# Evidence for longitudinally polarized W bosons in the electroweak production of same-sign WW pairs

**Erik Bachmann** on behalf of the analysis team

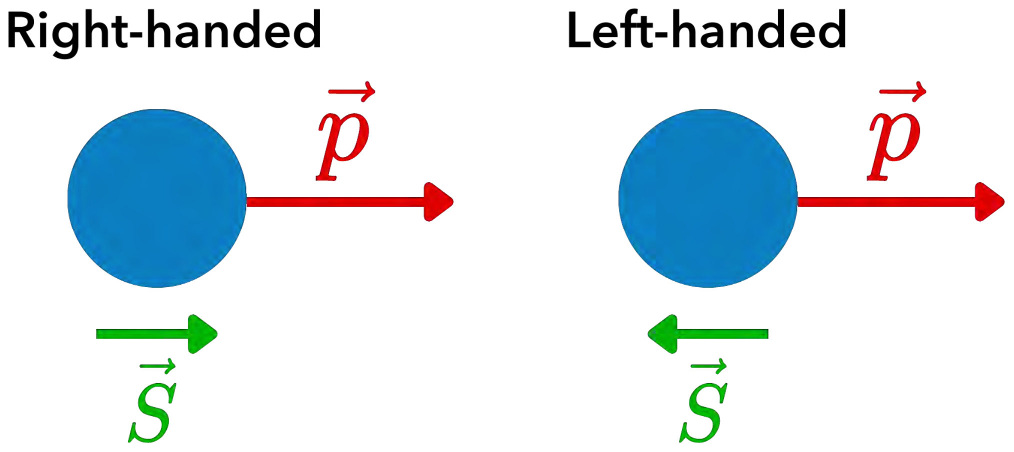
Institute of Nuclear and Particle Physics, Technische Universität Dresden

ATLAS-D Wuppertal, September 18, 2025

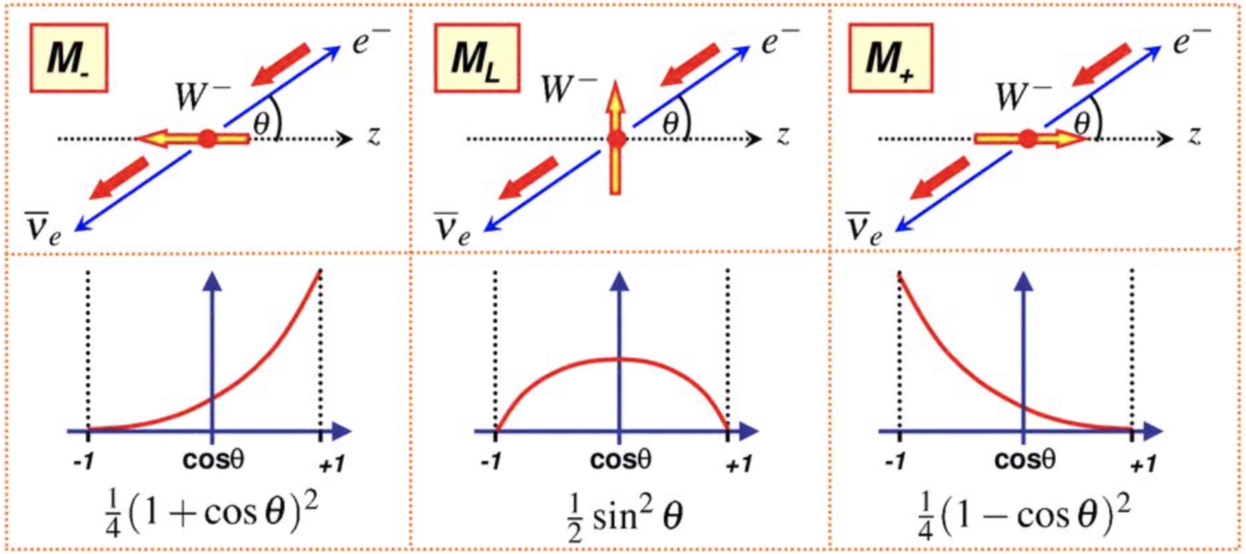
# Vector Boson Polarization

- Polarization: **alignment** of a particle's **spin** with its **momentum**
- Helicity: $h = \vec{S} \cdot \frac{\vec{p}}{|\vec{p}|}$
  - **Transverse (T):** $h = \pm 1$
  - **Longitudinal (L):** $h = 0$

  ⇒ **not Lorentz-invariant!**

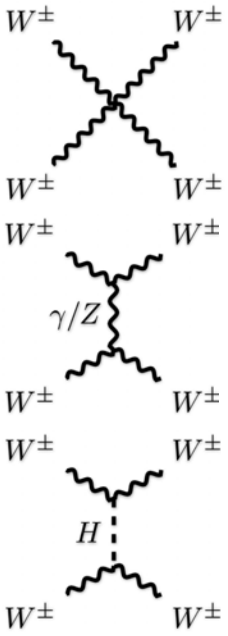


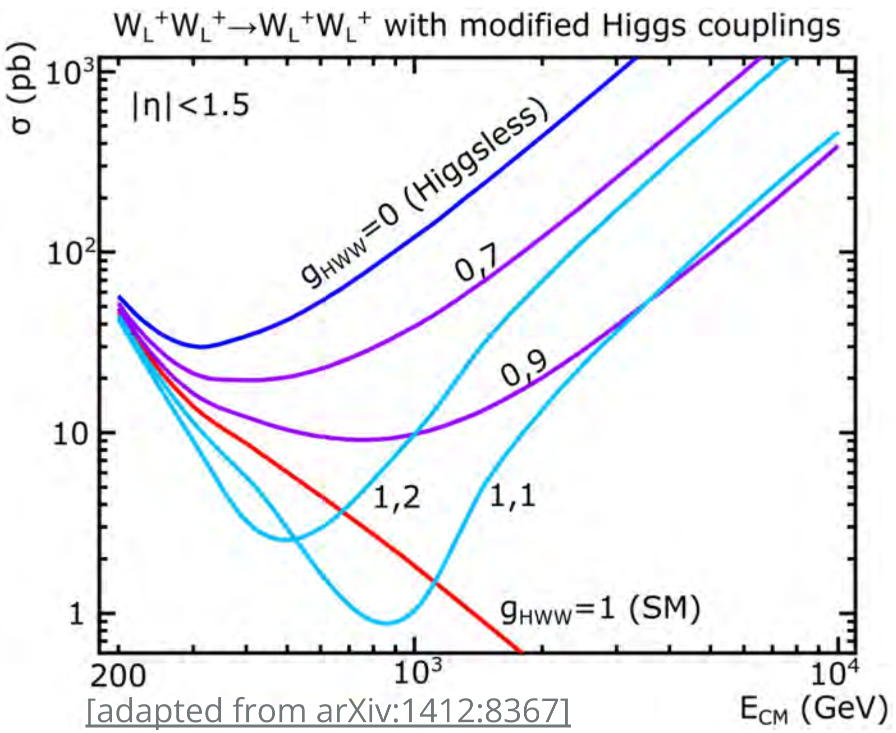

- Parity violation in weak interaction
  → **effects on decay kinematics**
- Decay angles cannot be reconstructed due to neutrinos
  → **use polarized theory predictions**



# Motivation

- Longitudinal polarization of electroweak gauge bosons is a direct consequence of the EWSB
- **Important test of the Higgs mechanism**
- Particularly interesting: **longitudinal VBS**



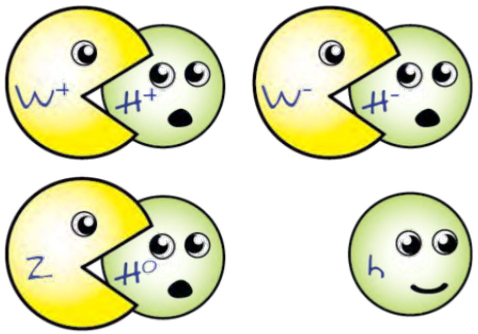

[adapted from arXiv:1412:8367]

**Higgs mechanism**

**Massive W and Z bosons**

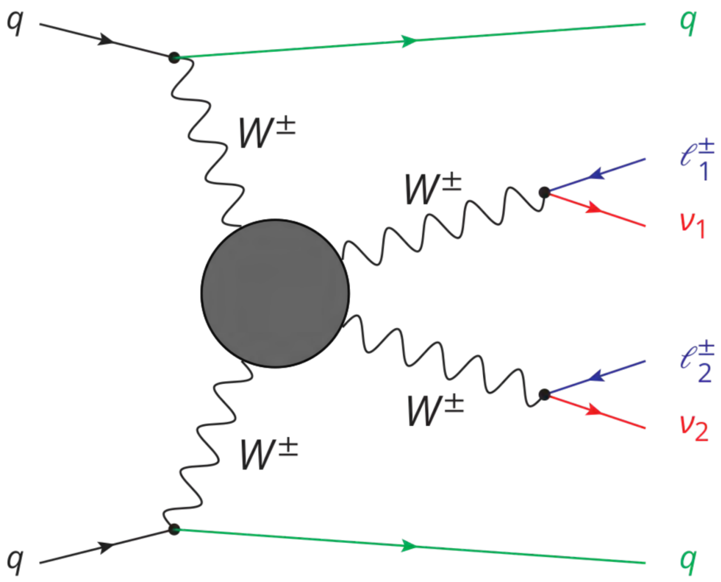
**Longitudinal polarization allowed**

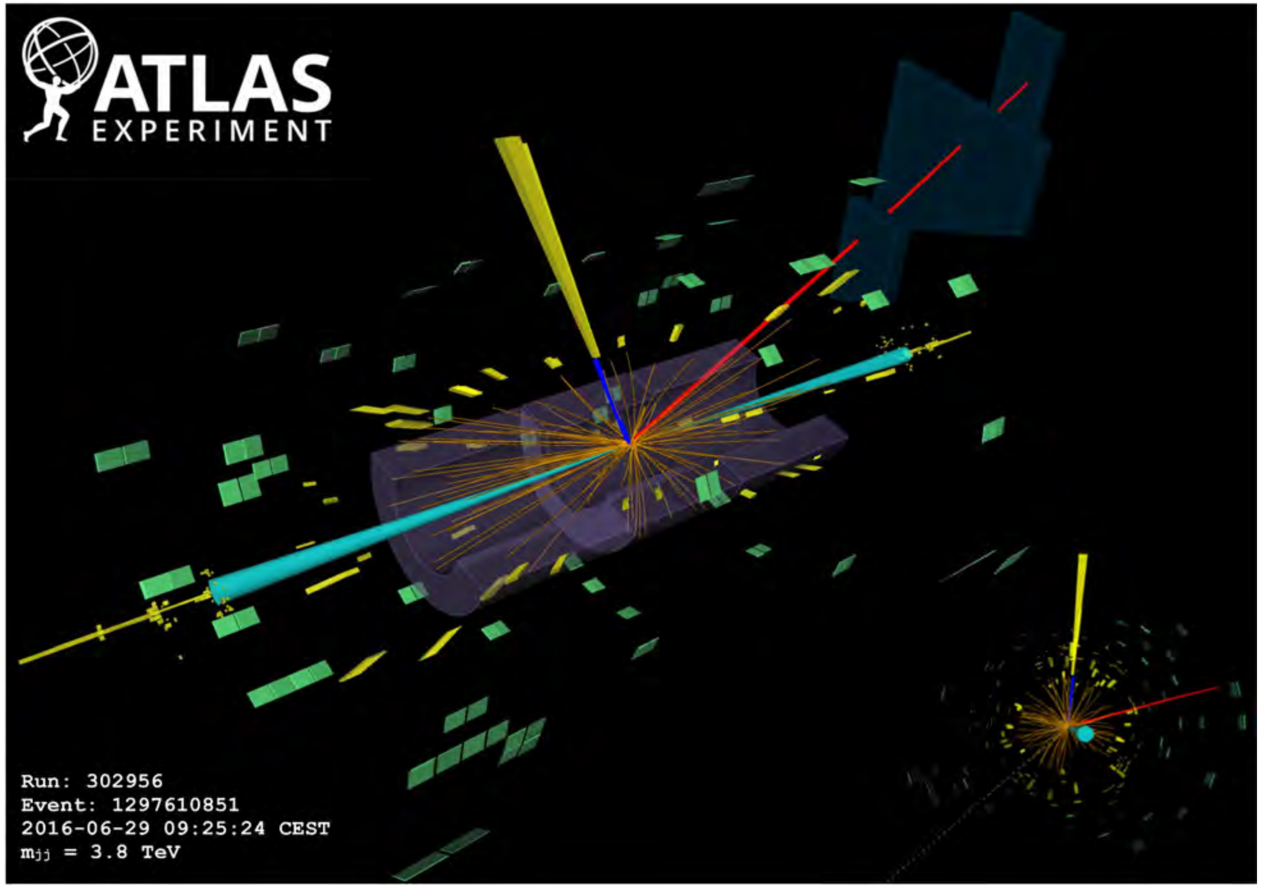
**Goldstone equivalence theorem**
"At high energy, longitudinal vector bosons are analogous to goldstone bosons"




[quantumdiaries.org]

# Same-sign WW scattering at the LHC

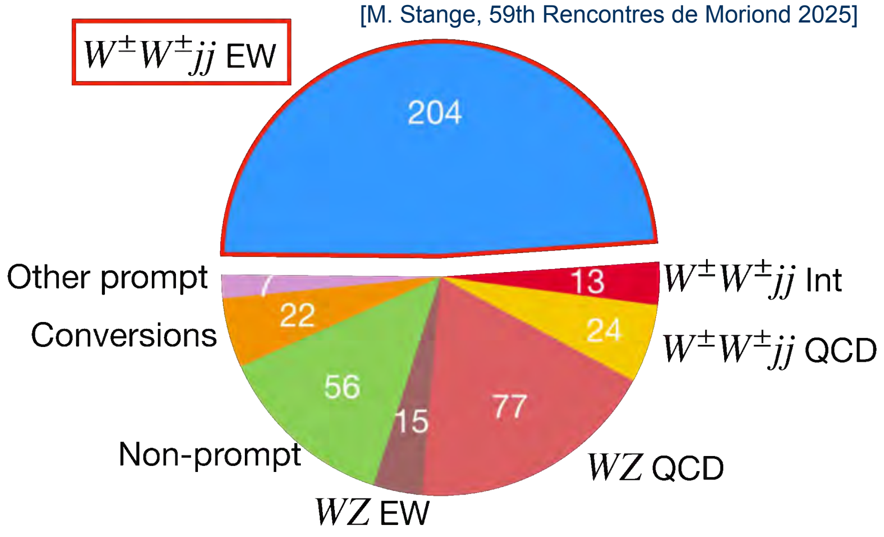


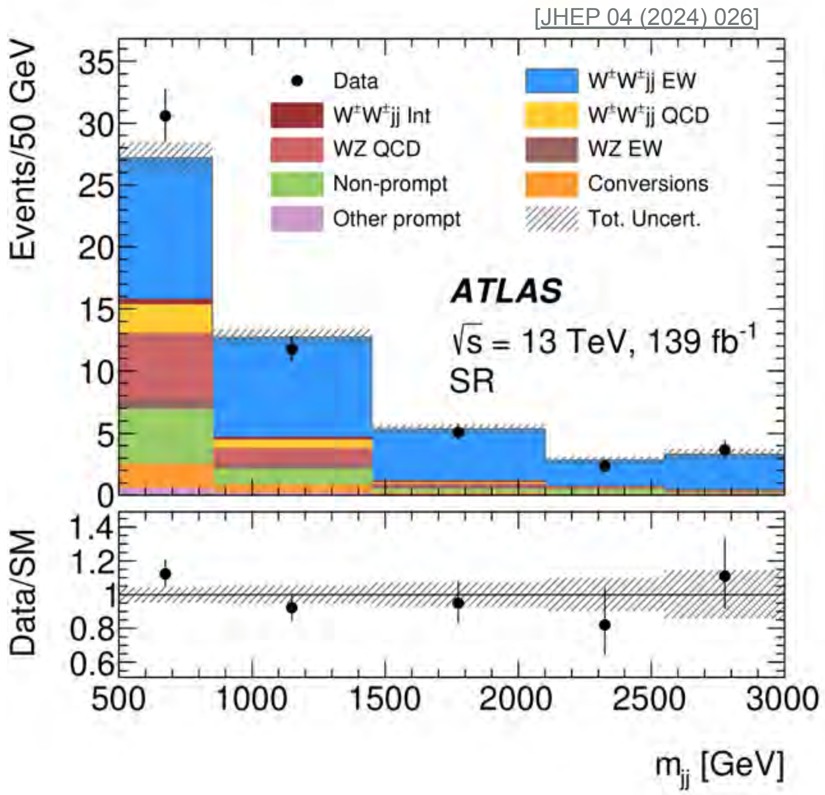



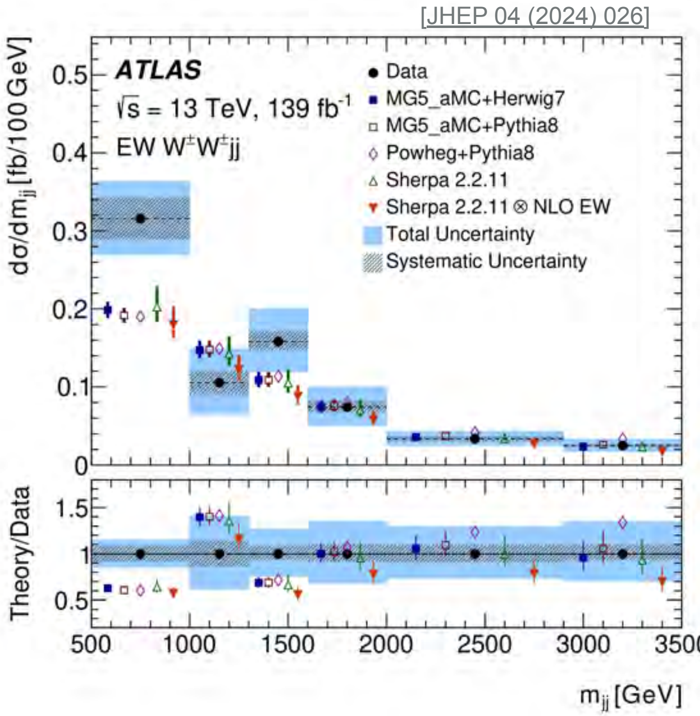

- Exactly **two same-charged leptons**
- At least **two well-separated jets** with $m_{jj} > 500\,\mathrm{GeV}$
- At least 30 GeV of **missing transverse momentum**

# Analysis baseline

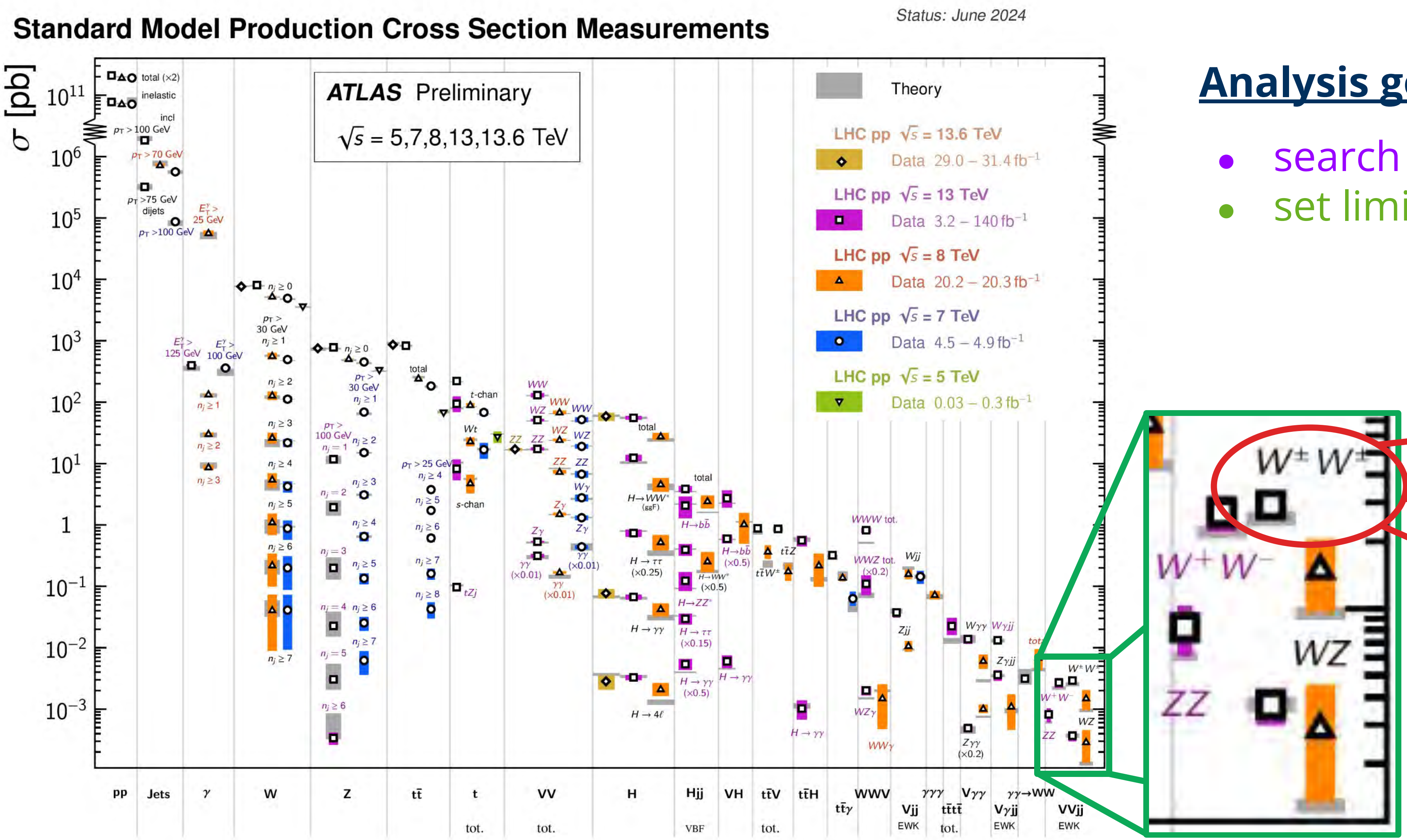
- Full Run 2 dataset (140 fb$^{-1}$)
- Unpolarized measurement of differential cross-section
  - **Cross-section measured with 10% accuracy**



# $W_L^\pm W_L^\pm jj$ is rare




## Analysis goals:

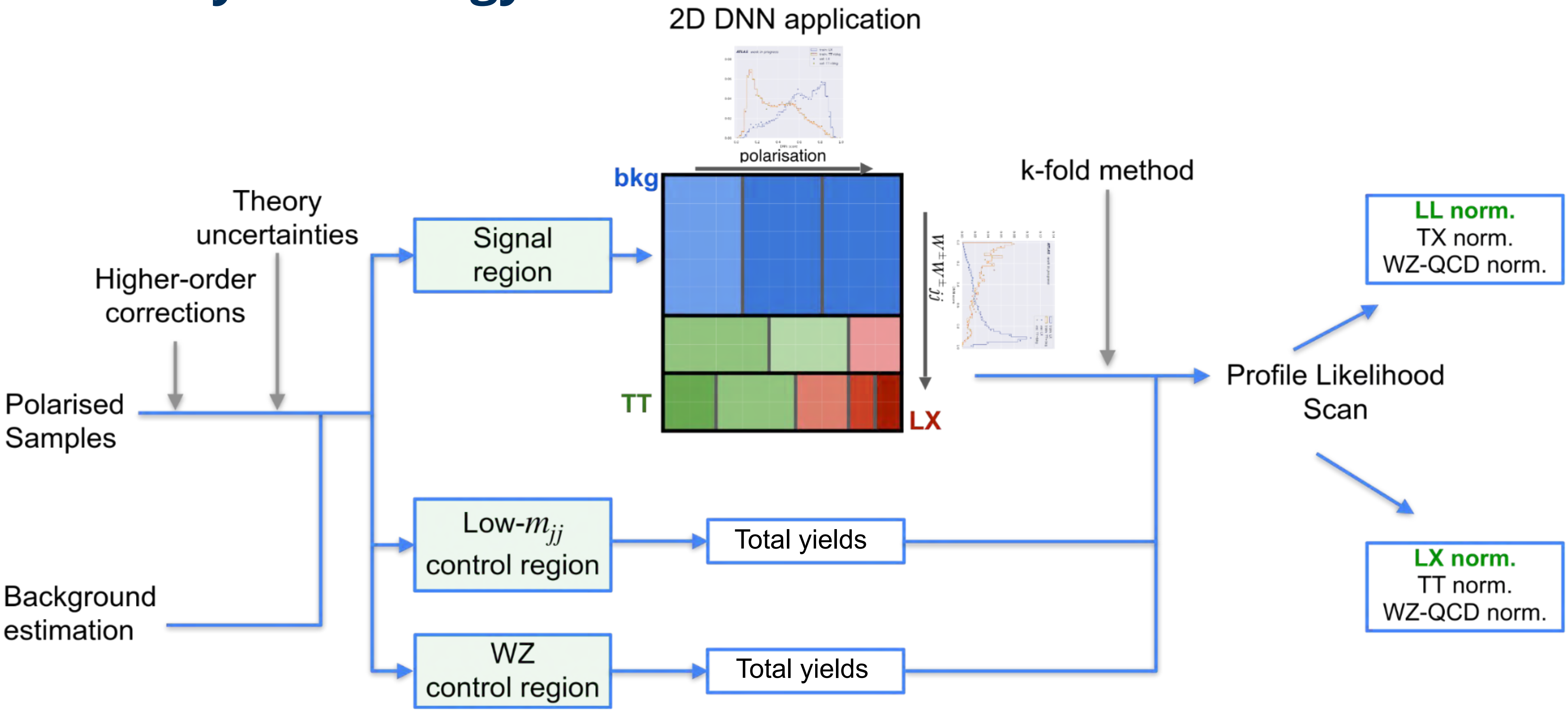
- search for $W_L^\pm W^\pm$
- set limit on $W_L^\pm W_L^\pm$

10% $W_L^\pm W_L^\pm$

29% $W_L^\pm W_T^\pm$

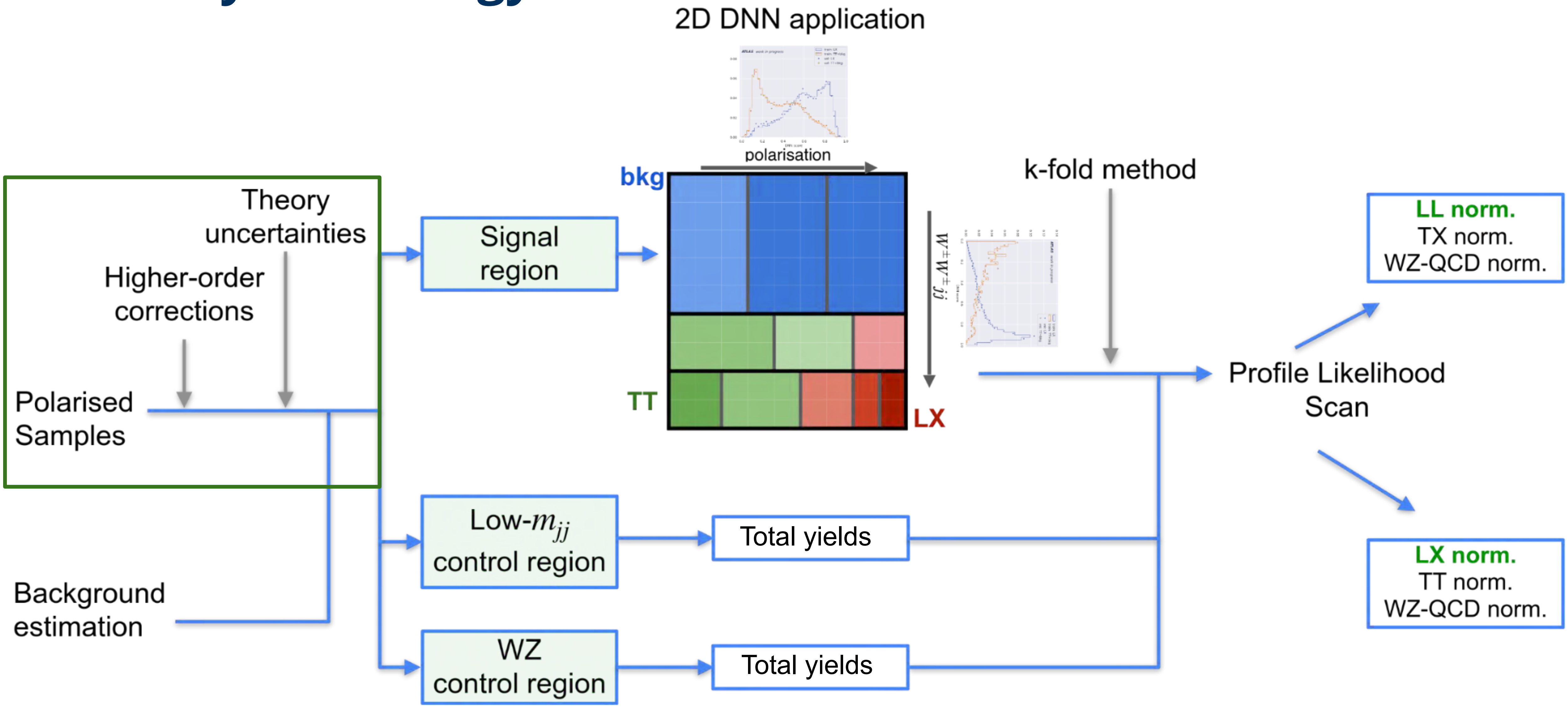
61% $W_T^\pm W_T^\pm$

# Analysis strategy



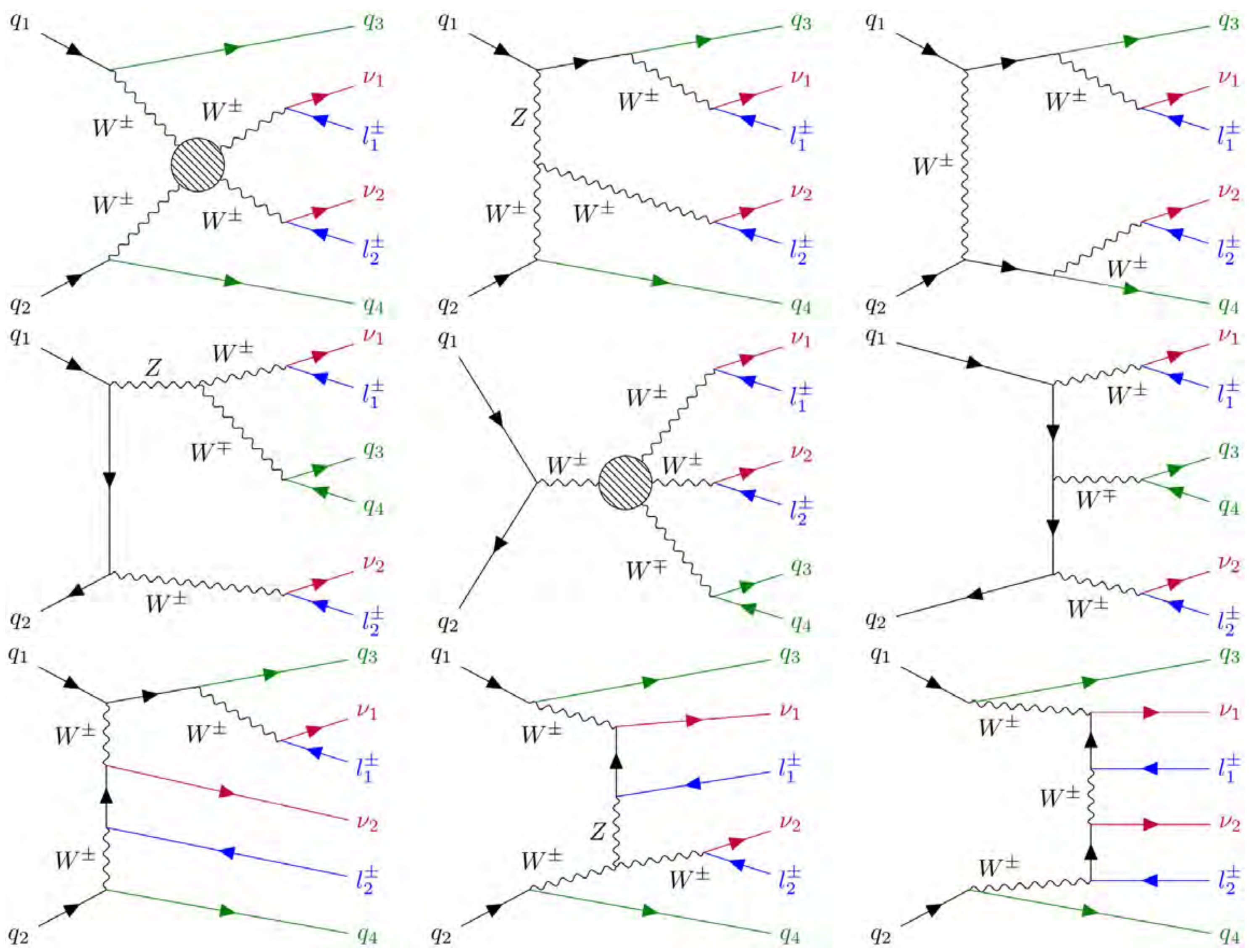

[M. Stange, SM approval meeting.]

# Analysis strategy

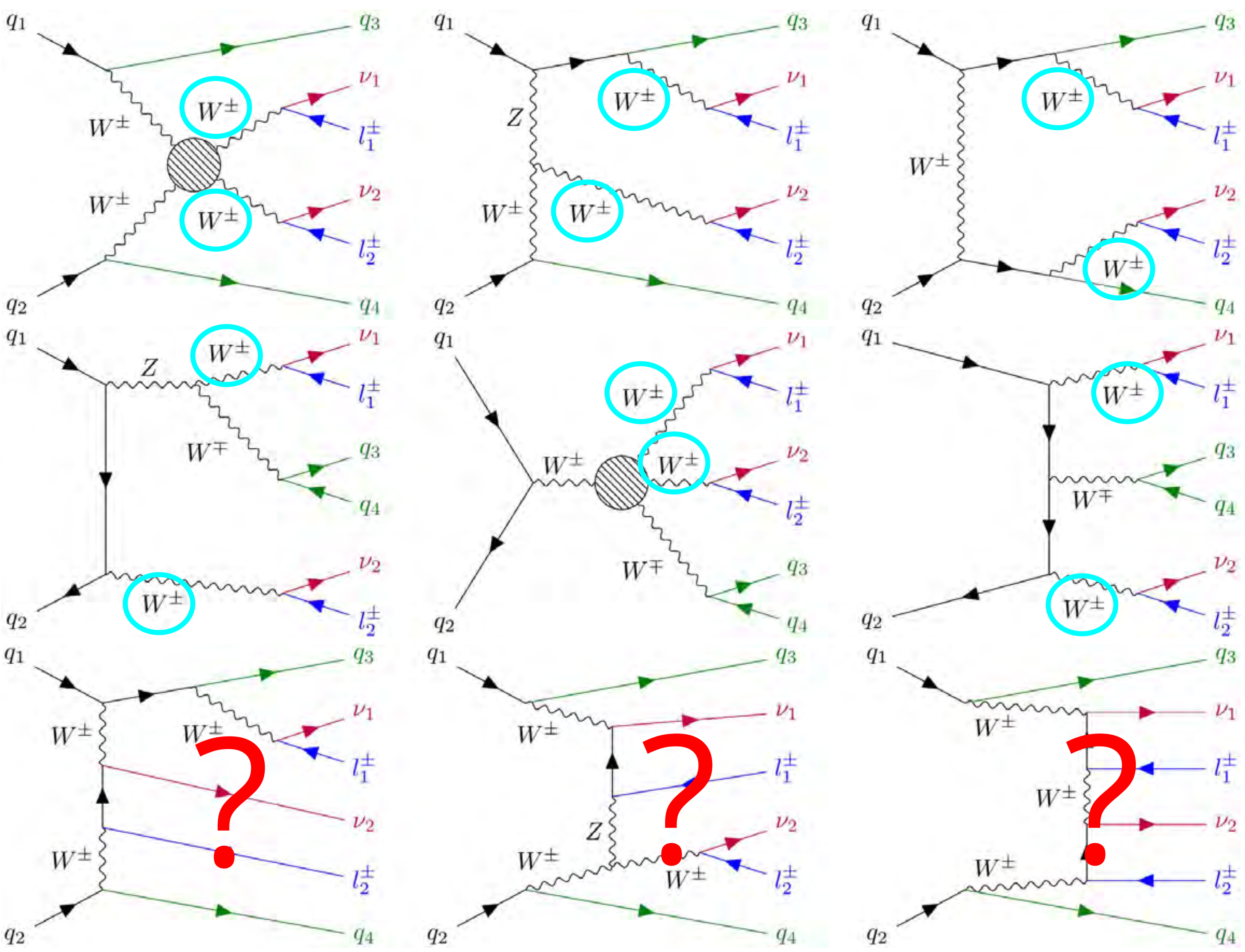



[M. Stange, SM approval meeting.]

# Polarized predictions for $\mathbf{W^{\pm}W^{\pm}jj}$



# Polarized predictions for $\mathbf{W^{\pm}W^{\pm}jj}$

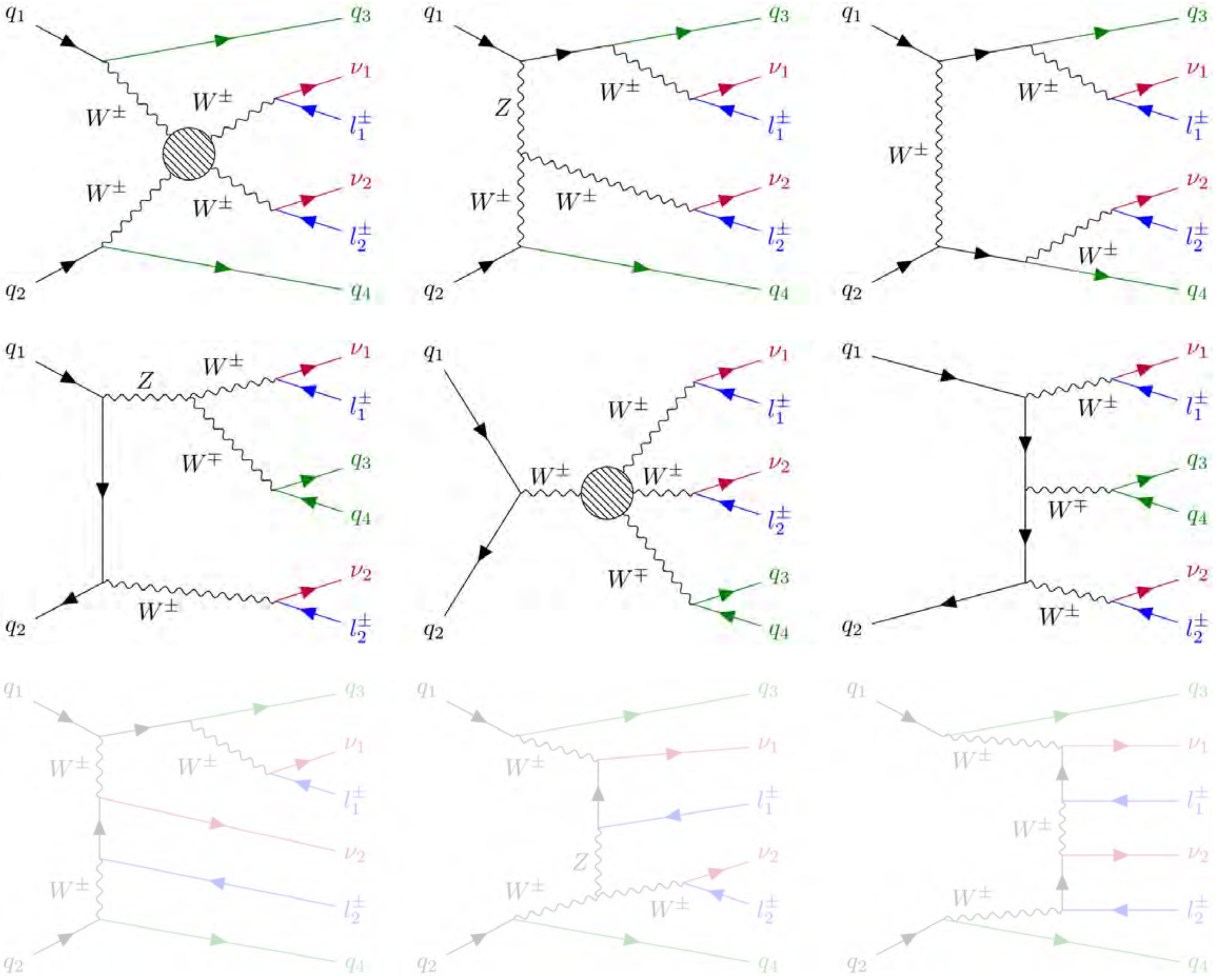


# Polarized predictions for $\mathbf{W^{\pm}W^{\pm}jj}$

- On-shell approximation necessary
  - Sherpa 3: **Narrow-Width Approx.**

$$\boxed{\frac{1}{\left(q^2 - m_V^2\right)^2 + \Gamma_V^2 m_V^2} \rightarrow \frac{\pi\delta\left(q^2 - m_V^2\right)}{\Gamma_V m_V}}$$

- Vector boson width set to zero to ensure gauge invariance

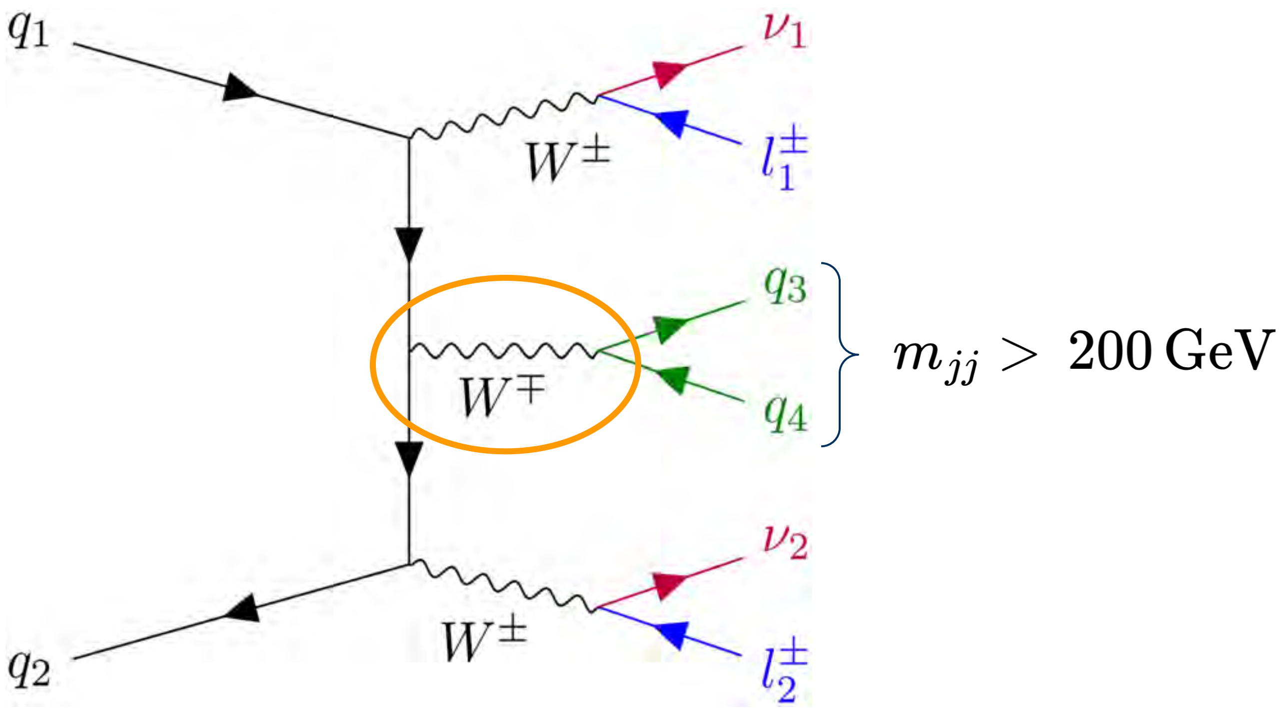


# Polarized predictions for $\mathbf{W^{\pm}W^{\pm}jj}$

- On-shell approximation necessary
  - Sherpa 3: **Narrow-Width Approx.**

$$\boxed{\frac{1}{\left(q^2 - m_V^2\right)^2 + \Gamma_V^2 m_V^2} \rightarrow \frac{\pi\delta\left(q^2 - m_V^2\right)}{\Gamma_V m_V}}$$

- Vector boson width set to zero to ensure gauge invariance
  $\rightarrow$ **divergence at $\mathbf{m_{qq} = m_W}$**



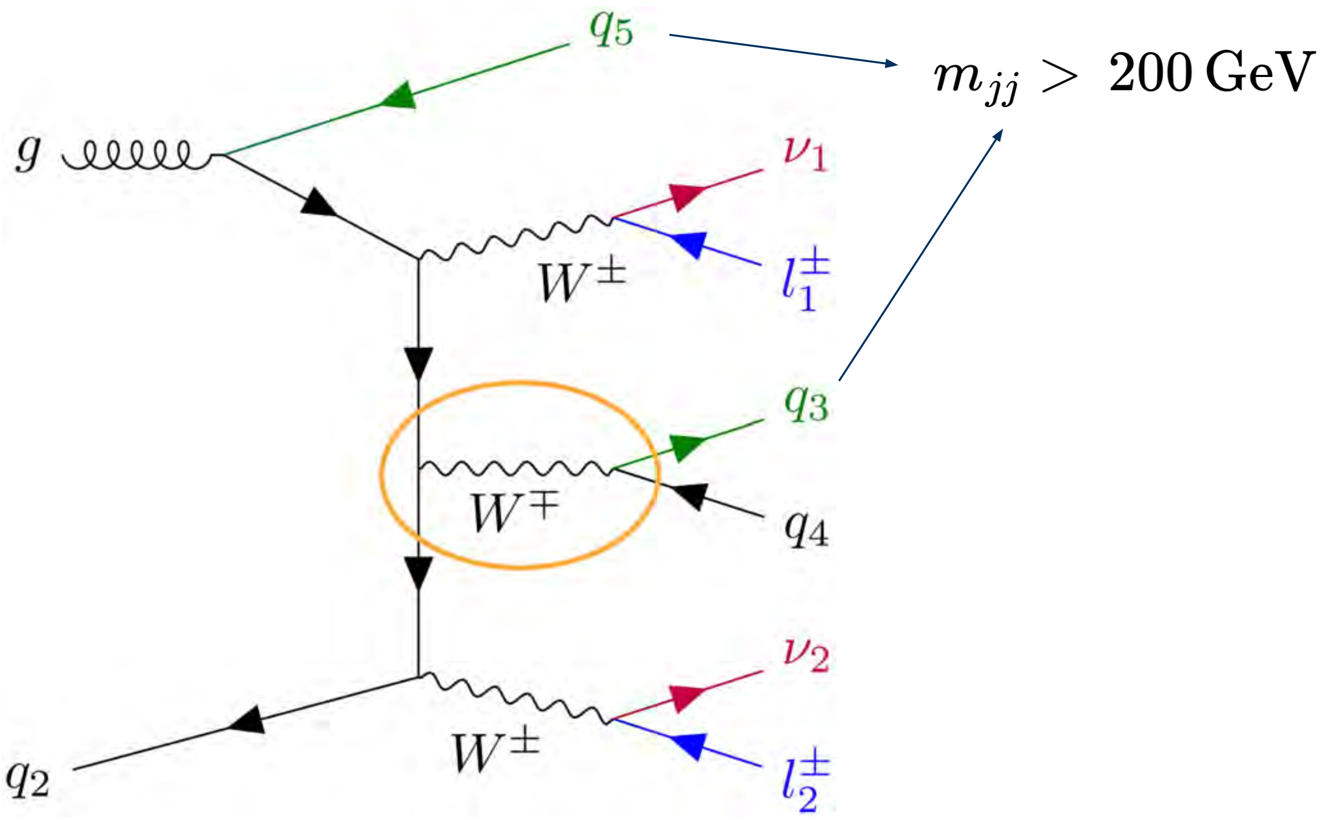
# Polarized predictions for $\mathbf{W^{\pm}W^{\pm}jj}$

- On-shell approximation necessary
  - Sherpa 3: **Narrow-Width Approx.**

$$\boxed{\frac{1}{\left(q^2 - m_V^2\right)^2 + \Gamma_V^2 m_V^2} \to \frac{\pi\delta\left(q^2 - m_V^2\right)}{\Gamma_V m_V}}$$

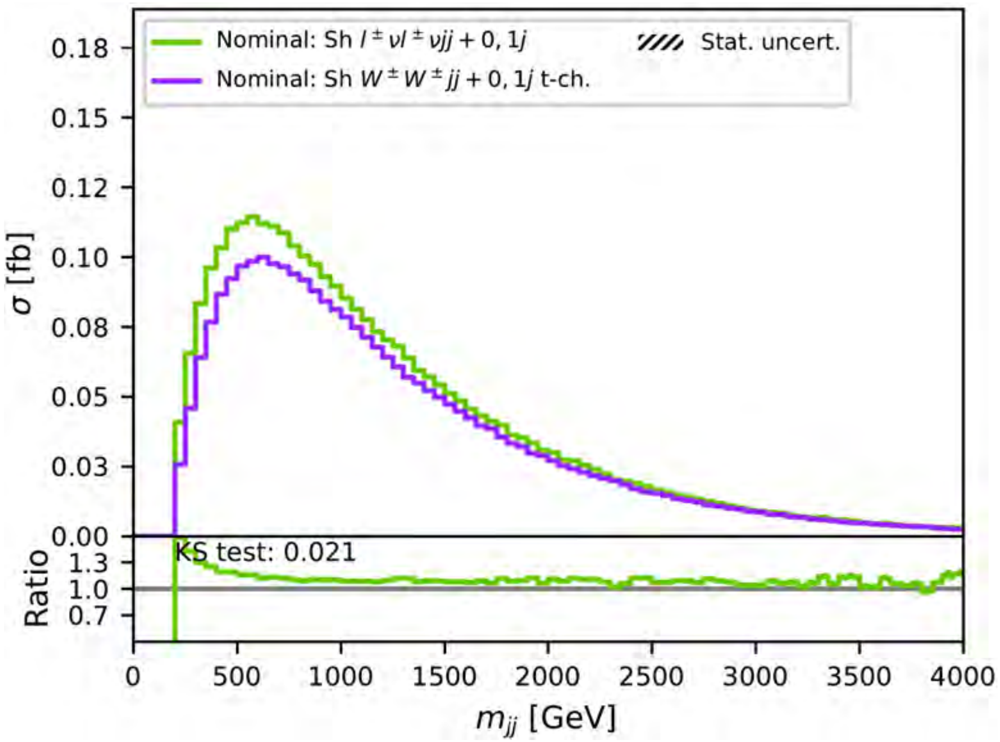
- Vector boson width set to zero to ensure gauge invariance
  $\rightarrow$ **divergence at $\mathbf{m_{qq} = m_W}$**

- **Simulate $\mathbf{W^{\pm}W^{\pm}jj + 0, 1j}$** to include NLO QCD effects
  - need **VBS-approximation** to suppress triboson diagrams



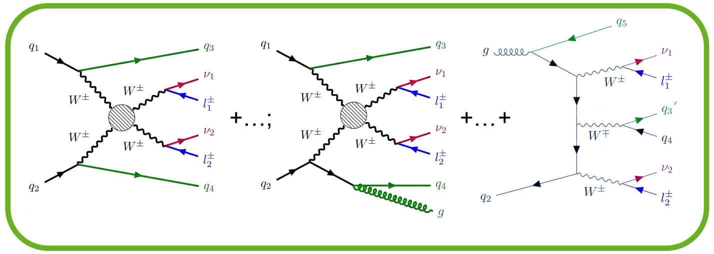
# Correction for missing triboson contribution

- Full **off-shell** $\boldsymbol{\ell^{\pm}\ell^{\pm}\nu\nu jj + 0, 1j}$ to correct missing triboson contribution

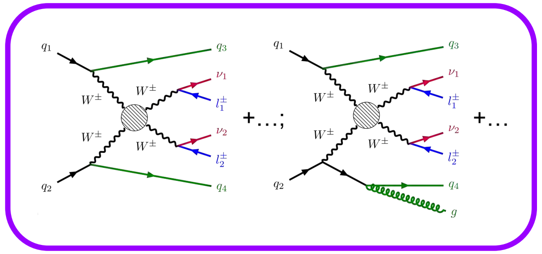


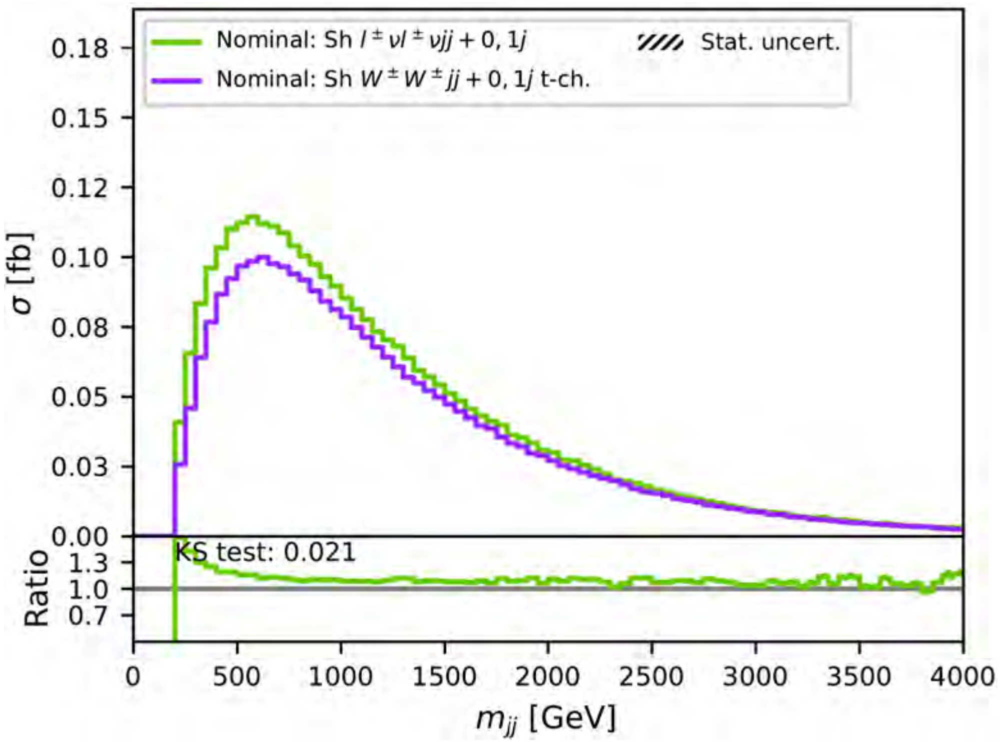
# Correction for missing triboson contribution

- Full **off-shell $\boldsymbol{\ell^{\pm}\ell^{\pm}\nu\nu jj + 0,1j}$** to correct missing triboson contribution
- **Multivariate reweighting** using deep neural network [arXiv:1907.08209]



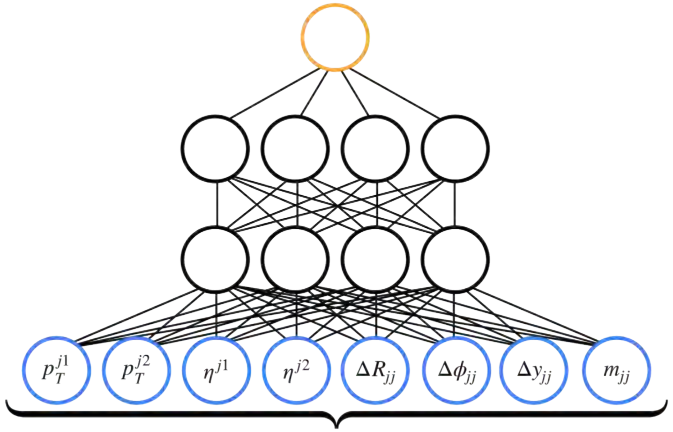
vs.






$$\boxed{\frac{p_1(x)}{p_2(x)} \approx \frac{\mathrm{DNN}(x)}{1-\mathrm{DNN}(x)}}$$

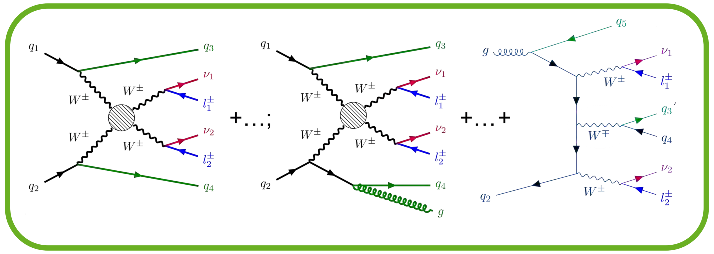
Output: $score \in (0,1)$



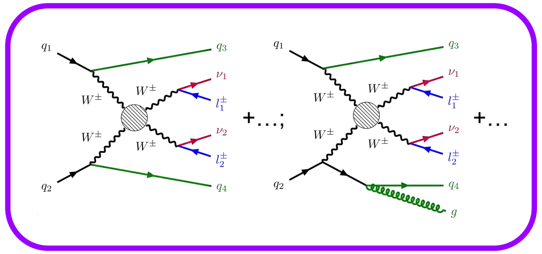

Input: di-jet kinematic [M. Stange, PhD defense]

# Correction for missing triboson contribution

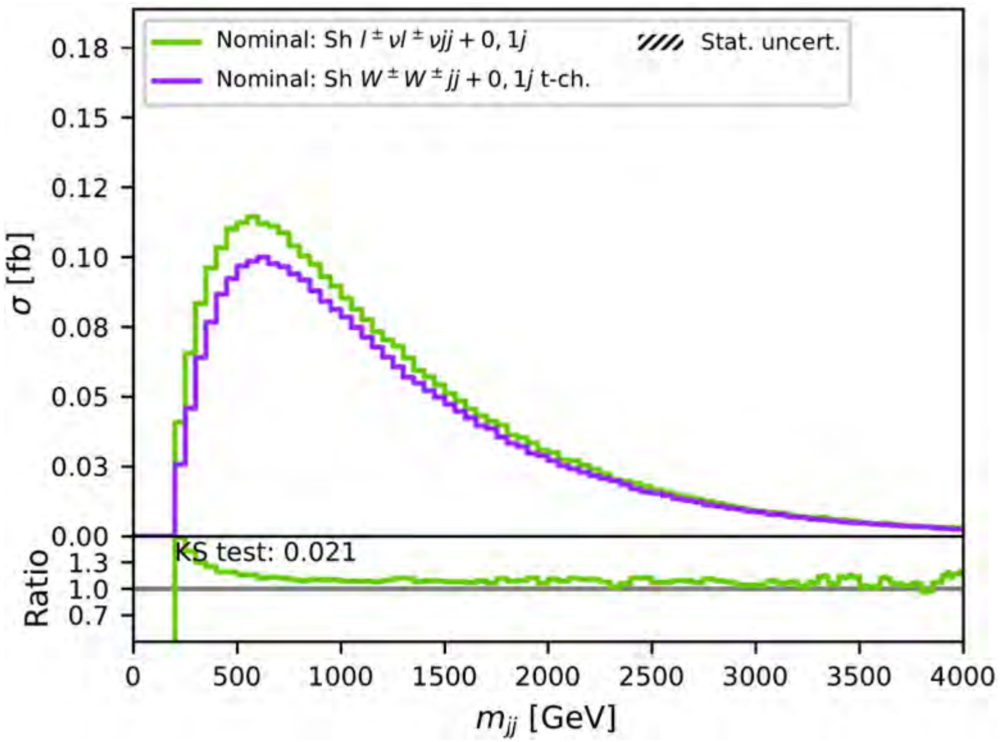
- Full **off-shell** $\boldsymbol{\ell^\pm\ell^\pm\nu\nu jj + 0,1j}$ to correct missing triboson contribution
- **Multivariate reweighting** using deep neural network [arXiv:1907.08209]

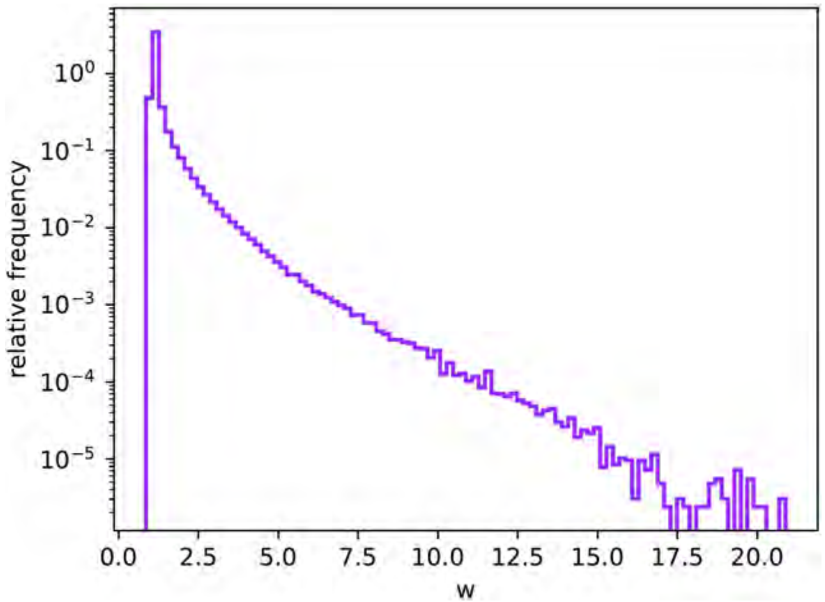


vs.



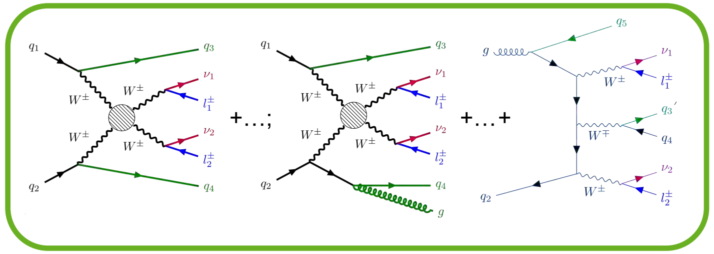
$$\frac{p_1(x)}{p_2(x)} \approx \frac{\mathrm{DNN}(x)}{1-\mathrm{DNN}(x)}$$



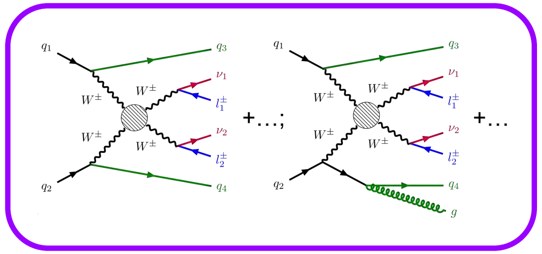



# Correction for missing triboson contribution

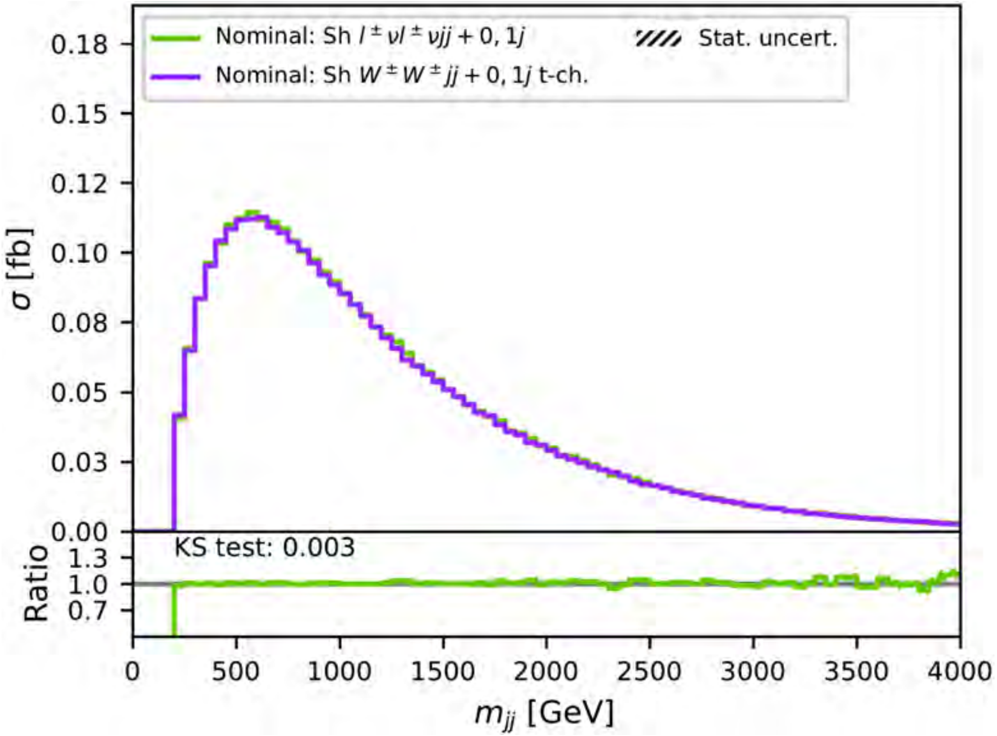
- Full **off-shell** $\boldsymbol{\ell^\pm\ell^\pm\nu\nu jj + 0,1j}$ to correct missing triboson contribution
- **Multivariate reweighting** using deep neural network [arXiv:1907.08209]

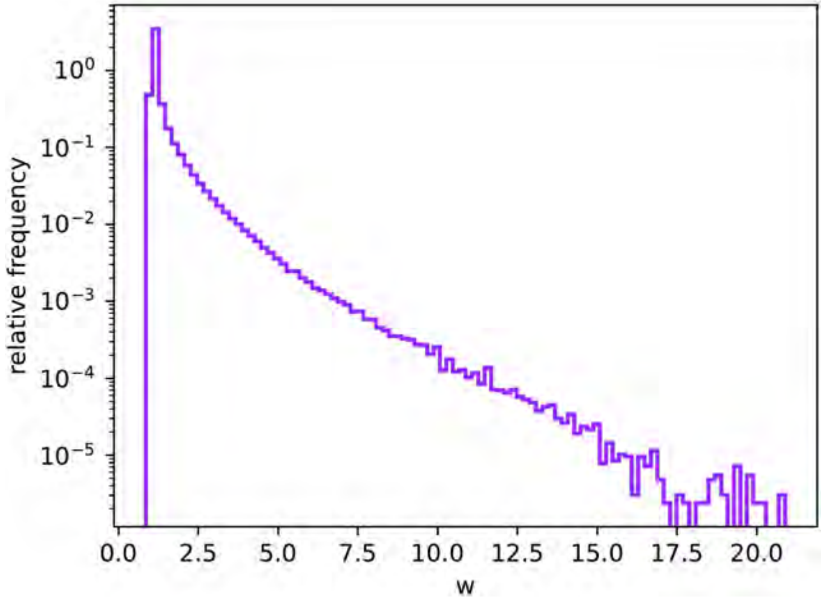


vs.



$$\boxed{\frac{p_1(x)}{p_2(x)} \approx \frac{\mathrm{DNN}(x)}{1-\mathrm{DNN}(x)}}$$



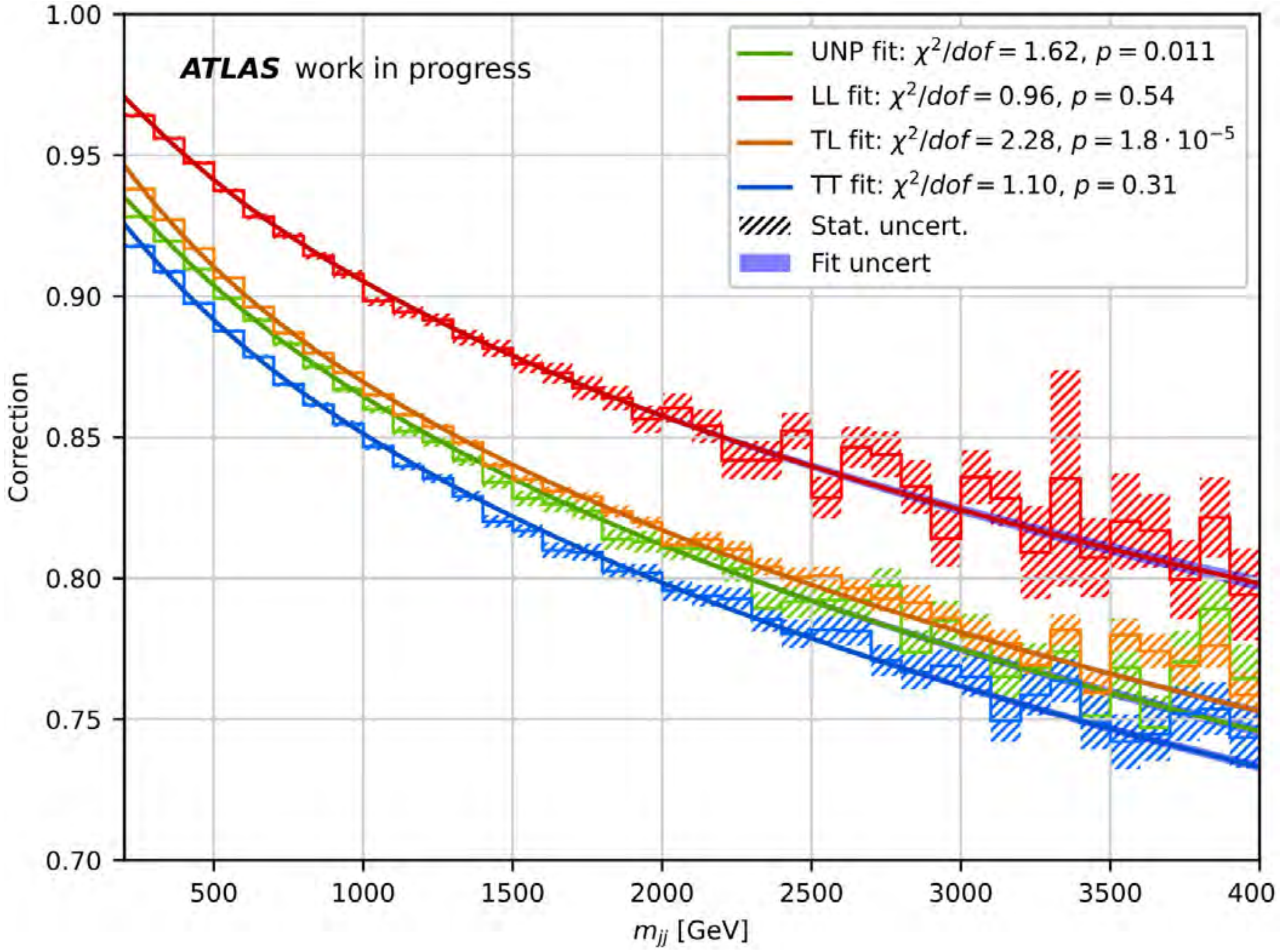



# Higher-order corrections – NLO EW

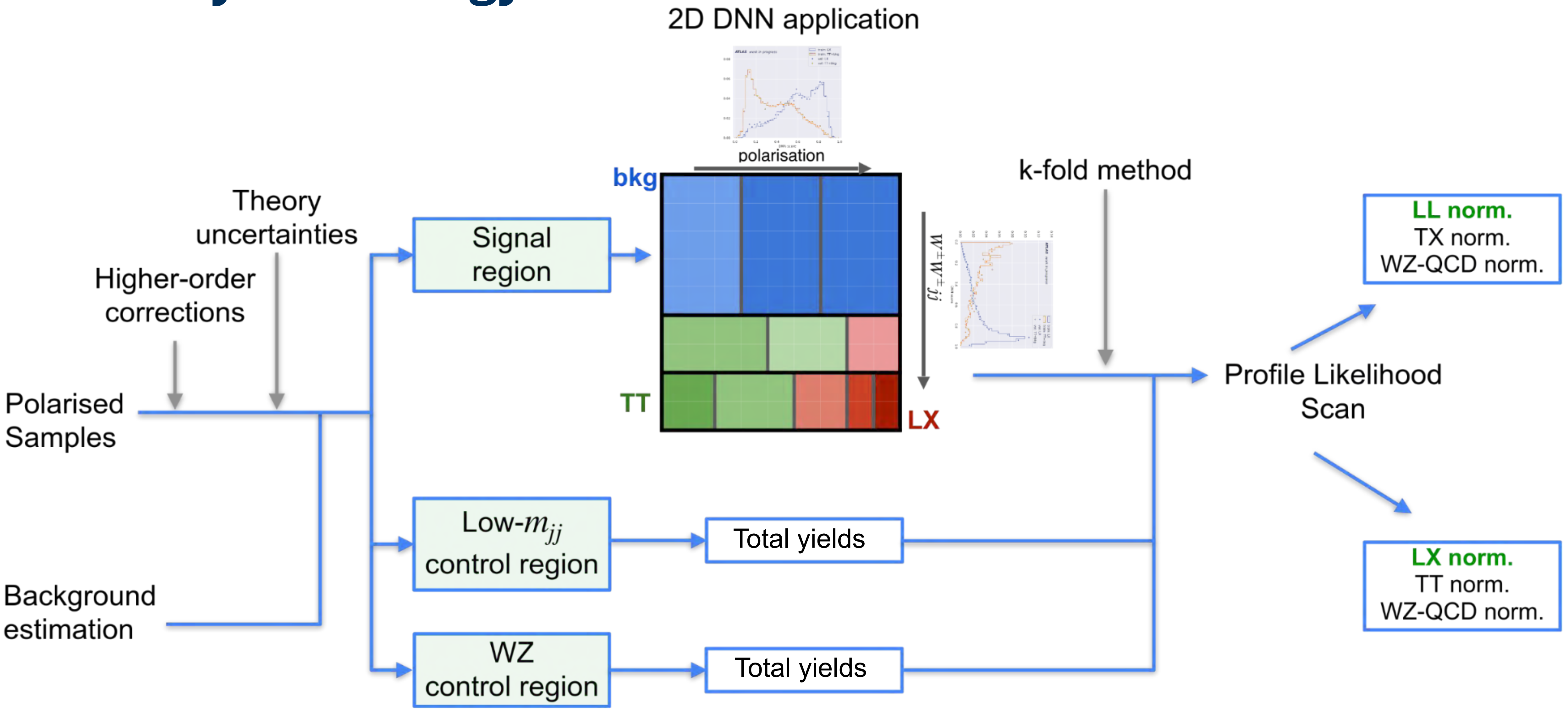
- **Polarized NLO EW corrections** provided by A. Denner et al. [arXiv:2409.03620]

- Perform separate fit for each polarization

$$f(m_{jj}) = p_0 + p_1 \ln\frac{m_{jj}}{\mathrm{GeV}} + p_2 \ln^2\frac{m_{jj}}{\mathrm{GeV}}$$

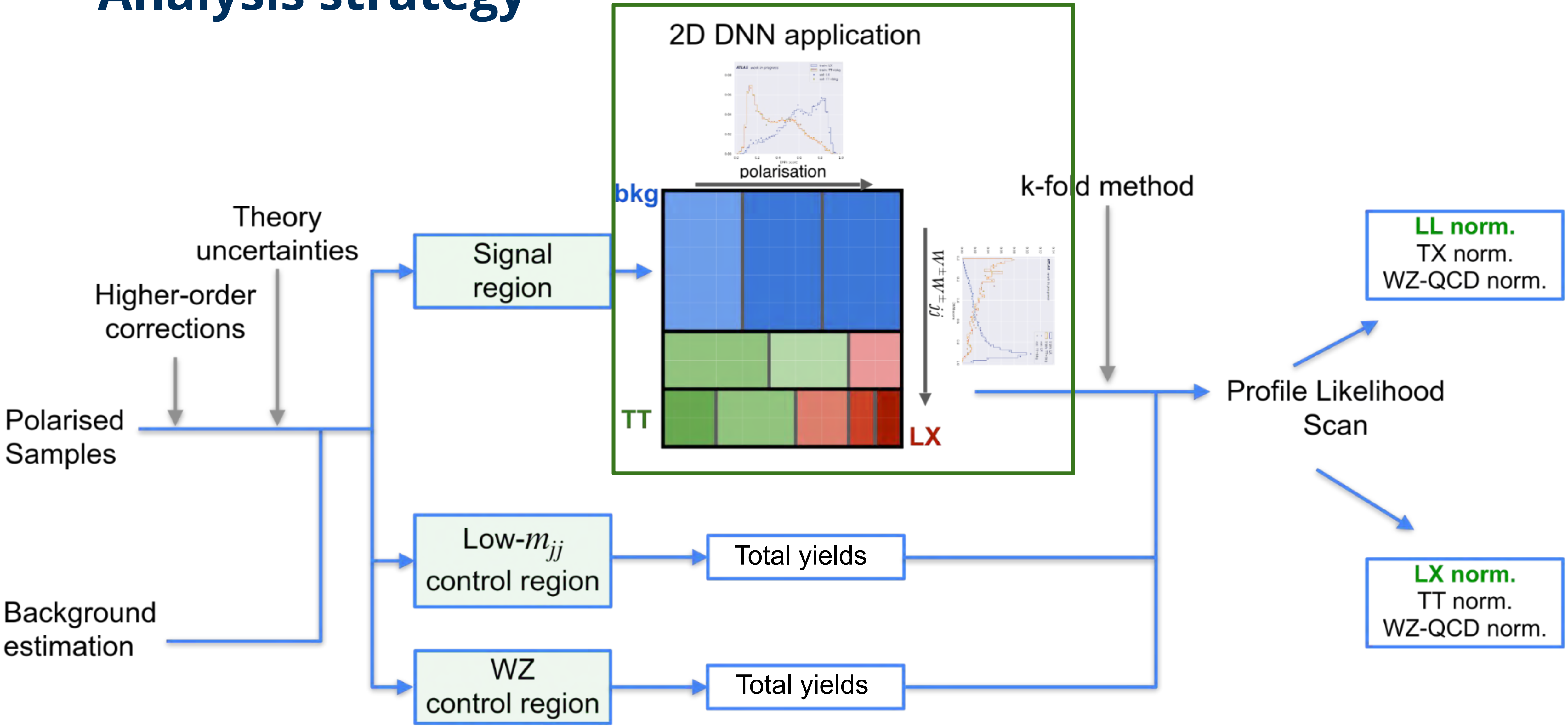


# Analysis strategy



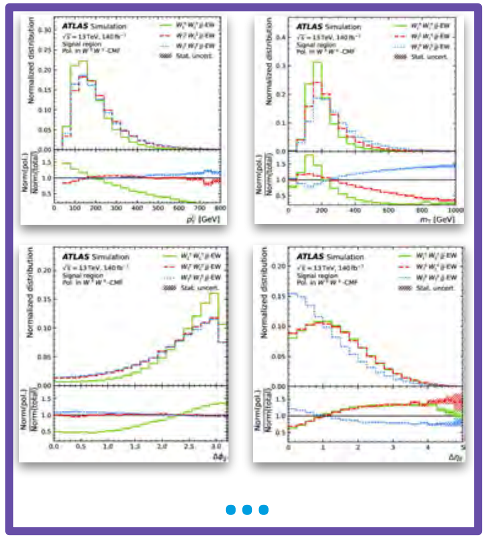

[M. Stange, SM approval meeting.]

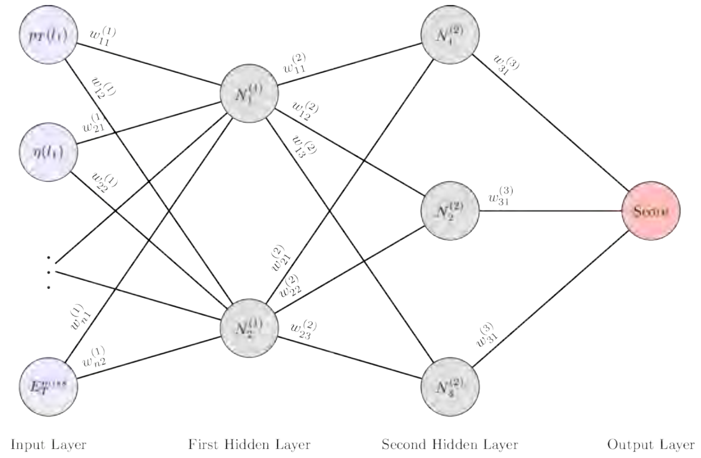
# Analysis strategy

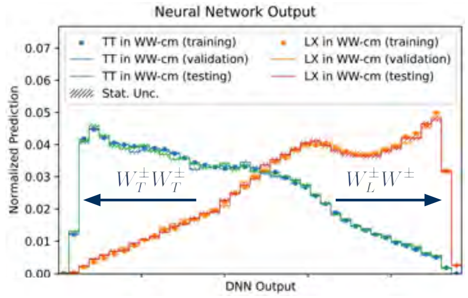



[M. Stange, SM approval meeting.]

# Signal discrimination – Overview

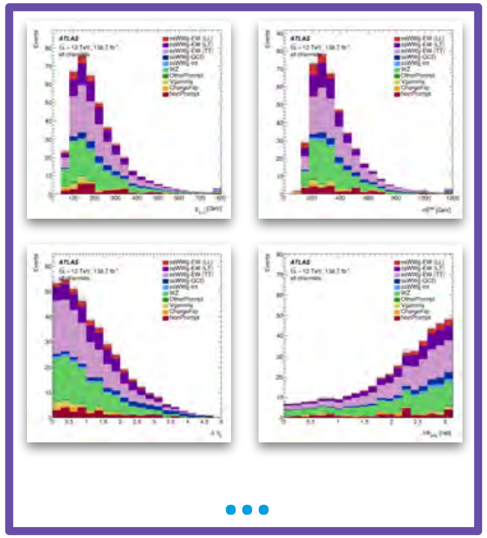


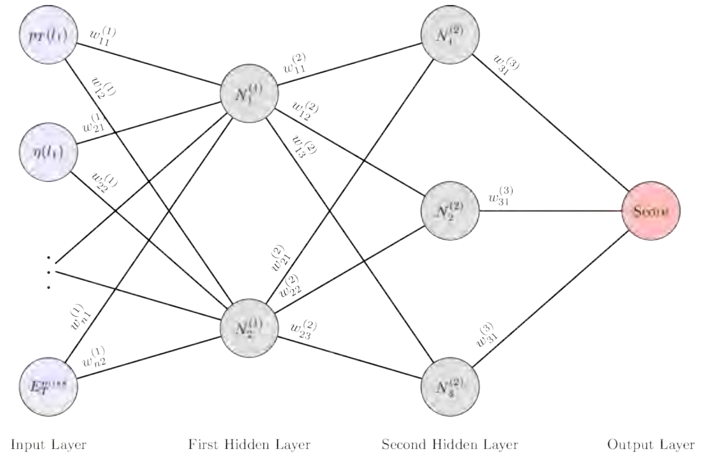


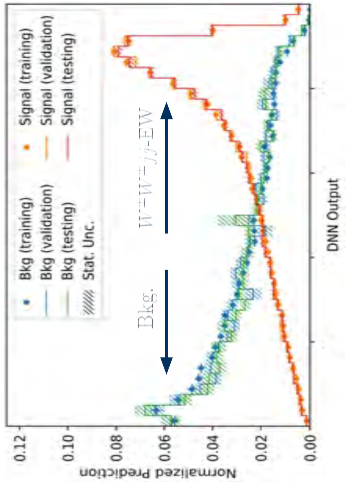



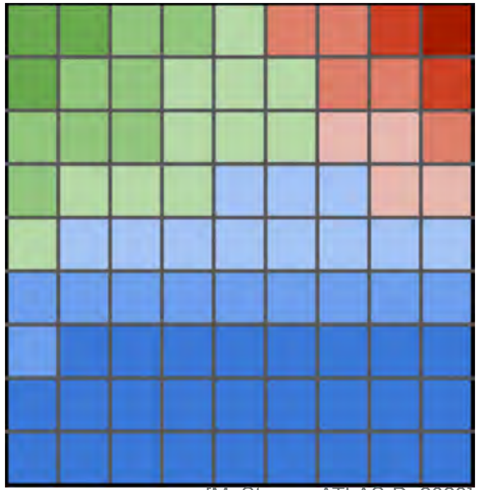

TT

LX





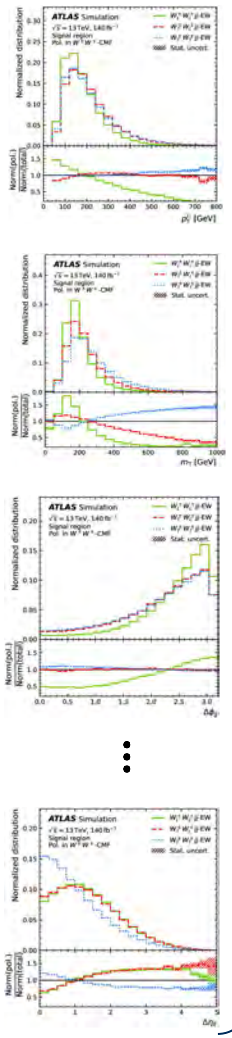



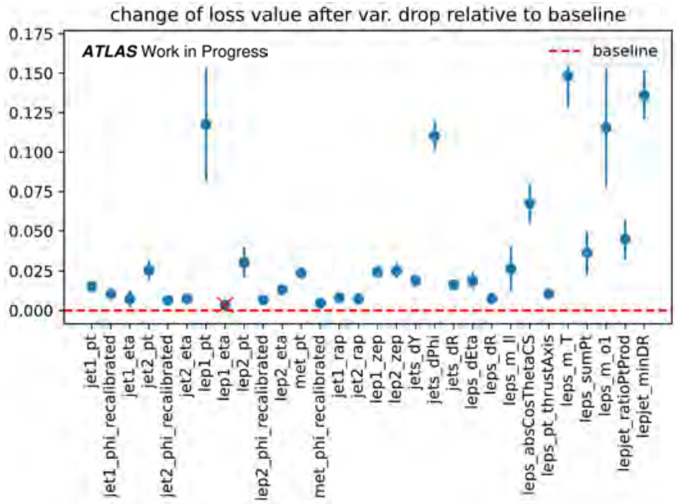
[M. Stange, ATLAS-D, 2023]

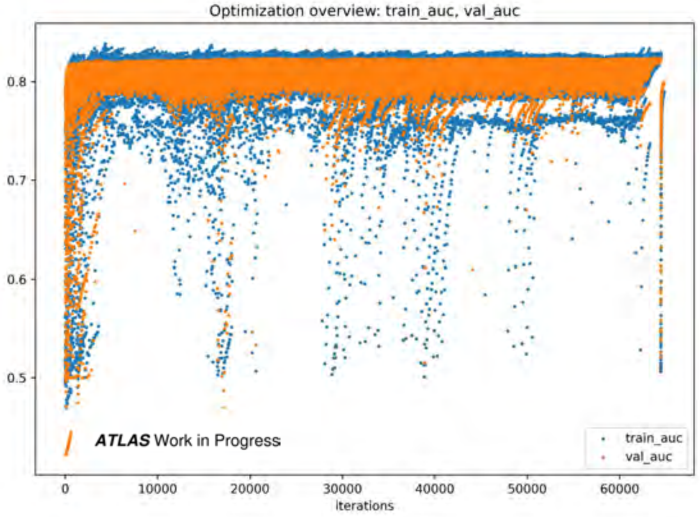
bkg

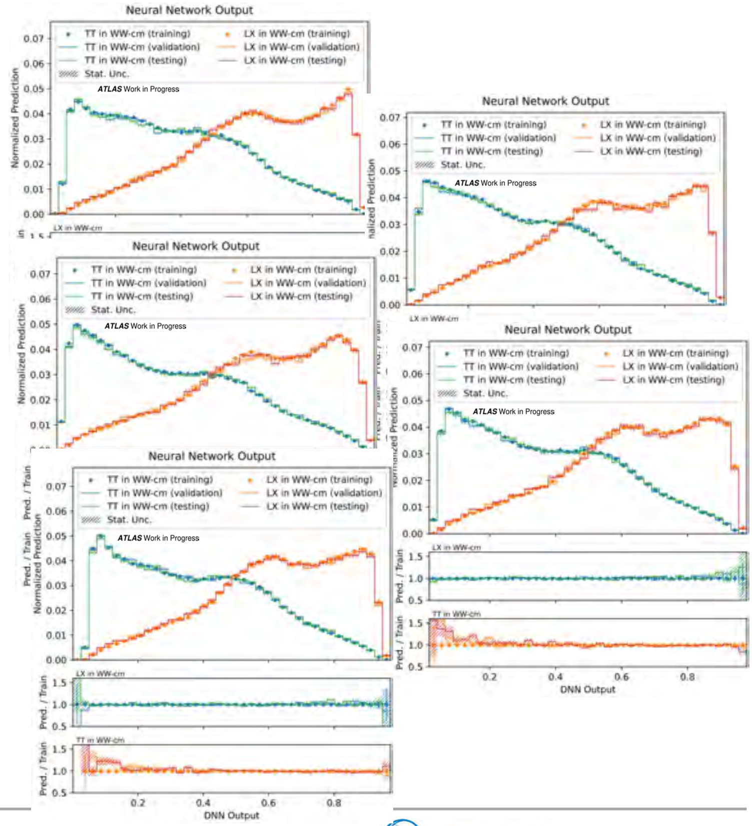
# Signal discrimination – DNN optimization with `optima-ml`



| Input Variable Selection | Hyperparameter optimization (Optuna) | Crossvalidation |
|---|---|---|
| • hyperparameter optimization with all input variables (Optuna)<br>• optimize input variables with fixed hyperparameters using backwards elimination | • find best hyperparameters with optimized input variables | • perform 5-fold cross-validation |



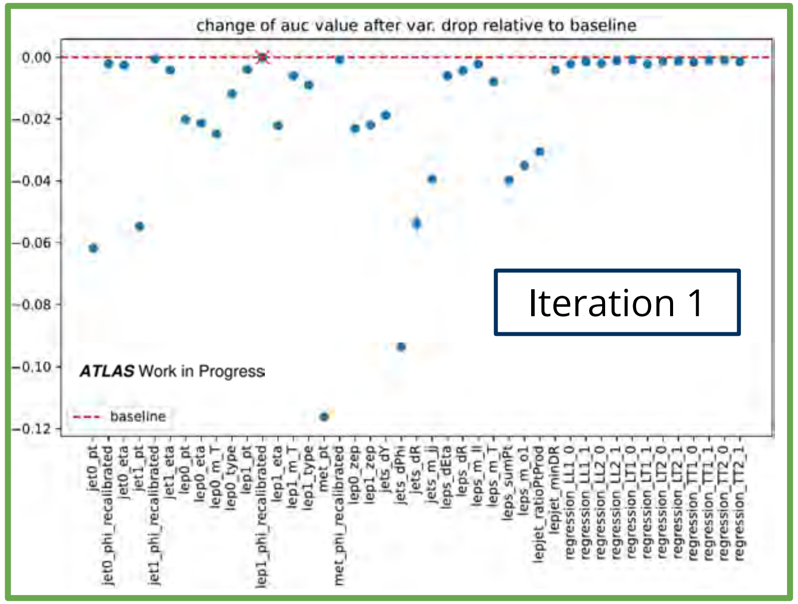



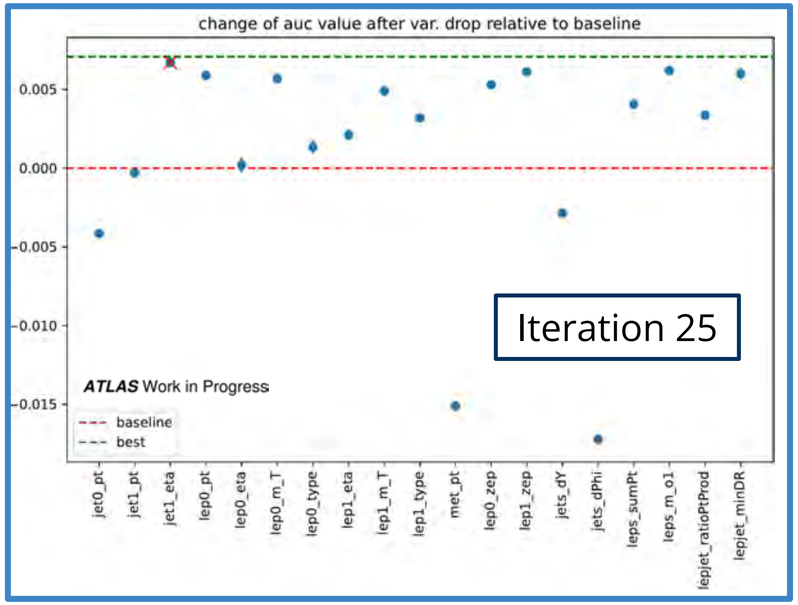



# Signal discrimination – DNN input variable selection

**Backwards Elimination**

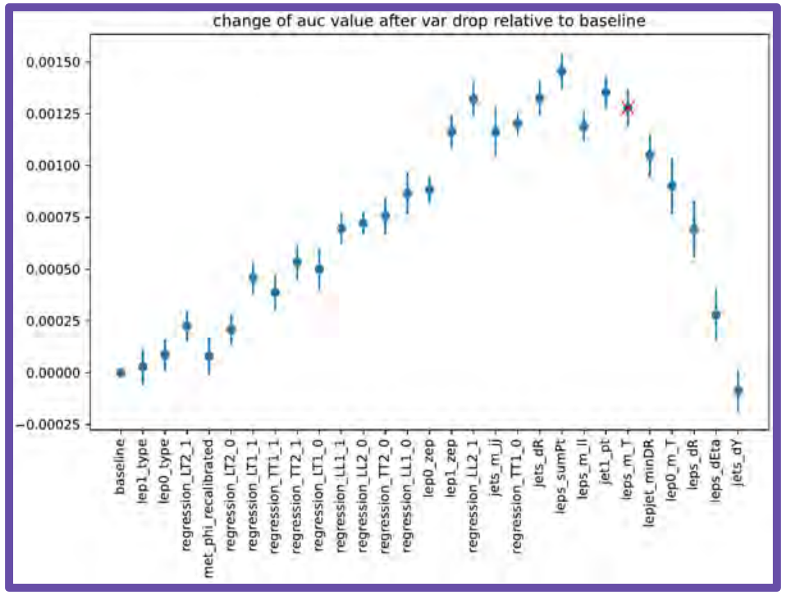
- Baseline ANN with all input variables
- Iteratively remove unimportant/redundant input variables
- Evaluate the importance of an input variable based on change of AUC value:
  - random shuffling between events
    → *fast, but no correlations*
  - remove from training set and retrain
    → *computationally expensive, but can identify redundant variables*
- Retrain the models at the end of each iteration



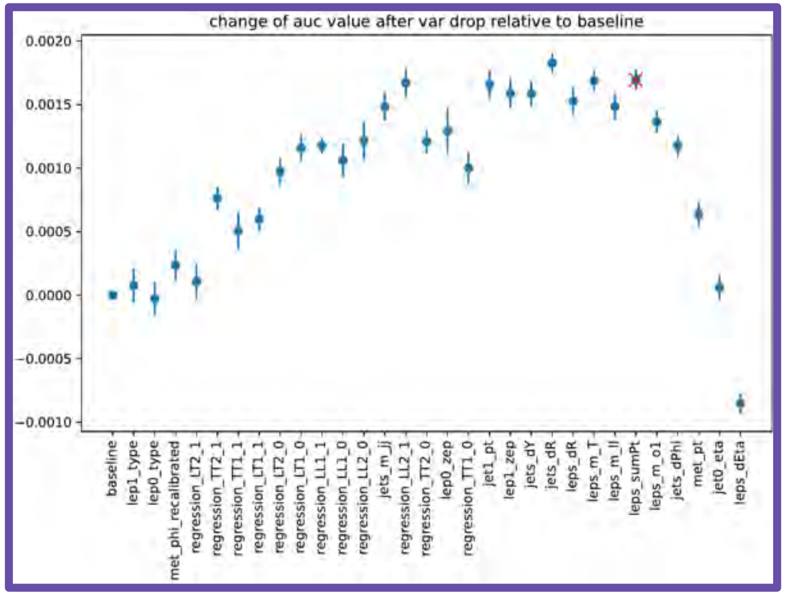



# Signal discrimination – DNN input variable selection

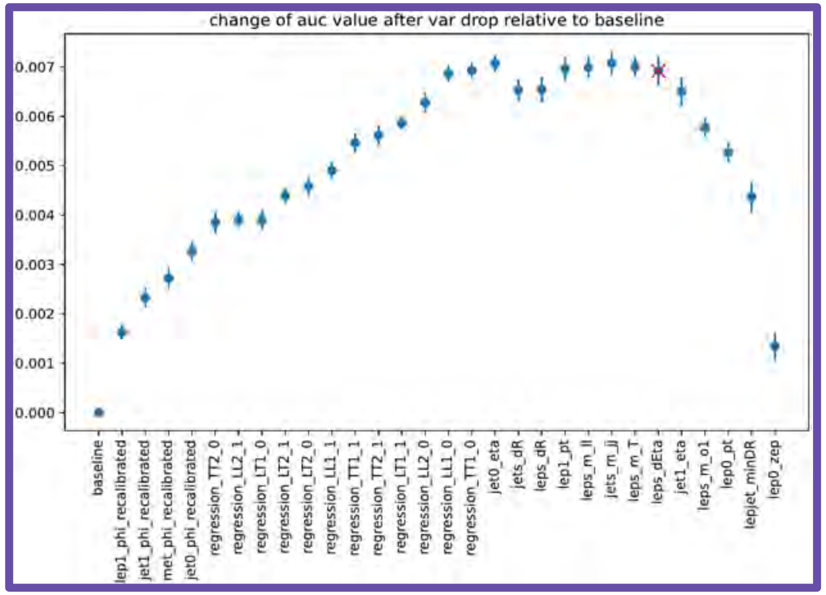
$W^\pm_L W^\pm_L$ vs. $W^\pm_T W^\pm$




$W^\pm_L W^\pm$ vs. $W^\pm_T W^\pm_T$




$W^\pm W^\pm jj$-EW vs. Bkg.



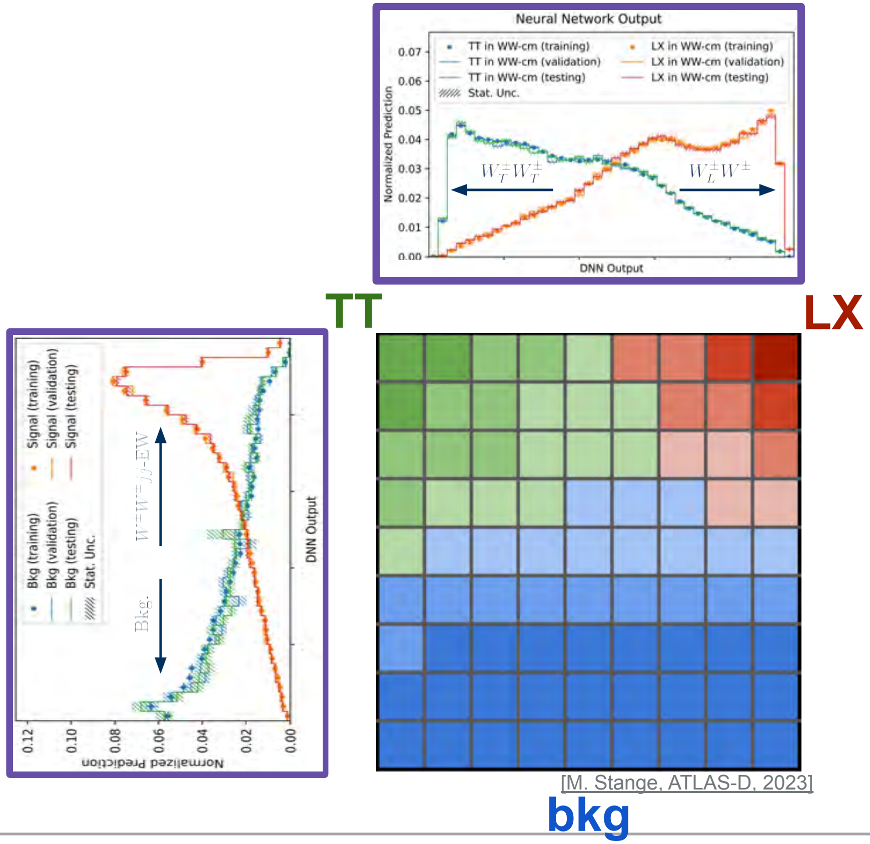
# Signal discrimination – Binning optimization

**Approximated likelihood $\mathcal{L}(\mu_L, \mu_T, \vec{\nu})$**

- $W_L^\pm W_L^\pm$ / $W_L^\pm W^\pm$ signal strength $\mu_L$
- $W_T^\pm W^\pm$ / $W_T^\pm W_T^\pm$ signal strength $\mu_T$
- normalization nuisance parameters $\vec{\nu}$

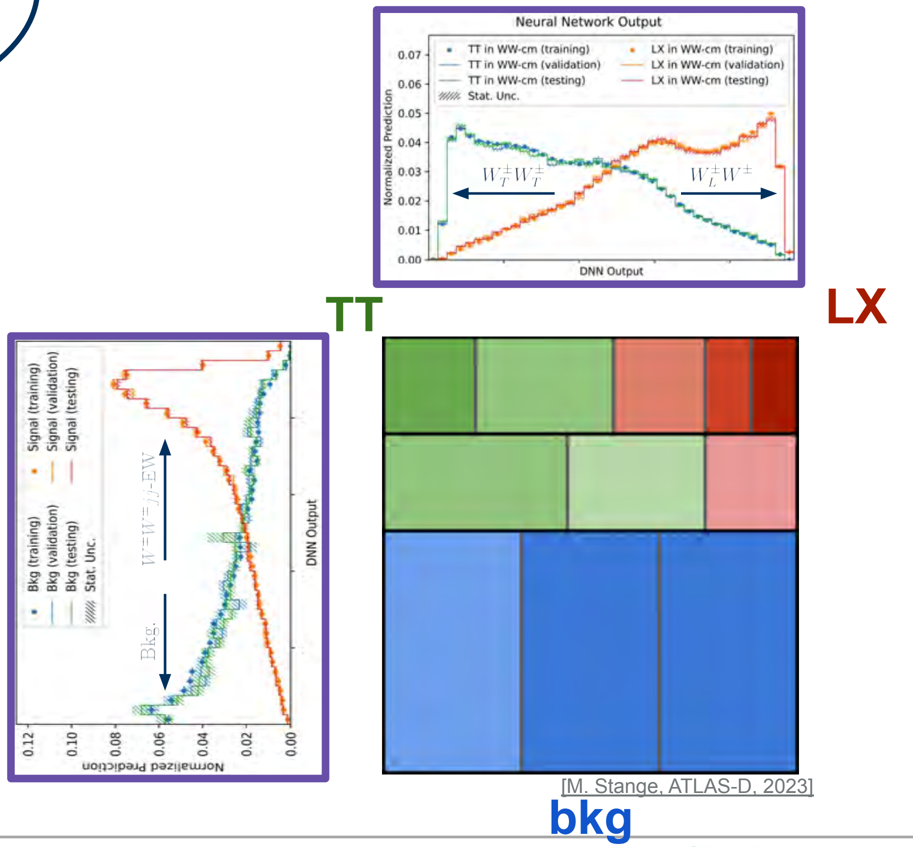
**Approximate significance**

- $q_{0,A} = -2\ln(\mathcal{L}(0, \hat{\hat{\mu}}_T, \hat{\hat{\vec{\nu}}})/\mathcal{L}(1,1,\vec{1}))$
- $Z \approx \sqrt{q_{0,A}}$




[M. Stange, ATLAS-D, 2023]

choose binning that maximizes significance



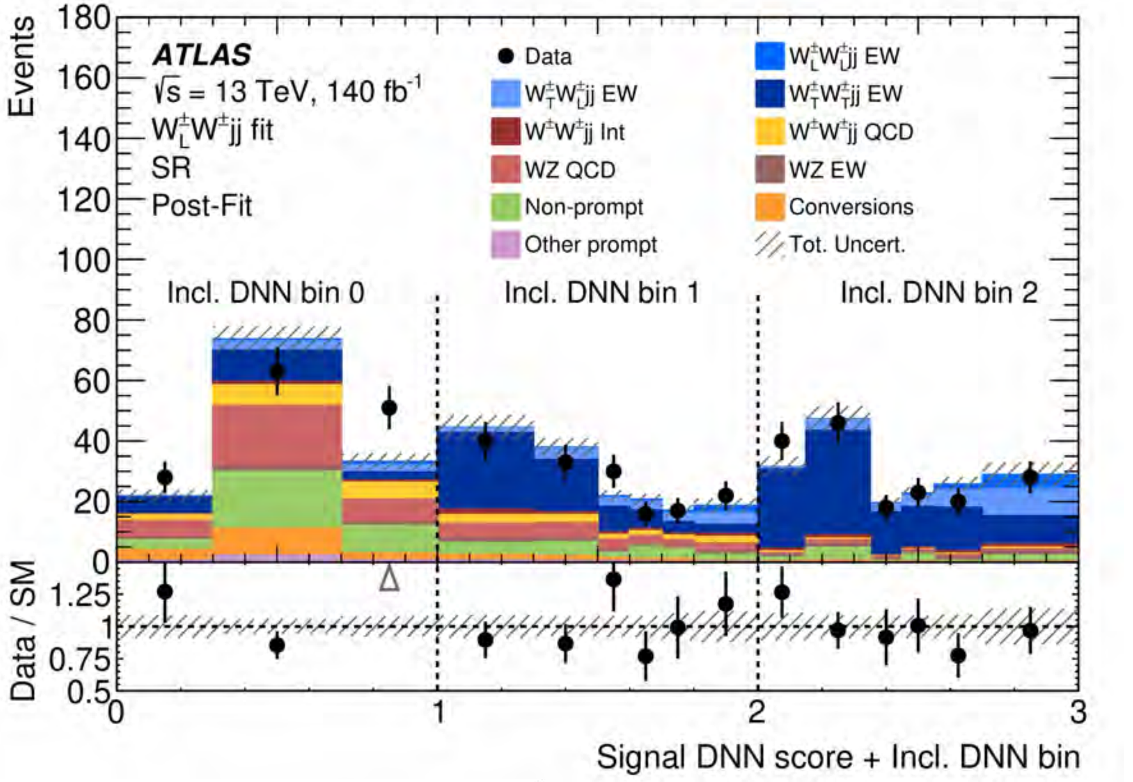

[M. Stange, ATLAS-D, 2023]

# Results – Single Boson polarization

- **Significance of 3.3 $\sigma$** for $W_L^\pm W^\pm jj$ (expected 4.0 $\sigma$)

⇒ **First evidence for longitudinal polarization in vector boson scattering!**

- Measured cross-section in **agreement with the Standard Model**
- Dominated by **statistical uncertainty**



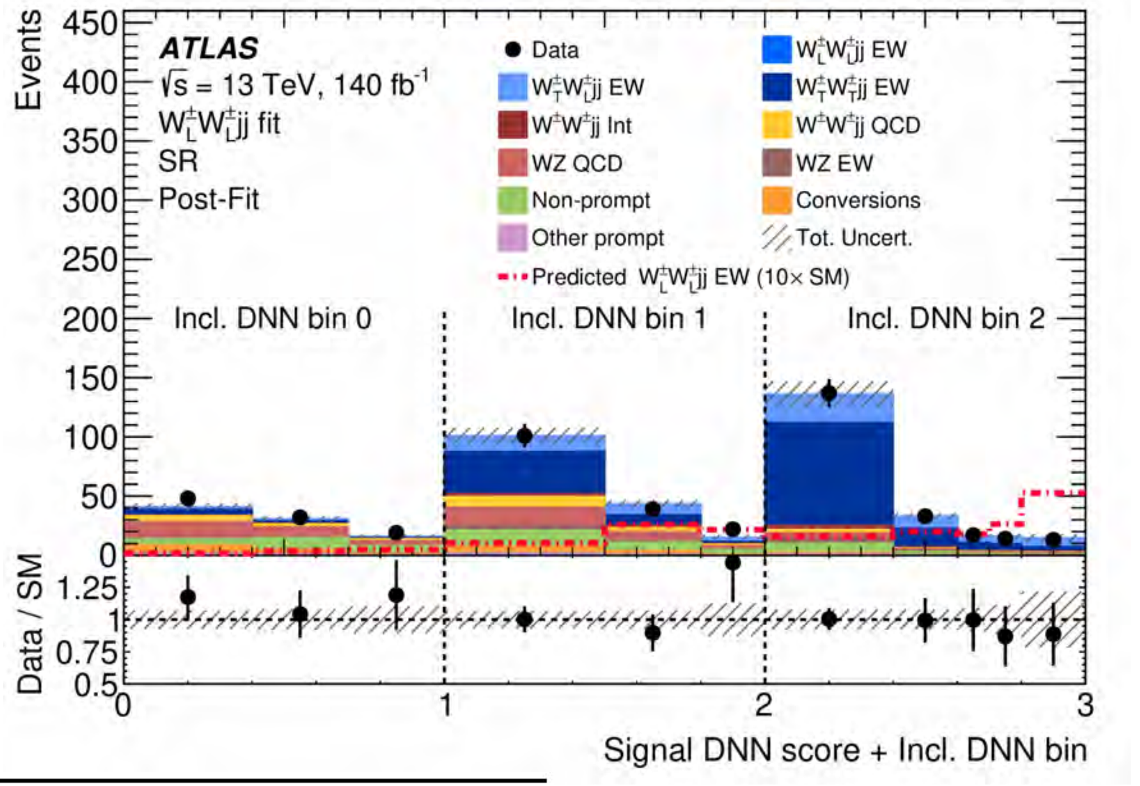

| Predicted $\sigma\mathcal{B}$ (fb) | Measured $\sigma\mathcal{B}$ (fb) | Uncertainty breakdown (fb) |
|---|---|---|
| $1.18 \pm 0.29$ | $0.88 \pm 0.30$ (tot.) | $\pm 0.28$ (stat.) $\pm 0.05$ (mod. syst.) $\pm 0.08$ (exp. syst.) |

# Results – Double Boson polarization

- **95% CL upper limit of 0.45 fb** (expected 0.70 fb)

⇒ **Most stringent limit on fully-longitudinally polarized $W^{\pm}W^{\pm}jj$-EW**

- Measured cross-section in **agreement with the Standard Model**
- Dominated by **statistical uncertainty**

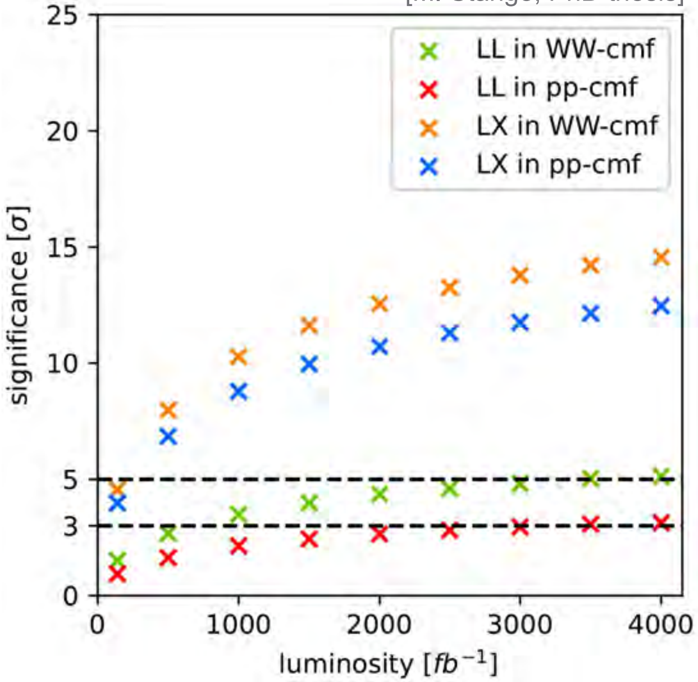



| Predicted $\sigma\mathcal{B}$ (fb) | Measured $\sigma\mathcal{B}$ (fb) | Uncertainty breakdown (fb) |
|---|---|---|
| $0.29 \pm 0.07$ | $0.01 \pm 0.21$ (tot.) | $\pm 0.20$ (stat.) $\pm 0.02$ (mod. syst.) $\pm 0.05$ (exp. syst.) |

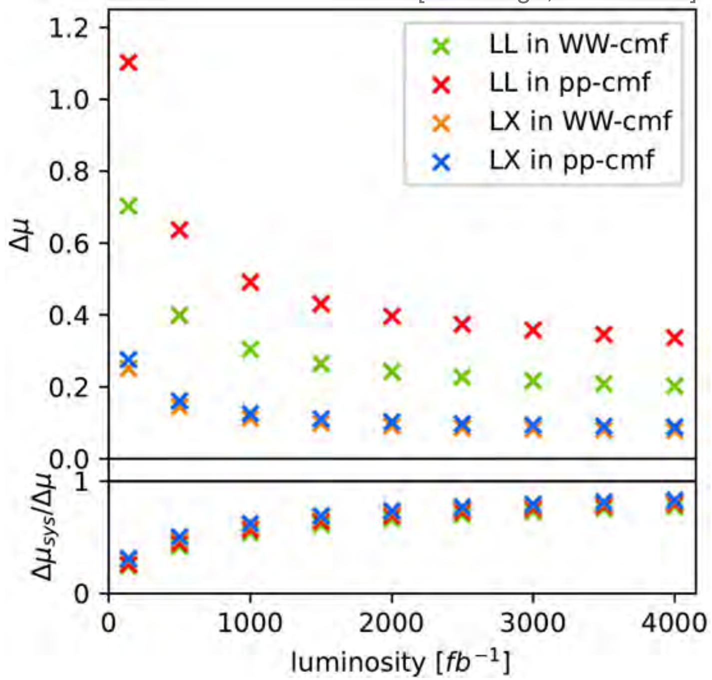
# Projections for Future LHC Runs

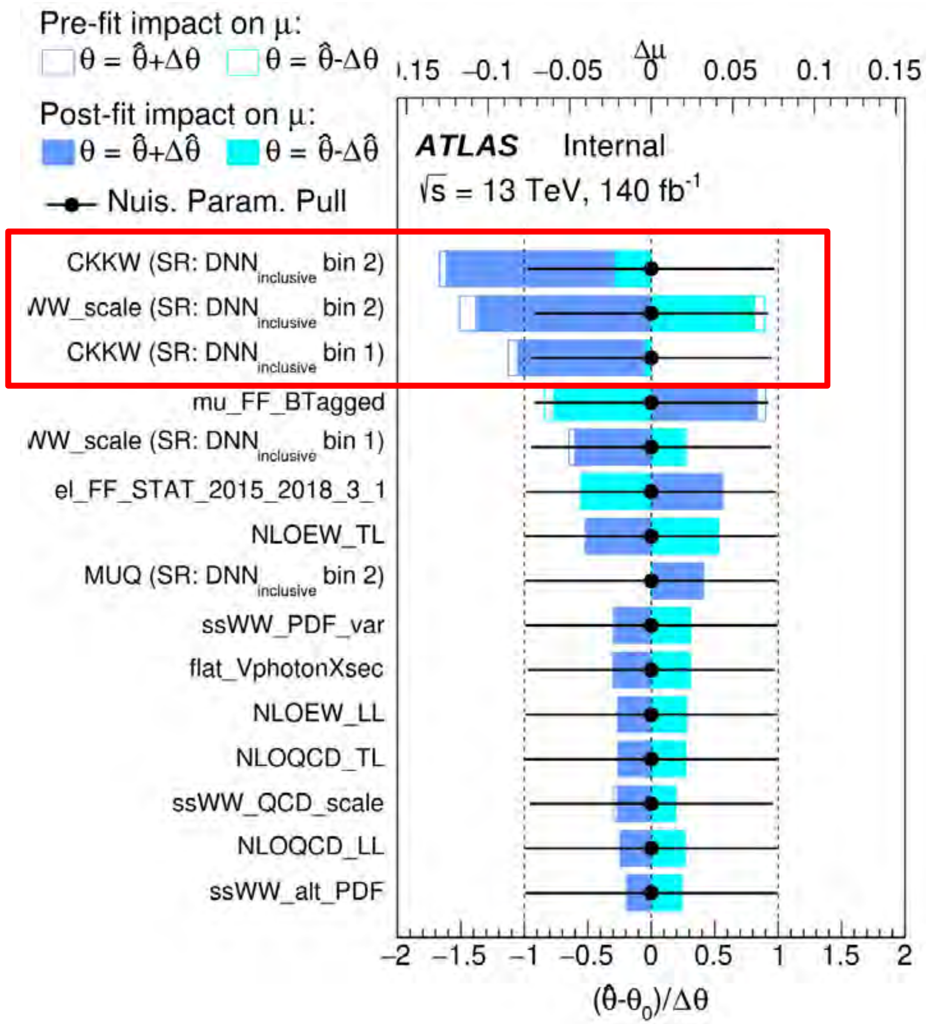
- Conservative estimate: no reduction in systematic uncertainties
- $\mathbf{W_L^\pm W^\pm}$ expected to be **observable with 500 fb$^{-1}$**
- **Potential observation of $\mathbf{W_L^\pm W_L^\pm}$ with 3000 fb$^{-1}$**
  - signal strength limited by **scale uncertainties**

[M. Stange, PhD thesis]




[M. Stange, PhD thesis]



# Conclusion

- State-of-the-art theory predictions
- Sophisticated neural network optimization techniques → published on PyPI (`optima-ml`)
- **First evidence for longitudinal polarization in vector boson scattering: 3.3 (4.0)$\sigma$**
- **Most stringent upper limit on $\mathbf{W^{\pm}_{L}W^{\pm}_{L}jj}$ production cross section: 0.45 (0.70) fb**
- Observation of $W^{\pm}_{L}W^{\pm}_{L}jj$ at the HL-LHC might be possible for ATLAS alone
    - **NLO QCD predictions needed to reduce scale uncertainties**

PHYSICAL REVIEW LETTERS 135, 111802 (2025)

Editors' Suggestion | Featured in Physics

**Evidence for Longitudinally Polarized $W$ Bosons in the Electroweak Production of Same-Sign $W$ Boson Pairs in Association with Two Jets in $pp$ Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector**

G. Aad *et al.*[*]
(ATLAS Collaboration)

(Received 17 March 2025; accepted 9 July 2025; published 10 September 2025)

Physics — VIEWPOINT

**Probing the Higgs Mechanism with Particle Collisions and AI**

A deep neural network has proven essential in confirming a key prediction of one of the standard model's cornerstones.

By Salvatore Rappoccio

# Backup

# How to try OPTIMA?

- Published on PyPI:
  - pip install optima-ml[keras]
  - pip install optima-ml[lightning]
- Source code and detailed usage instructions available on CERN gitlab: gitlab.cern.ch/atlas-germany-dresden-vbs-group/optima
- API documentation at optima-docs.docs.cern.ch
- Roadmap:
  - Support for Keras 3 → Cross-framework support
  - Support for HTCondor

**Any feedback is welcome! Send me a mail at erik.bachmann@tu-dresden.de**

atlas-germany-dres... / **OPTIMA**

## OPTIMA: an Optimization Platform for Tuning Input Variables and Model Parameters

OPTIMA is a framework to perform highly parallelized hyperparameter optimization and input variable selection of arbitrary Keras or Lightning neural networks for supervised learning tasks.

### Table of Contents

# Signal discrimination – DNN training

- Training data extracted from SR
  - NLO QCD and EW corrections to polarized samples (1D QCD, unpol. EW correction)
  - WZ QCD normalization from CR
- low- and high-level observables + regression DNN outputs as inputs
- non-linear scaling to get ~Gaussian shape

| process | raw events | predicted events |
|---|---|---|
| EW6 LL $WW$-CM | 215008 | 18.29 |
| EW6 TL $WW$-CM | 170387 | 58.88 |
| EW6 TT $WW$-CM | 456456 | 124.50 |
| EW6 LL p-CM | 19673 | 11.49 |
| EW6 TL p-CM | 21471 | 67.84 |
| EW6 TT p-CM | 46320 | 123.07 |
| ssWW EWK MGH7 | 201488 | 206.52 |
| ssWW QCD | 25375 | 24.05 |
| ssWW Int | 79805 | 7.57 |
| WZ EWK | 8273 | 14.95 |
| WZ QCD Sherpa222 | 27800 | 82.75 → 28.50 |
| WZ QCD Sherpa2212 | 22199 | 76.79 → 28.50 |
| top | 2192 | 5.02 |
| ZZ | 2880 | 2.51 |
| chFlip | 8231 | 10.10 |

| Kinematics | Descriptions | scaling |
|---|---|---|
| | Low-level variables | |
| $p_T^{l1}$ | $p_T$ of the leading lepton | $\log_{10}(x)$ |
| $\eta^{l1}$ | $\eta$ of the leading lepton | $x$ |
| $l1$ type | type of the leading lepton | $x$ |
| $p_T^{l2}$ | $p_T$ of the subleading lepton | $\log_{10}(x)$ |
| $\eta^{l2}$ | $\eta$ of the subleading lepton | $x$ |
| $\phi^{l2} - \phi^{l1}$ | Recalibrated $\phi$ of the subleading lepton | $x$ |
| $l2$ type | type of the subleading lepton | $x$ |
| $p_T^{j1}$ | $p_T$ of the leading jet | $\log_{10}(x)$ |
| $\eta^{j1}$ | $\eta$ of the leading jet | $x$ |
| $\phi^{j1} - \phi^{l1}$ | Recalibrated $\phi$ of the leading jet | $x$ |
| $p_T^{j2}$ | $p_T$ of the subleading jet | $\log_{10}(x)$ |
| $\eta^{j2}$ | $\eta$ of the subleading jet | $x$ |
| $\phi^{j2} - \phi^{l1}$ | Recalibrated $\phi$ of the subleading jet | $x$ |
| $p_T^{miss}$ | Missing transverse momentum | $\log_{10}(x)$ |
| $\phi(p_T^{miss}) - \phi^{l1}$ | Recalibrated $\phi$ of the missing transverse energy | $x$ |
| | High-level variables | |
| $Z_{l1}^{\star}$ | Zeppenfeld variable of the leading lepton | $\sqrt{x}$ |
| $Z_{l2}^{\star}$ | Zeppenfeld variable of the subleading lepton | $\sqrt{x}$ |
| $M_T^{l1,MET}$ | Transverse mass of the leading lepton and $p_T^{miss}$ | $\sqrt{x}$ |
| $M_T^{l2,MET}$ | Transverse mass of the subleading lepton and $p_T^{miss}$ | $\sqrt{x}$ |
| $\Delta R_{l1,l2}$ | $\Delta R$ between the two leading leptons | $x$ |
| $\Delta\eta_{l1,l2}$ | $\Delta\eta$ between the two leading leptons | $\sqrt{x}$ |
| $M_{l1,l2}$ | Invariant mass of the two leading leptons | $\log_{10}(x)$ |
| $p_T^{l1,l2}$ | $p_T$ of the dilepton system | $\sqrt{x}$ |
| $M_T^{l1,l2,MET}$ | Transverse mass of the dilepton system | $\sqrt{x}$ |
| $M_{o1}^{l1,l2,MET}$ | Early-projected massless invariant mass of dilepton system and missing transverse energy | $\sqrt{x}$ |
| $\Delta R_{j1,j2}$ | $\Delta R$ between the two leading jets | $x$ |
| $\Delta y_{j1,j2}$ | $\Delta y$ between the two leading jets | $x$ |
| $M_{j1,j2}$ | Invariant mass of the two leading jets | $\log_{10}(x)$ |
| $\Delta\phi_{j1,j2}$ | $\Delta\phi$ between the two leading jets | $x$ |
| $(p_T^{l1} \cdot p_T^{l2})/(p_T^{j1} \cdot p_T^{j2})$ | $p_T$ ratio of leptons and jets | $\log_{10}(x+0.02)$ |
| $min(\Delta R_{l1/l2,j1/j2})$ | Minimal $\Delta R$ between the leptons and jets | $x$ |

# OPTIMA

## Hyperparameter optimization

**Hyperparameter search algorithm**

hyperparameters

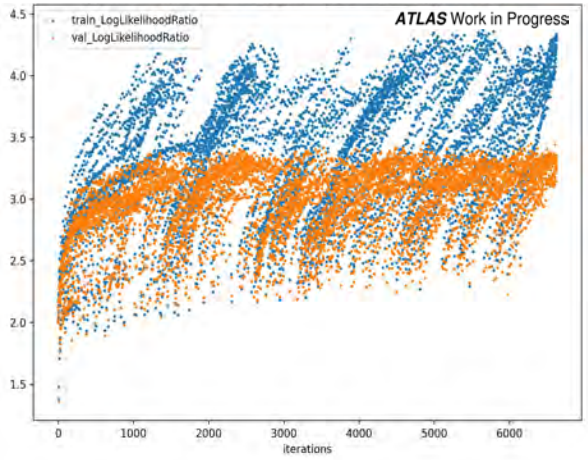
**trials**

callback

***After every epoch***
- calculate target and auxiliary metrics
- evaluate overtraining constraints
- Early Stopping

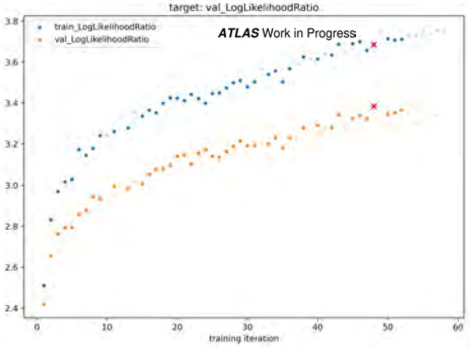
target and auxiliary metrics (every epoch)

**metric values of all trials**

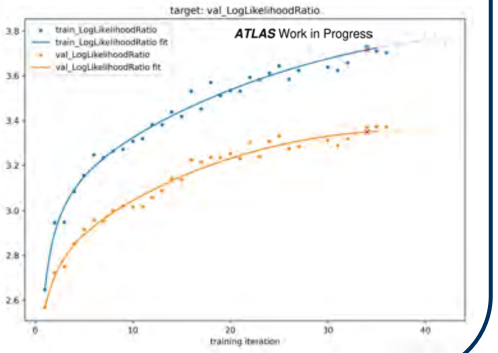



**evaluation**

best value



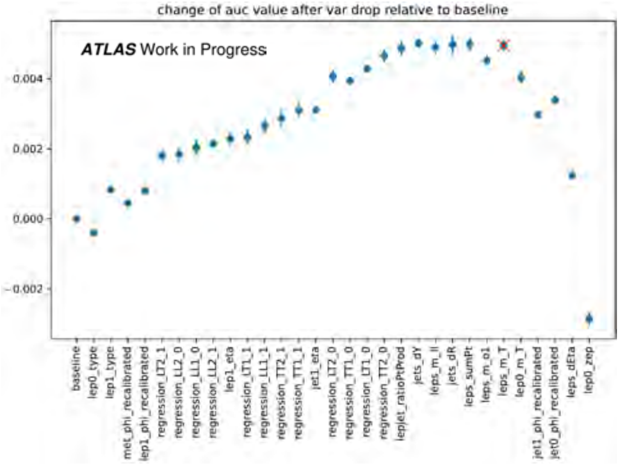

best fit



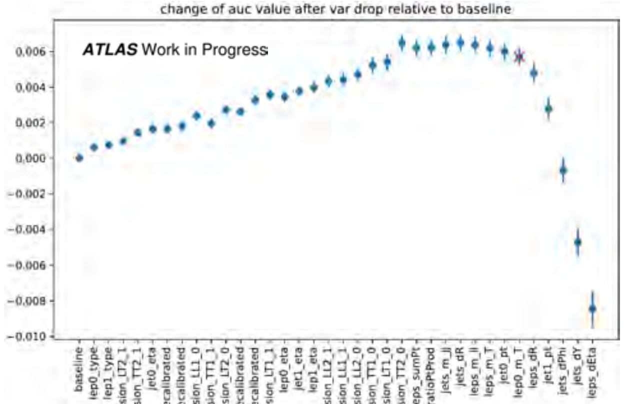

**K-fold cross-validation**

# Signal discrimination – DNN input variable selection

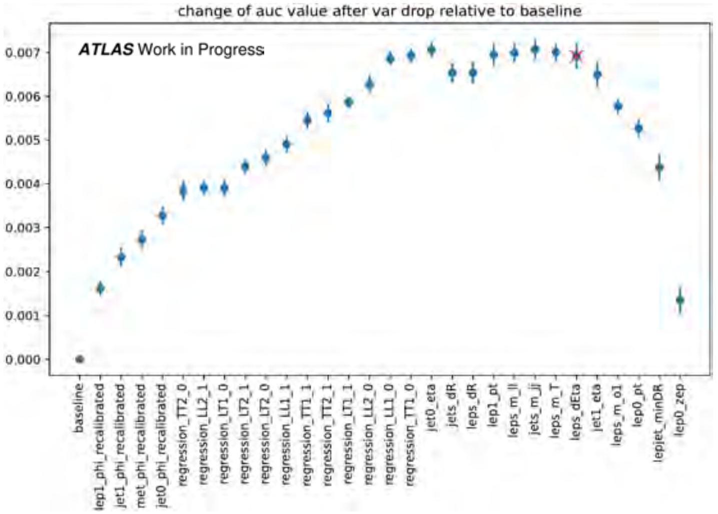
## LL vs TX in p-cmf



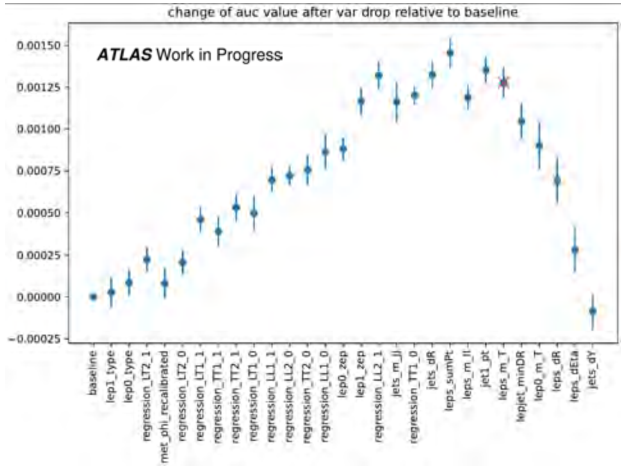

## LX vs TT in p-cmf



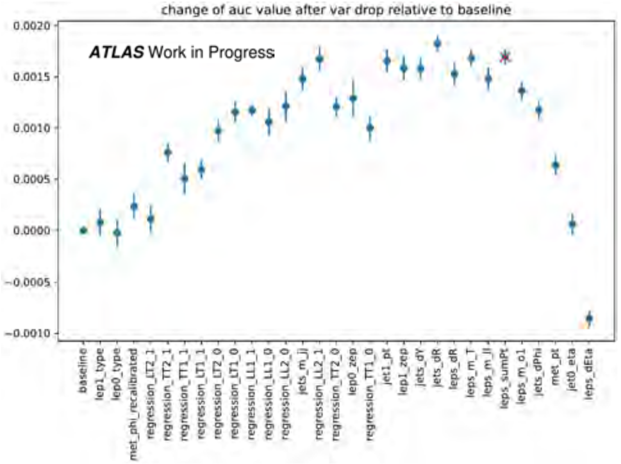

## ssWW vs Bkg




## LL vs TX in WW-cmf

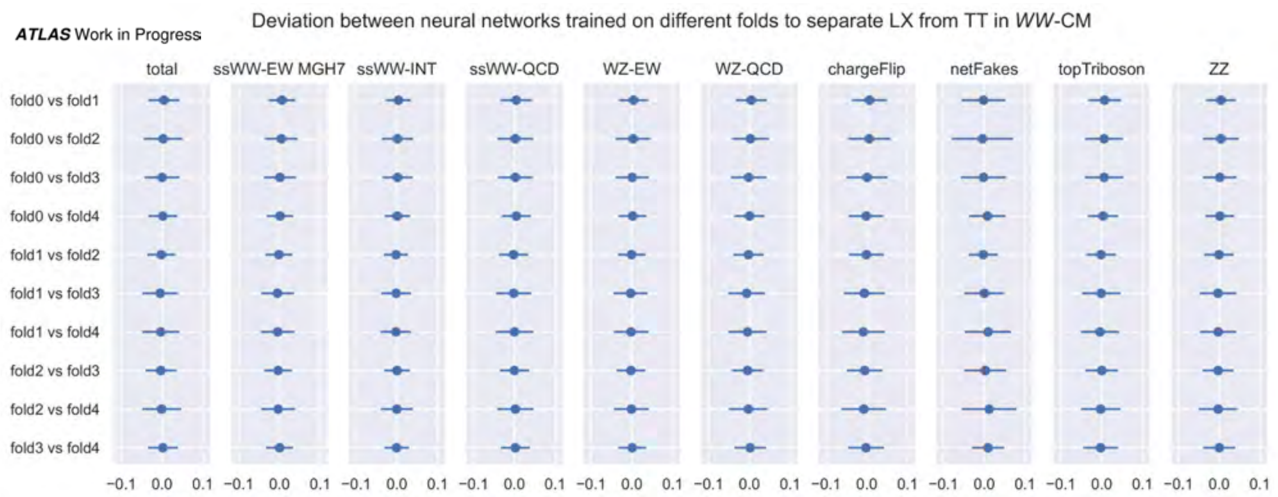


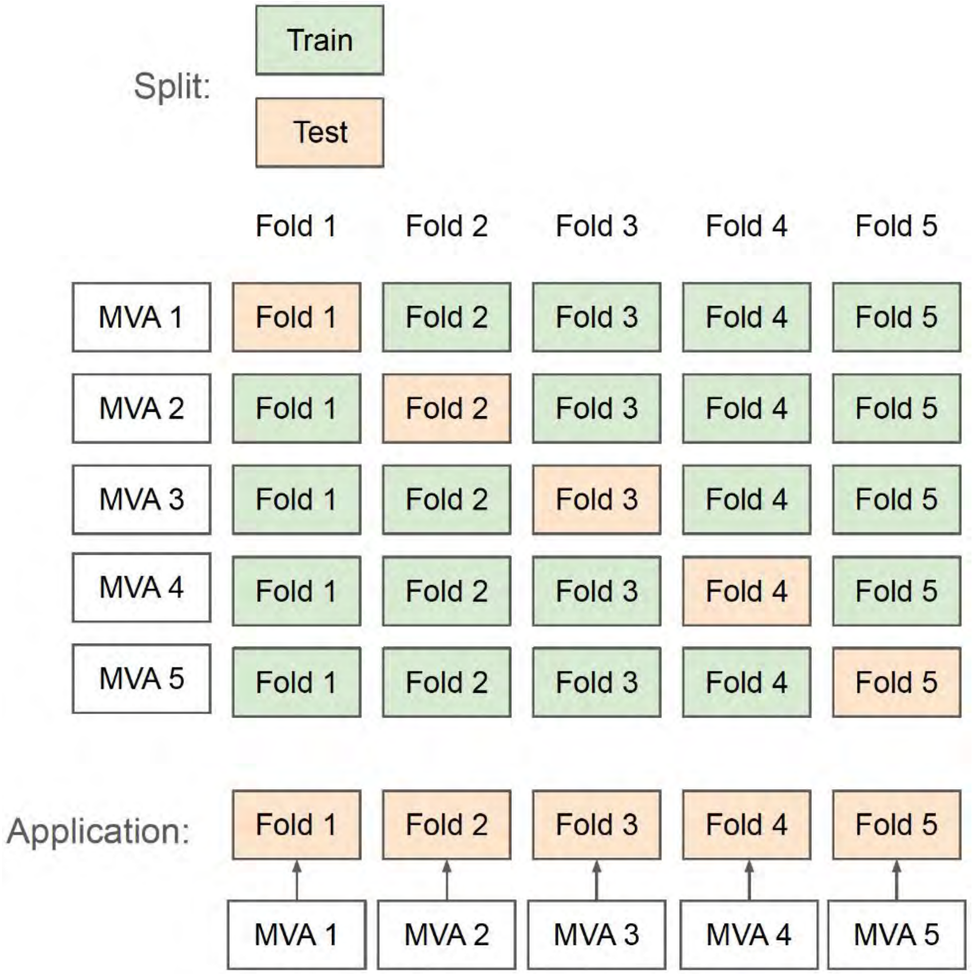

## LX vs TT in WW-cmf



# Signal discrimination – k-fold method

- MVA will always perform better on training data than on new, unseen data
  $\rightarrow$ **Apply MVA only on unseen data to avoid mismodelling**
- Use 5-fold splitting to train 5 MVAs on different splits
  $\rightarrow$ 80% of available data used to train each model
- During application, apply each model only on the testing dataset




[M. Stange, EB meeting.]