# Employing Matrix Elements with Neural Networks to Search for Higgs Self-coupling

DPG Spring Conference Karlsruhe
7.3.2024

Christoph Ames

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

FSP ATLAS
Erforschung von Universum und Materie

# Matrix element method

LMU
LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Weight of decay mode X in an event: $W_X = \left| \langle \Phi_f | \hat{H}_X | \Phi_i \rangle_{d\Psi} \right|^2$

- Integrate over all initial states
- Hamilton operator for chosen decay mode X
- Select a final state (particle types, four-momenta)
- Physically allowed phase space

➡ Weight describes likelihood that decay mode X is contained in the event

# Machine learning

$$W_X = \left| \langle \Phi_f | \hat{H}_X | \Phi_i \rangle_{d\Psi} \right|^2$$

Problem: Very high number of initial states in hadron collider (parton density functions)
➡ Integration requires large amount of computing time

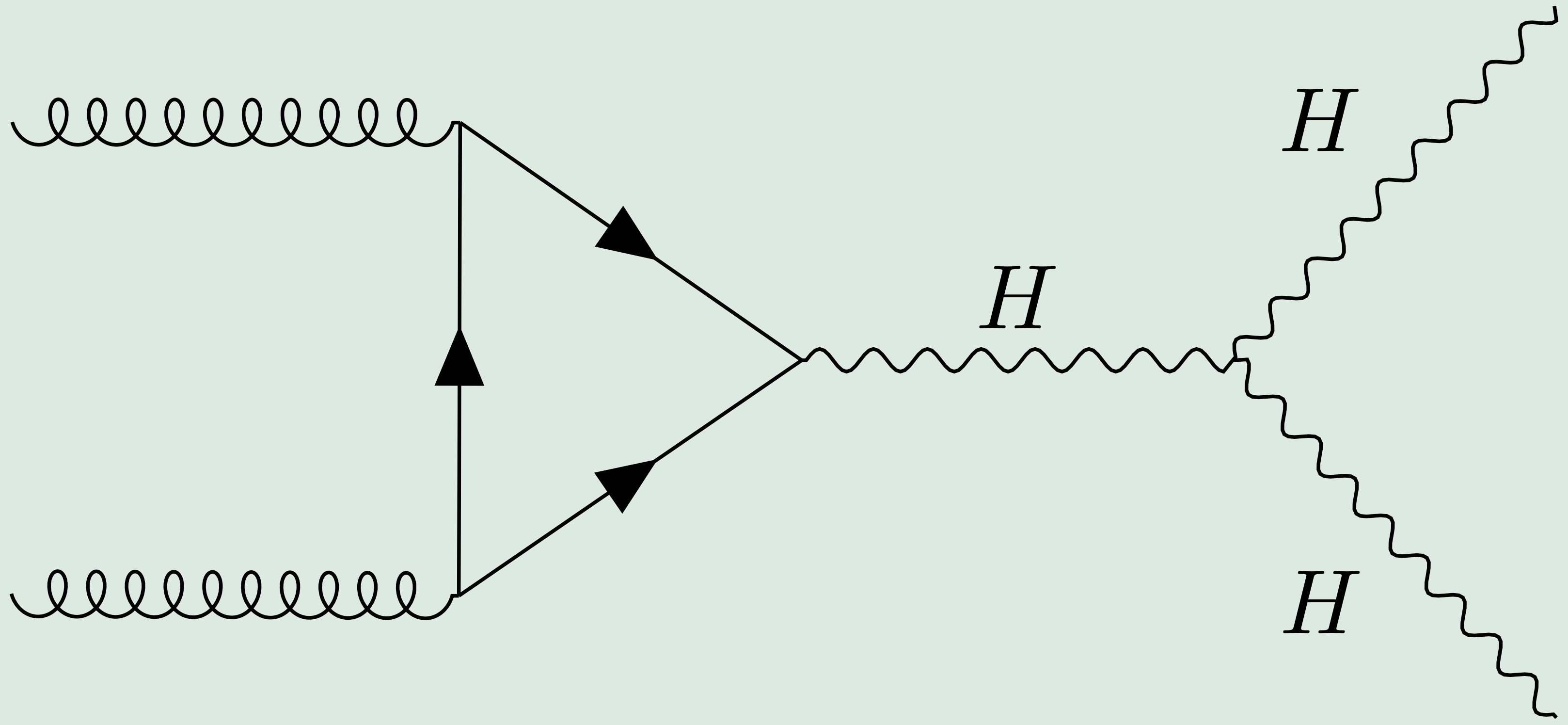
Solution: Use machine learning to estimate the weight based on a given final state
➡ No integration required, much quicker

# Objective of the analysis

Objective: accelerate the matrix element method using machine learning

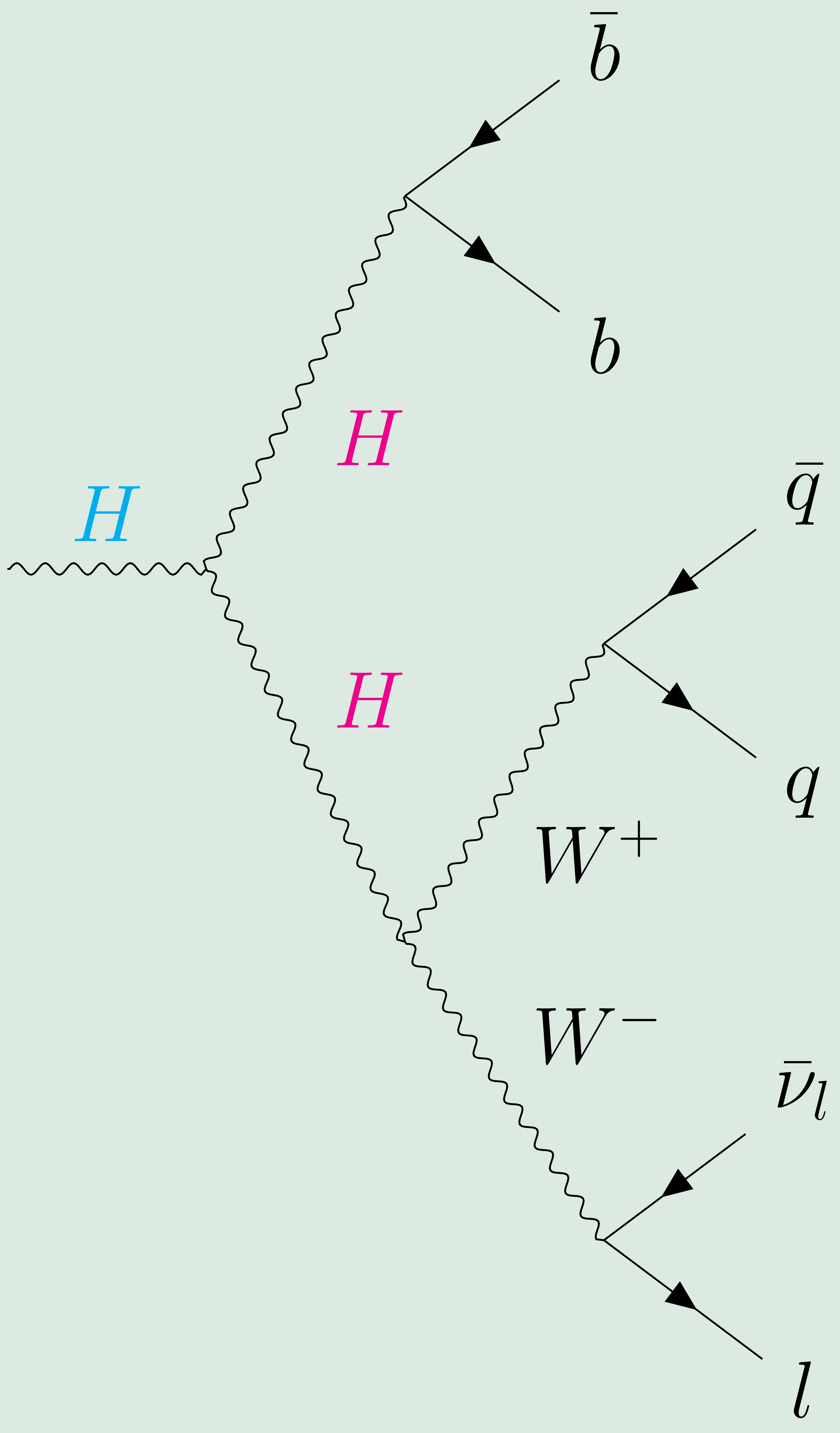
If successful, this concept could be used to pre-filter a large portion of background events for future analyses

Test analysis: search for Higgs self-coupling (heavily suppressed)



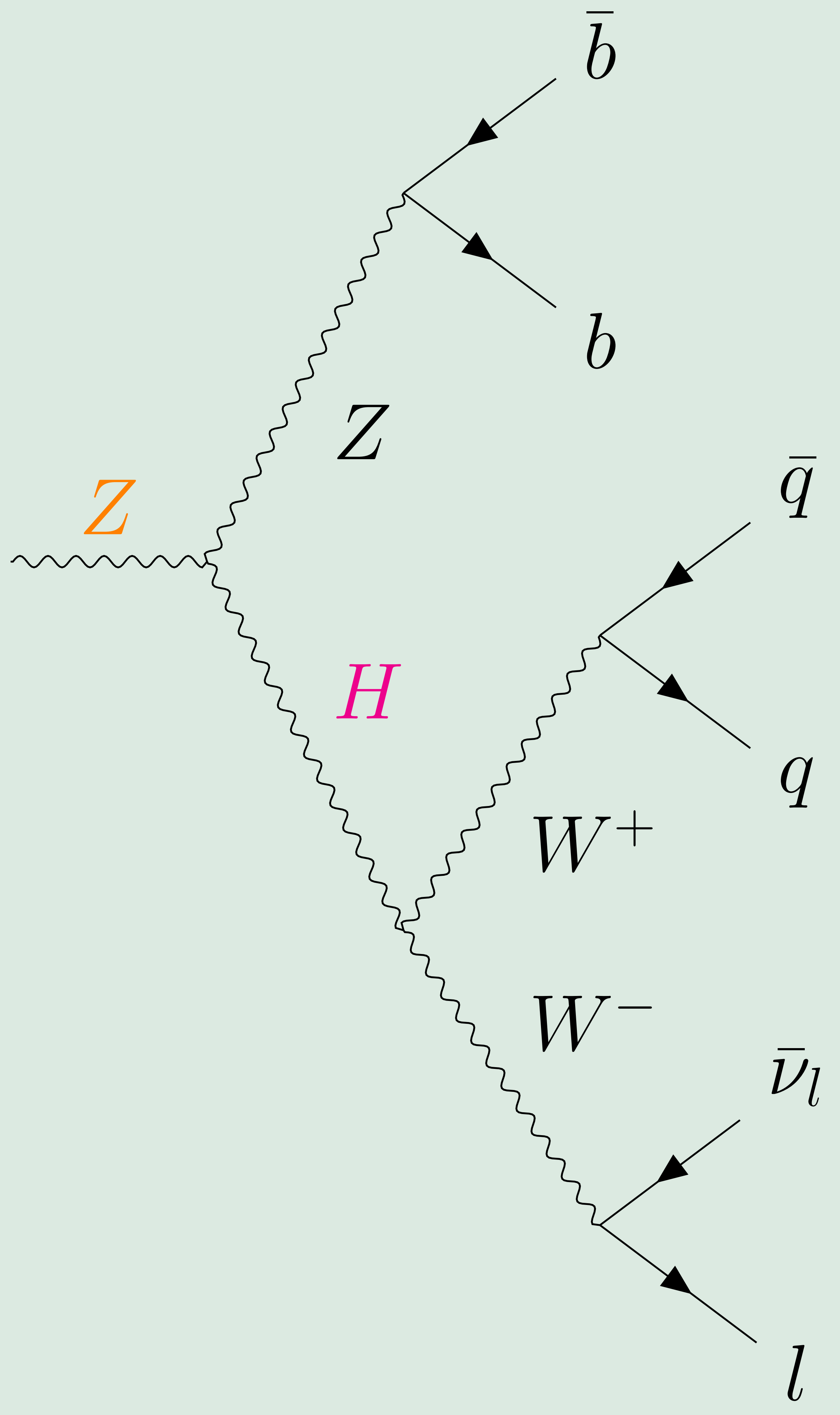
# Signal decay channel

- Off-shell Higgs boson decays to two on-shell Higgs bosons via self-coupling
- Predicted by the Standard Model, but not yet detected
- Simulation cross section $gg \rightarrow H \rightarrow HH$: 38fb
  - Run3: ~300fb$^{-1}$ ➡ Expect around 1600 events



# Background decay channel

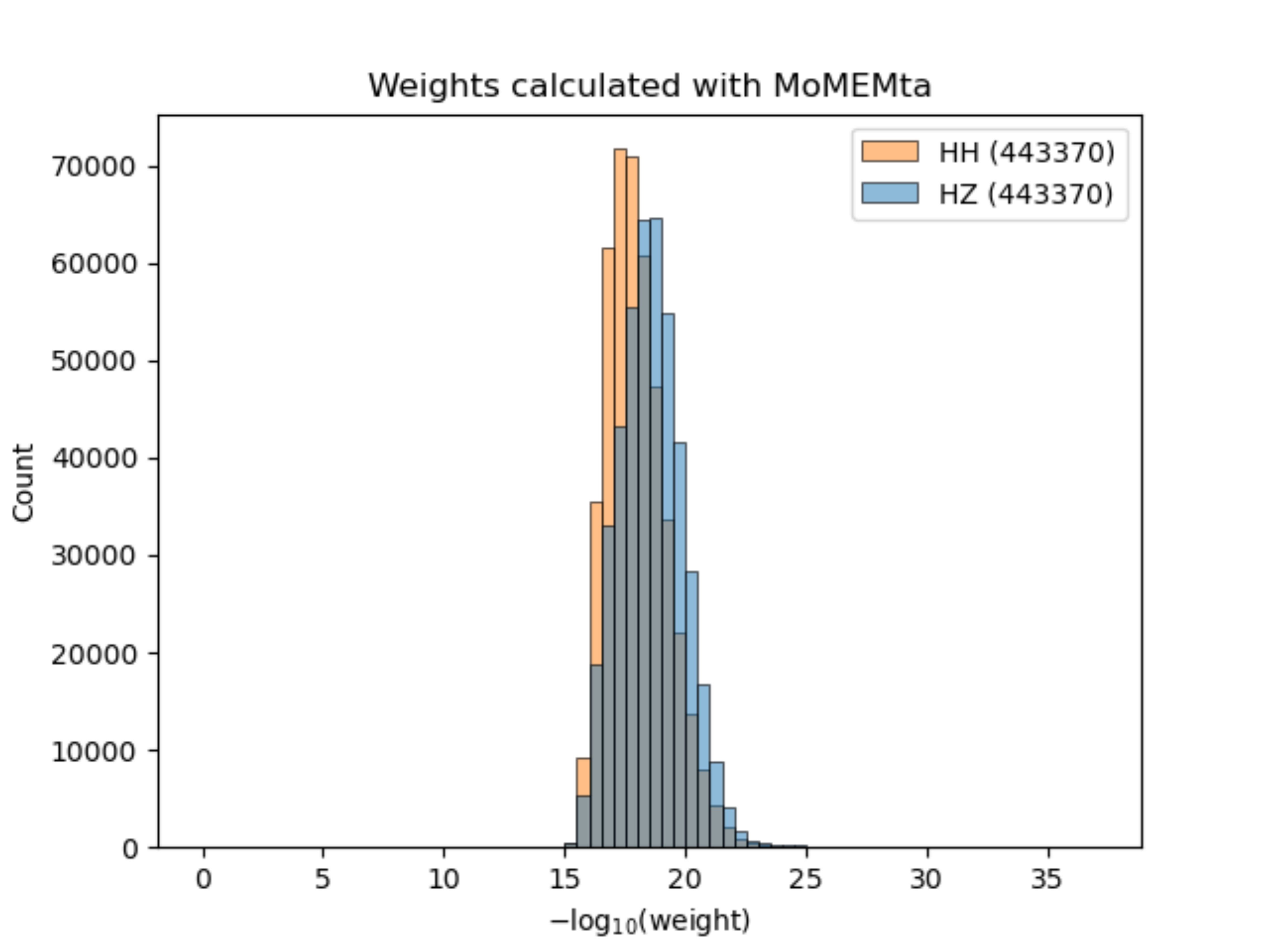
- Off-shell Z boson emits an on-shell Higgs boson
- Same final state
- Simulation cross section $gg \to Z \to HZ$: 51fb



# Data simulation, weight calculation

aMC@NLO, POWHEG, MadGraph, FastJet: Simulate event, reconstruct jets

MoMEMta: Uses aMC@NLO to compute matrix element weights
- Main input: four-momenta of the end-state particles
- Performs parameter transformations to reduce number of integration variables

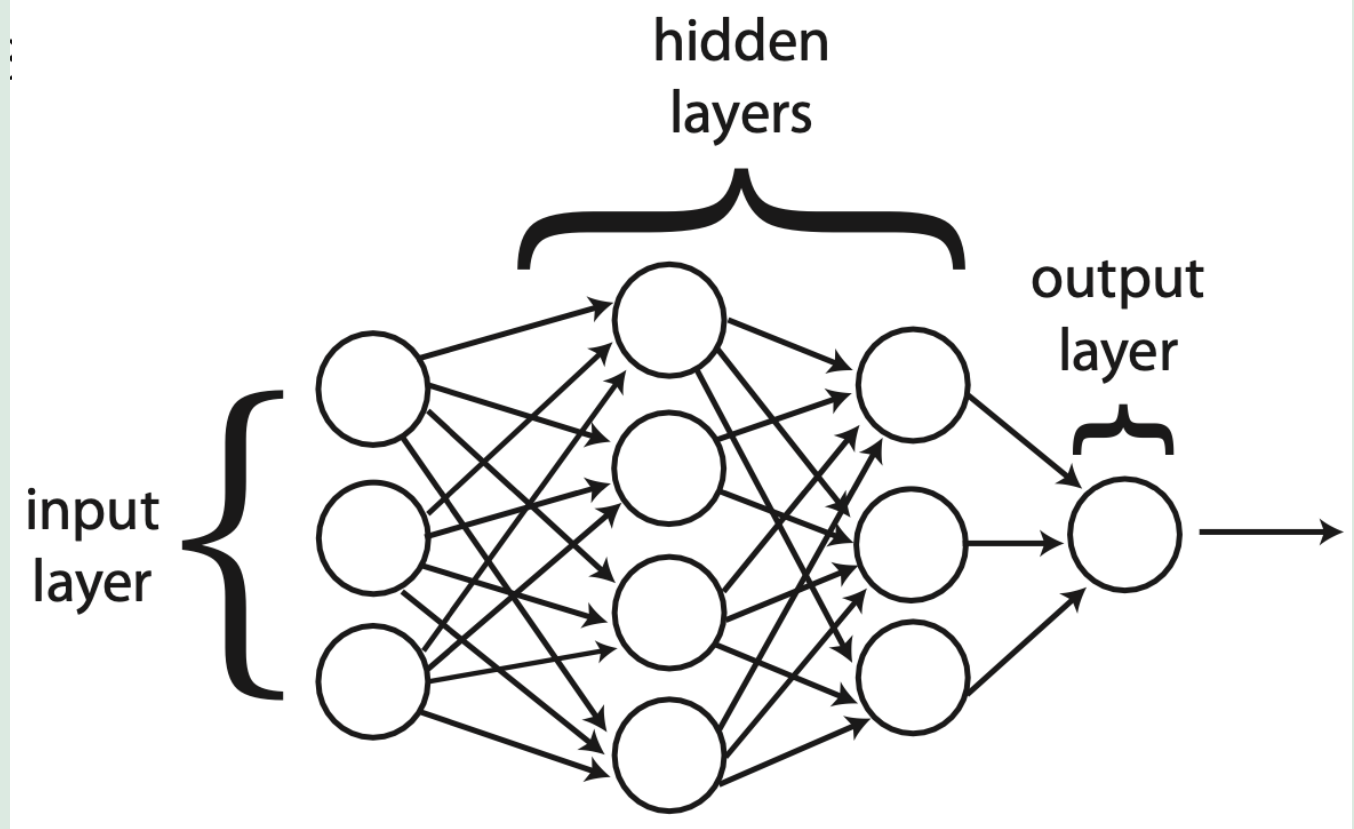


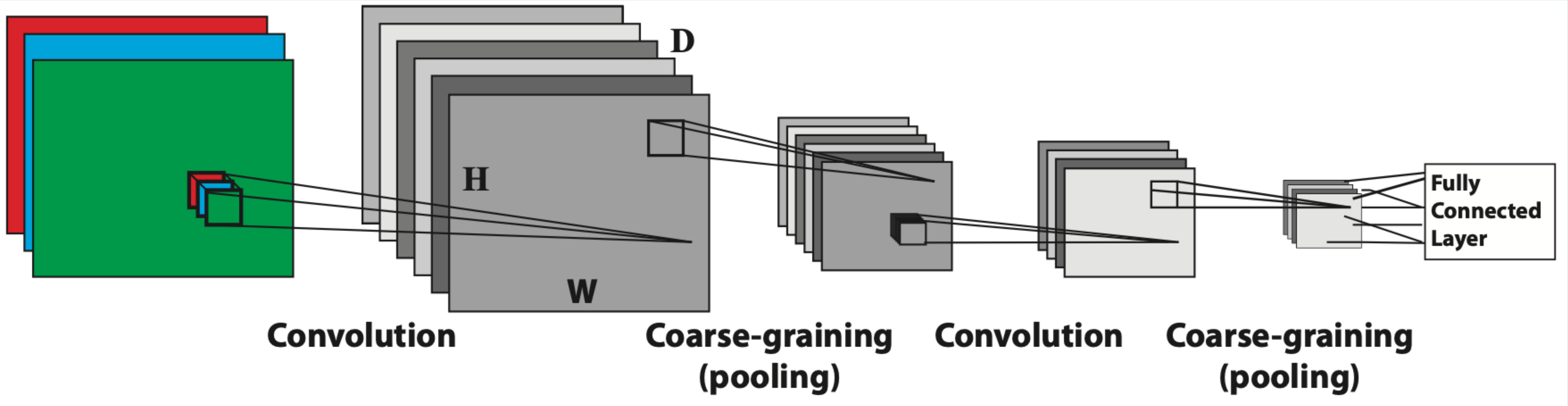
# Neural networks

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Comparing two types of neural networks:

- Feed-forward (right)
- Convolutional (bottom)




images from: 10.1016/j.physrep.2019.03.001

# Evaluating the networks

Regression: Network predicts a value, not a class

➡ How should that be evaluated?

$$R^2 = 1 - \frac{\sum_{(y,\hat{y})} (y - \hat{y})^2}{\sum_{y} (y - \bar{y})^2}$$

true value

predicted value

true mean

If $R^2 = 1$: Network perfectly predicts values

If $R^2 = 0$: Network might as well always predict the mean value

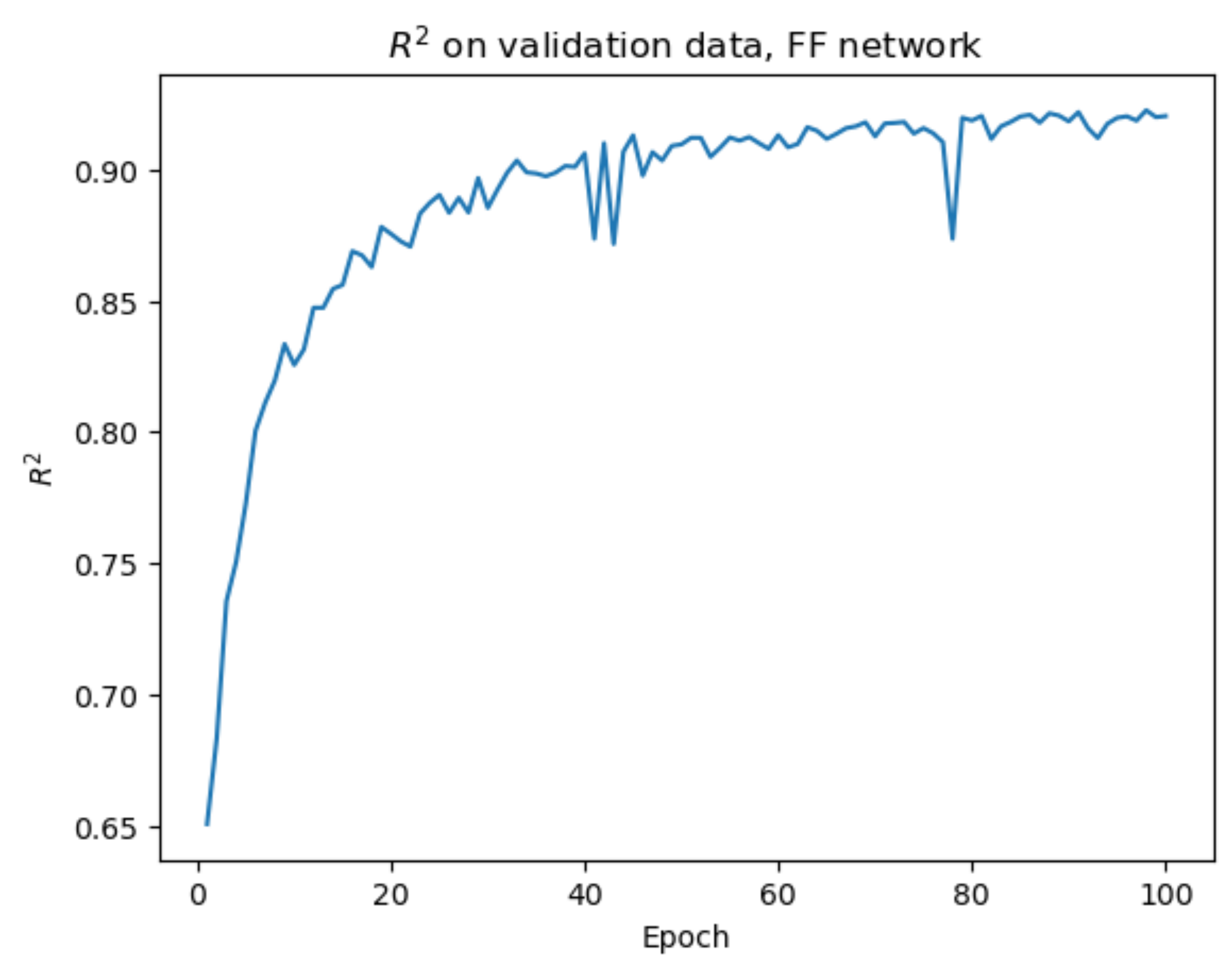
If $R^2 < 0$: Consider network as not knowing anything

Whether a value for $R^2 > 0$ is good or not is arbitrary!

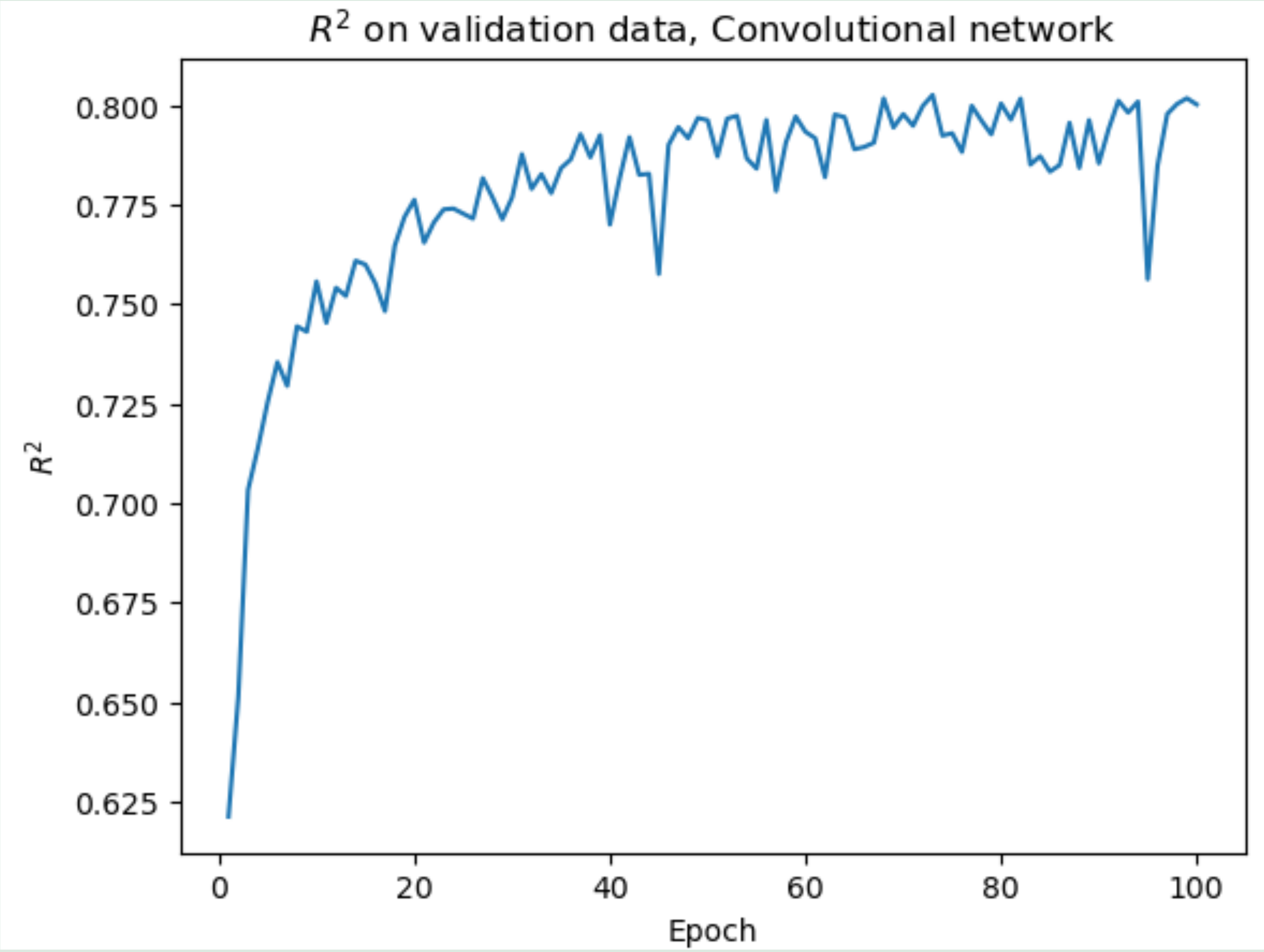
# Results

- around 443000 events each for signal and background

Feed-forward

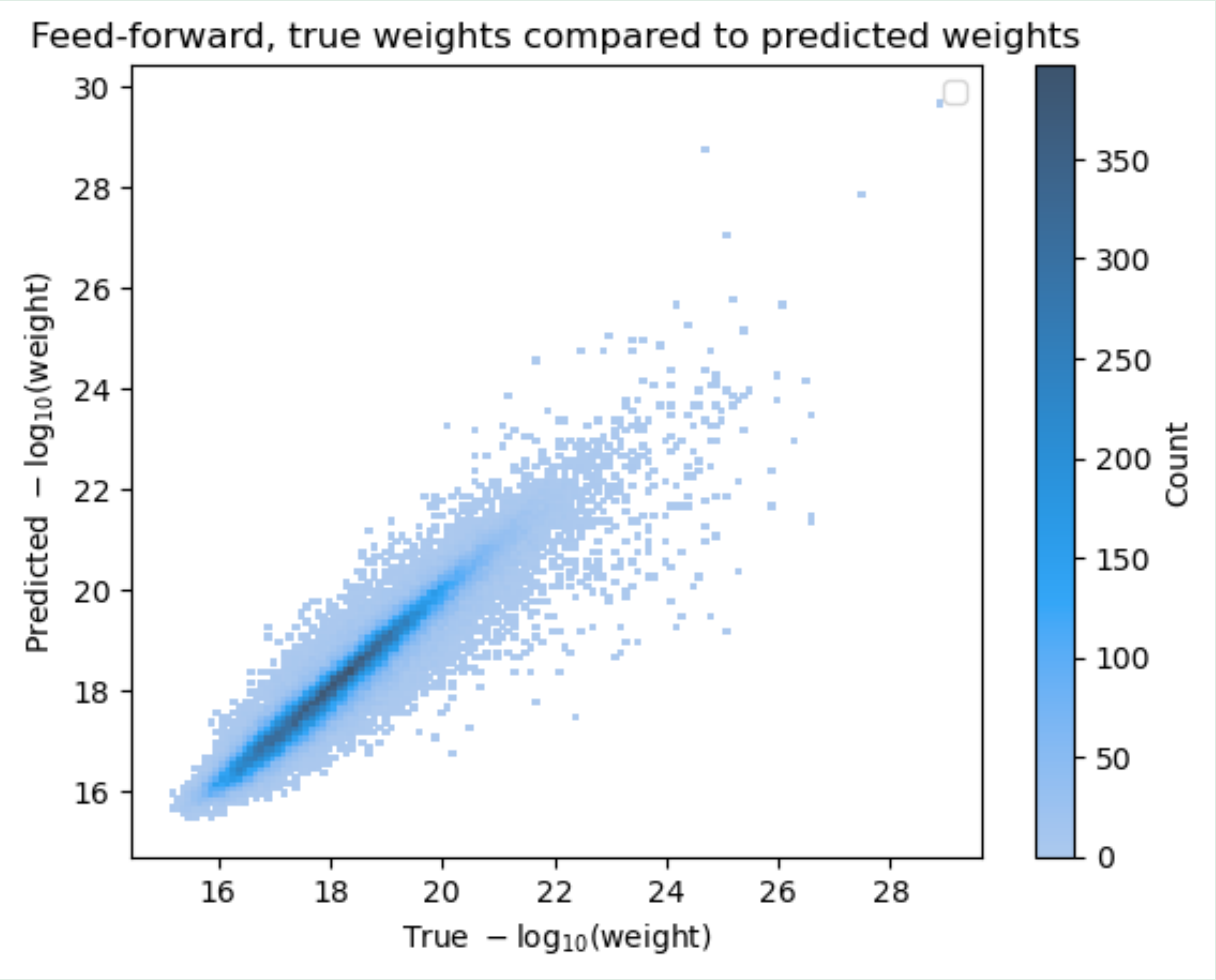



Convolutional

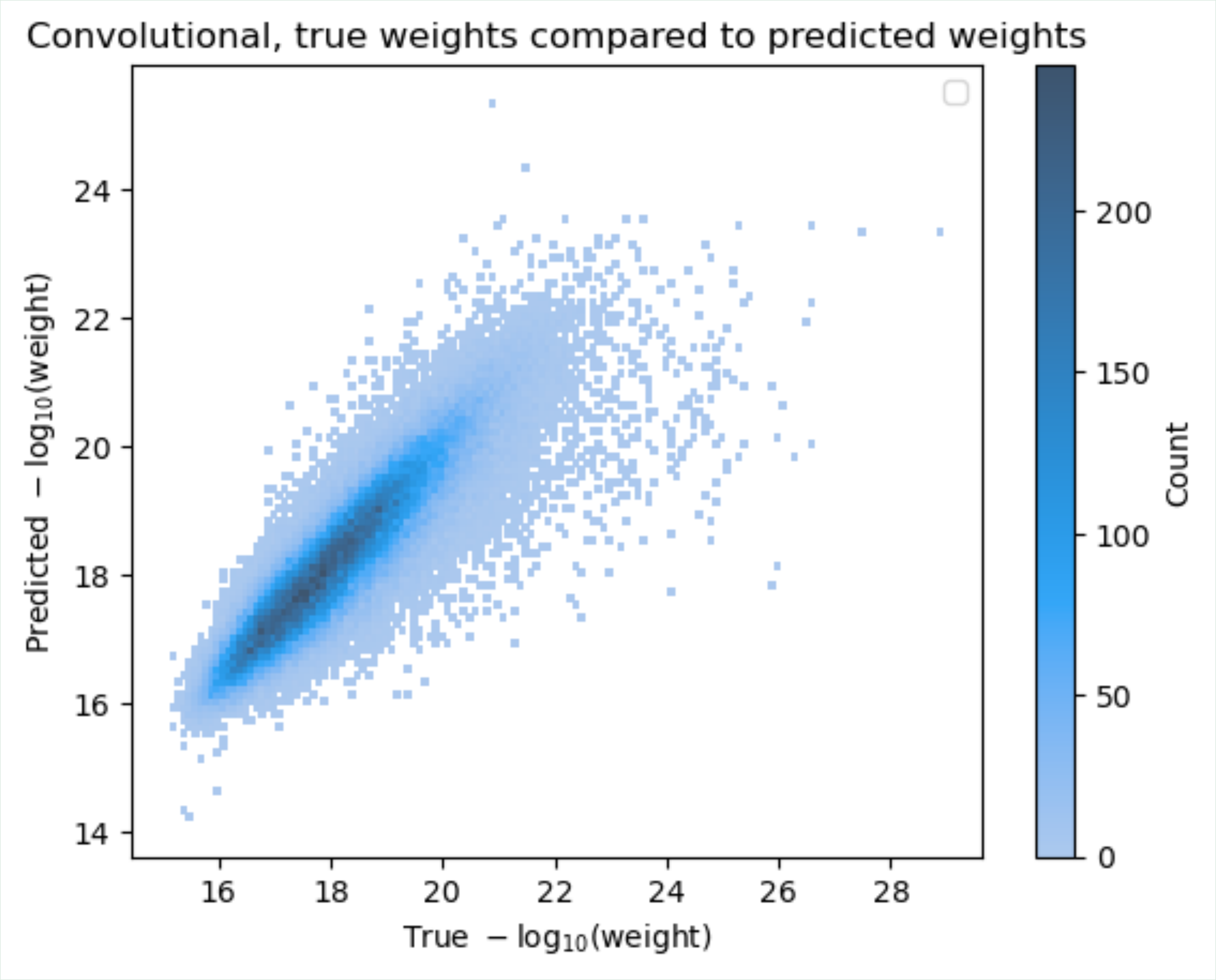


# Results

## Feed-forward




## Convolutional



# Summary and outlook

$$W_X = \left| \langle \Phi_f | \hat{H}_X | \Phi_i \rangle_{d\Psi} \right|^2$$

- Use machine learning to accelerate the matrix element method
- Test analysis: Higgs self-coupling
- Results of neural networks look promising, feed-forward better so far

Outlook: Improve neural networks

# Back-up

# Data simulation and jet reconstruction

POWHEG: Simulation of initial processes: $gg \to HH, gg \to HZ$

MadGraph5: Simulate rest of event:

- Connect gluons to initial protons via parton density functions
- Calculate additional gluons via initial state radiation
- Determine decay modes, final state radiation, parton showers and hadronisation

FastJet: b and quark jets

- Jets constructed with $R = 0.4$
- Pruning: $z_{cut} = 0.1, D_{cut} = \frac{m_q}{p_T}$

# Neural networks

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

## Feed-forward

| $R^2$ | Layers | Batch Size | Learning Rate | Momentum | Weight Decay | Dampening | Nesterov |
|---|---|---|---|---|---|---|---|
| 0.917 | [320, 160, 80, 40] | 32 | $10^{-4}$ | 0.99 | $10^{-8}$ | 0.1 | False |
| 0.902 | [320, 160, 80, 40] | 32 | $10^{-4}$ | 0.90 | $10^{-8}$ | 0.1 | False |
| 0.885 | [120, 60, 30, 15] | 32 | $10^{-4}$ | 0.99 | $10^{-8}$ | 0.1 | False |