Search for a rare process:
Higgs boson production with a single top quark
Separation of tH and ttH
Cut based vs. neural-network based selection

Finn von Ameln

supervised by Florian Mausolf, Prof. Dr. Johannes Erdmann

Netzwerk Teilchenwelt - CERN Project weeks

03rd November 2023

- discovered in 2012
- short lifetime = “instant” decay
- only reconstructed by decay particles
- multiple decay channels, different branching ratios

for our dataset: 2 files, gg→tHq / ttH→yy

gluon fusion main production mode

tHq, ttH rare

tHq not yet observed

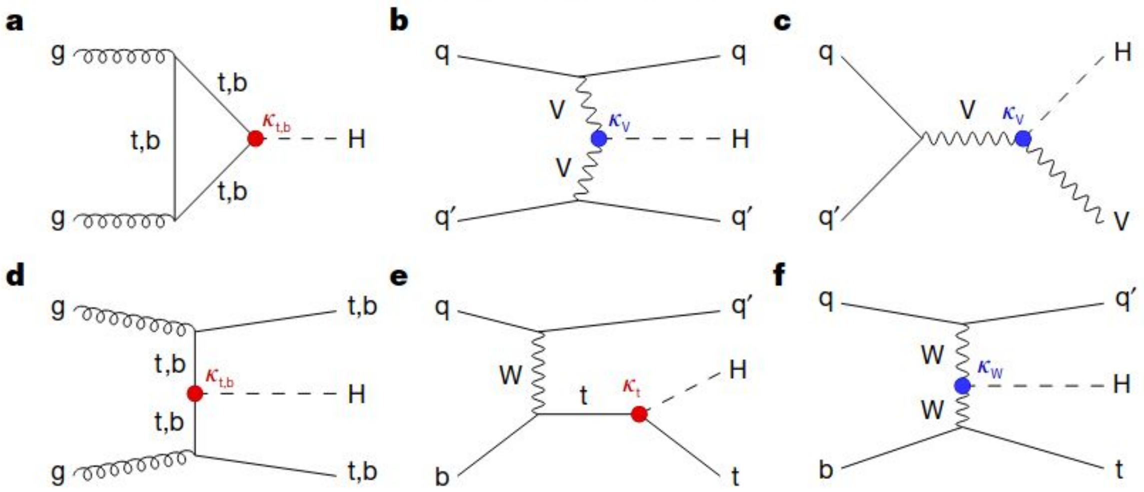


fig. 1: The production modes of the Higgs boson

d = ttH, e,f = tHq

- The Higgs “gives” mass to fermions by Yukawa coupling
- mass t = 172.76 (GeV)
- tHq process, H can originate from t/W
- interfering destructively in SM (hence a rare process)
- test Yukawa coupling in SM
- if Yukawa coupling different from SM = constructive?
    - then tHq more frequent

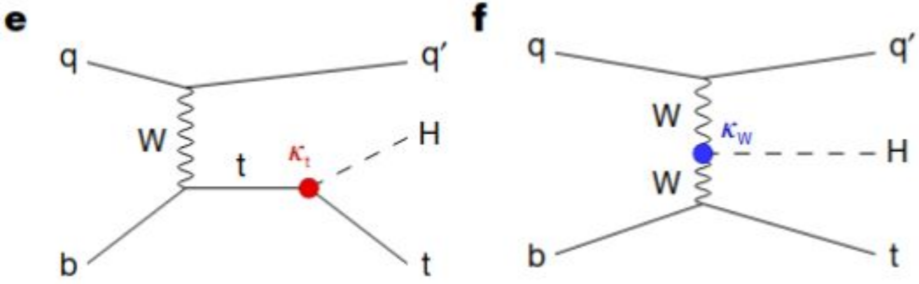


fig. 2: The Higgs production with a single top quarks

how can we see this process in a detector?

t→ Wb

W→ lepton(charged)ν / W→2q,

(H → γγ in our dataset)

taking ttH = background

similar to tH but much more frequent

(+good background distinction with γγ)

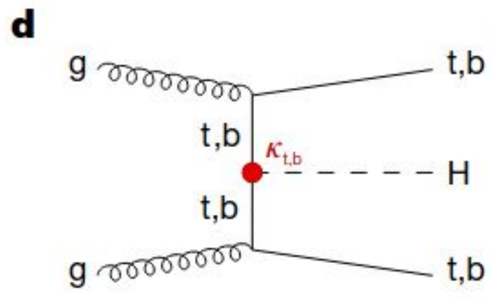


fig. 3: The Higgs production with a pair of top quarks

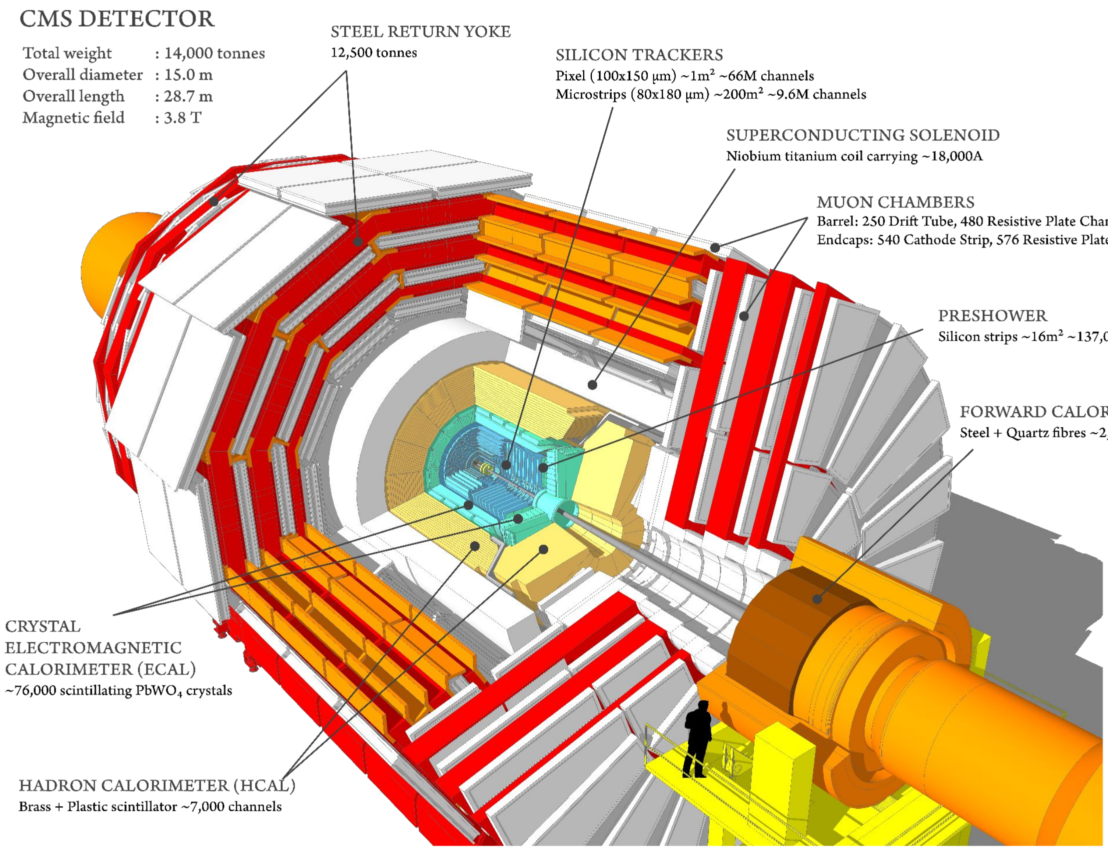


fig. 4: The CMS detector

- excellent invariant mass resolution
- small branching ratio (0.23%)

measurement (Run 2):

- we now see the signal and background
- fitting →mass peak
- lower = fitting signal subtracting background
- tHq = clearer but less events

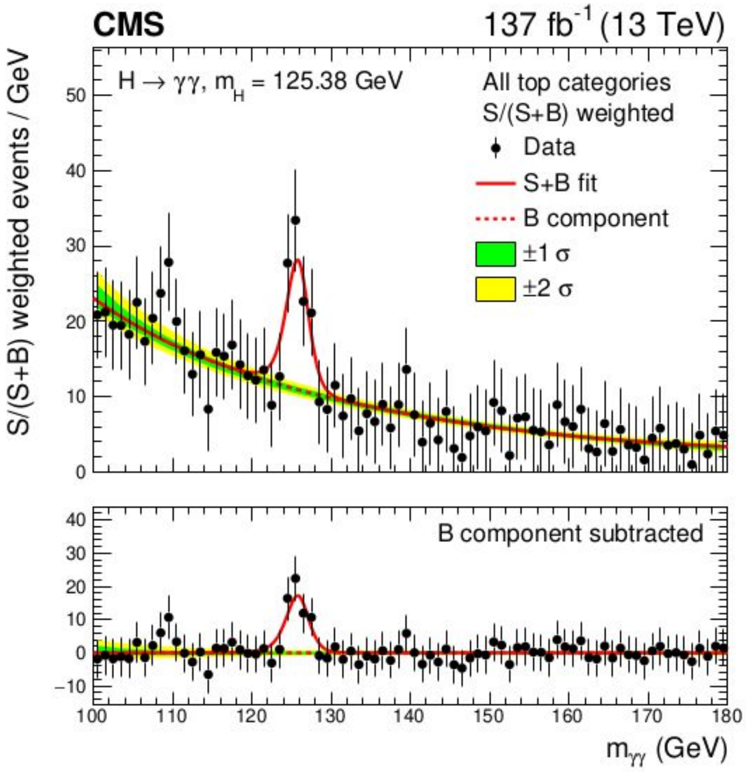
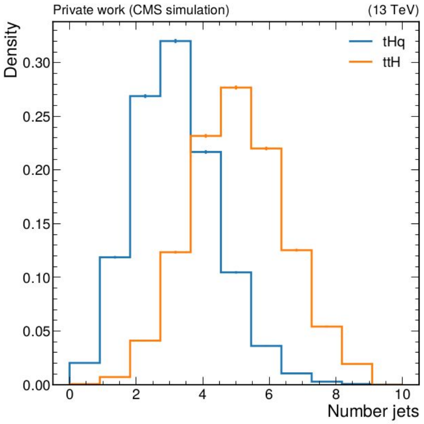
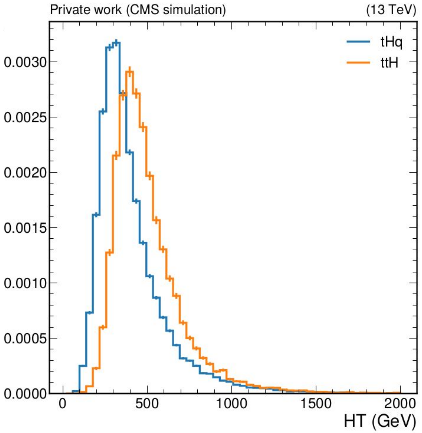
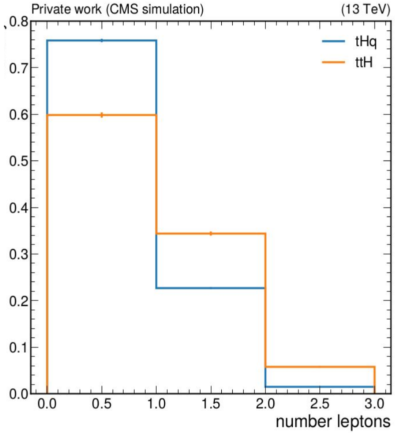


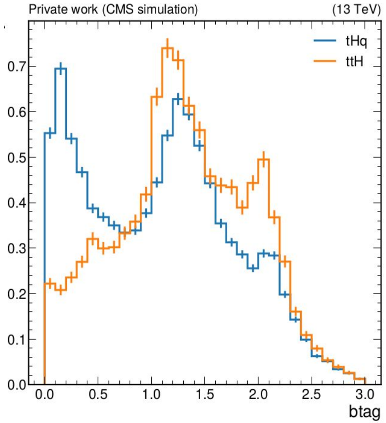
fig. 5: CMS invariant mass results for all top categories

Cut based selection









jet cut: number jets <5

[% remaining]

tHq: 85.87

ttH: 36.45

ht cut: ht <500 GeV

sum of all pt’s

[% remaining]

tHq: 66.61

ttH: 22.15

lepton cut: number leptons <2

[% remaining]

tHq: 65.59

ttH: 20.74

btag cut: btag <1

sum 3 highest btag scores/event

[% remaining]

tHq: 44.03

tH: 7.68

information flows from input to output

multiple layers of interconnected neurons

each layer processing and transforming data to produce an output

!optimizing by training

Nodes = individual processing units

Weights = adjustable parameters through training

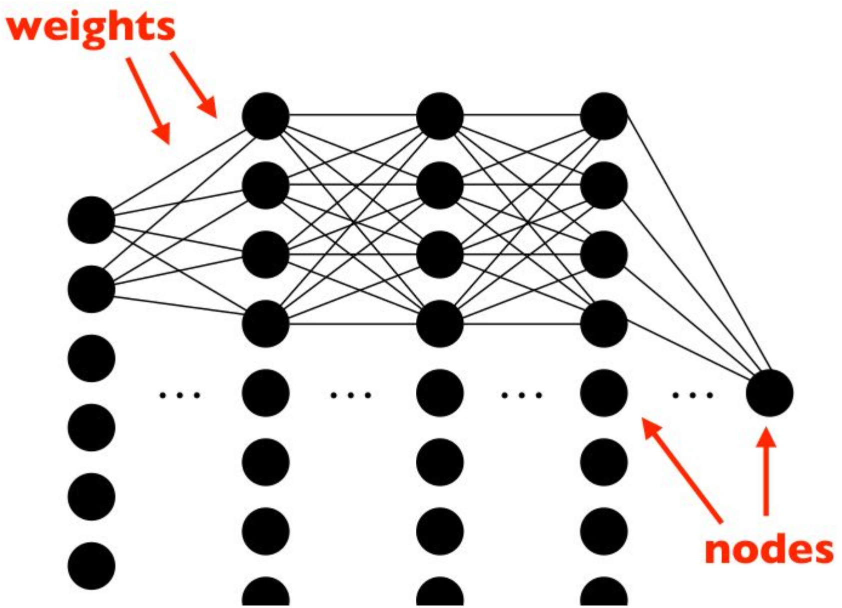


fig. 6: Feed-forward-neural-network structure

events overall = 100.000

training data = 30.000

test data =20.000

amount nodes = 20,20,10

batch size = 1024 beschreiben

epochs = 50 beschreiben

(binary classification problem:

loss =binary cross entropy

adam optimizer

activation = sigmoid, relu)

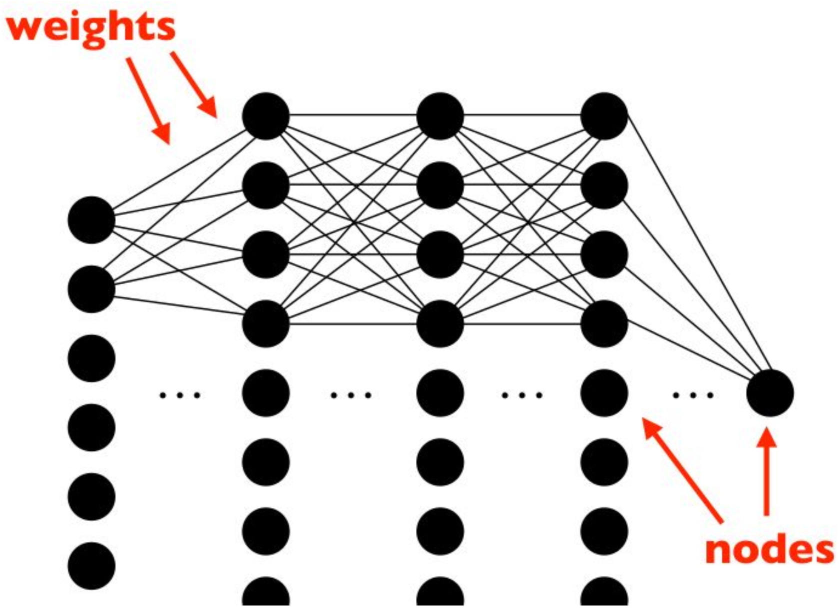
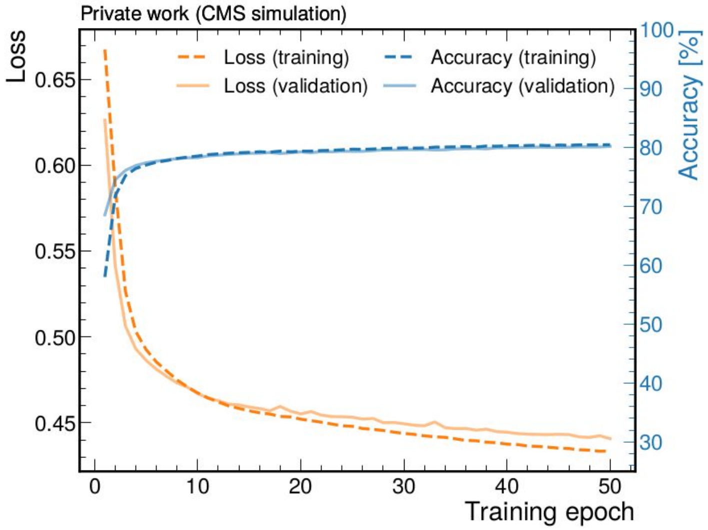
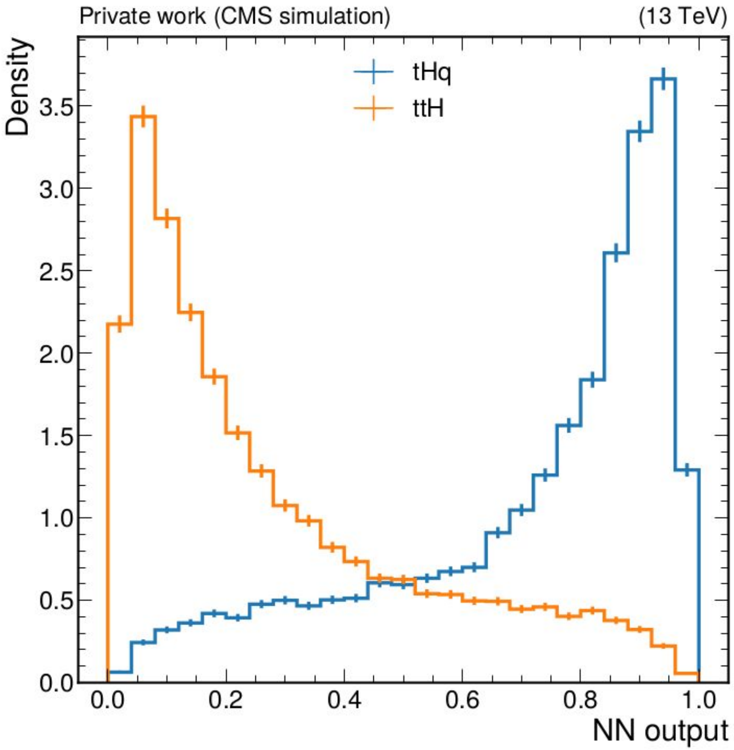


fig. 6: Feed-forward-neural-network structure





the nn can now separate ttH from tHq (functional approximation)

accuracy = events correctly classified

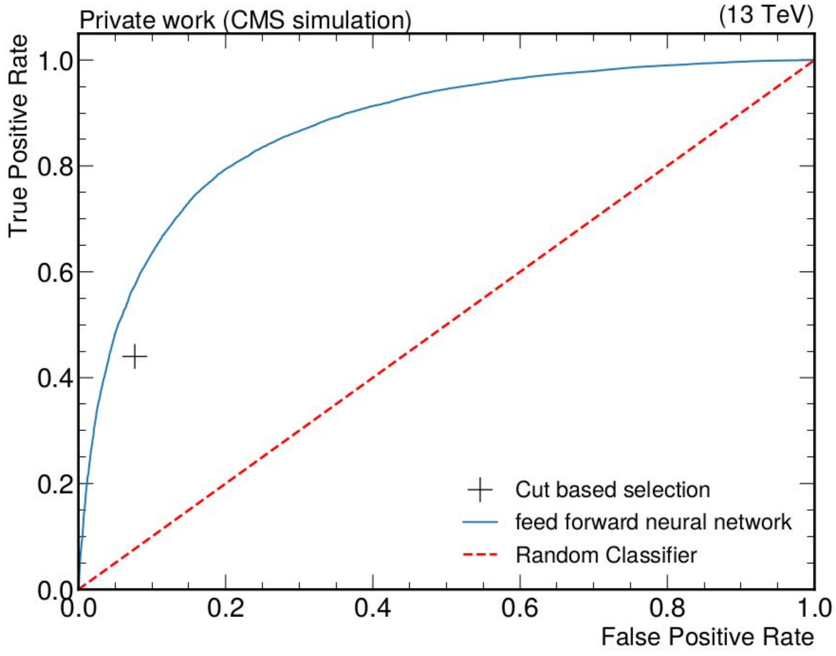
Efficiency comparison - final results

comparing cut based vs nn:

(different inputs!)

The nn's cuts are more efficient

"higher" curve = better



there's still room for improvement:

-changing parameters

-adding more kinematic variables

-increasing amount events

Finn von Ameln

Special thanks to Florian and Johannes,

any feedback is welcome!

Finn von Ameln

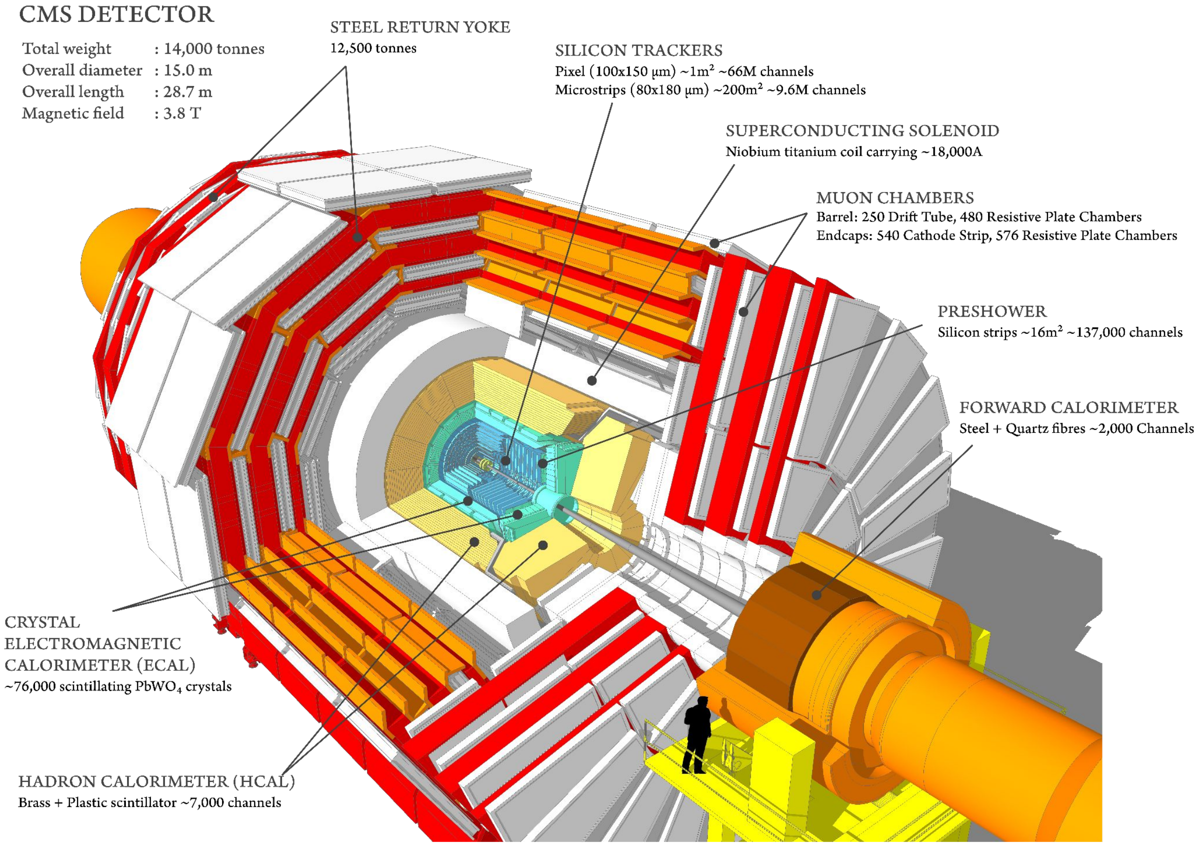
supervised by Florian Mausolf, Prof. Dr. Johannes Erdmann

Netzwerk Teilchenwelt - CERN Project weeks

03rd November 2023

Backup slides

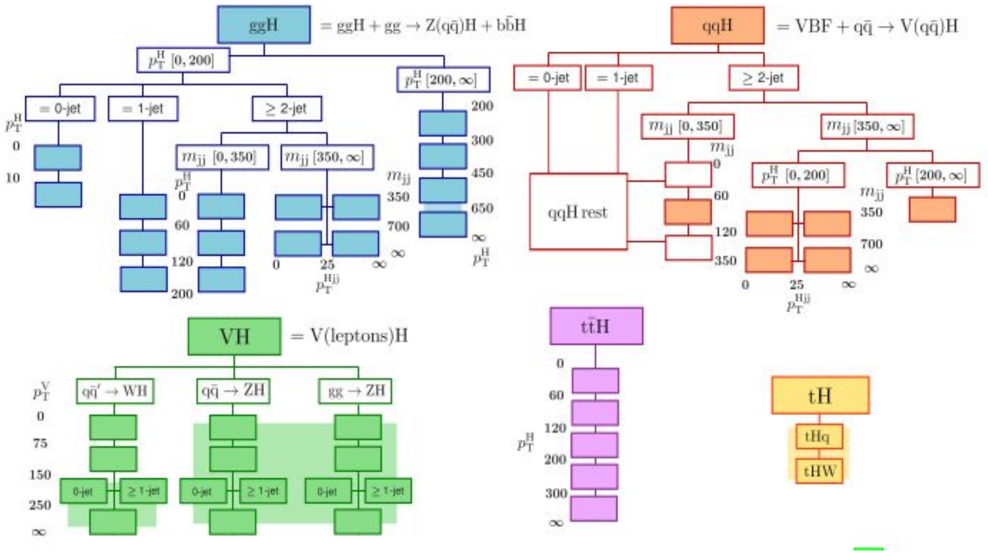
- solenoid magnet -bendig particles
  identify charge, measure momentum
- silicon Tracker - identifying tracks
  (paths taken by charged particles)
- ECAL, HCAL - measuring energy
  hadronic, electronic
- muon chambers - detecting muons
  (passing through ecal)



Kinematic characteristics

- Current research: STXS-framework
  - Higgs production modes are defined base upon kinematic regions
  - allowing for easier separation

- Selecting kinematic variables to filter events
  - (some) STXS conditions implemented in python
  - (photon pre-selection already included)
  - azimuthal angle, pseudorapidity, …



Finn von Ameln