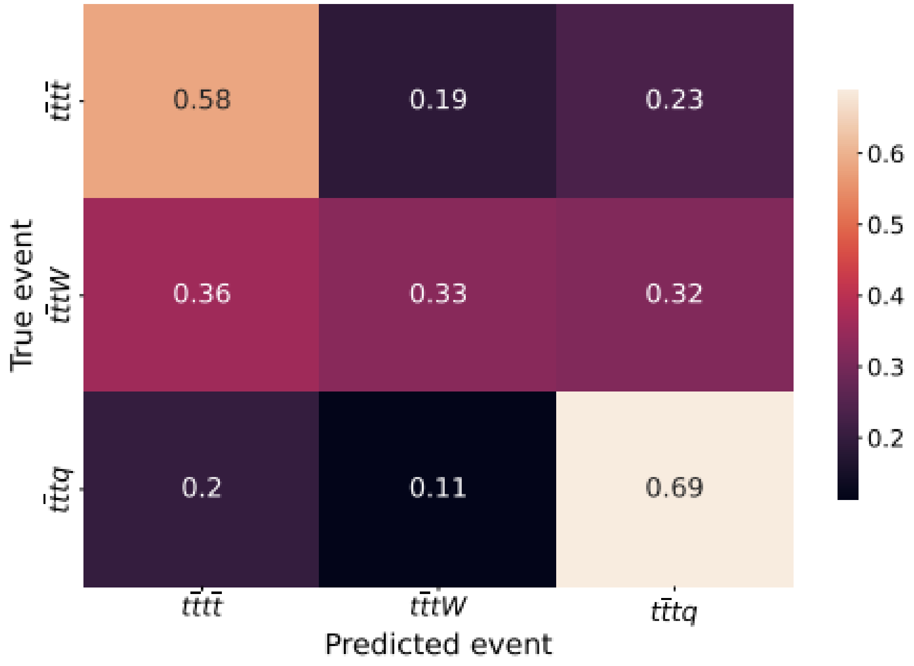
(a) Binary SR-2ℓ model.



(b) Binary SR-3ℓ model.



(c) Triple SR-2ℓ model.

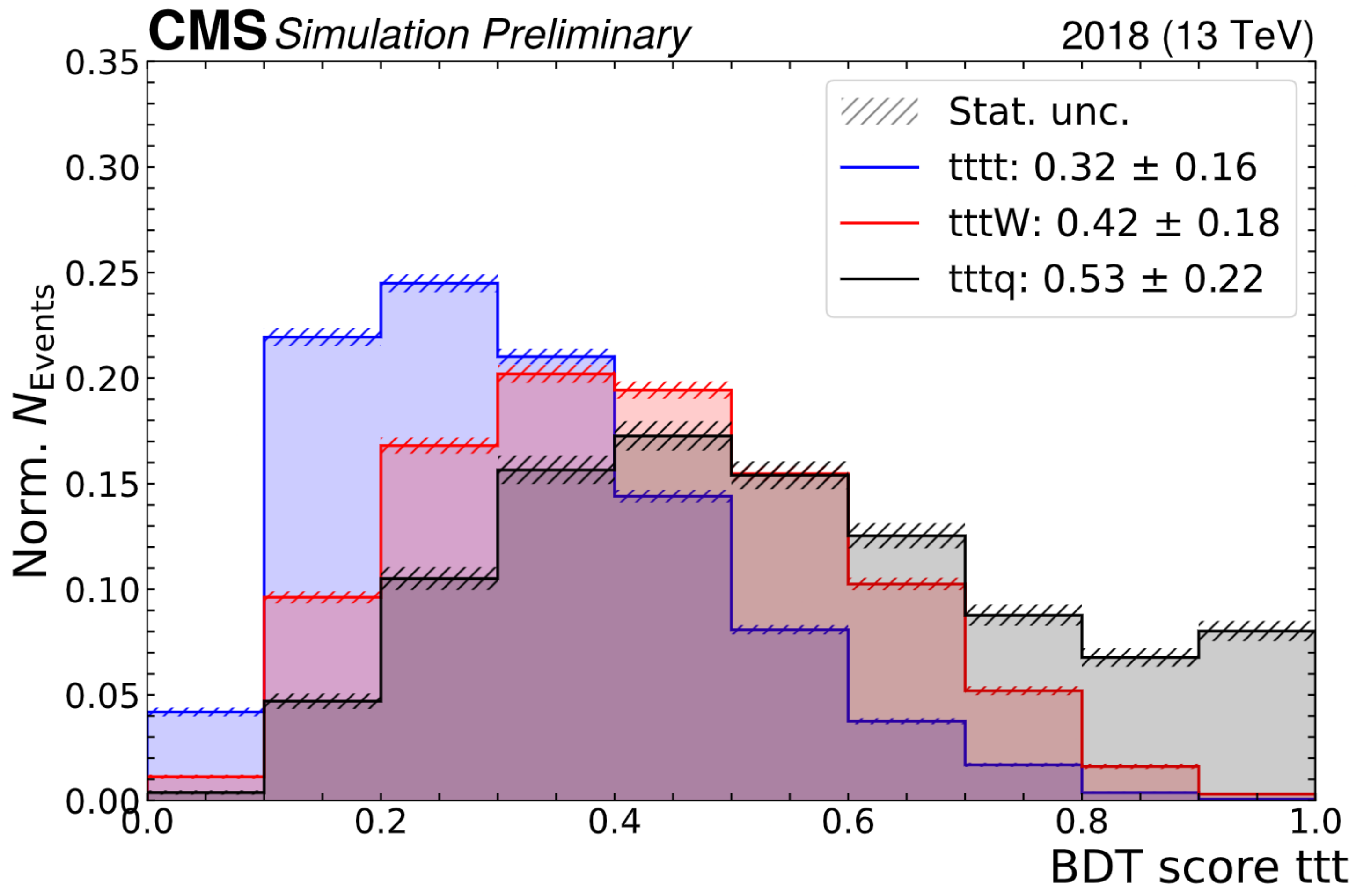


(d) Triple SR-3ℓ model.

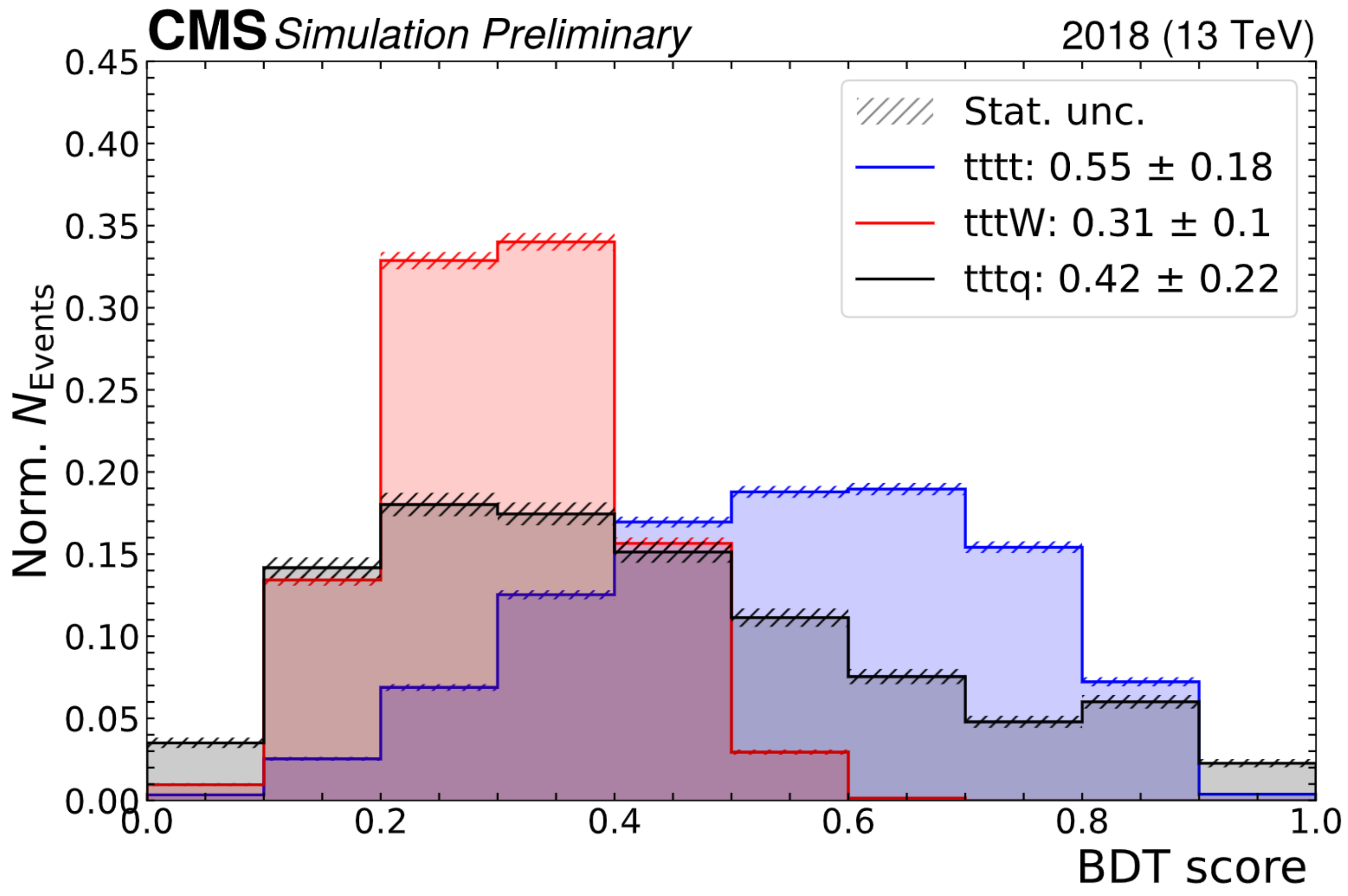
GHENT UNIVERSITY

BDT scores

Binary SR-2l



Triple SR-2l

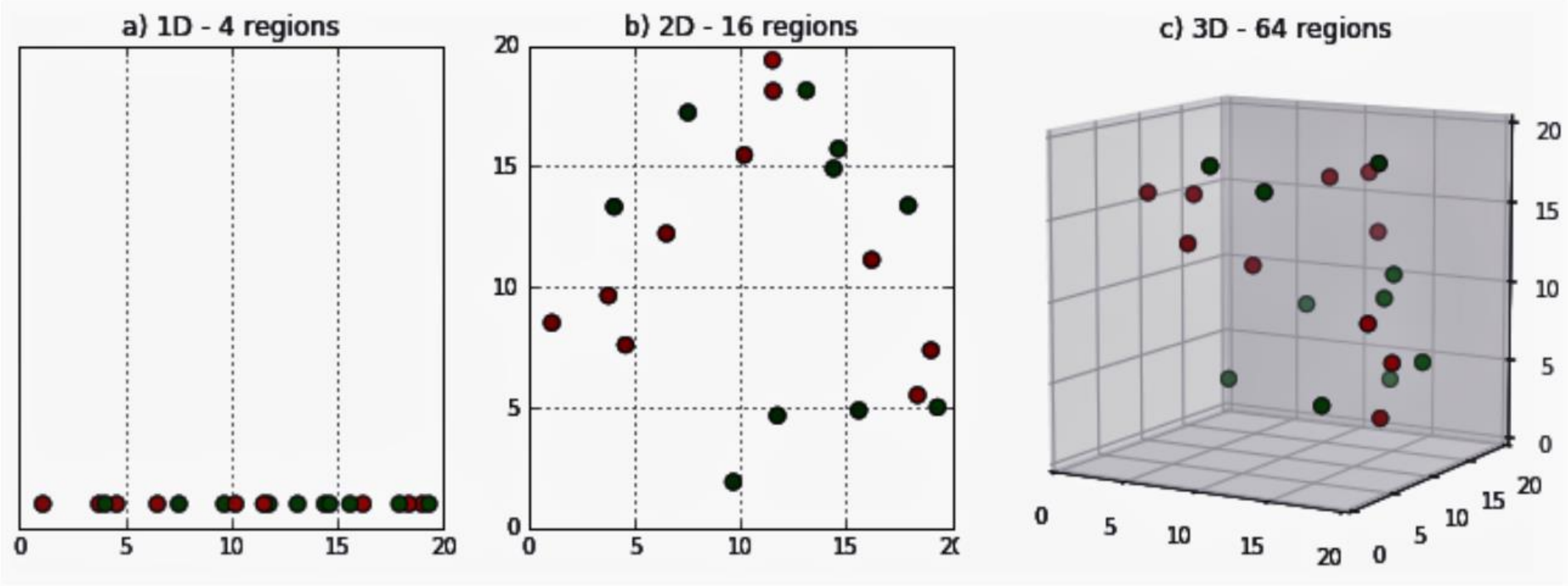


GHENT UNIVERSITY

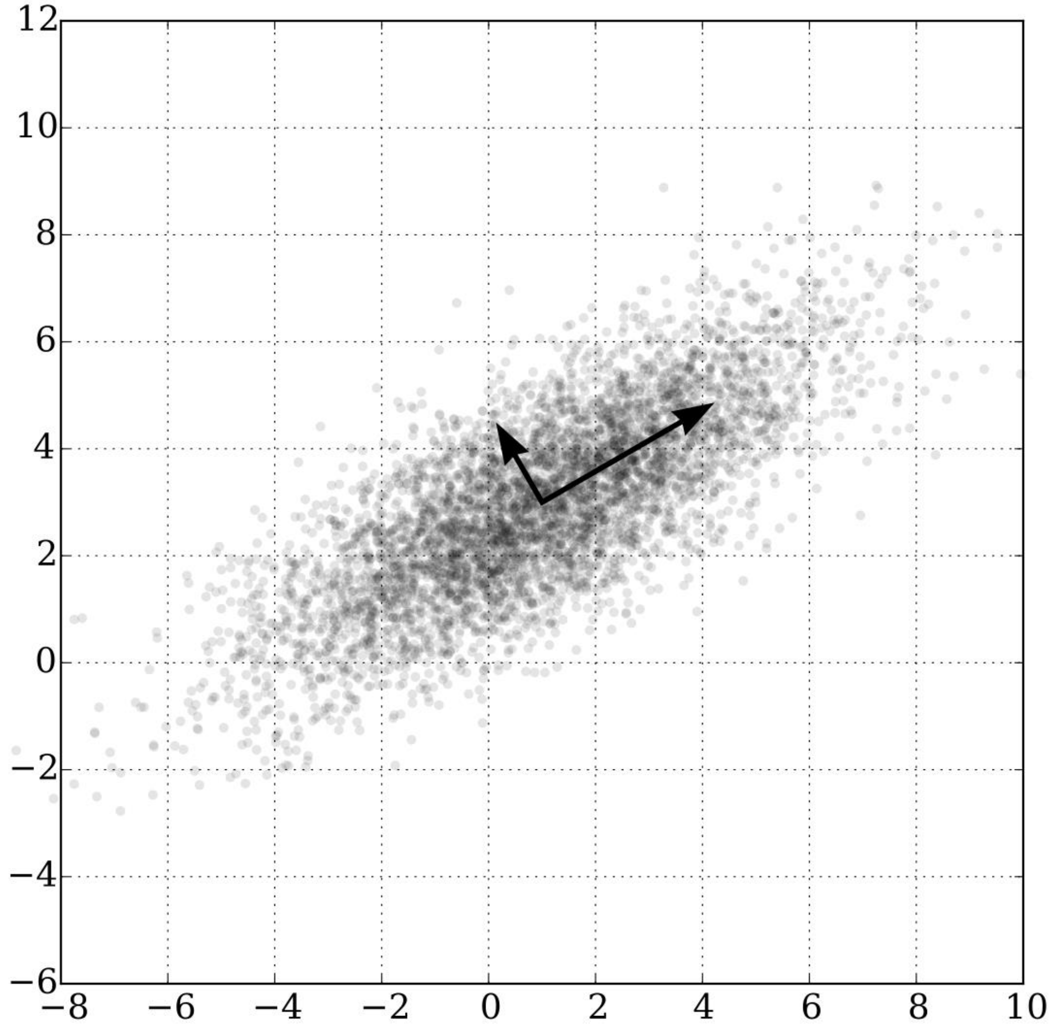
PCA & LDA

- Curse of dimensionality:
  Higher dimensions → Lower statistical power

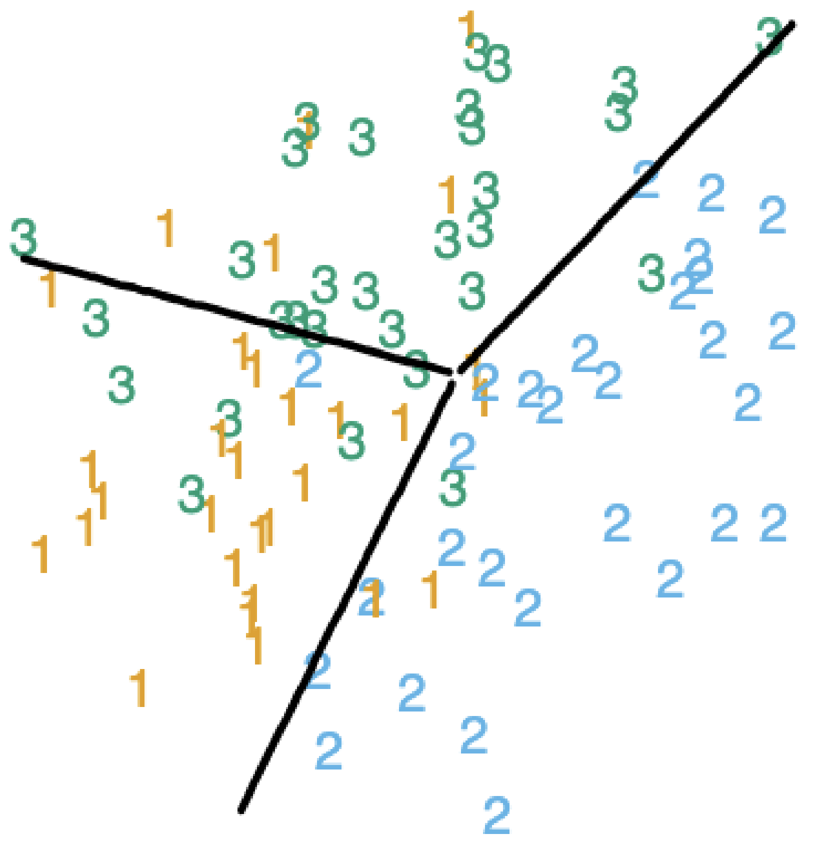
- Dimensionality reduction via PCA & LDA
  - PCA: Maximum variability of the data
  - LDA: Maximum separability of the classes



PCA

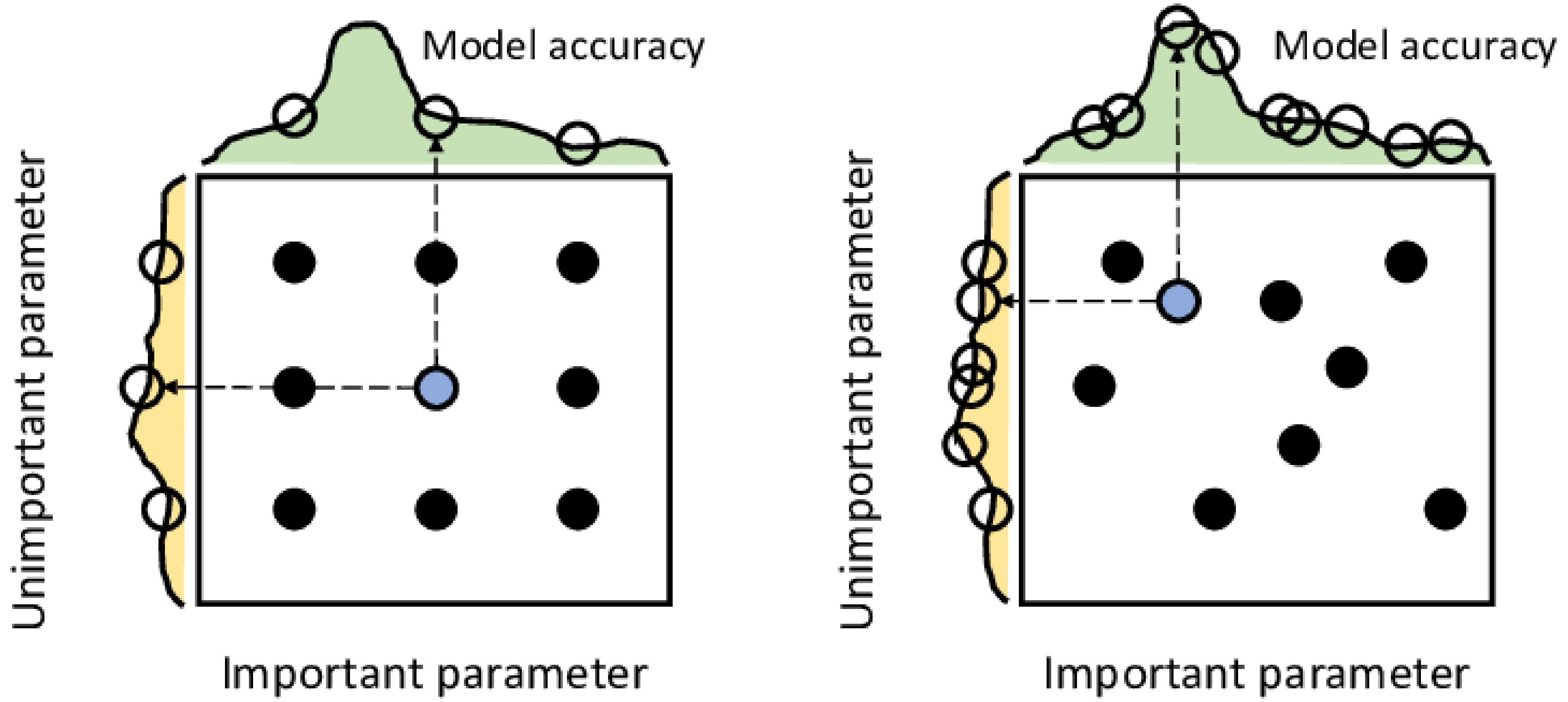


LDA



GHENT UNIVERSITY

Grid vs Random search



(a) (b)

Log Loss function

$$-\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log\left(p\left(y_i\right)\right) + \left(1 - y_i\right) \cdot \log\left(1 - p\left(y_i\right)\right)$$

GHENT UNIVERSITY

Feature importance

- Feature importance formula:

$$\hat{I_j^2}(T) = \sum_{t=1}^{J-1} \hat{i_t^2} 1(v_t = j)$$

Summation over nonterminal nodes t of the J-terminal node tree T, v_t is the splitting variable associated with node t

- Terminal region splitting R into two subregions (R_l, R_r) with empirical improvement in squared error:

$$i^2(R_l, R_r) = \frac{w_l w_r}{w_l + w_r}(\bar{y}_l - \bar{y}_r)^2$$

where $\bar{y}_l$, $\bar{y}_r$ are the left and right daughter response means respectively, and w_l, w_r are the corresponding sums of the weights