Applications of Machine Learning Techniques

at the ATLAS collaboration

> string_data Workshop

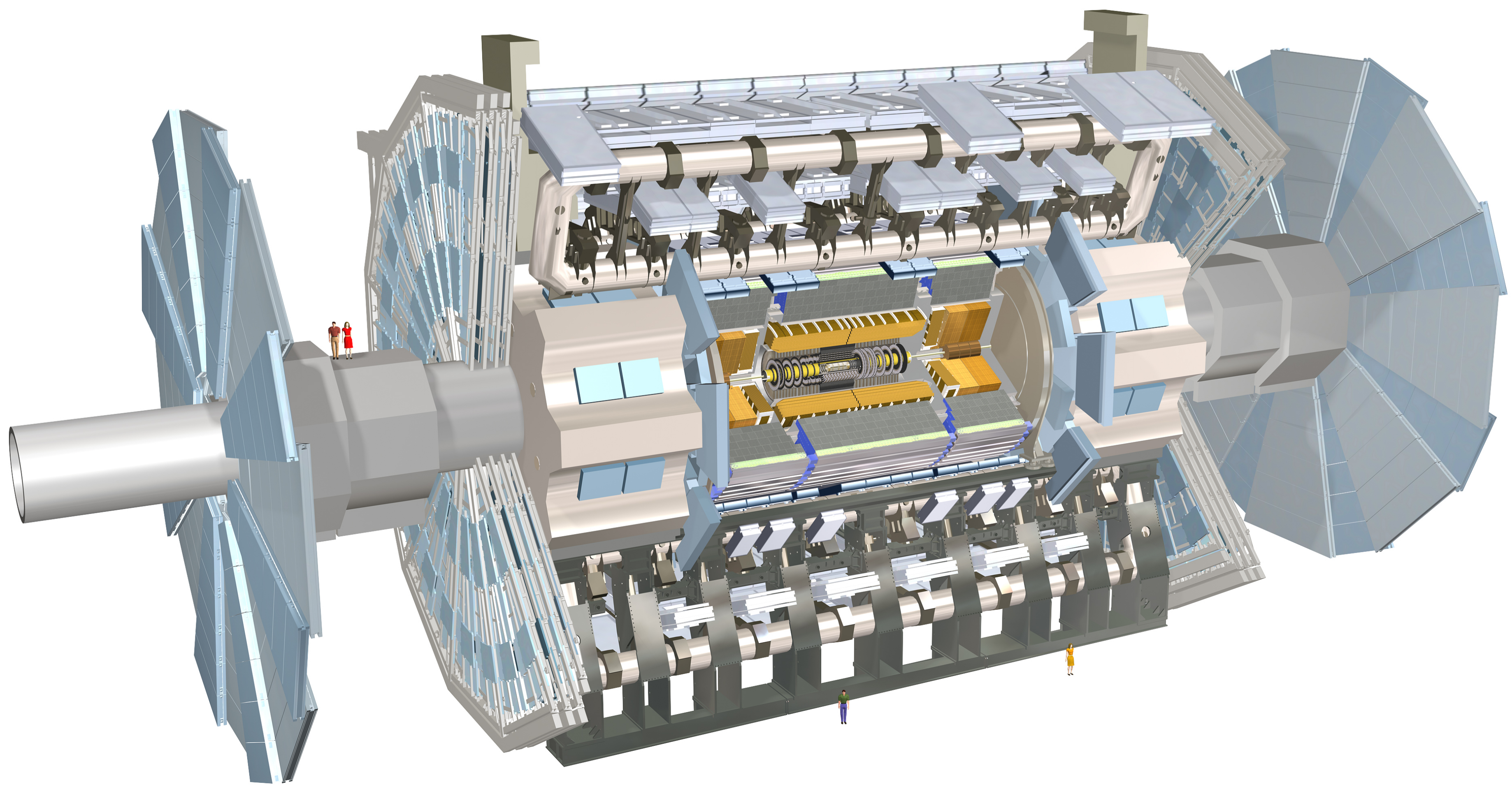
David Handl

27th March 2018

Bundesministerium
für Bildung
und Forschung

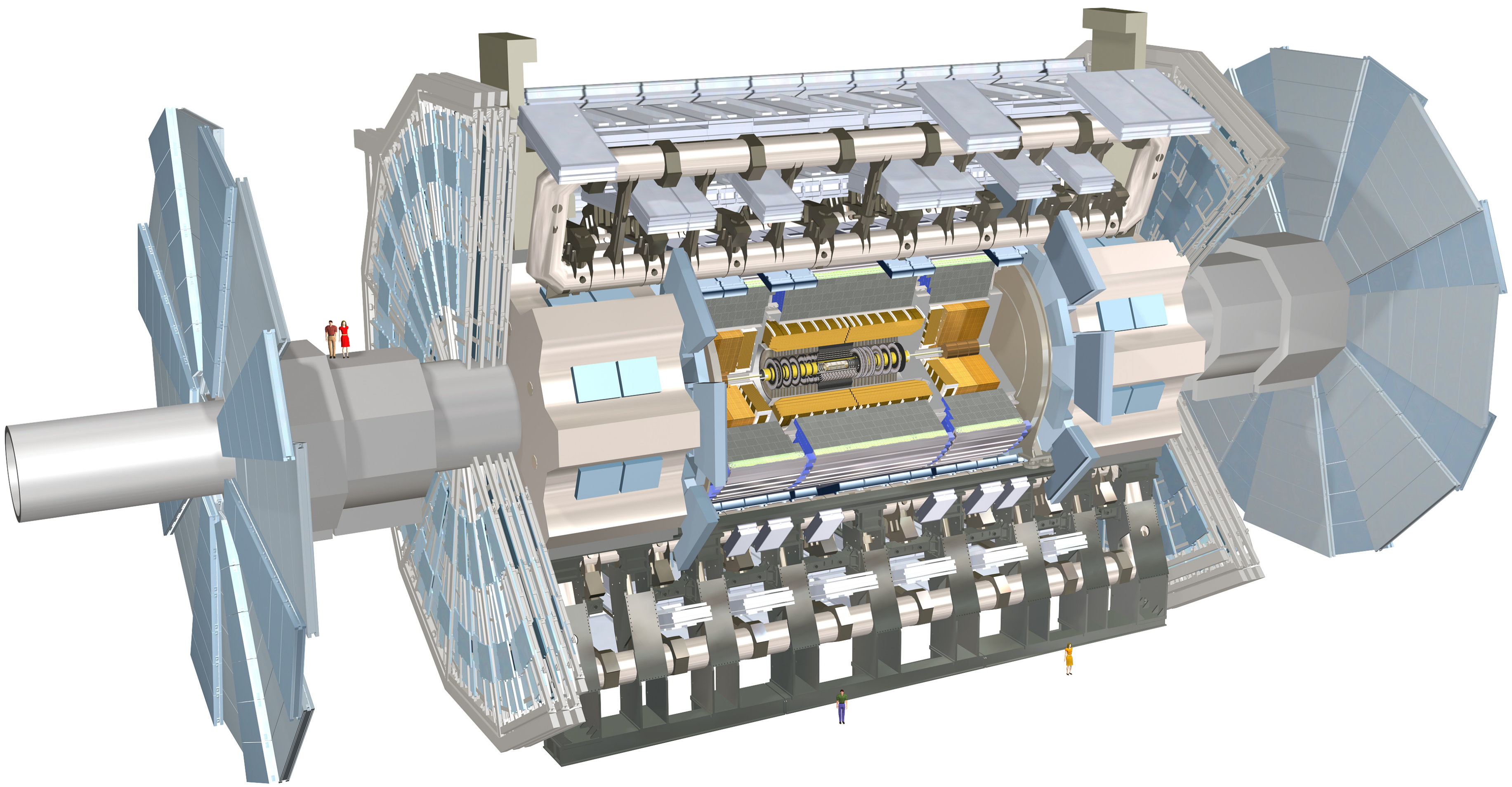
FSP 103
ATLAS LMU

The ATLAS experiment



The ATLAS experiment

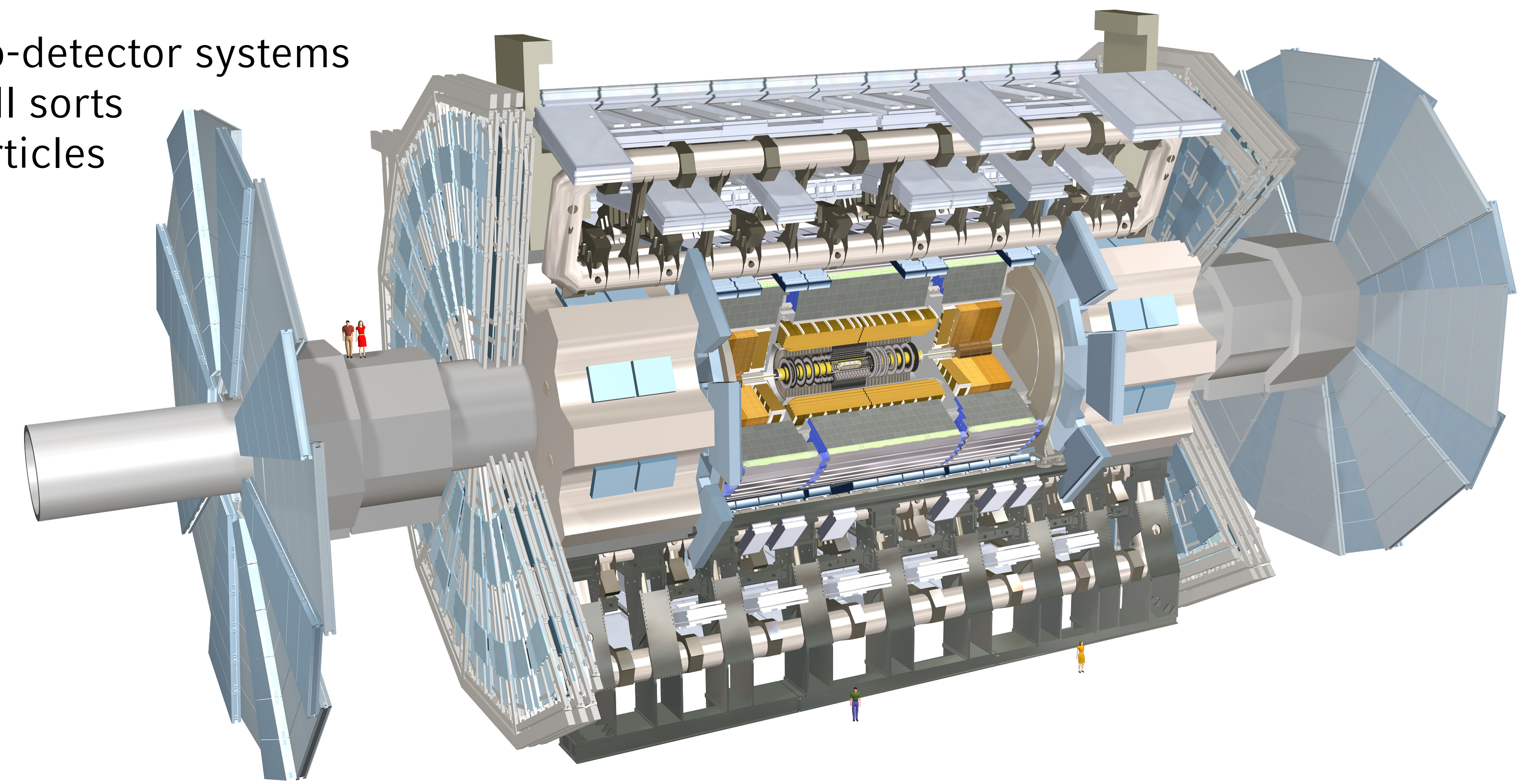
General purpose detector, collecting the data from particle collisions from the Large Hadron Collider (LHC)



The ATLAS experiment

General purpose detector, collecting the data from particle collisions from the Large Hadron Collider (LHC)

Multiple sub-detector systems to identify all sorts of stable particles

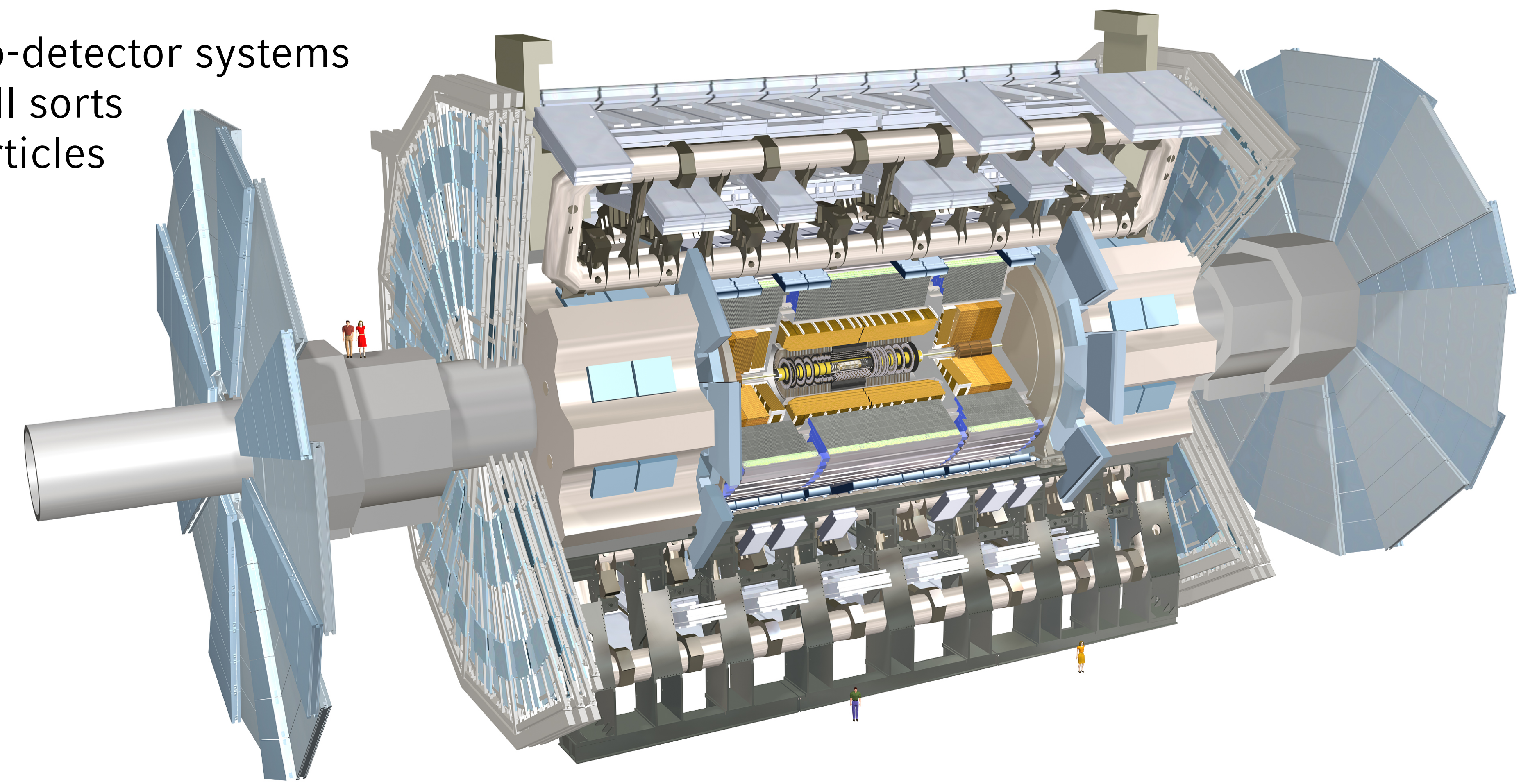


The ATLAS experiment

General purpose detector, collecting the data from particle collisions from the Large Hadron Collider (LHC)

Measuring the properties of the Standard Model (SM) particles at highest precision

Multiple sub-detector systems to identify all sorts of stable particles



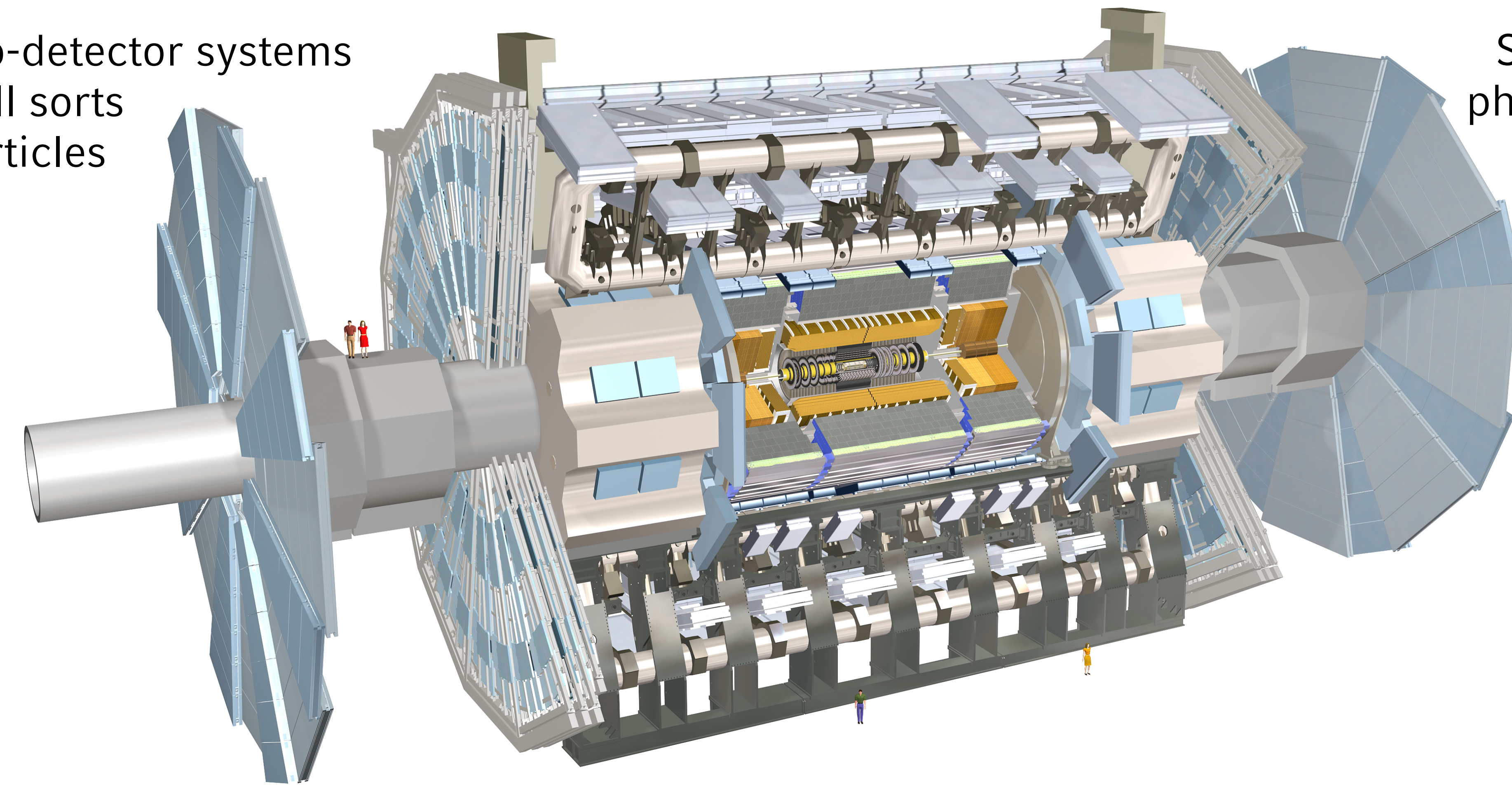
The ATLAS experiment

General purpose detector, collecting the data from particle collisions from the Large Hadron Collider (LHC)

Measuring the properties of the Standard Model (SM) particles at highest precision

Multiple sub-detector systems to identify all sorts of stable particles

Searching for new phenomena beyond the SM



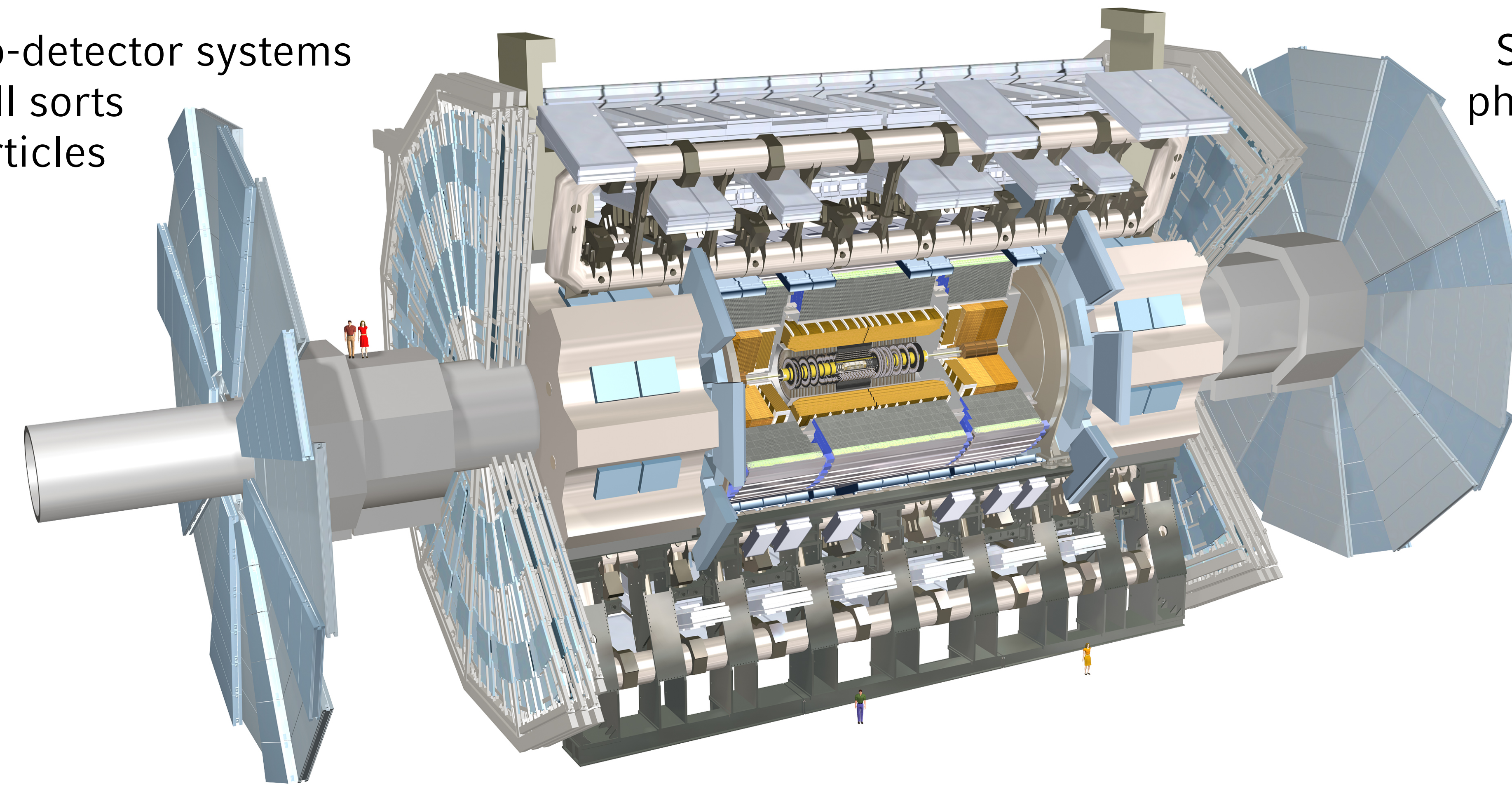
The ATLAS experiment

General purpose detector, collecting the data from particle collisions from the Large Hadron Collider (LHC)

Measuring the properties of the Standard Model (SM) particles at highest precision

Multiple sub-detector systems to identify all sorts of stable particles

Searching for new phenomena beyond the SM



Collision

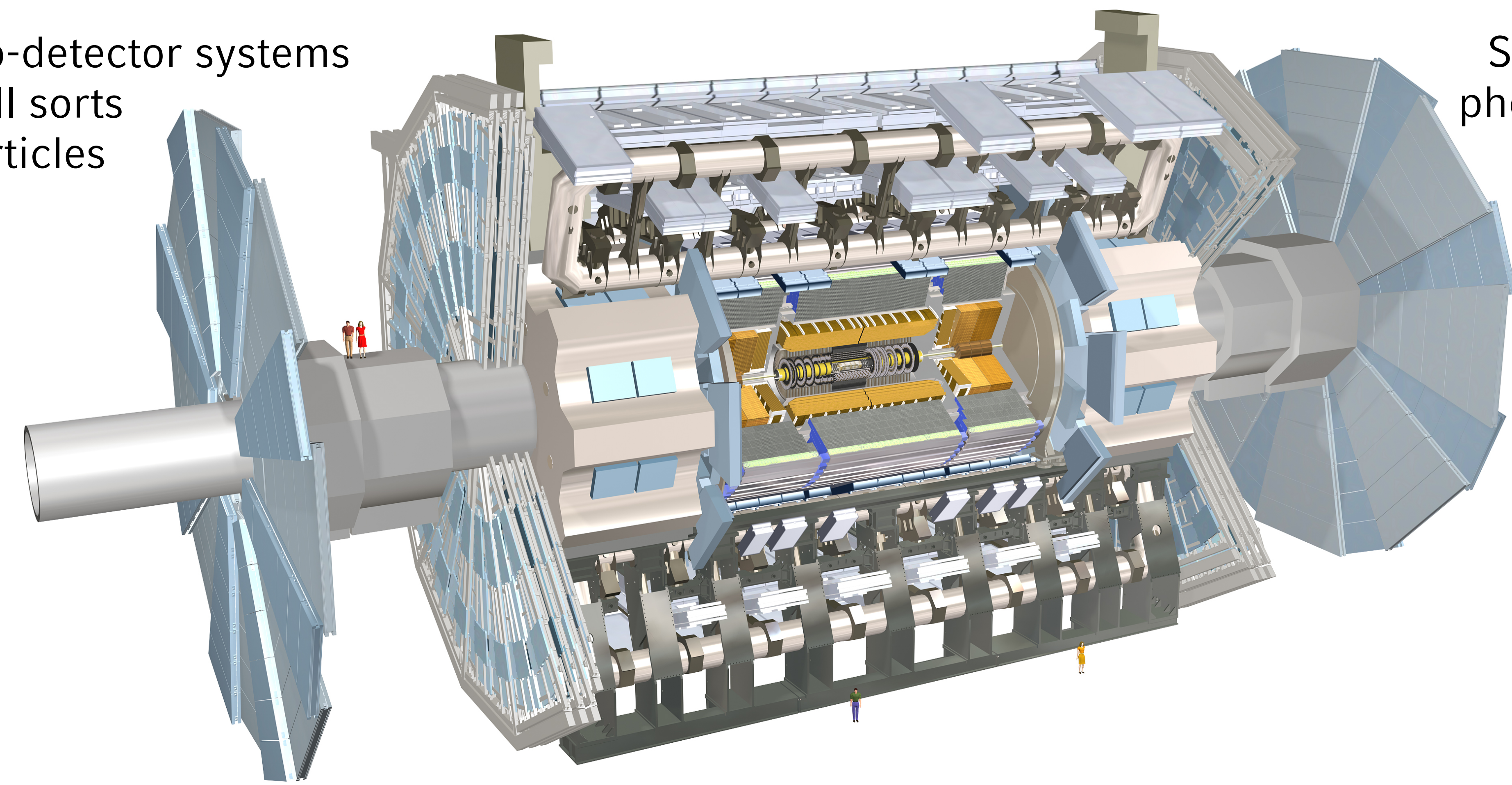
The ATLAS experiment

General purpose detector, collecting the data from particle collisions from the Large Hadron Collider (LHC)

Measuring the properties of the Standard Model (SM) particles at highest precision

Multiple sub-detector systems to identify all sorts of stable particles

Searching for new phenomena beyond the SM



Collision → *Data processing & reconstruction*

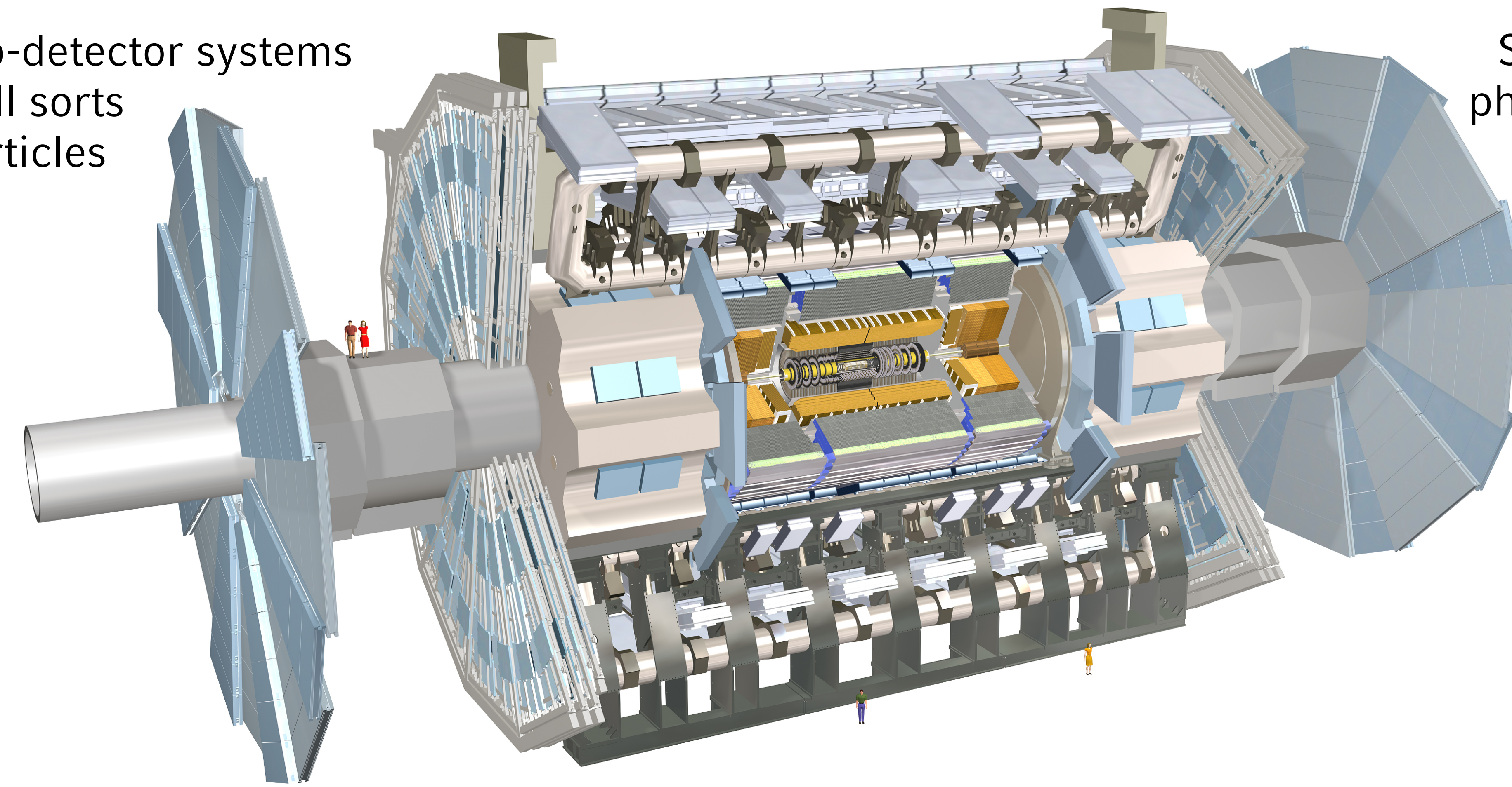
The ATLAS experiment

General purpose detector, collecting the data from particle collisions from the Large Hadron Collider (LHC)

Measuring the properties of the Standard Model (SM) particles at highest precision

Multiple sub-detector systems to identify all sorts of stable particles

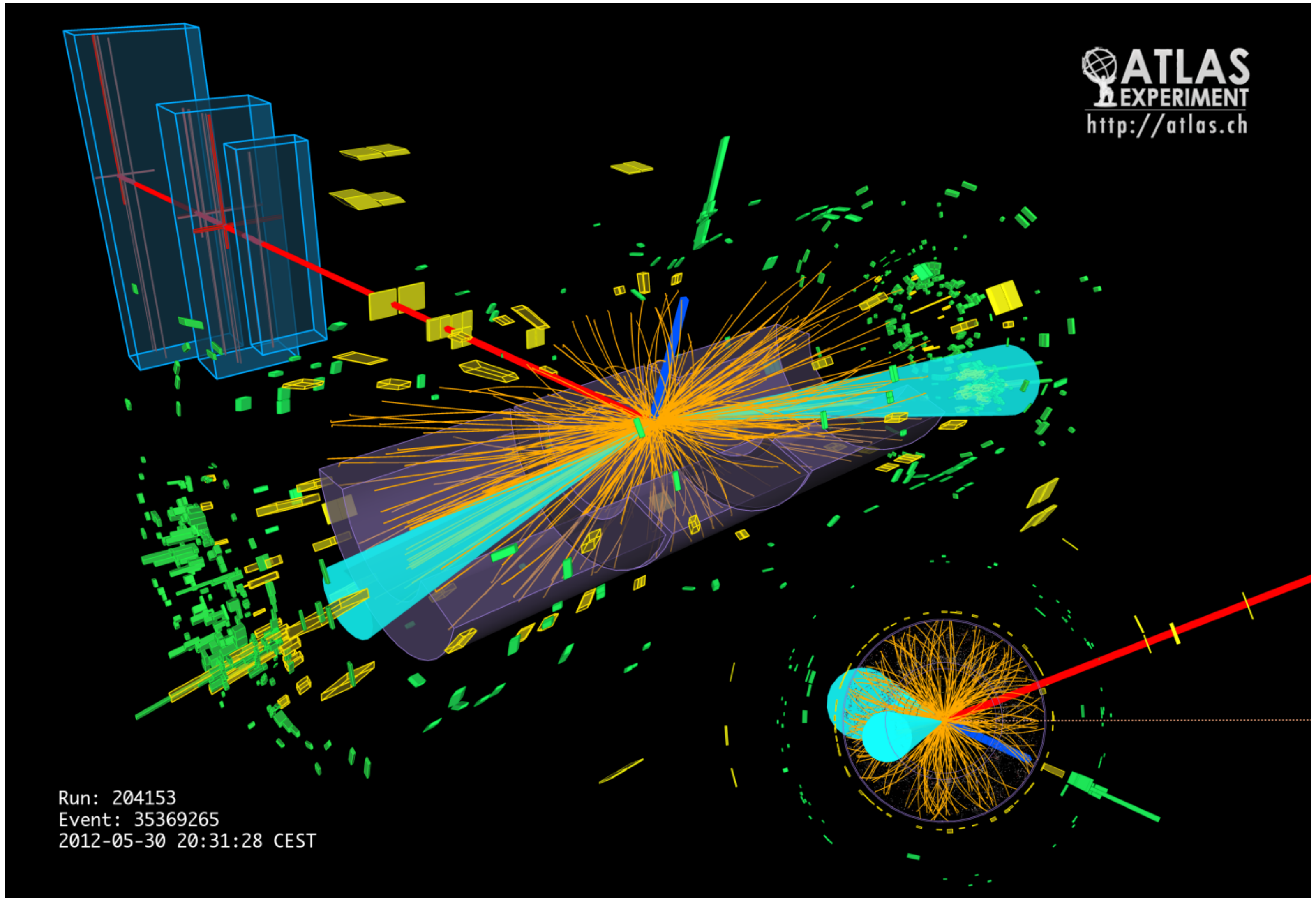
Searching for new phenomena beyond the SM



Collision → *Data processing & reconstruction* → *Data analysis*

A different view

H→ττ candidate event from 2012

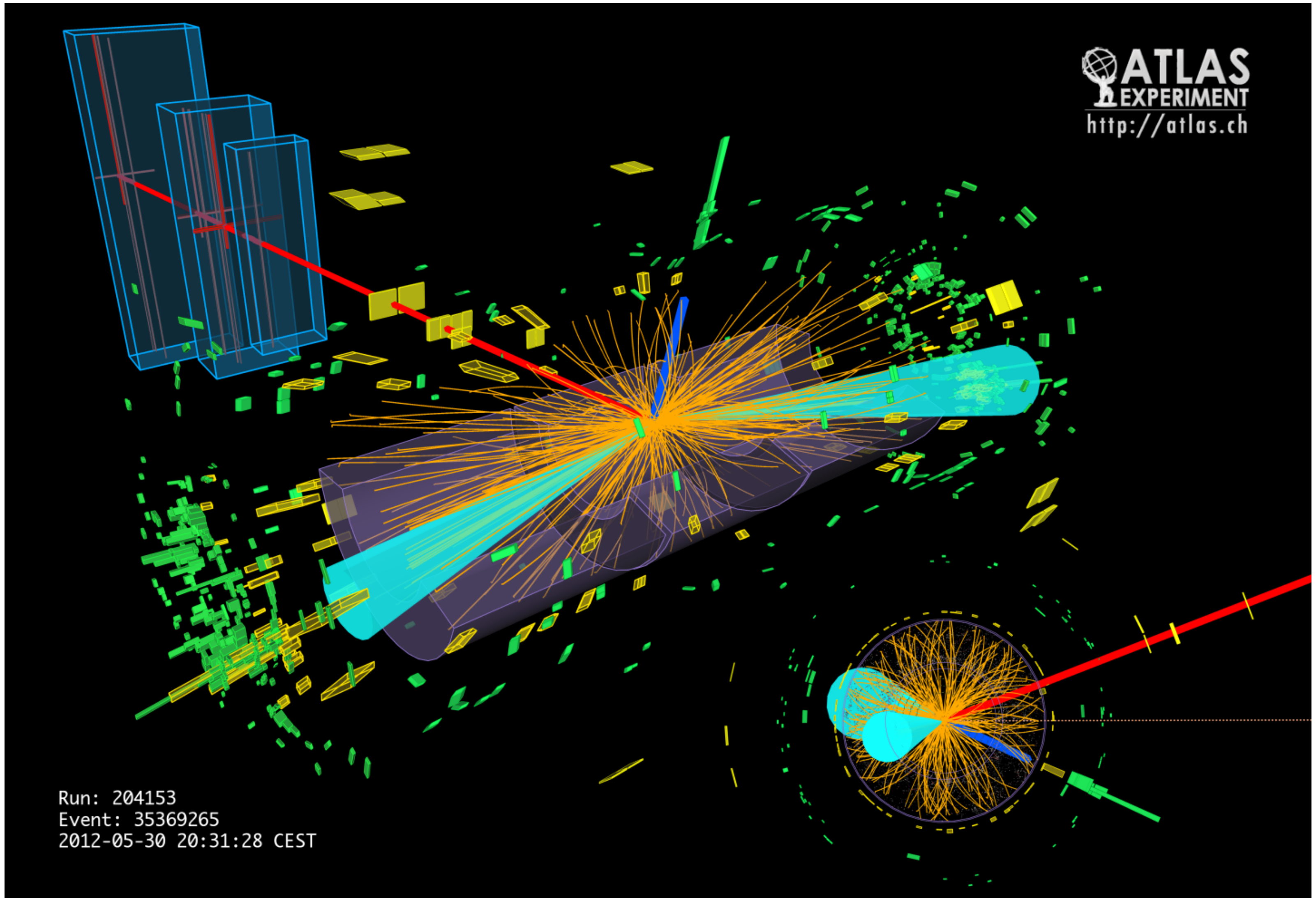


- Sub-detector systems **record various** types of **signatures**
- Aim to **reconstruct objects** such as particle tracks or calorimeter showers, etc.
- **Combine information** to distinguish and identify different particles (muons, electrons, photons, jets, *missing transverse energy*(E_T^{miss}))

Collision → ***Data processing & reconstruction*** → ***Data analysis***

A different view

H→ττ candidate event from 2012



- Sub-detector systems **record various** types of **signatures**
- Aim to **reconstruct objects** such as particle tracks or calorimeter showers, etc.
- **Combine information** to distinguish and identify different particles (muons, electrons, photons, jets, *missing transverse energy* (E_T^{miss}))

Massive amount of data have to be processed and analysed

Collision → *Data processing & reconstruction* → *Data analysis*

Where does machine learning (ML) fit in?

Data processing & reconstruction

Data analysis

Where does machine learning (ML) fit in?

Data processing & reconstruction

Data analysis

Comprehensible performance

Sufficient size of datasets

- Lepton identification
- W-boson and top quark identification
- b-quark tagging
- Quark/gluon jet tagging
- …

Where does machine learning (ML) fit in?

Data processing & reconstruction

Comprehensible performance

Sufficient size of datasets

- Lepton identification
- W-boson and top quark identification
- b-quark tagging
- Quark/gluon jet tagging
- …

Data analysis

Extensive validation studies required to understand results

Possibly statistical limitations

- Searches for/measurements of the Higgs boson
- Searches for new physics
- …

Where does machine learning (ML) fit in?

Data processing & reconstruction

Comprehensible performance

Sufficient size of datasets

- Lepton identification **(BDT)**
- W-boson and top quark identification **(BDT, DNN)**
- b-quark tagging **(RNN, DNN, BDT)**
- Quark/gluon jet tagging **(CNN)**
- …

Data analysis

Extensive validation studies required to understand results

Possibly statistical limitations

- Searches for/measurements of the Higgs boson **(BDT)**
- Searches for new physics **(BDT)**
- …

Machine learning

Interdisciplinary field of **computer science, statistics** and **probability theory**

Mathematical model mapping a set of input values to output values

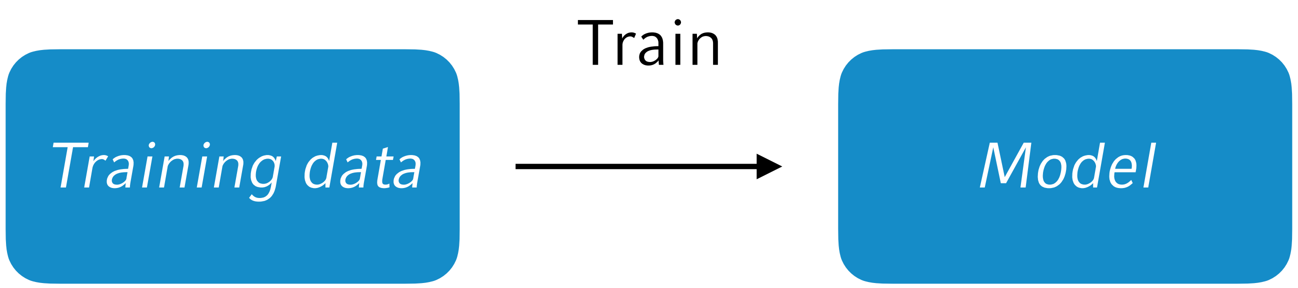
Estimation of a statistical model from data (**learning**)

Machine learning

Interdisciplinary field of **computer science, statistics** and **probability theory**

Mathematical model mapping a set of input values to output values

Estimation of a statistical model from data (**learning**)

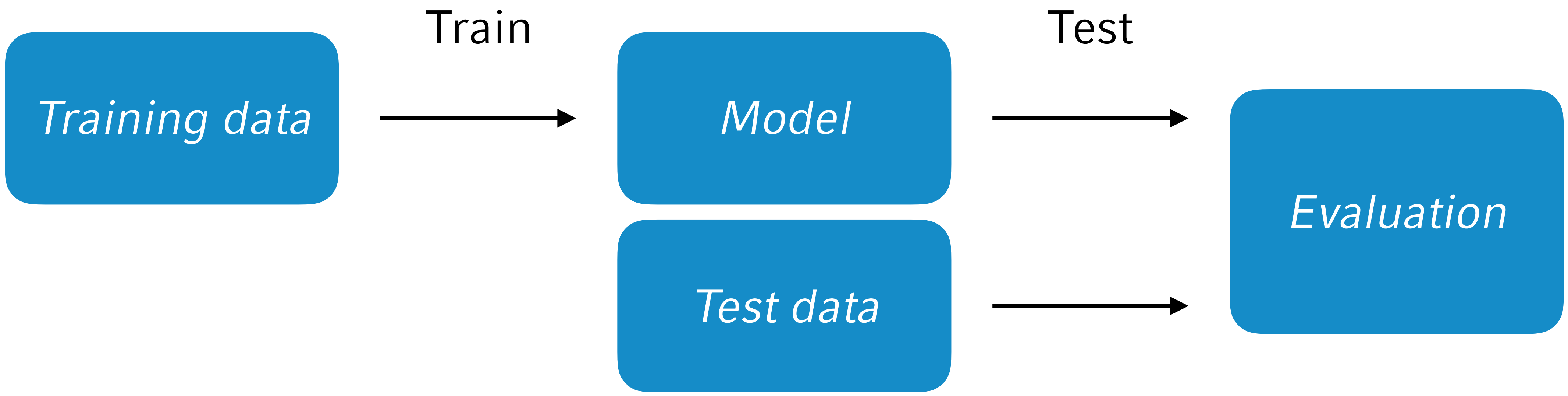


Machine learning

Interdisciplinary field of **computer science, statistics** and **probability theory**

Mathematical model mapping a set of input values to output values

Estimation of a statistical model from data (**learning**)

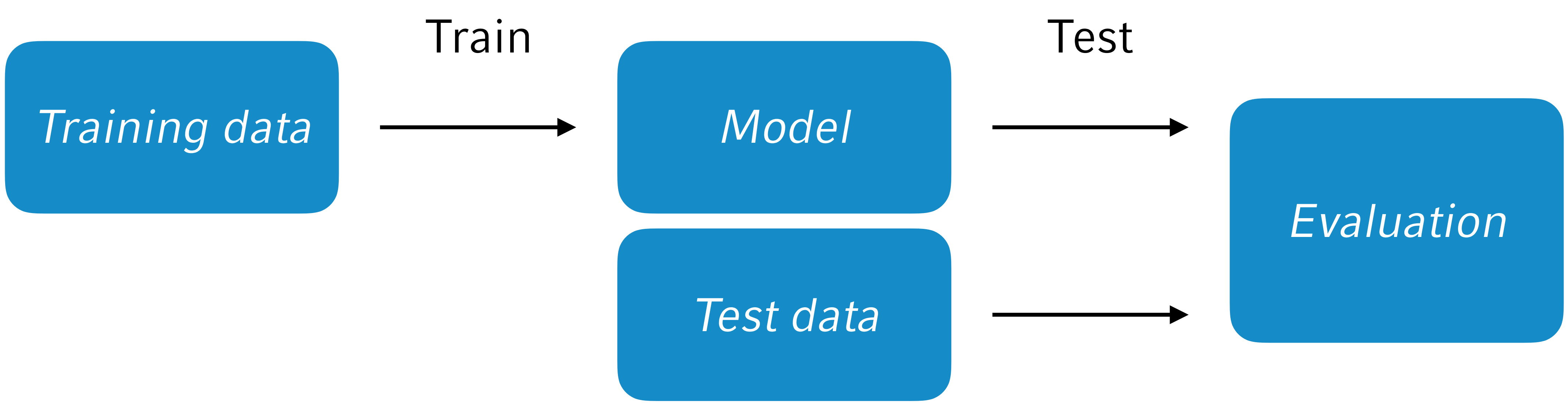


Machine learning

Interdisciplinary field of **computer science, statistics** and **probability theory**

Mathematical model mapping a set of input values to output values

Estimation of a statistical model from data (**learning**)

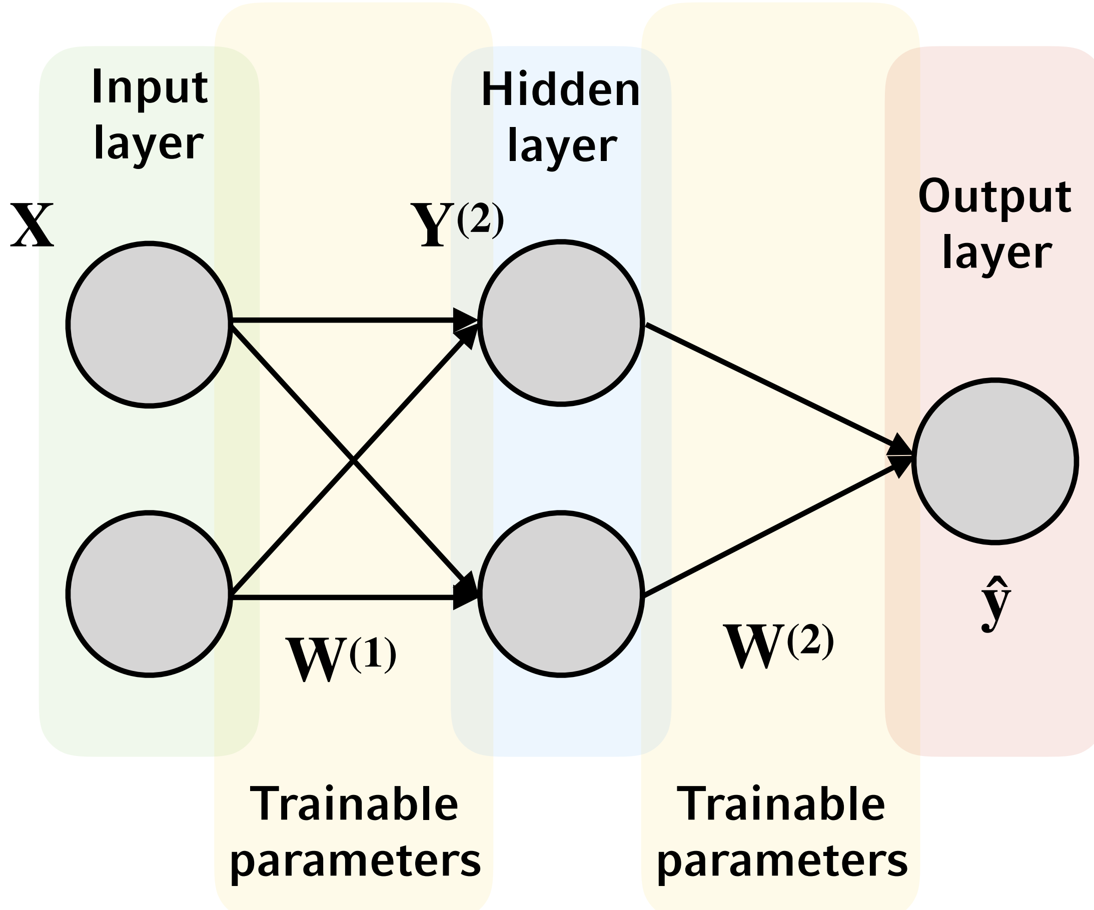


Make **predictions** on new data based on the estimated statistical model

Widely applied in HEP: **Analyis, Computing, Reconstruction, Triggering, etc.**

Multiple architectures available depending on the use case (boosted decision trees, neural networks, convolutional networks, …)

Neural networks (NN)



$$\mathbf{Y}^{(2)} = \mathbf{X}\mathbf{W}^{(1)}$$

$$\mathbf{a}^{(2)} = f(\mathbf{Y}^{(2)})$$

$$\mathbf{Y}^{(3)} = \mathbf{a}^{(2)}\mathbf{W}^{(2)}$$

$$\hat{\mathbf{y}} = f(\mathbf{Y}^{(3)})$$

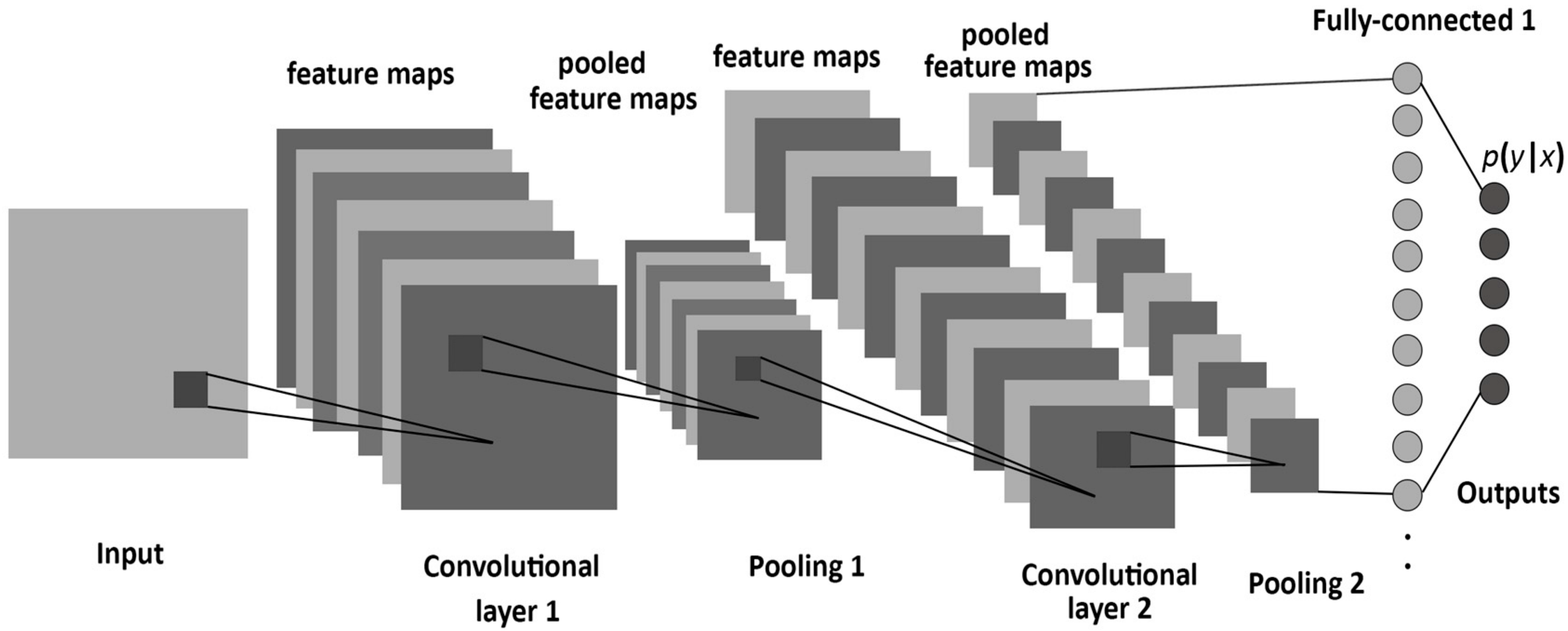
Form the basis of modern algorithms

Mapping an **n**-dimensional **input** to a **m**-dimnesional **output** by matrix multiplication

f indicates the activation of a single neuron (sigmoid, tanh, ReLU, …)

By optimising the weights the predicted output is optimised

Deep Convolutional Neural Network (CNN)



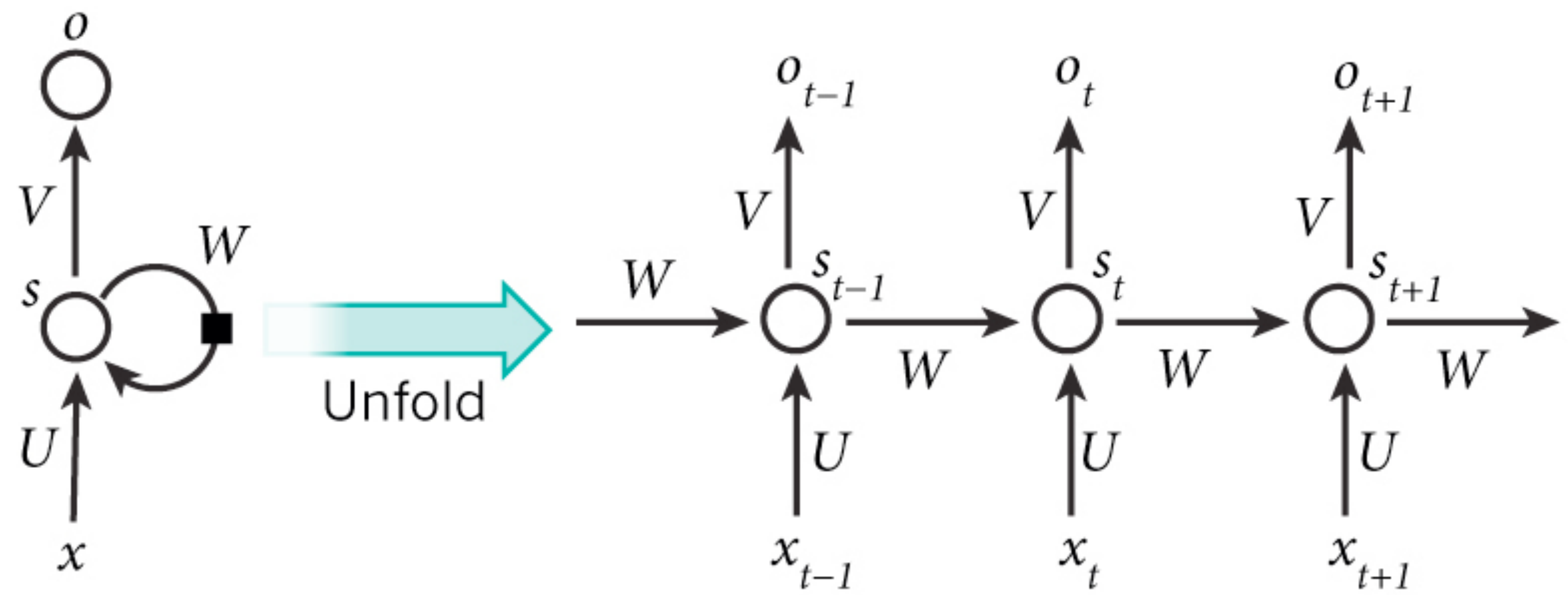
Entropy 2017, *19*(6), 242; doi:10.3390/e19060242

Processing data of a grid-like topology (e.g. 2-d images)

Convolutional layers are organised in feature maps (e.g. indicating different properties)

Pooling layer creating an '*invariance to local translations*'

Recurrent Neural Network (RNN)



'Deep Learning', doi:10.1038/nature14539

Map an input sequence onto an output sequence

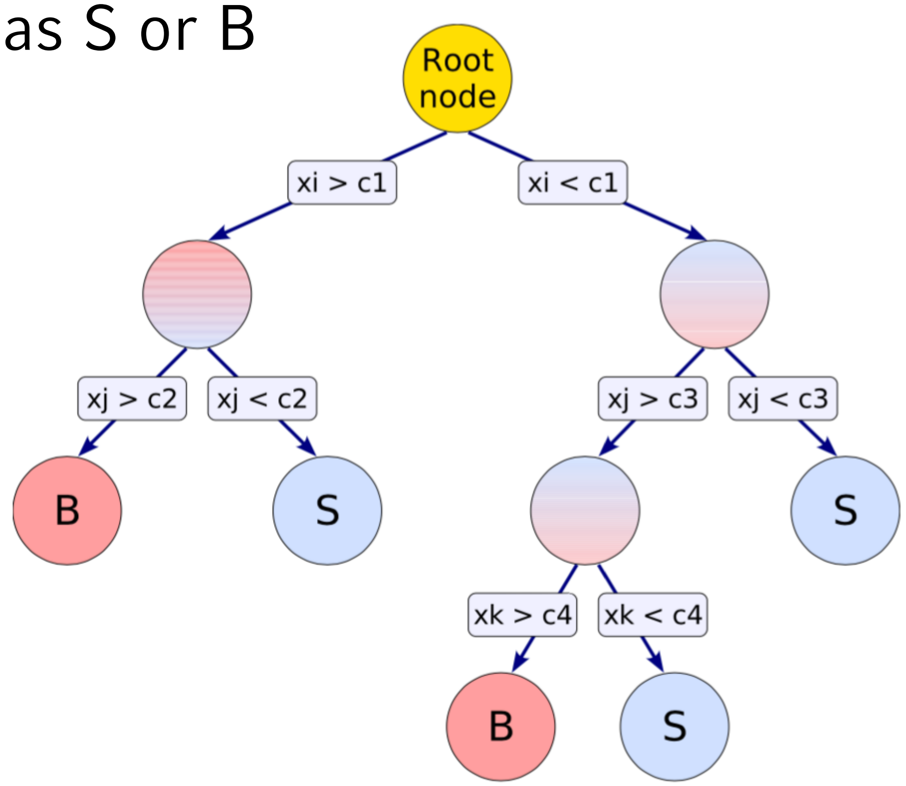
Neurons get inputs from other neurons at different time steps

Possible to process sequences of variable size

Boosted Decision Trees (BDT)

Sequential application of cuts, final nodes classify an event as S or B

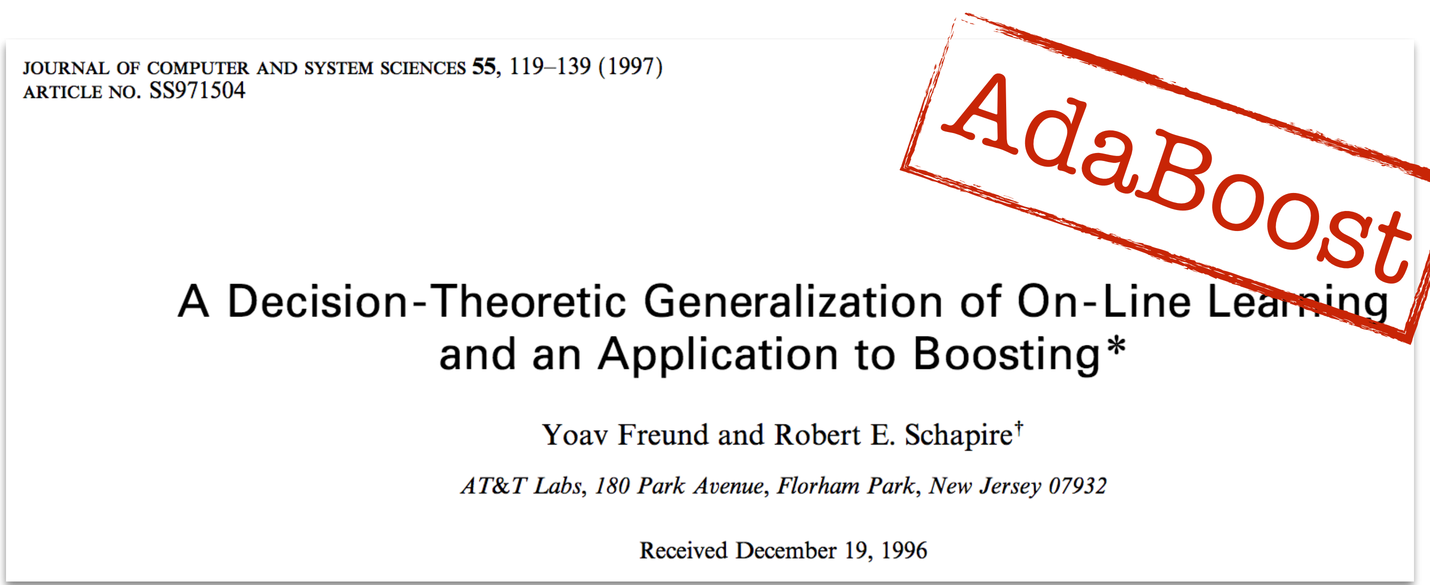
- Easy to interpret and visualise
- Weak variables are ignored (doesn't deteriorate the performance)
- But also very sensitive to statistical fluctuations in training data



For each variable **find** the best **partition** ("cut"), and repeat with each subsequent node

<u>Boosted Decision Trees (1996)</u>

- Build highly effective classifiers by combining a large number of mediocre ones



JOURNAL OF COMPUTER AND SYSTEM SCIENCES **55**, 119–139 (1997)
ARTICLE NO. SS971504

AdaBoost

A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*

Yoav Freund and Robert E. Schapire†

AT&T Labs, 180 Park Avenue, Florham Park, New Jersey 07932

Received December 19, 1996

ML techniques for
reconstruction

Quark versus gluon jets with jet images

ATL-PHYS-PUB-2017-017

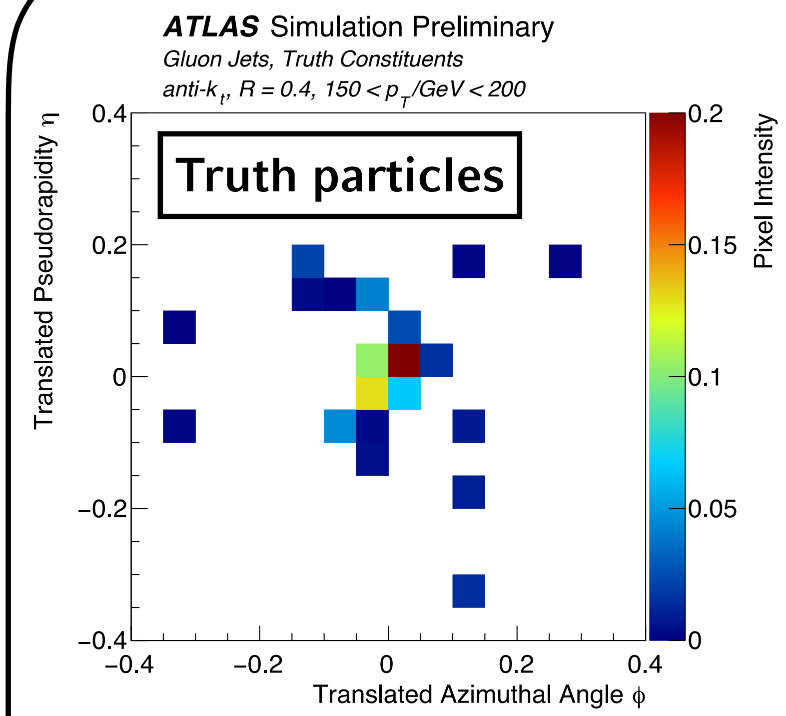
- Quarks and gluons **"hadronize"** in the detector and **form a jet**
- Differentiating between quark-/gluon-initiated jets has broad applicability in measurements and searches
- Full detector simulation based on **rotated, Lorentz boosted** and **normalised** fixed size grids **(jet images)**
- **CNN** utilises entire jet radiation pattern

Quark versus gluon jets with jet images

Gluon Jet

ATL-PHYS-PUB-2017-017

- Quarks and gluons **"hadronize"** in the detector and **form a jet**
- Differentiating between quark-/gluon-initiated jets has broad applicability in measurements and searches
- Full detector simulation based on **rotated, Lorentz boosted** and **normalised** fixed size grids **(jet images)**
- **CNN** utilises entire jet radiation pattern

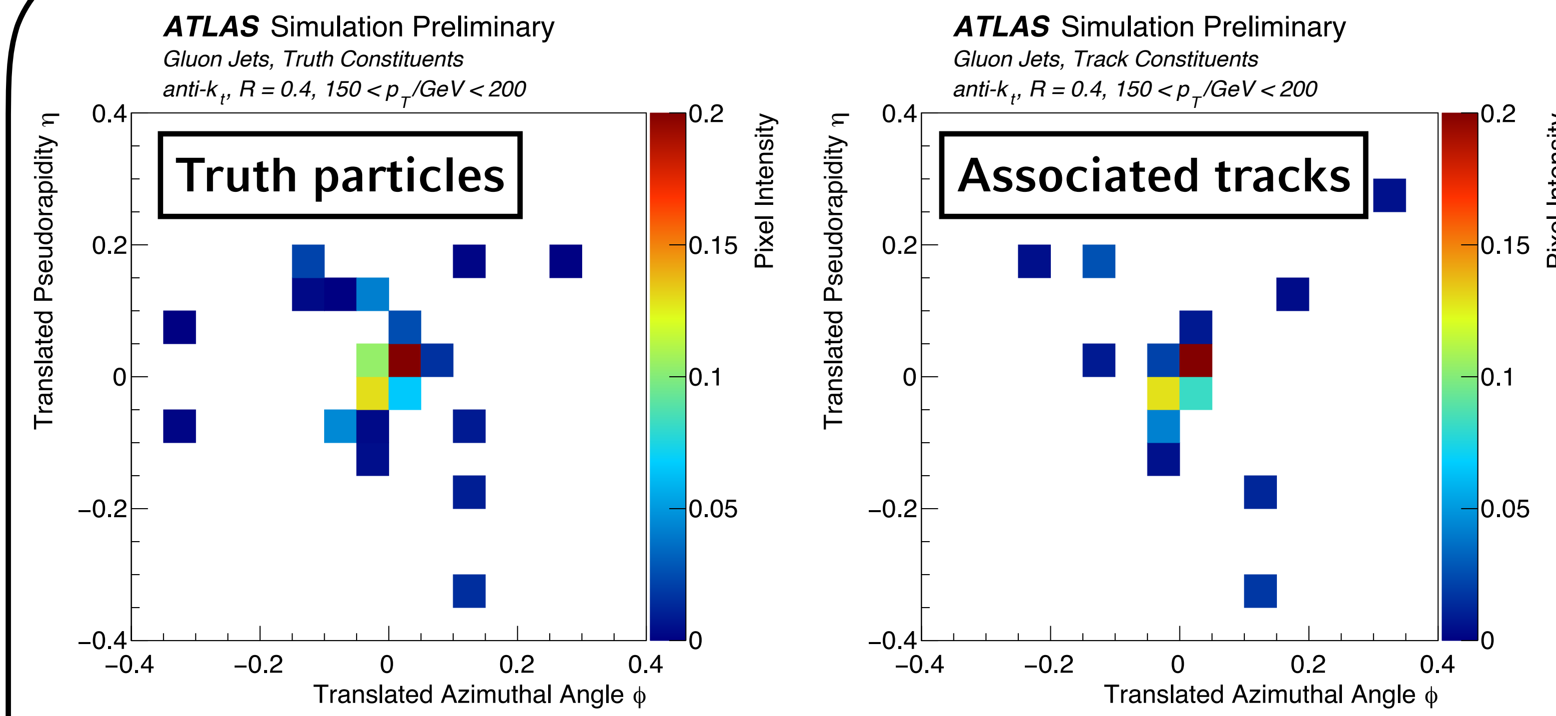


Quark versus gluon jets with jet images

ATL-PHYS-PUB-2017-017

- Quarks and gluons **"hadronize"** in the detector and **form a jet**
- Differentiating between quark-/gluon-initiated jets has broad applicability in measurements and searches
- Full detector simulation based on **rotated, Lorentz boosted** and **normalised** fixed size grids **(jet images)**
- **CNN** utilises entire jet radiation pattern

Gluon Jet

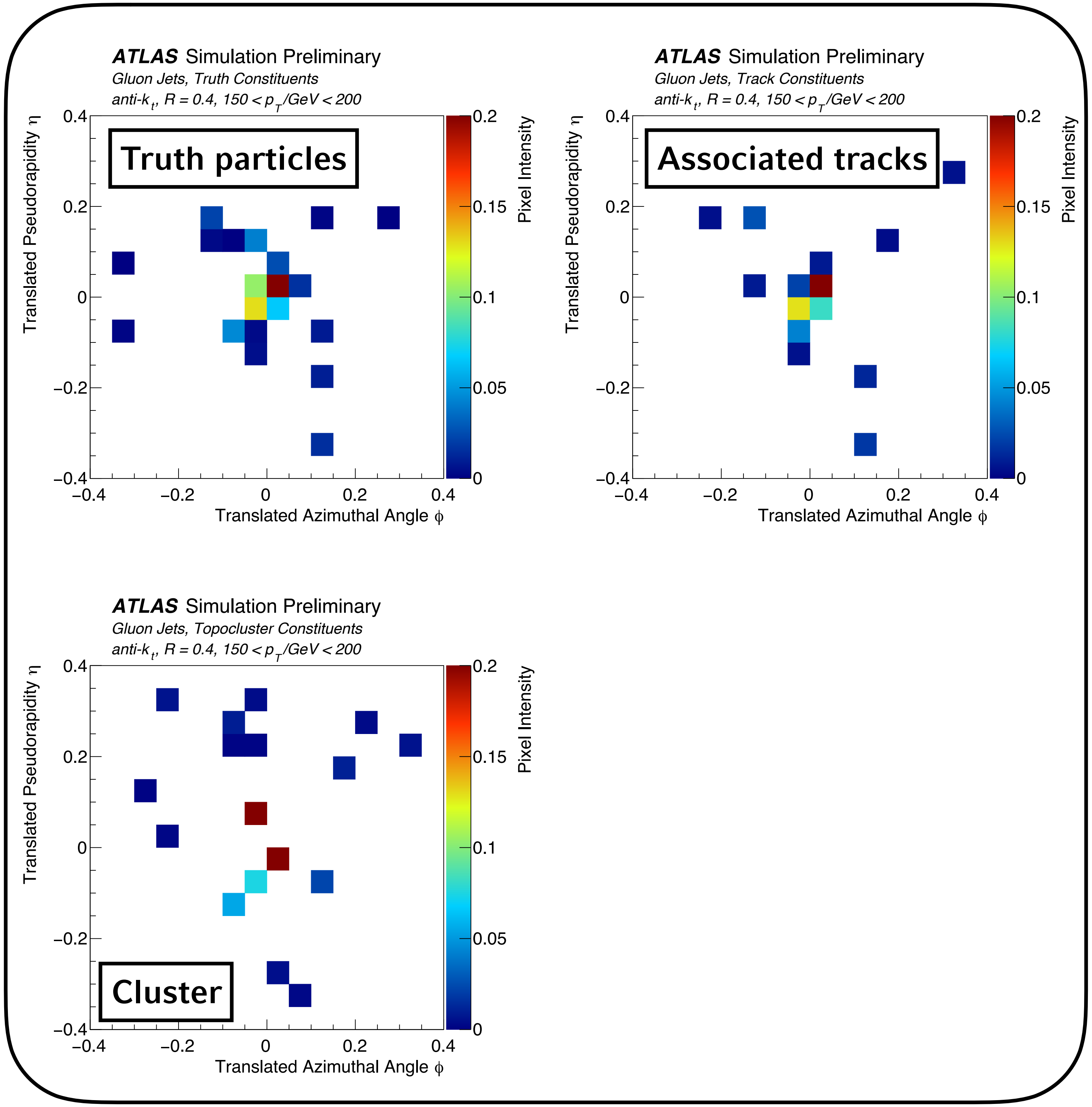


Quark versus gluon jets with jet images

[ATL-PHYS-PUB-2017-017]

- Quarks and gluons **“hadronize“** in the detector and **form a jet**
- Differentiating between quark-/gluon-initiated jets has broad applicability in measurements and searches
- Full detector simulation based on **rotated, Lorentz boosted** and **normalised** fixed size grids **(jet images)**
- **CNN** utilises entire jet radiation pattern

Gluon Jet

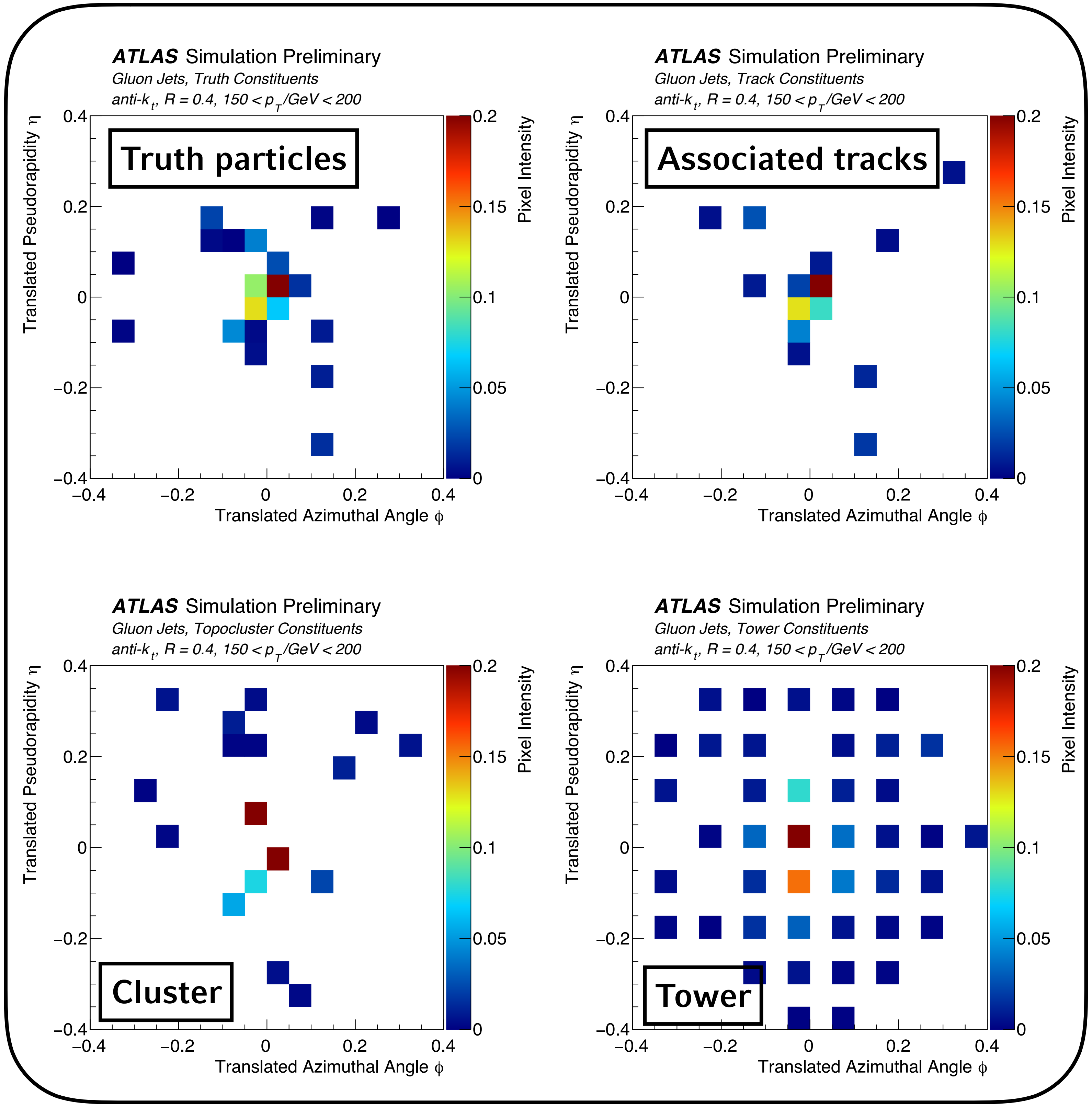


Quark versus gluon jets with jet images

ATL-PHYS-PUB-2017-017

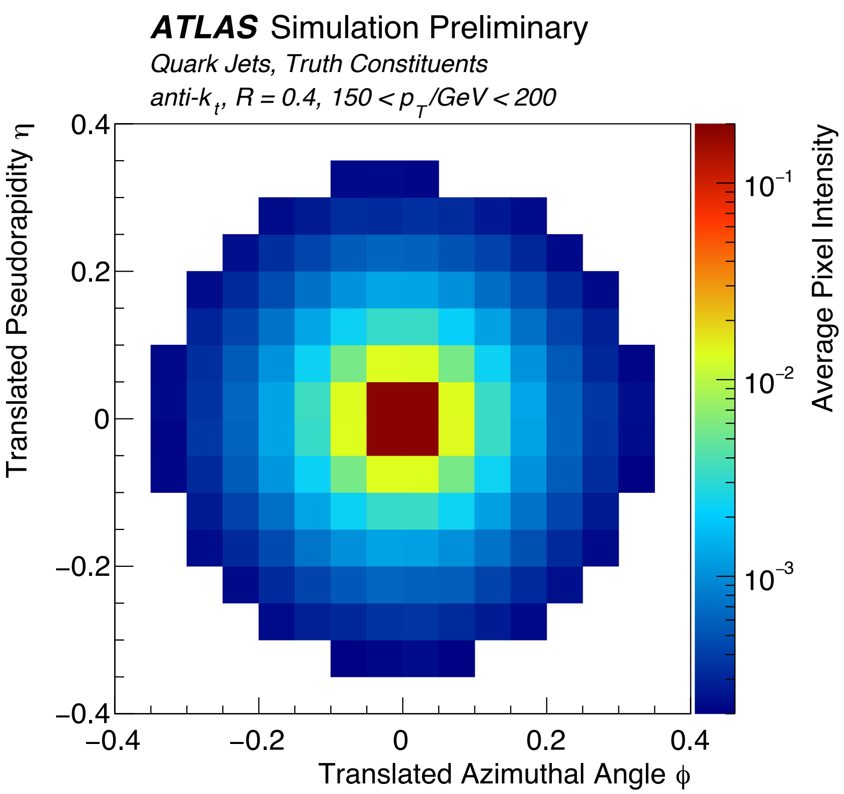
- Quarks and gluons **"hadronize"** in the detector and **form a jet**
- Differentiating between quark-/gluon-initiated jets has broad applicability in measurements and searches
- Full detector simulation based on **rotated, Lorentz boosted** and **normalised** fixed size grids **(jet images)**
- **CNN** utilises entire jet radiation pattern

Gluon Jet

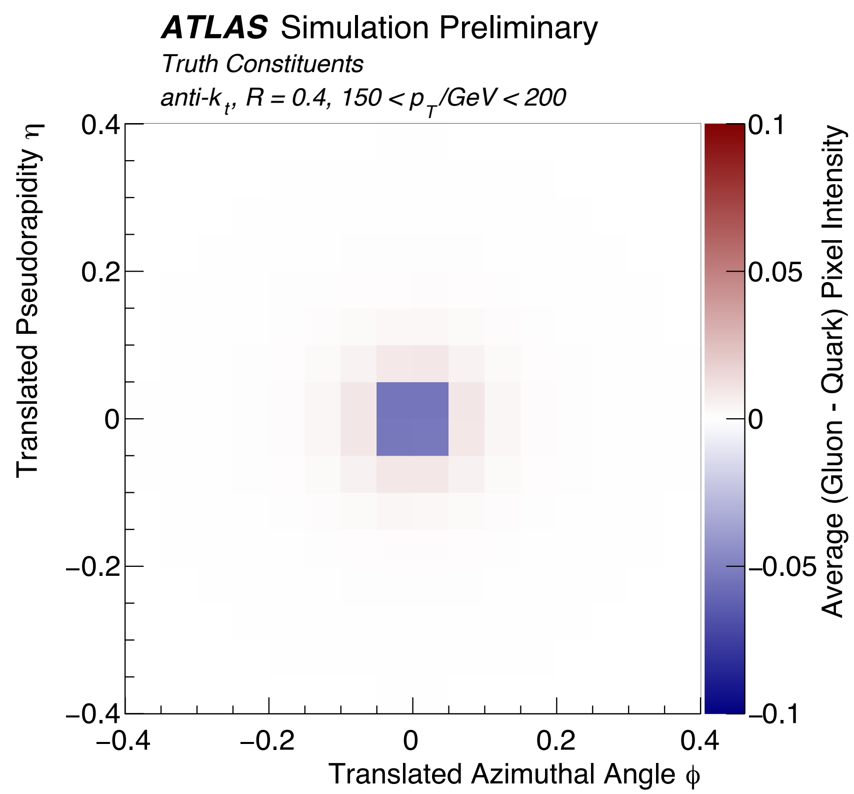


Average jet images

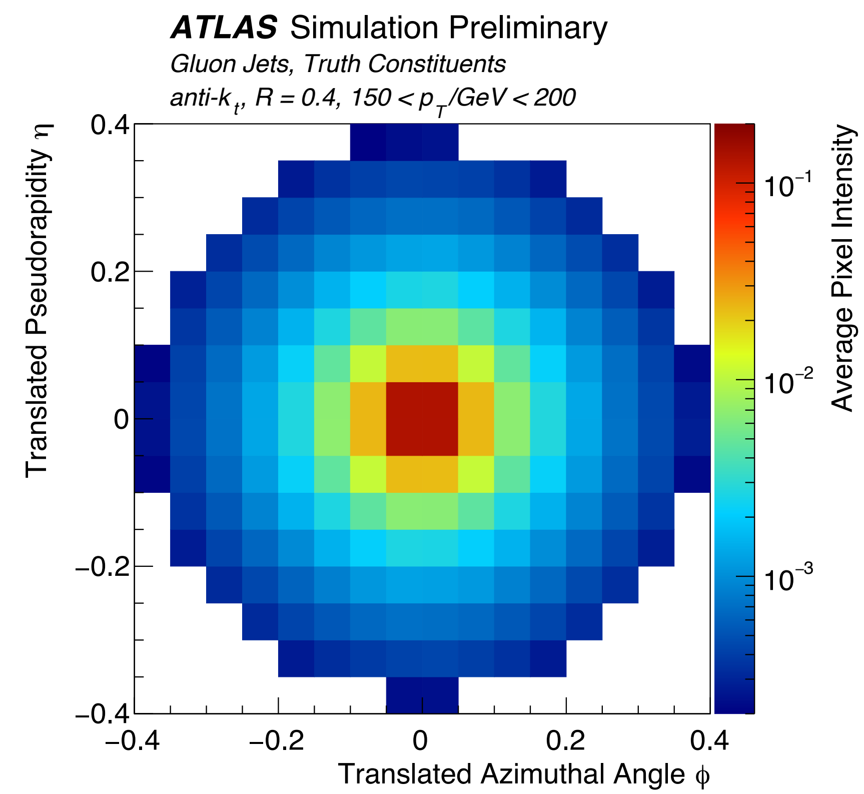
ATL-PHYS-PUB-2017-017



Quark, truth particles



Difference of these images

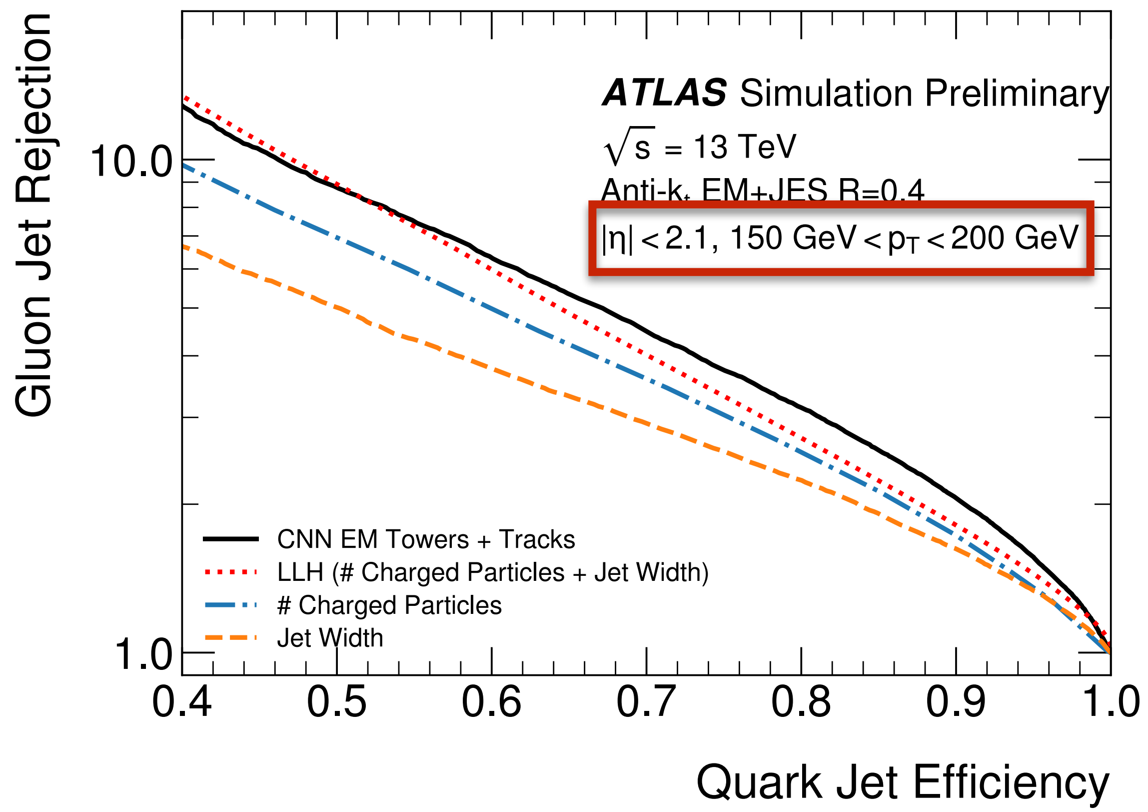


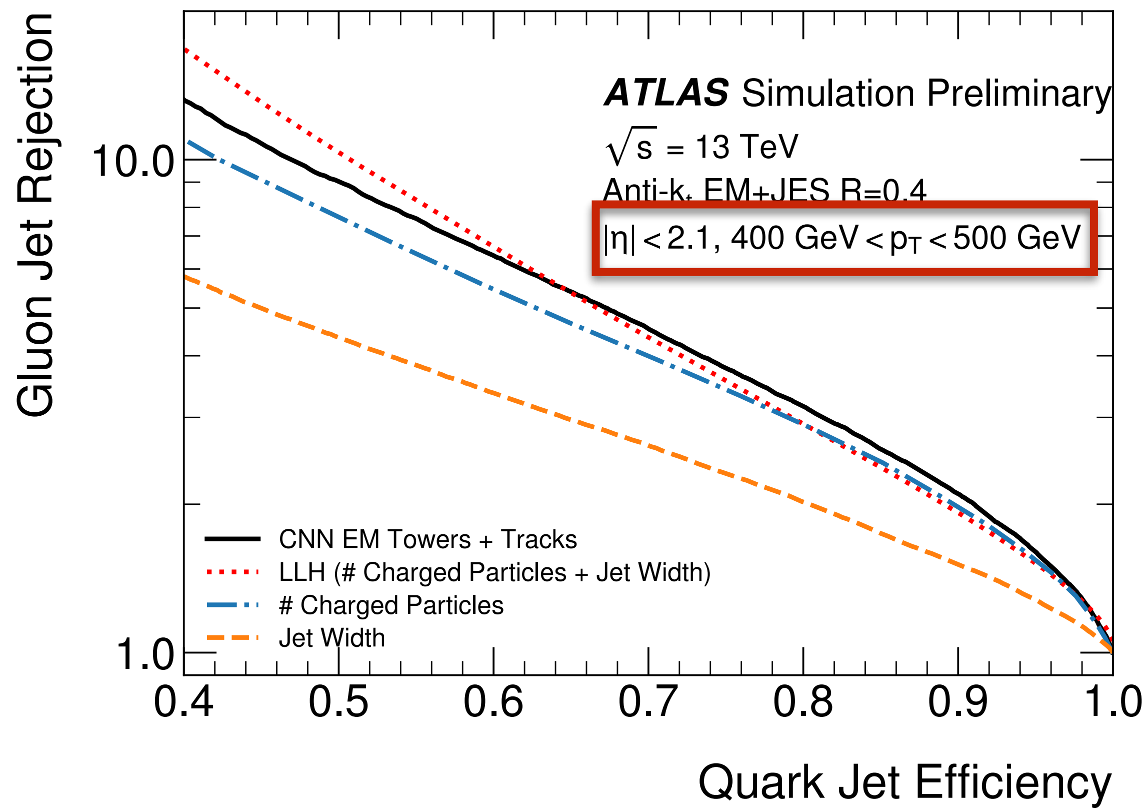
Gluon, truth particles

- **Gluon** jets tend to have **more constituents** and a **broader radiation pattern**
- Average is also determined for the other 3 types of images

Results of quark/gluon classification

ATL-PHYS-PUB-2017-017





- Two types of images are stacked and classification is performed
- CNN based tagging algorithm shows similar performance than individual physically motivated observables
- Further improvements are under investigation

Identifying b-tagged jets with RNN's

ATL-PHYS-PUB-2017-003

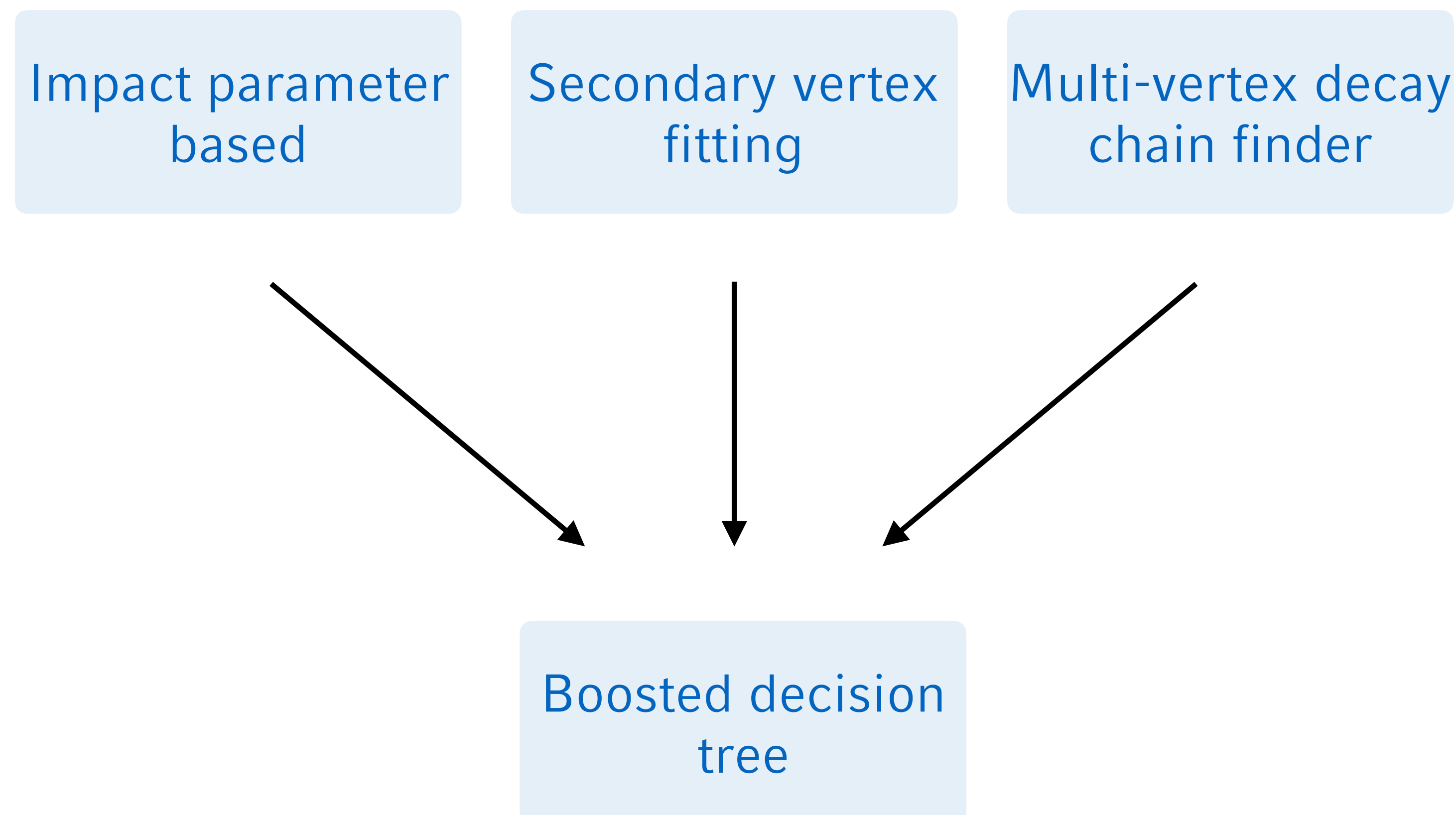
- Important for precise SM measurements ($H \rightarrow bb$) as well as exploring new physics
- Aim to **separate jets** containing a **b-hadron** from jets initiated by **lighter quark** flavours
- Classify 4 categories:
  b-, c-, light-hadrons and **hadronic τ** decays
- b-hadrons travel a few mm before decaying → secondary vertex

Identifying b-tagged jets with RNN’s

[ATL-PHYS-PUB-2017-003]

- Important for precise SM measurements ($H \to bb$) as well as exploring new physics
- Aim to **separate jets** containing a **b-hadron** from jets initiated by **lighter quark** flavours
- Classify 4 categories:
  b-, c-, light-hadrons and **hadronic τ** decays
- b-hadrons travel a few mm before decaying → secondary vertex

ATLAS default high-level algorithm



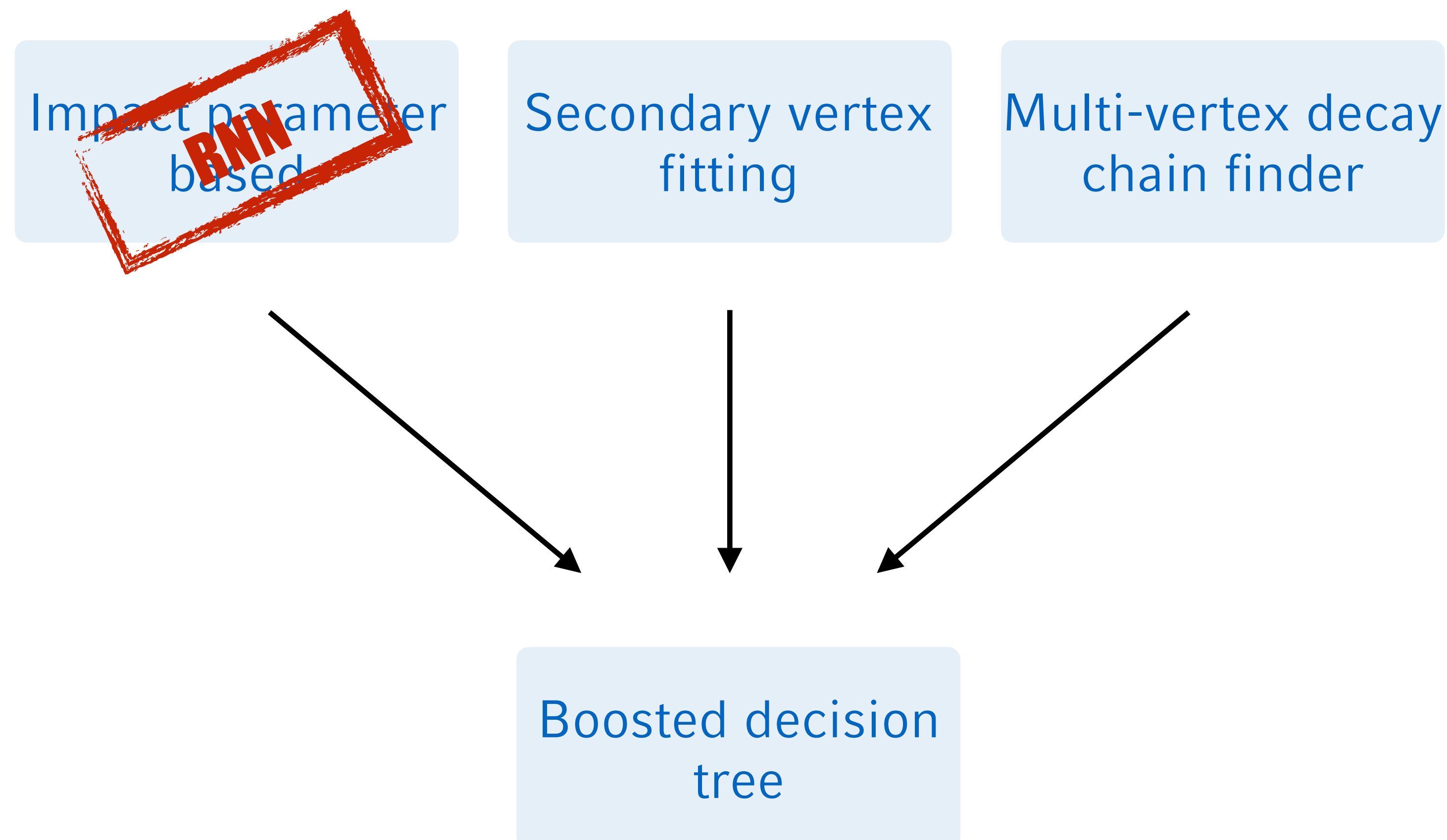
Identifying b-tagged jets with RNN's

ATL-PHYS-PUB-2017-003

- Important for precise SM measurements ($H \rightarrow bb$) as well as exploring new physics
- Aim to **separate jets** containing a **b-hadron** from jets initiated by **lighter quark** flavours
- Classify 4 categories:
  b-, c-, light-hadrons and **hadronic τ** decays
- b-hadrons travel a few mm before decaying → secondary vertex

Novel approach uses track properties as input to **RNN**

ATLAS default high-level algorithm

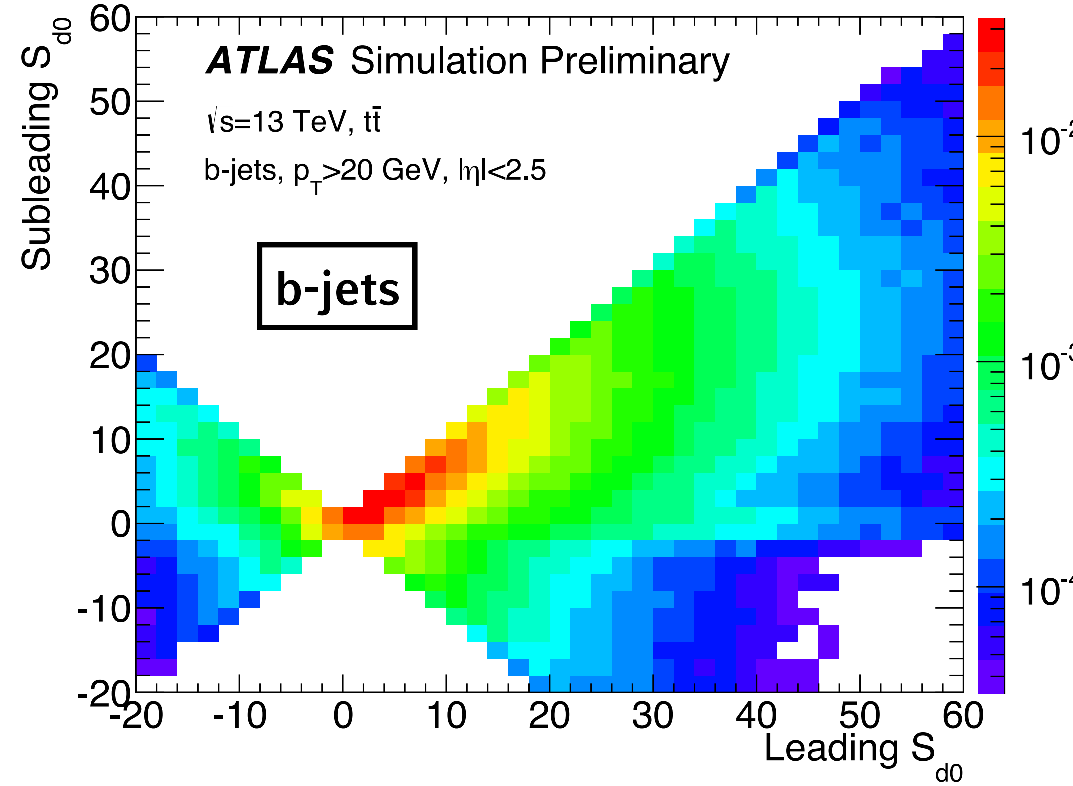


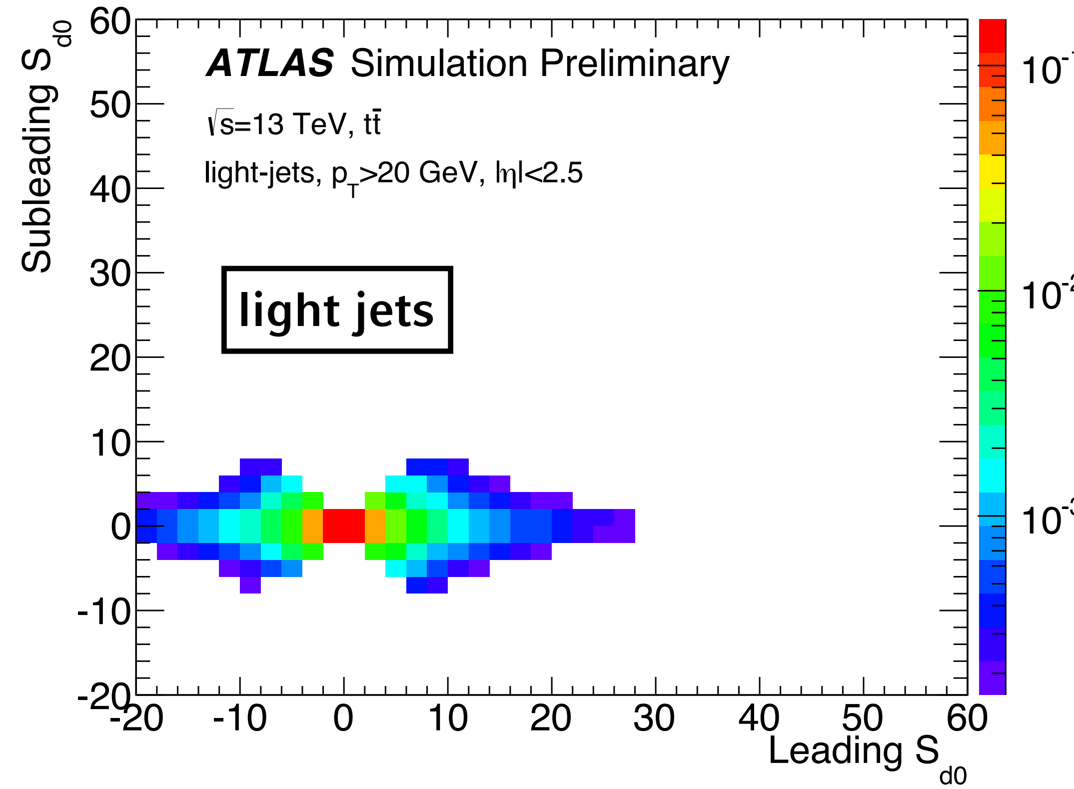
Identifying b-tagged jets with RNN’s

[ATL-PHYS-PUB-2017-003]()

- RNN replaces the impact parameter based algorithm
- Default impact parameter based algorithm builds a discriminant from a likelihood method
- RNN directly learns sequential dependencies (in this case multiple tracks according to a jet)

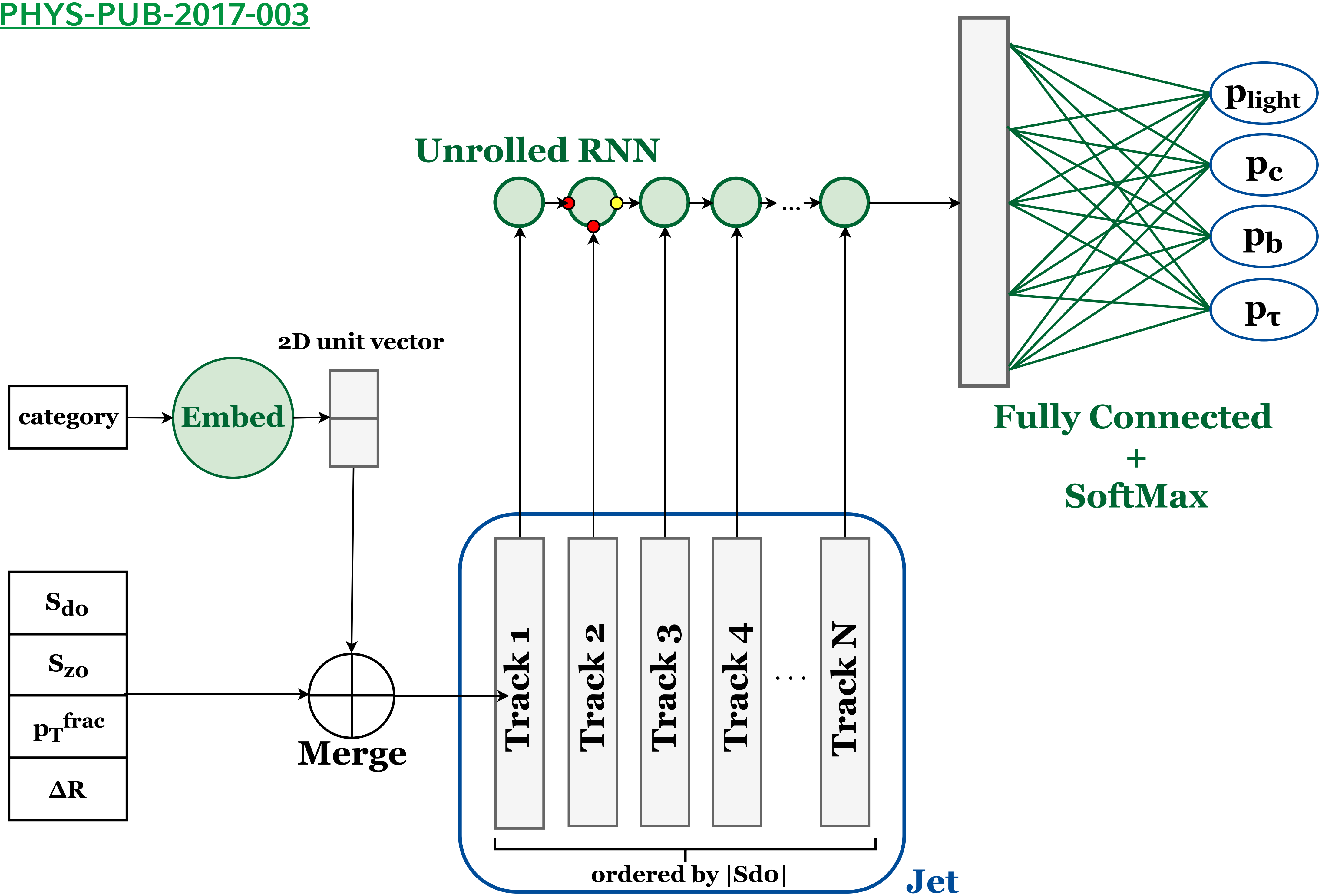
e.g. plots indicate the difference of the transverse impact parameter of the tracks associated to either a b-jet or a light flavour jet





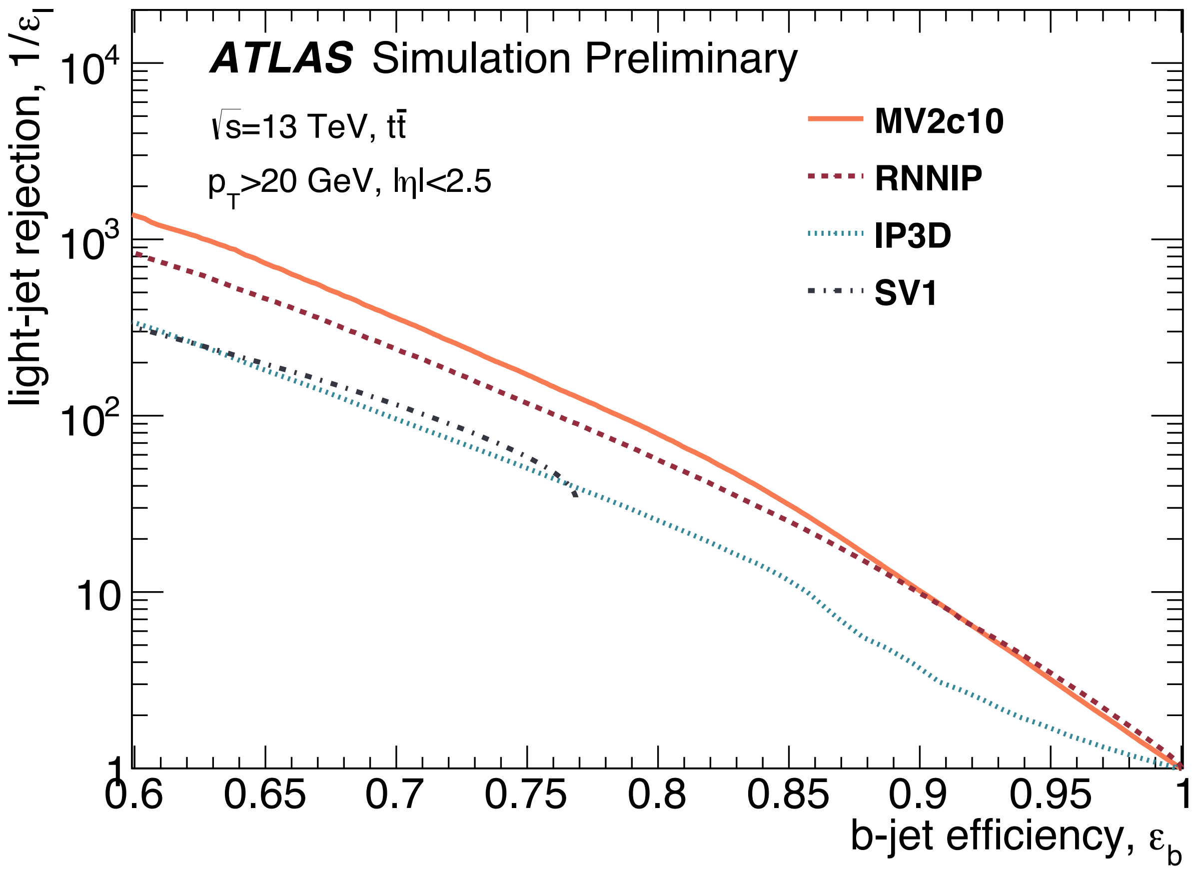
Network architecture

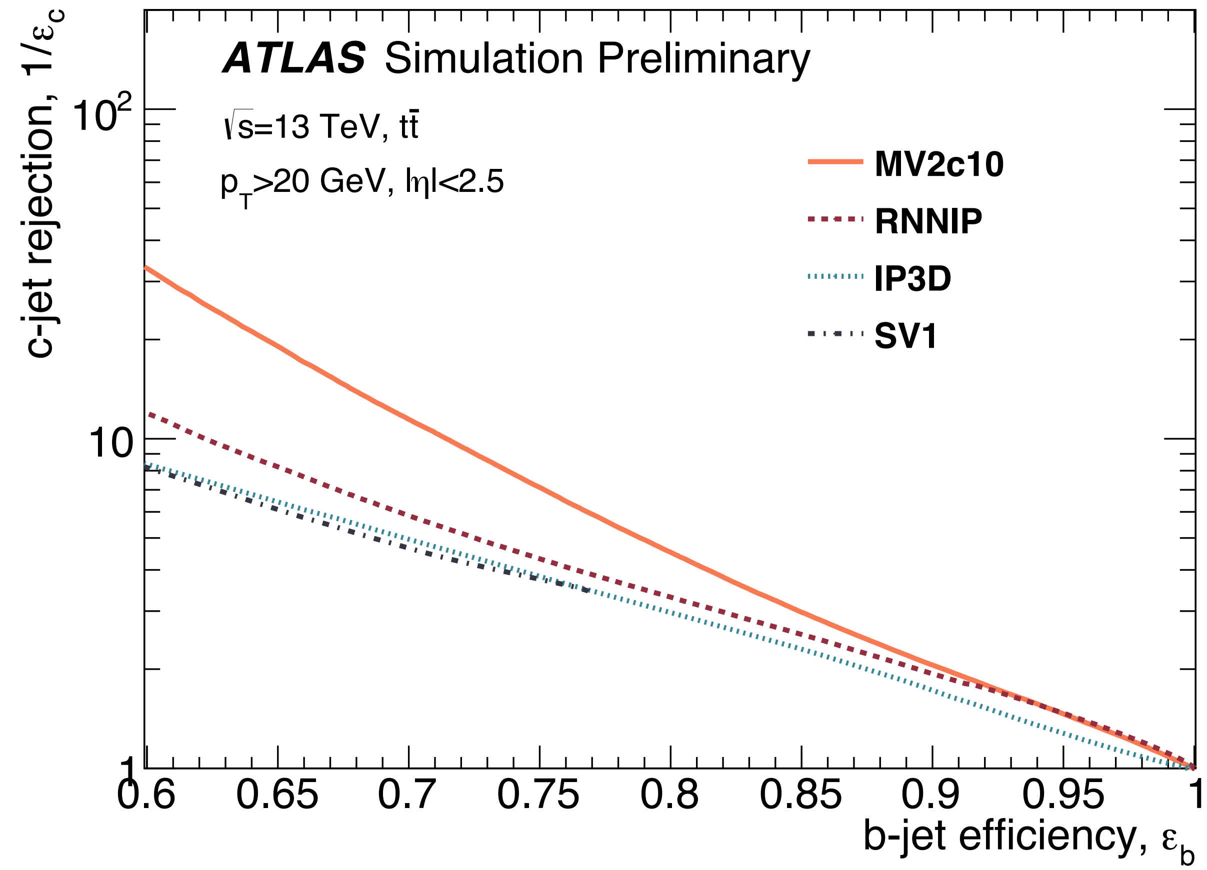
ATL-PHYS-PUB-2017-003



Performance of recurrent classifier

ATL-PHYS-PUB-2017-003





- Default high-level algorithm (MV2c10) also depicted as an upper limit on the performance
- Recurrent classifier (RNNIP) outperforms the current standard approach (IP3D)
- Additional studies ongoing to further improve the tagging efficiency

ML techniques in

data analyses

Multivariate techniques in physics analyses

Algorithms aim to **separate** particular events from each other
(e.g. hypothetical supersymmetric signature from SM background)

ML techniques to search for new particles is **relatively novel**

There are also potential risks that have to be reduced or completely avoided

Particular emphasis on following categories:

Multivariate techniques in physics analyses

Algorithms aim to **separate** particular events from each other
(e.g. hypothetical supersymmetric signature from SM background)

ML techniques to search for new particles is **relatively novel**

There are also potential risks that have to be reduced or completely avoided

Particular emphasis on following categories:

Statistics	***Optimisation***	***Validation***	***Systematics***

Multivariate techniques in physics analyses

Algorithms aim to **separate** particular events from each other
(e.g. hypothetical supersymmetric signature from SM background)

ML techniques to search for new particles is **relatively novel**

There are also potential risks that have to be reduced or completely avoided

Particular emphasis on following categories:

Statistics

- How many training events?
- Relation between trainable parameters and training events
- Ideas to increase statistics?

Optimisation

- Algorithm?
- Figure of merit?
- Input variables
- Trainable parameters
- Avoid overtraining

Validation

- Modeling of input?
- What does the algorithm learn?
- Correlations?
- Modeling in CR and VR

Systematics

- Fluctuations of syst. MC?
- Uncertainties for different algorithms?
- How to increase statistics?
- Ideas for mitigation?

Multivariate techniques in physics analyses

Algorithms aim to **separate** particular events from each other
(e.g. hypothetical supersymmetric signature from SM background)

ML techniques to search for new particles is **relatively novel**

There are also potential risks that have to be reduced or completely avoided

Particular emphasis on following categories:

Statistics

- How many training events?
- Relation between trainable parameters and training events
- Ideas to increase statistics?

Optimisation

- Algorithm?
- Figure of merit?
- Input variables
- Trainable parameters
- Avoid overtraining

Validation

- Modeling of inputs?
- What does the algorithm learn?
- Correlations?
- Modeling in CR and VR

Systematics

- Fluctuations of syst. MC?
- Uncertainties for different algorithms?
- How to increase statistics?
- Ideas for mitigation?

No tight order of these categories — makes it more complex

Multivariate techniques in physics analyses

Algorithms aim to **separate** particular events from each other
(e.g. hypothetical supersymmetric signature from SM background)

ML techniques to search for new particles is **relatively novel**

There are also potential risks that have to be reduced or completely avoided

Particular emphasis on following categories:

Statistics

- How many training events?
- Relation between trainable parameters and training events
- Ideas to increase statistics?

Optimisation

- Algorithm?
- Figure of merit?
- Input variables
- Trainable parameters
- Avoid overtraining

Validation

- Modeling of inputs?
- What does the algorithm learn?
- Correlations?
- Modeling in CR and VR

Systematics

- Fluctuations of syst. MC?
- Uncertainties for different algorithms?
- How to increase statistics?
- Ideas for mitigation?

No tight order of these categories — makes it more complex

Understanding what the algorithm learns is vital!

How much training data is needed

"Get as much as you can!" — every Data Scientist always

No accurate answer! — Strongly depends on complexity of problem and learning algorithm

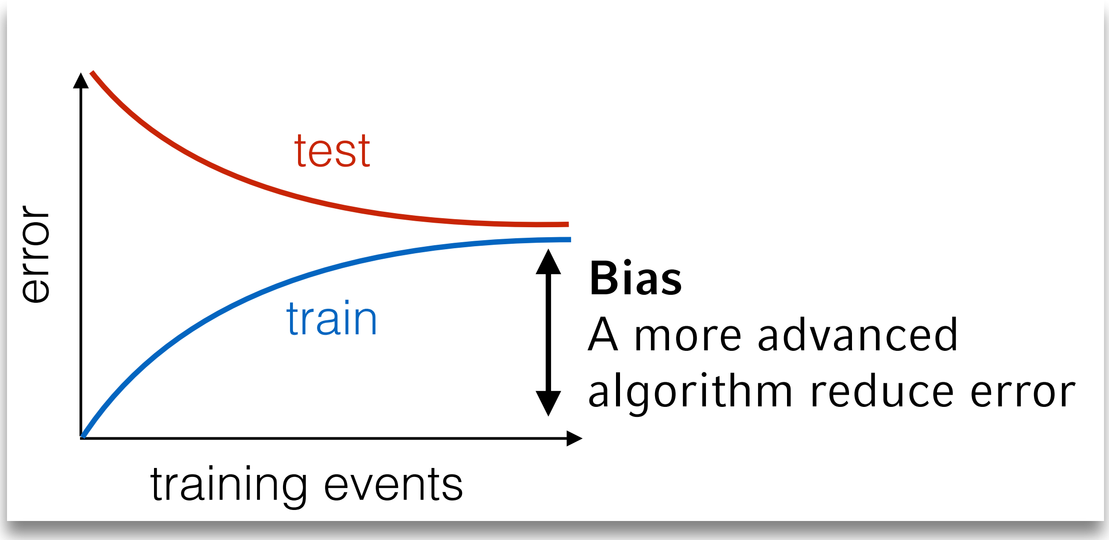
How much training data is needed

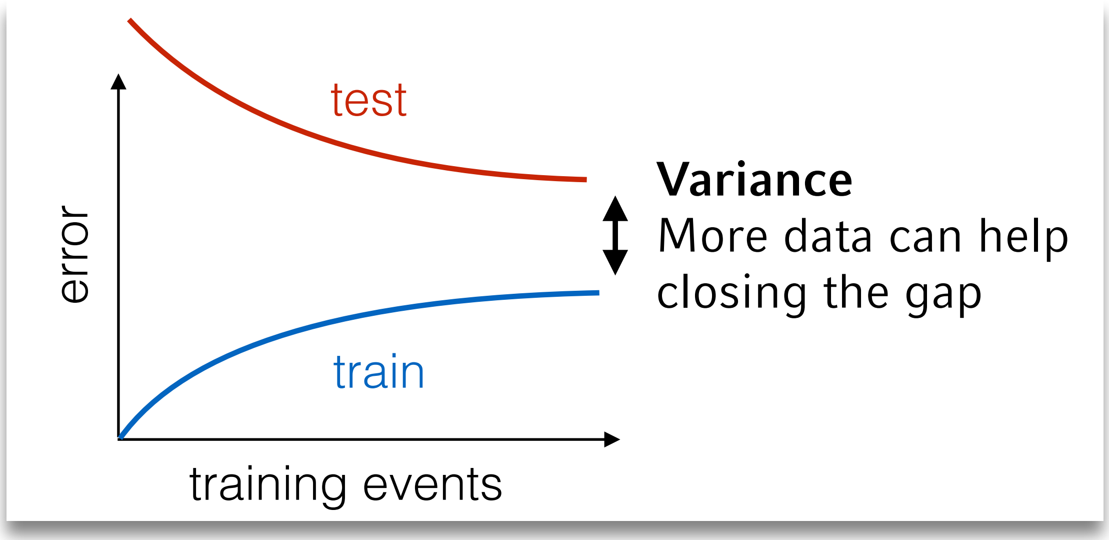
"Get as much as you can!" — every Data Scientist always

No accurate answer! — Strongly depends on complexity of problem and learning algorithm

Rules and tools that help to tackle this question:

- Perform learning curves
  (Error function can be mean squared error or any other)





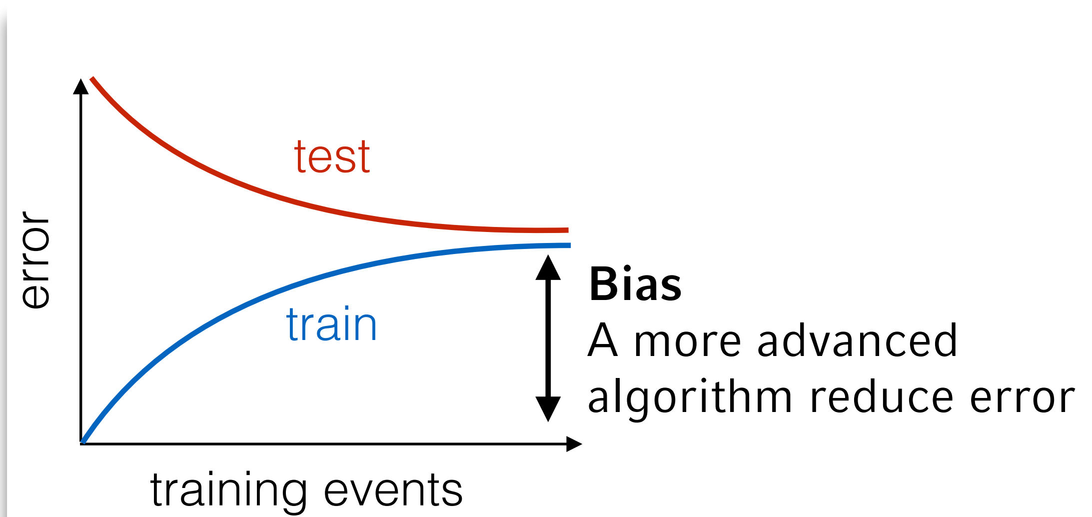
How much training data is needed

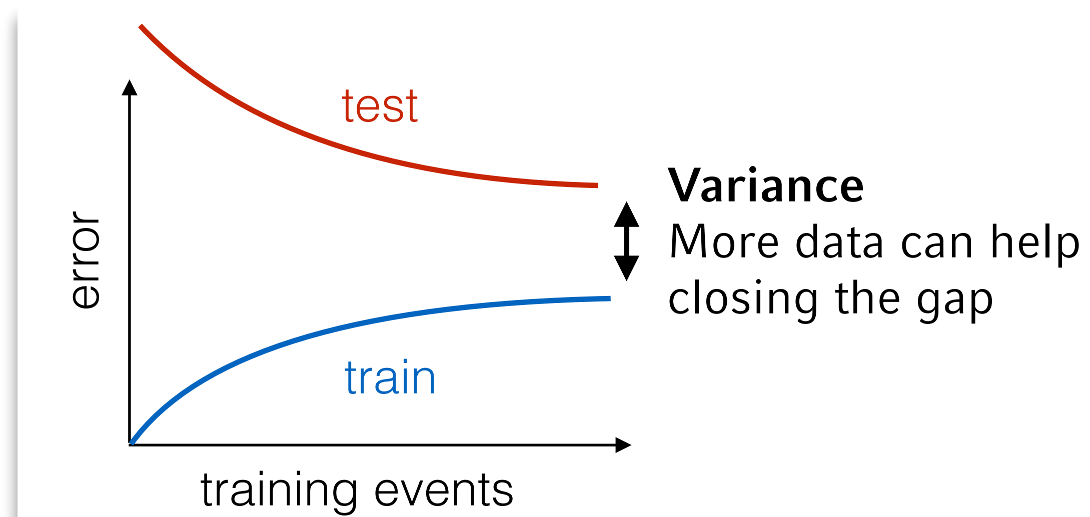
"Get as much as you can!" — every Data Scientist always

No accurate answer! — Strongly depends on complexity of problem and learning algorithm

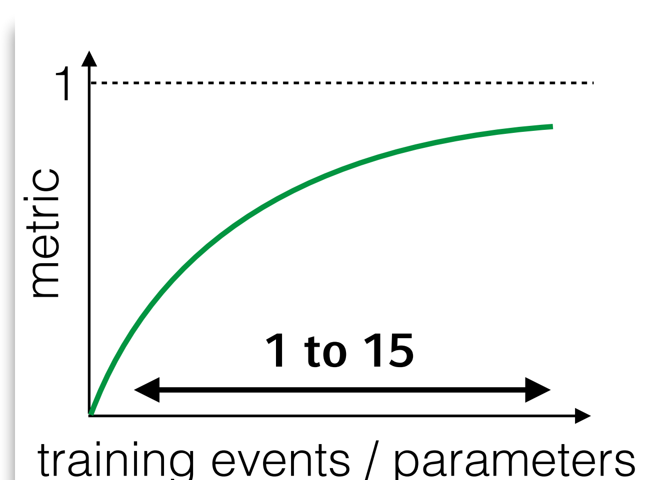
Rules and tools that help to tackle this question:

- Perform learning curves
  (Error function can be mean squared error or any other)





- 10-1 x more training data than trainable parameters
- Chosen metric also depends on the problem



💡 Metrics are parameters to evaluate algorithm performance

Associated production of Higgs boson with top quark pairs

arxiv:1712.08891

Higgs boson discovery by the **ATLAS** and **CMS** collaborations was a crucial milestone

Measuring Yukawa interactions are important, which account for fermion masses

So far, only the decay $\boldsymbol{H \to \tau\tau}$ has been observed and evidence of $\boldsymbol{H \to bb}$ has been found

Coupling of the **Higgs** boson to **top quark** could be sensitive to effects beyond the SM

Direct measurement can be achieved via the process $\mathbf{gg/q\bar{q} \to t\bar{t}H}$

Associated production of Higgs boson with top quark pairs

arxiv:1712.08891

Higgs boson discovery by the **ATLAS** and **CMS** collaborations was a crucial milestone

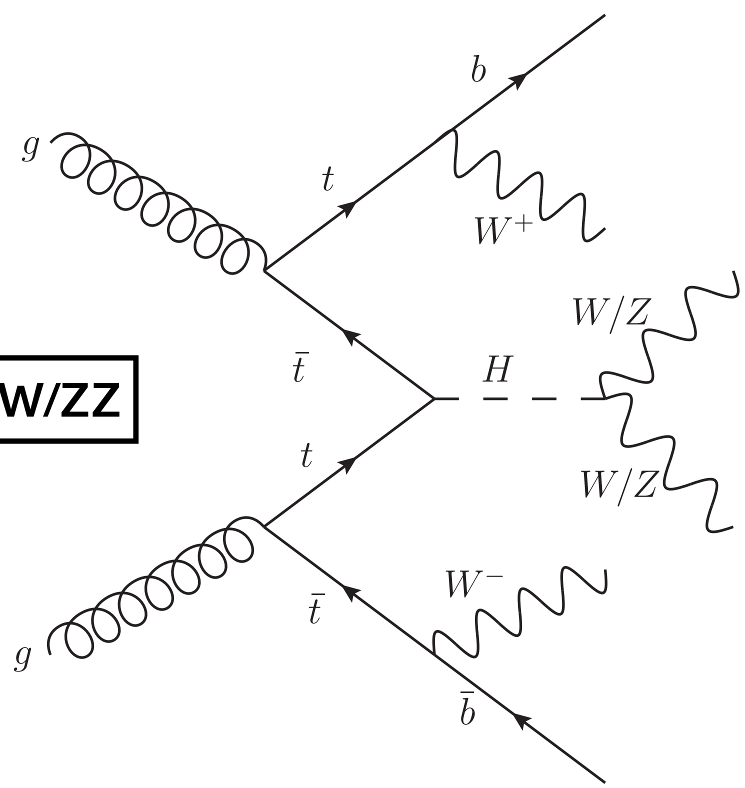
Measuring Yukawa interactions are important, which account for fermion masses

So far, only the decay $\boldsymbol{H \rightarrow \tau\tau}$ has been observed and evidence of $\boldsymbol{H \rightarrow bb}$ has been found

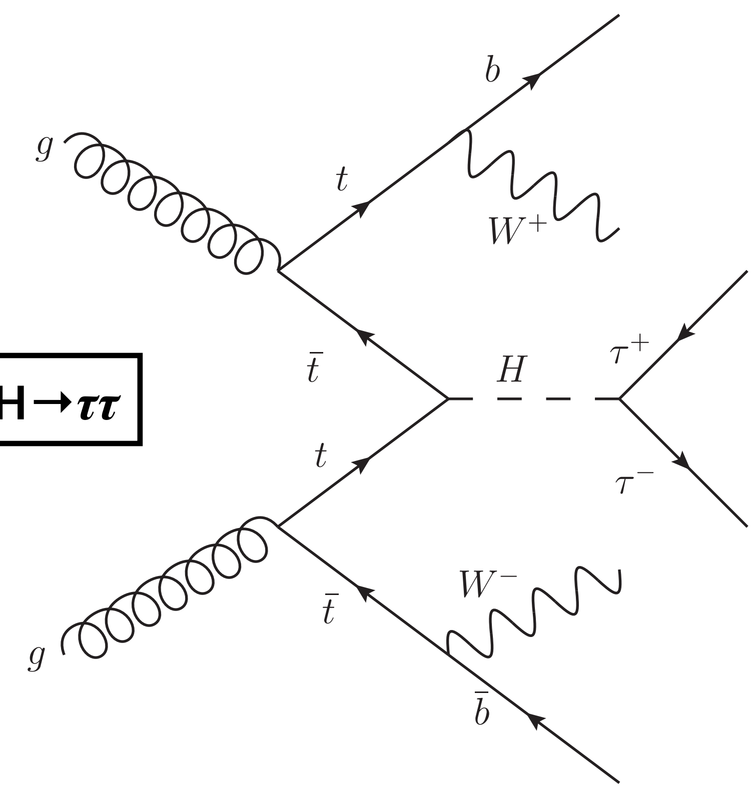
Coupling of the **Higgs** boson to **top quark** could be sensitive to effects beyond the SM

Direct measurement can be achieved via the process **gg/q$\bar{\text{q}}$→t$\bar{\text{t}}$H**

$\boxed{\mathbf{t\bar{t}H \rightarrow WW/ZZ}}$



$\boxed{\mathbf{t\bar{t}H \rightarrow \tau\tau}}$



Analysis strategy ($t\bar{t}H$)

arxiv:1712.08891

$t\bar{t}H$ production cross section is very **small** compared to SM background

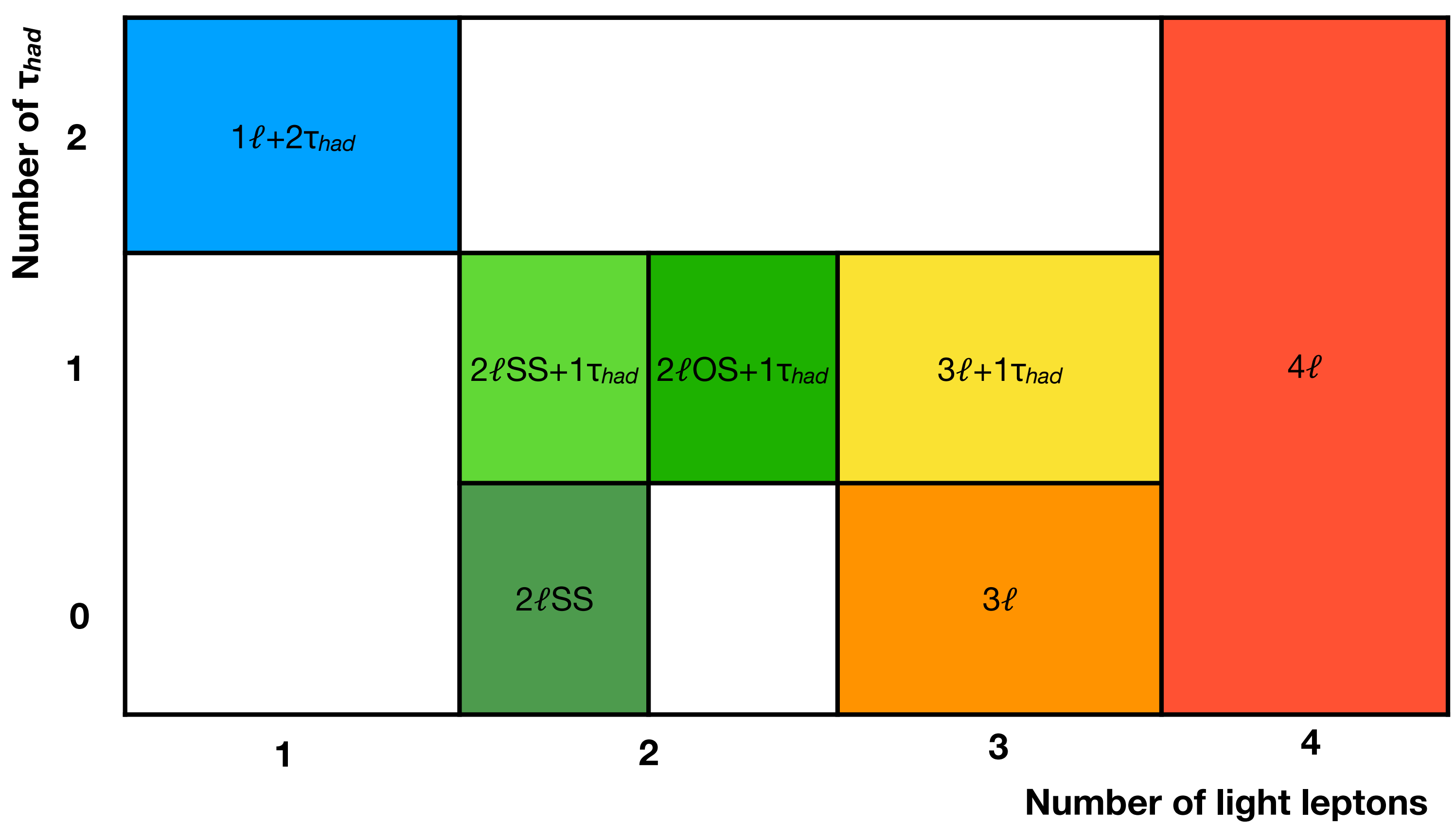
Extensive search strategy has been performed with many different final states:
2 - 4 lepton final states considering **electrons**, **muons** and hadronically decaying **taus**

Analysis strategy ($t\bar{t}H$)

arxiv:1712.08891

$t\bar{t}H$ production cross section is very **small** compared to SM background

Extensive search strategy has been performed with many different final states:
2 - 4 lepton final states considering **electrons**, **muons** and hadronically decaying **taus**



Analysis strategy ($t\bar{t}H$)

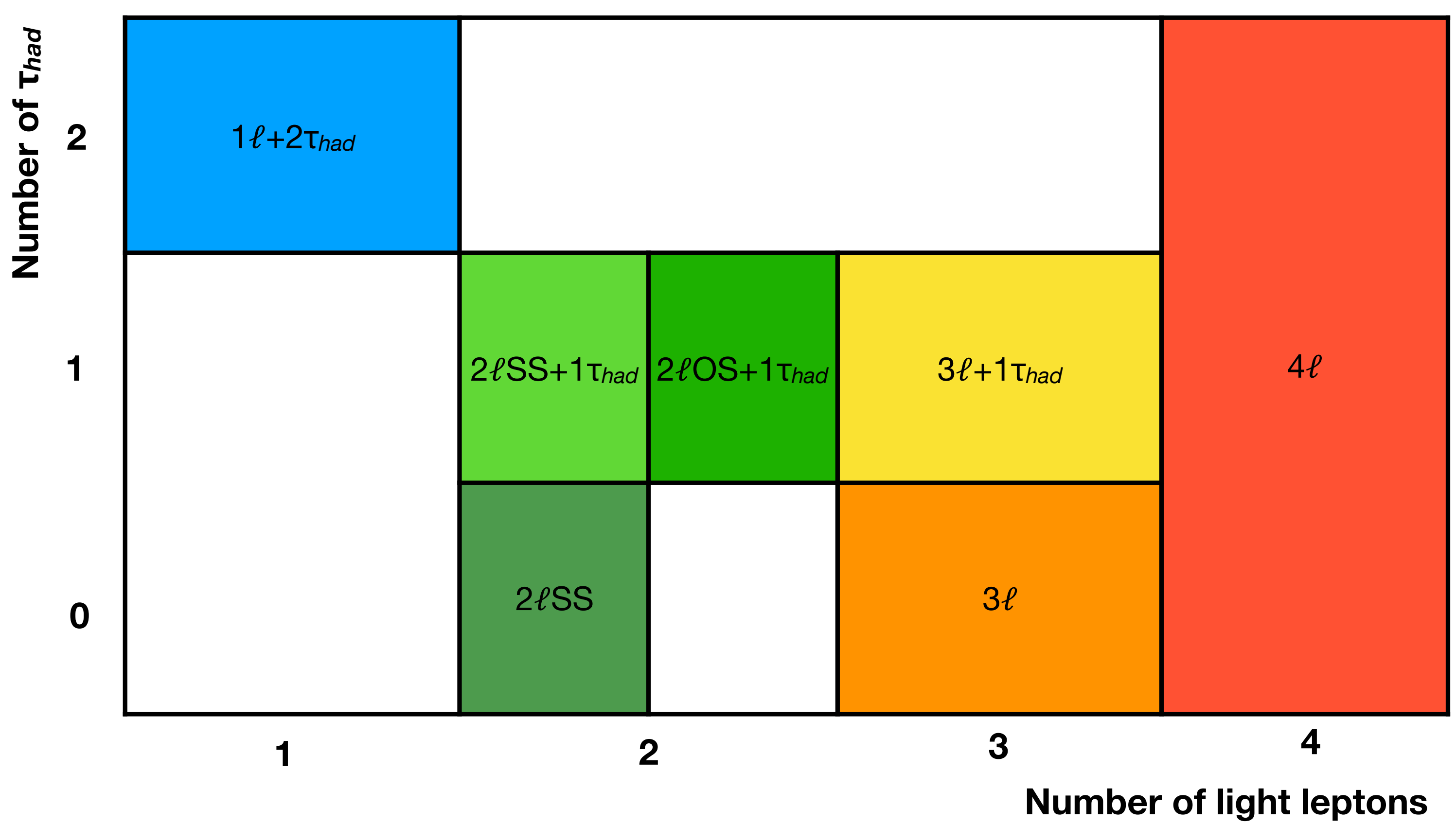
arxiv:1712.08891

Each of these "channels" are further splitted:

- into **control regions (CR)** for background estimations
- into **signal regions (SR)** with enhanced sensitivity

In total **332 030 events** are selected in data — **91! expected signal events**

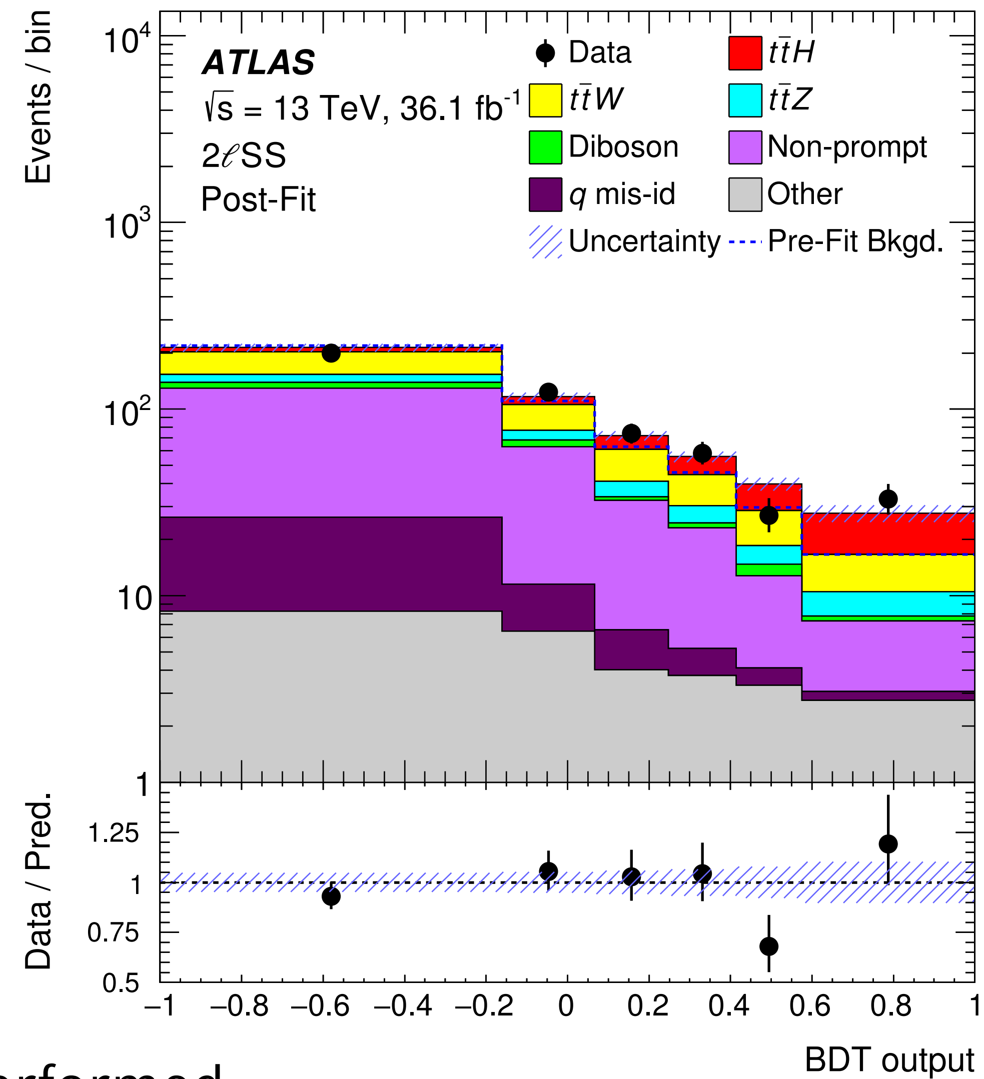
All channels perform BDT's to further improve the signal sensitivity



Results ($t\bar{t}H$)

arxiv:1712.08891

A **maximum-likelihood fit** is performed **simultaneously on all search regions** to extract the ttH cross section normalised to SM prediction



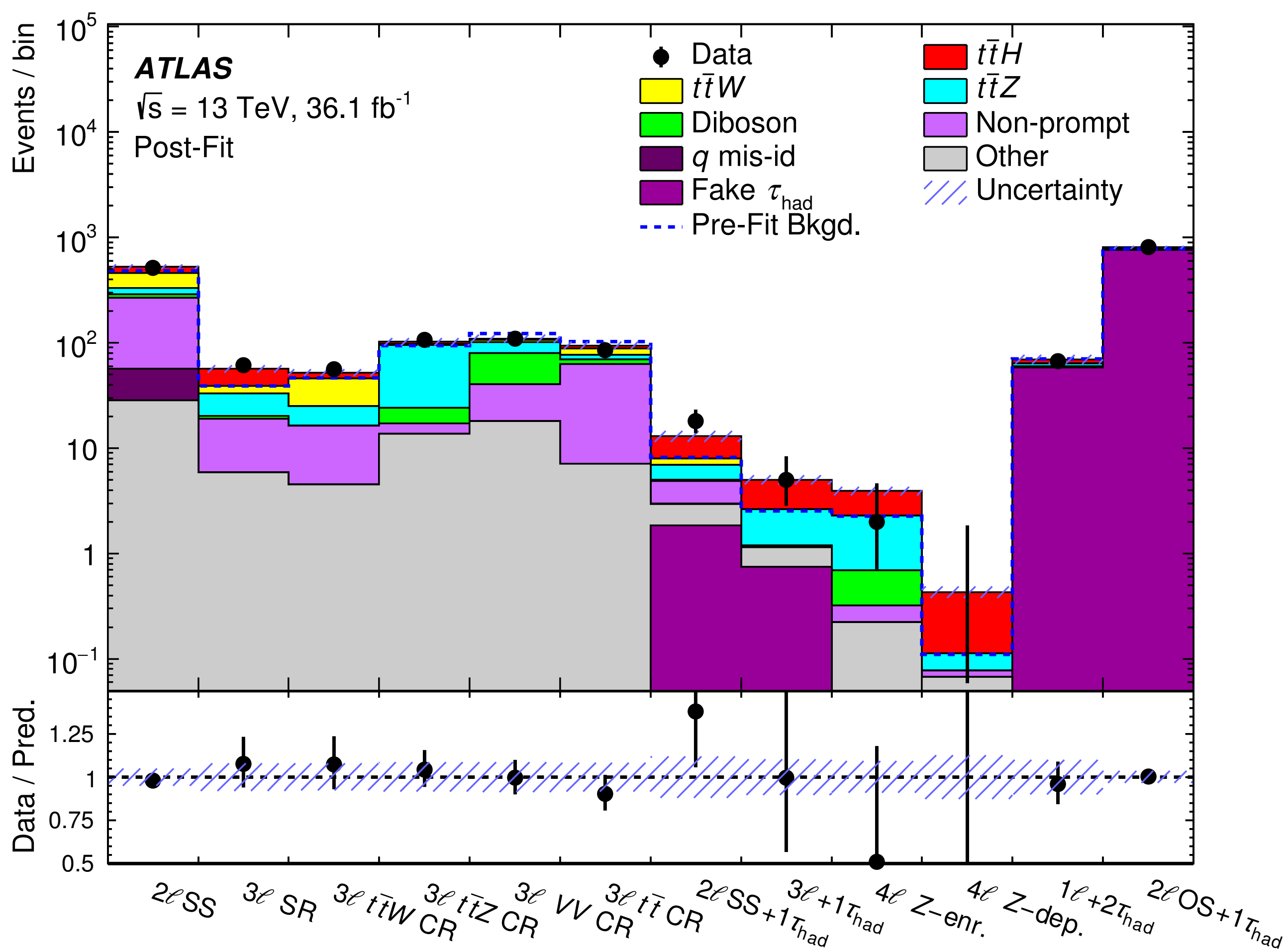
Example: **2 lepton channel**

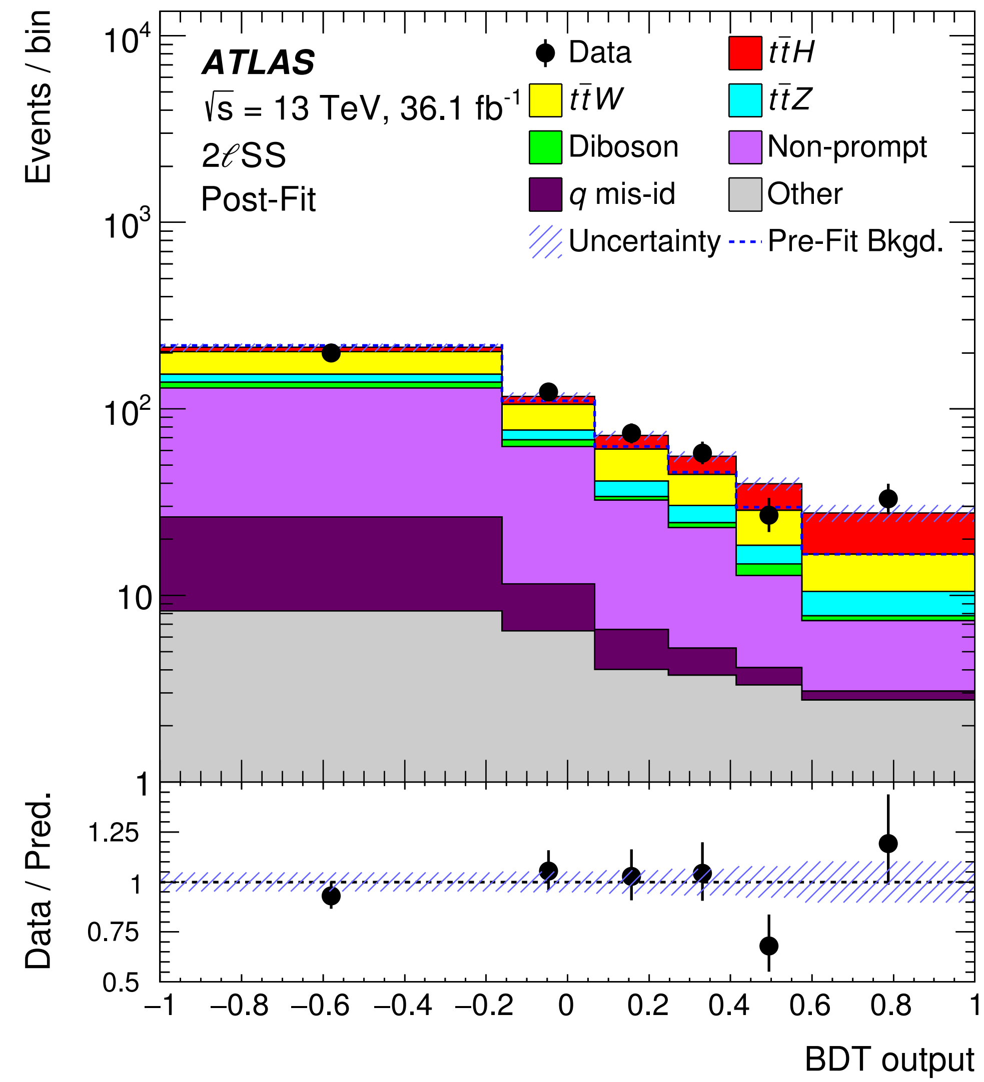
- Extensive optimisation and validation studies performed
- Modeling of the input variables
- Understanding correlation of input to the BDT output
- Study the bins size to enhance sensitivity and/or to keep the remaining backgrounds under control
- Very good agreement between data and SM prediction observed

Results (tt̄H)

[arxiv:1712.08891](https://arxiv.org/abs/1712.08891)

A **maximum-likelihood fit** is performed **simultaneously on all search regions** to extract the ttH cross section normalised to SM prediction





Excess of events over the SM prediction is found with an observed significance of **4.1 standard deviations**

→ **first evidence** of associated production of **Higgs boson** and **top quark pair**

Searching for supersymmetry with ML techniques

- Supersymmetry is a popular theory for physics beyond the SM
- Provides solutions to important open questions (hierarchy problem, dark matter, GUT, ...)
- Basic principle is a **symmetry** between **bosons** and **fermions**

Searching for supersymmetry with ML techniques

Supersymmetry is a popular theory for physics beyond the SM

Provides solutions to important open questions (hierarchy problem, dark matter, GUT, ...)

Basic principle is a **symmetry** between **bosons** and **fermions**

Example: **Searching for scalar top quarks**

Top ($\tilde{t}_1$) and **bottom squarks** are **superpartners** of **top** and **bottom quarks**

Naturalness arguments suggest a relatively **light $\tilde{t}_1$**

$\tilde{t}_1$ can be **produced** at **LHC**

Searching for supersymmetry with ML techniques

Supersymmetry is a popular theory for physics beyond the SM

Provides solutions to important open questions (hierarchy problem, dark matter, GUT, …)

Basic principle is a **symmetry** between **bosons** and **fermions**

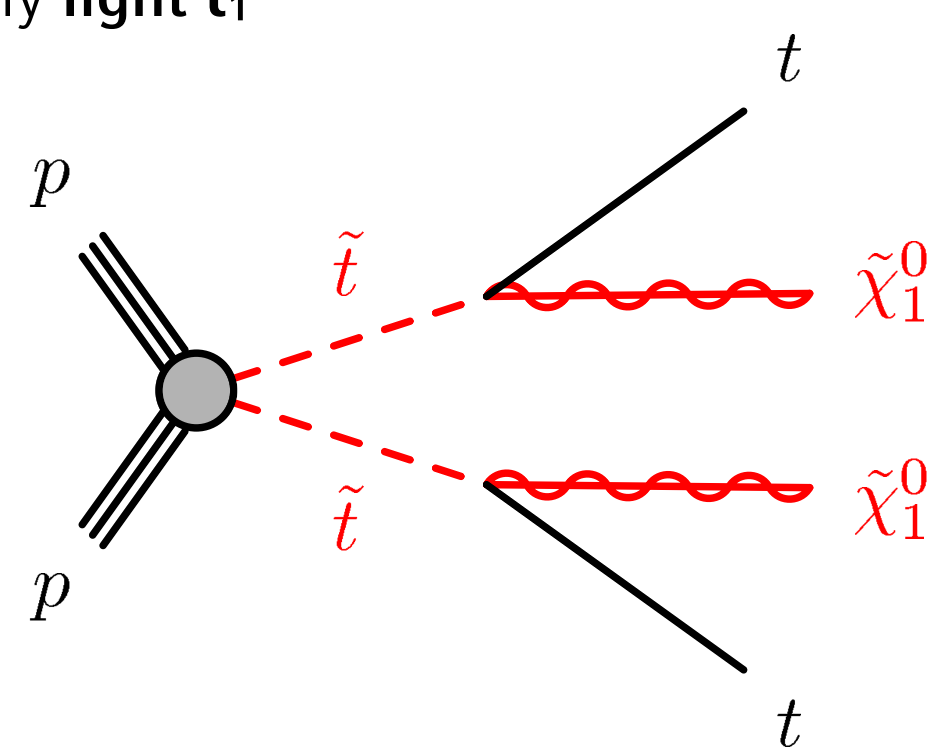
<u>Example: **Searching for scalar top quarks**</u>

Top ($\tilde{t}_1$) and **bottom squarks** are **superpartners** of **top** and **bottom quarks**

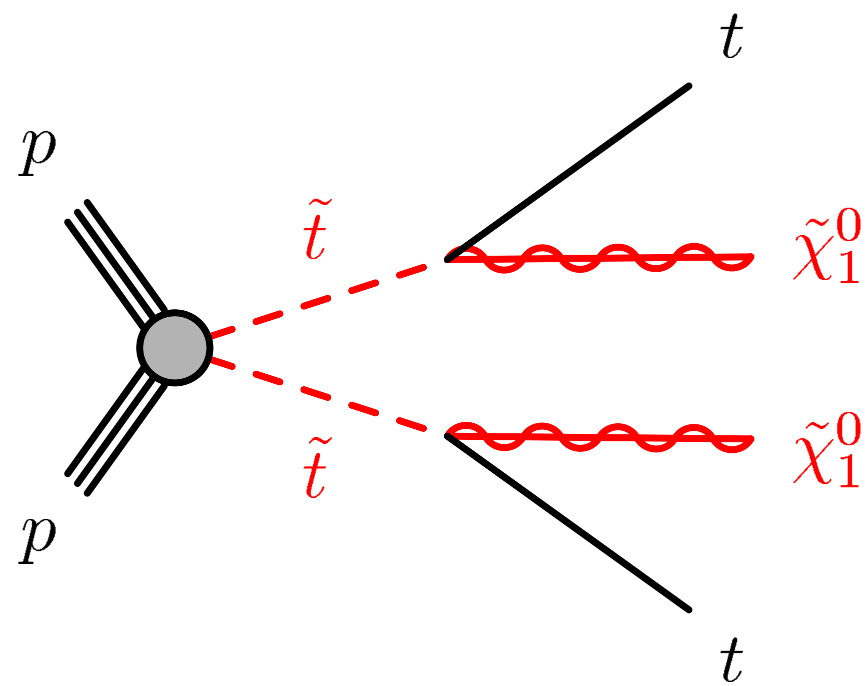
Naturalness arguments suggest a relatively **light $\tilde{t}_1$**

$\tilde{t}_1$ can be **produced** at **LHC**

Current searches derive **mass limits** in terms of simplified models



Scalar top pair production



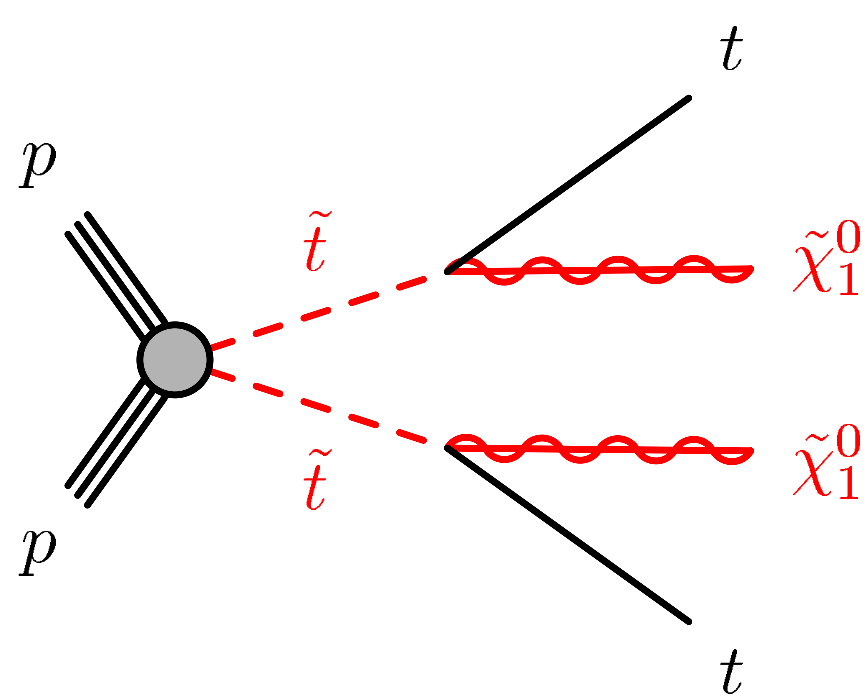
Simplified model

- Direct **stop** pair **production**
- mass splitting $\boldsymbol{\Delta}$**m** $\equiv m_{\tilde{t}1} - m_{\tilde{\chi}1}$
- Neutralinos $\tilde{\chi}_1$ produce **large** $\mathbf{E_T^{miss}}$

Difficult to **distinguish** from **$t\bar{t}$ bkg**

- **Similar final state**, except large E_T^{miss}

Scalar top pair production



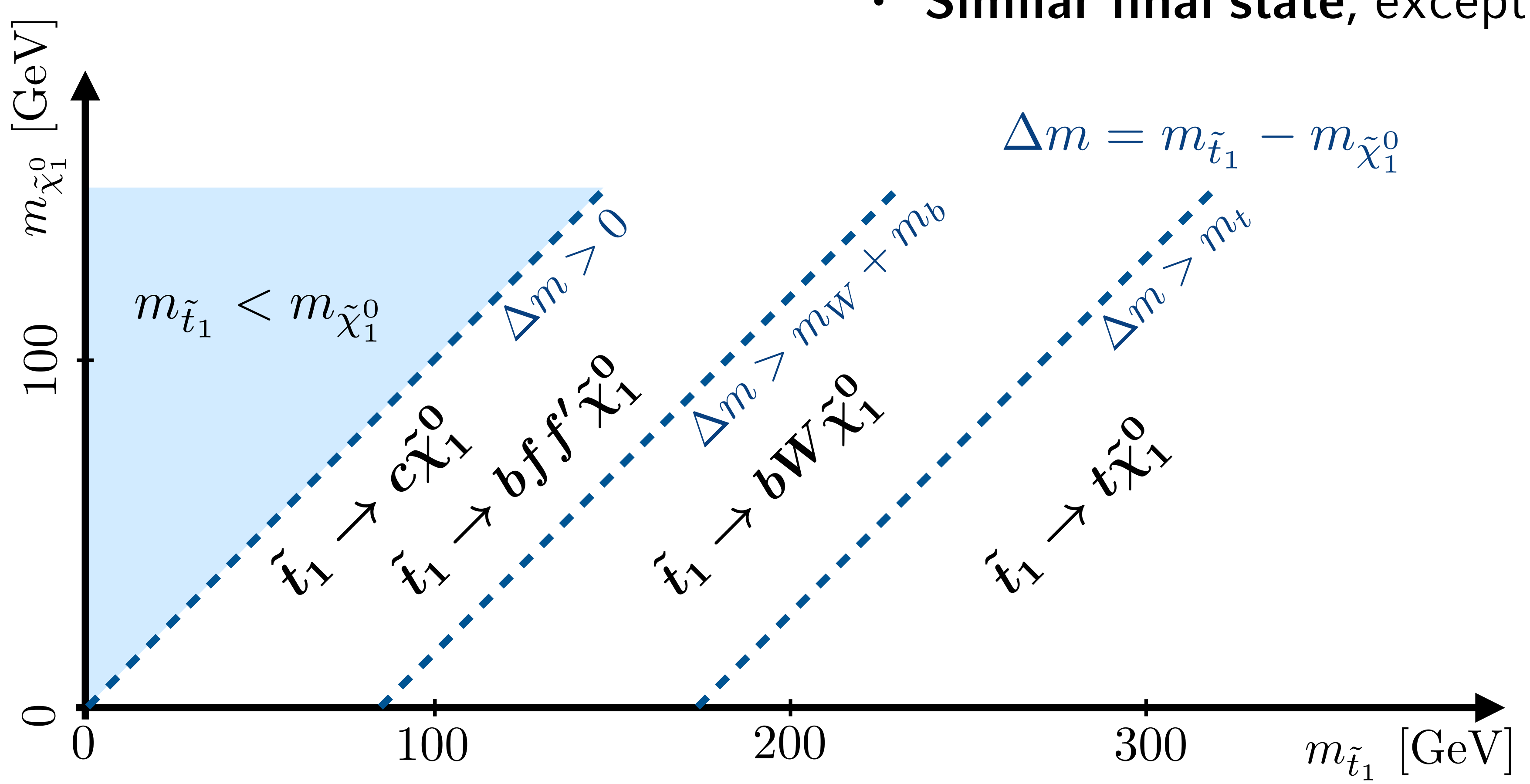
Simplified model

- Direct **stop** pair **production**
- mass splitting **Δm** ≡ $m_{\tilde{t}1}$ - $m_{\tilde{\chi}1}$
- Neutralinos $\tilde{\chi}_1$ produce **large E_T^{miss}**

Difficult to **distinguish** from **t$\bar{t}$ bkg**

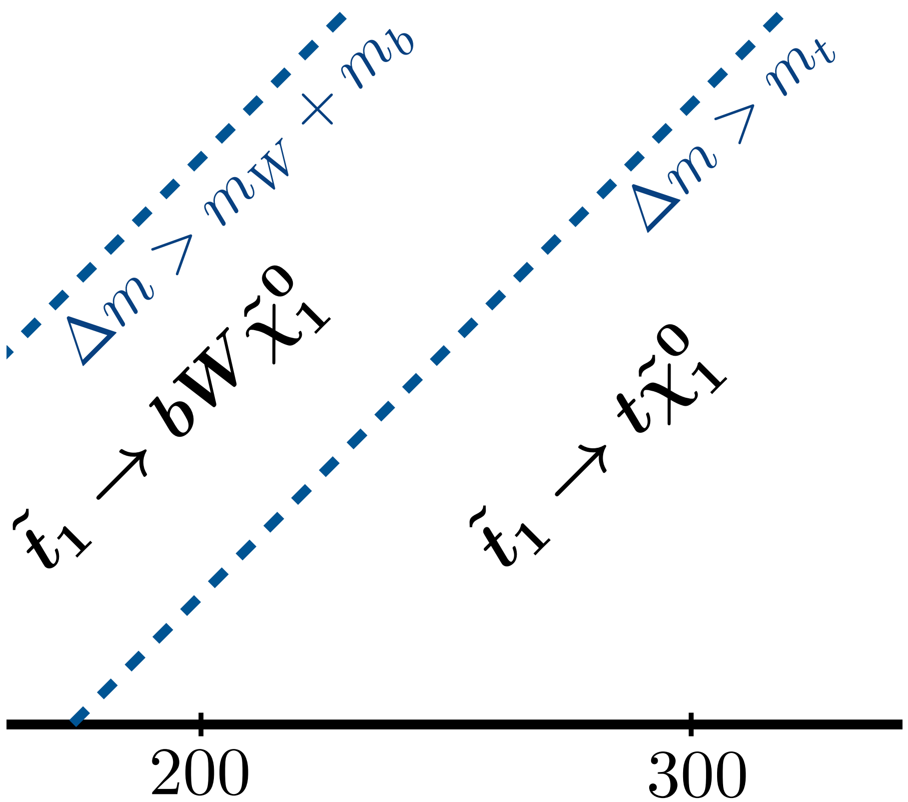
- **Similar final state**, except large E_T^{miss}

Mass plane of top squark (x) - neutralino (y)



Search for scalar top quarks in one lepton final states

arxiv:1711.11520



- Along the diagonal $\mathbf{\Delta m} \equiv m_{\tilde{t}1} - m_{\tilde{\chi}1} \sim \mathbf{m_t}$ the decay is **identical** to top quark pair production

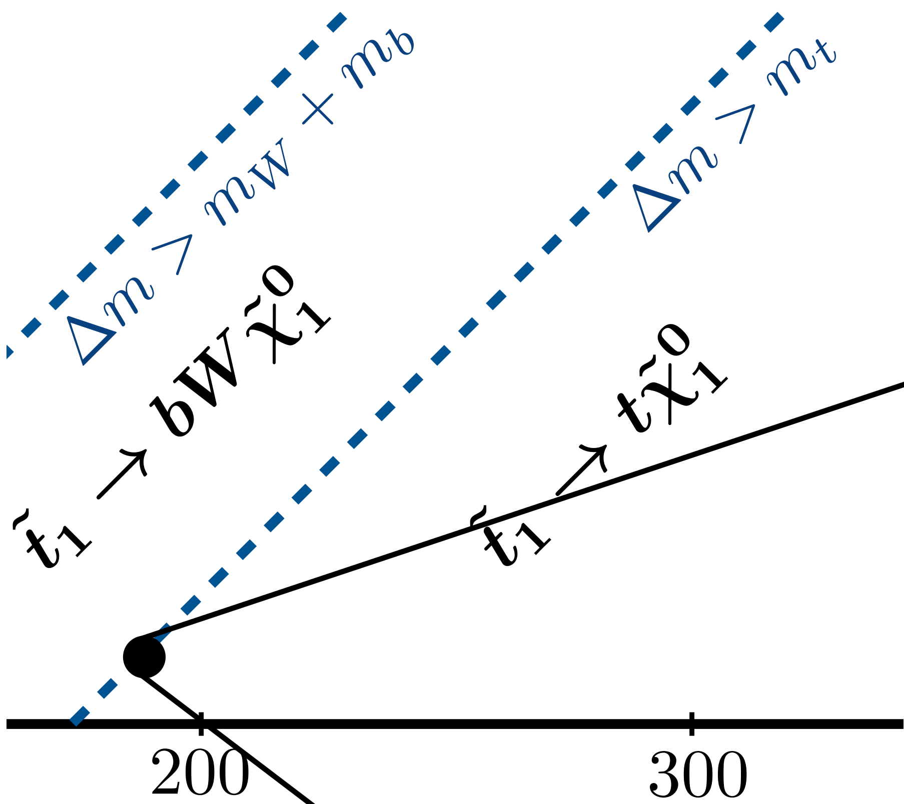
→ Analysis performs **3 independent BDTs** along the diagonal line

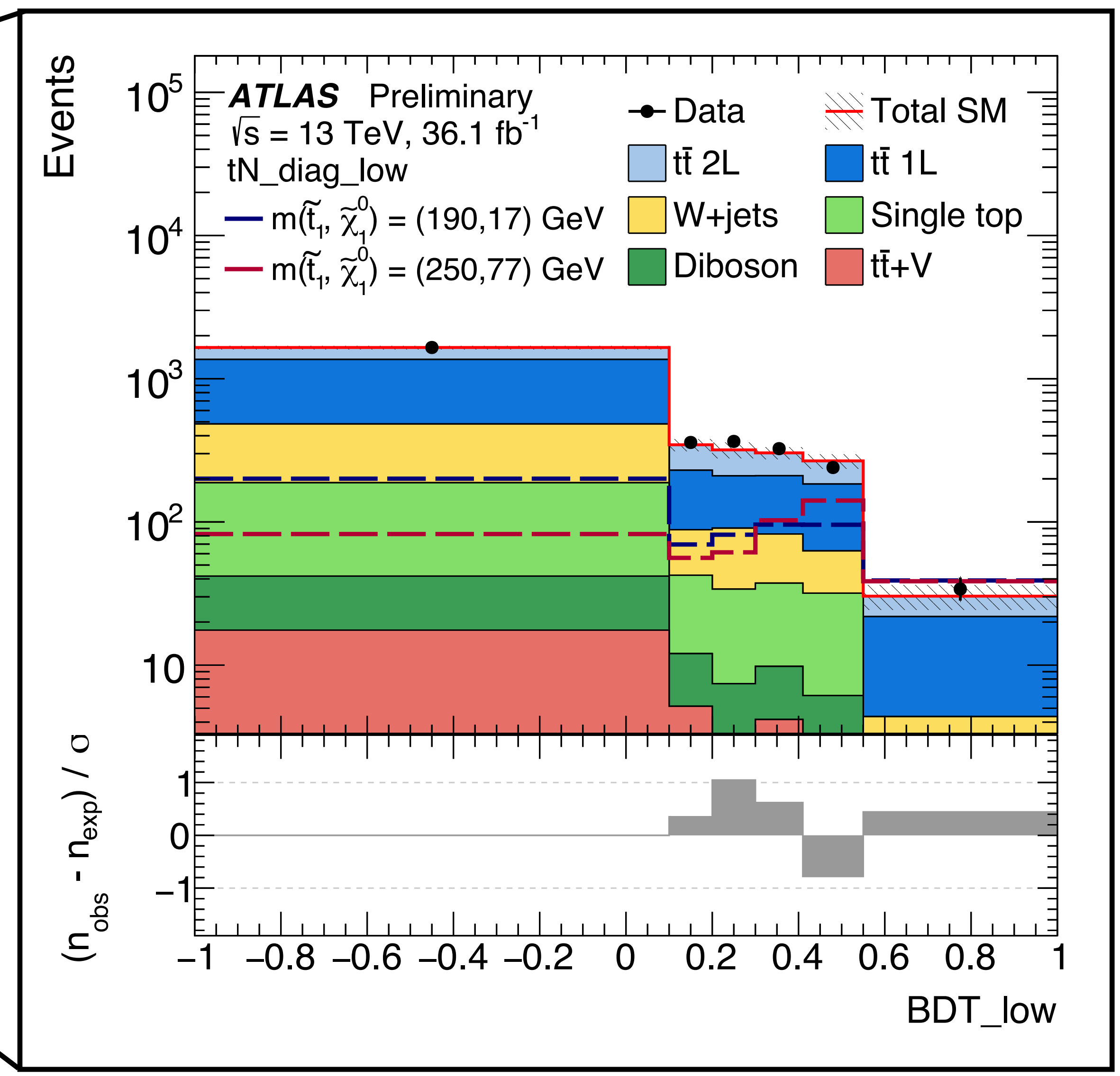
Search for scalar top quarks in one lepton final states

arxiv:1711.11520

- Along the diagonal **Δm** ≡ $m_{\tilde{t}1}$ - $m_{\tilde{\chi}1}$ ~ $\mathbf{m_t}$ the decay is **identical** to top quark pair production

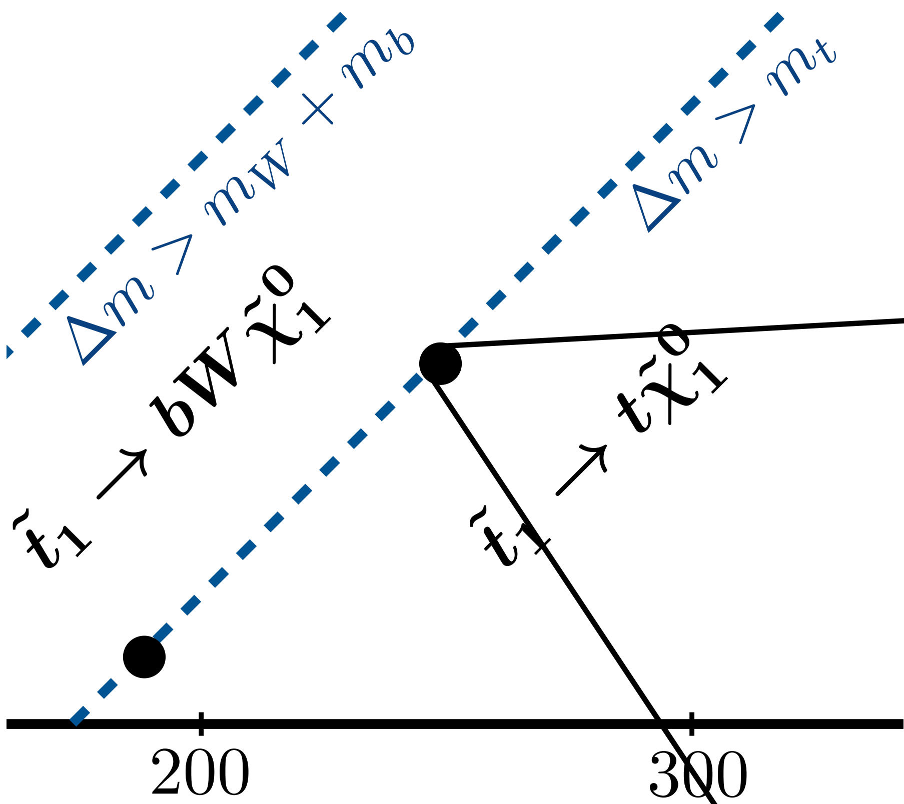
→ Analysis performs **3 independent BDTs** along the diagonal line





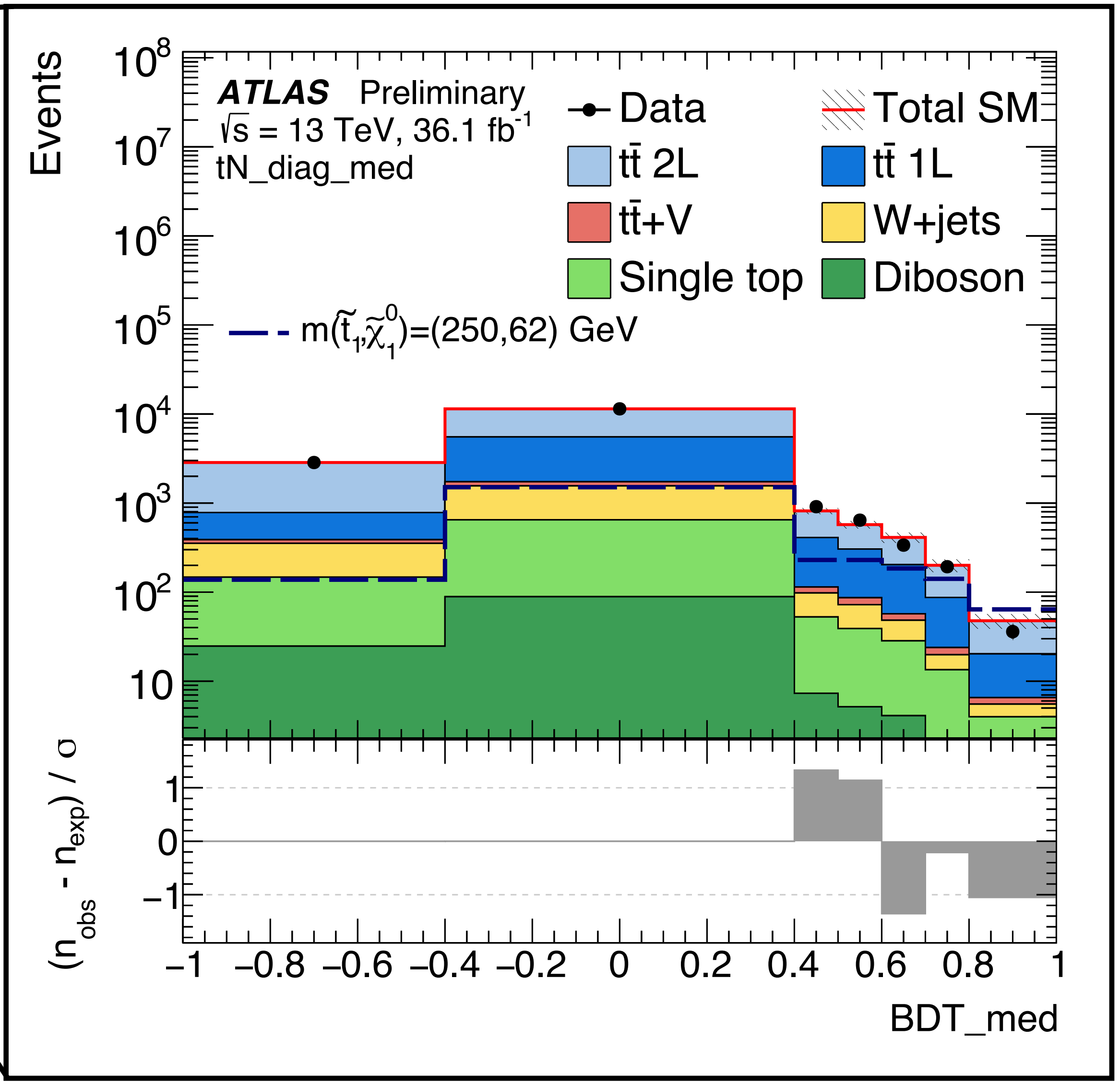
Search for scalar top quarks in one lepton final states

arxiv:1711.11520



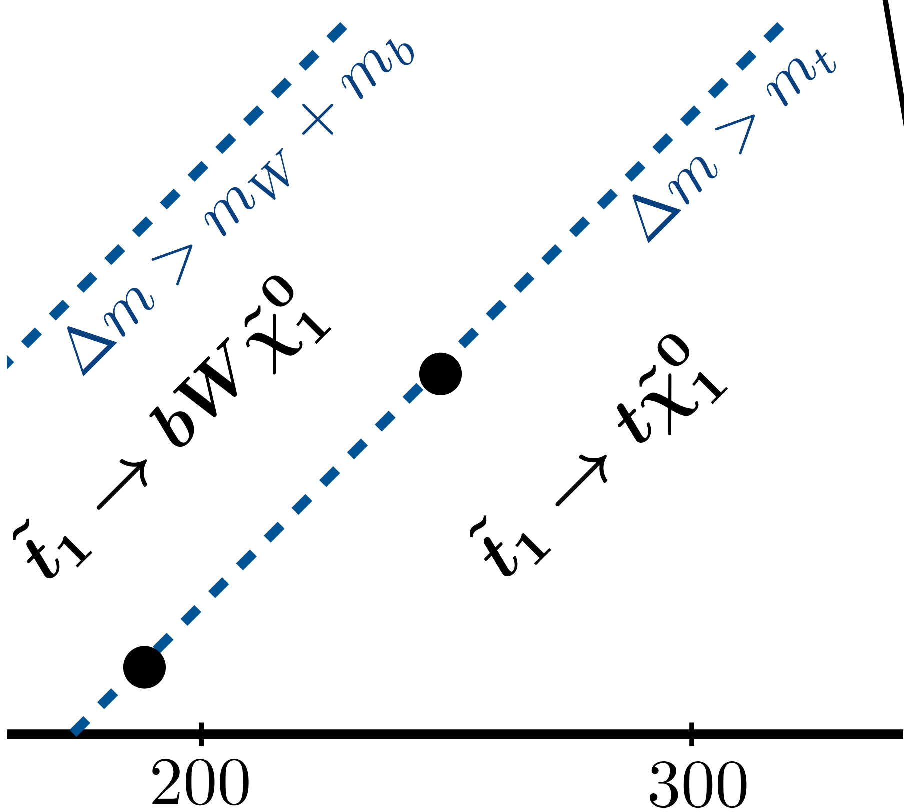
- Along the diagonal $\mathbf{\Delta m} \equiv m_{\tilde{t}1} - m_{\tilde{\chi}1} \sim \mathbf{m_t}$ the decay is **identical** to top quark pair production

→ Analysis performs **3 independent BDTs** along the diagonal line



Search for scalar top quarks in one lepton final states

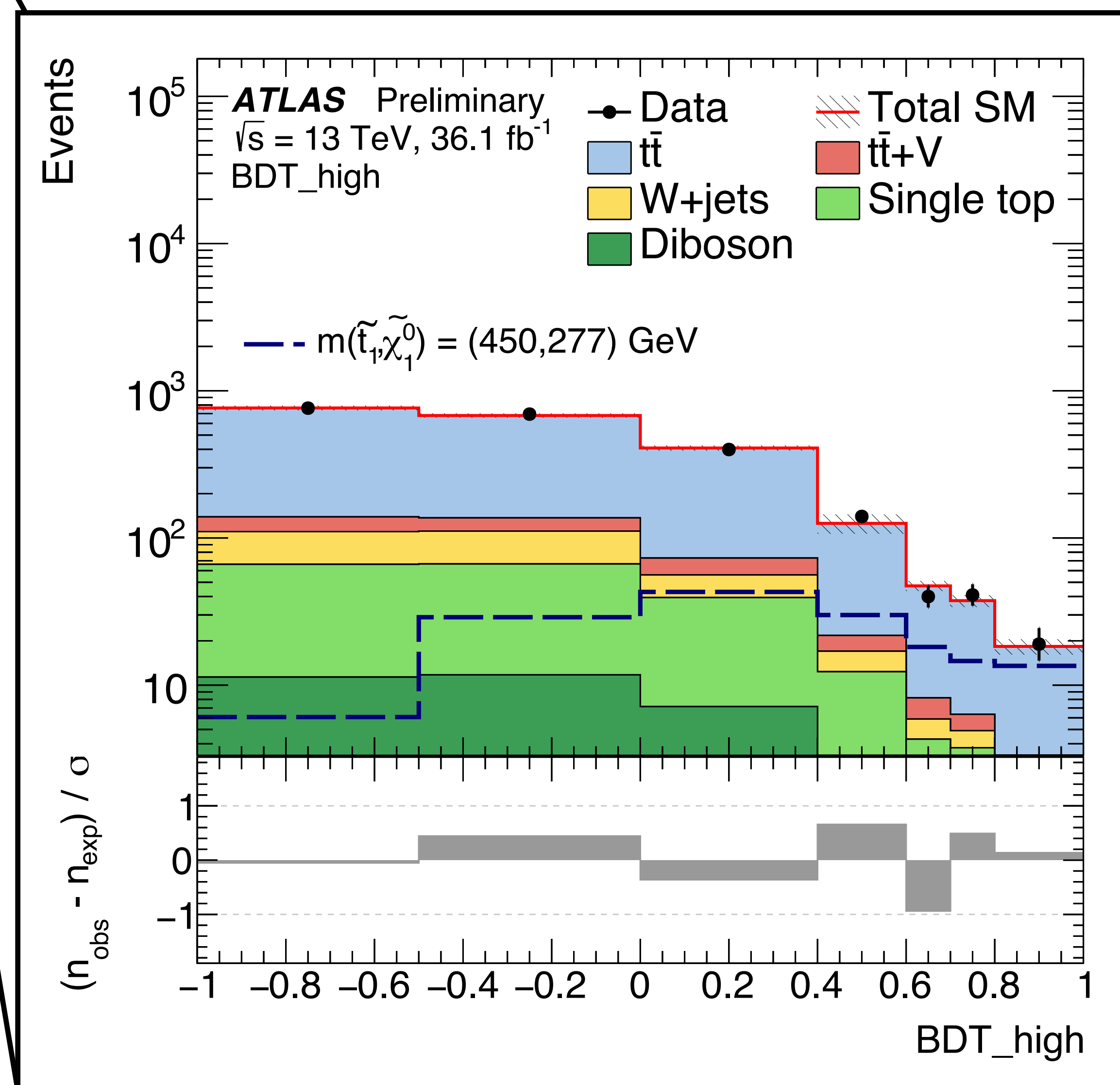
arxiv:1711.11520



- Dominant SM **background** is **estimated in data** with **low** output score
- **Signal region** defined by **large** output score
- Likelihood fit is performed in signal region

- Along the diagonal $\mathbf{\Delta m} \equiv m_{\tilde{t}1} - m_{\tilde{\chi}1} \sim \mathbf{m_t}$ the decay is **identical** to top quark pair production

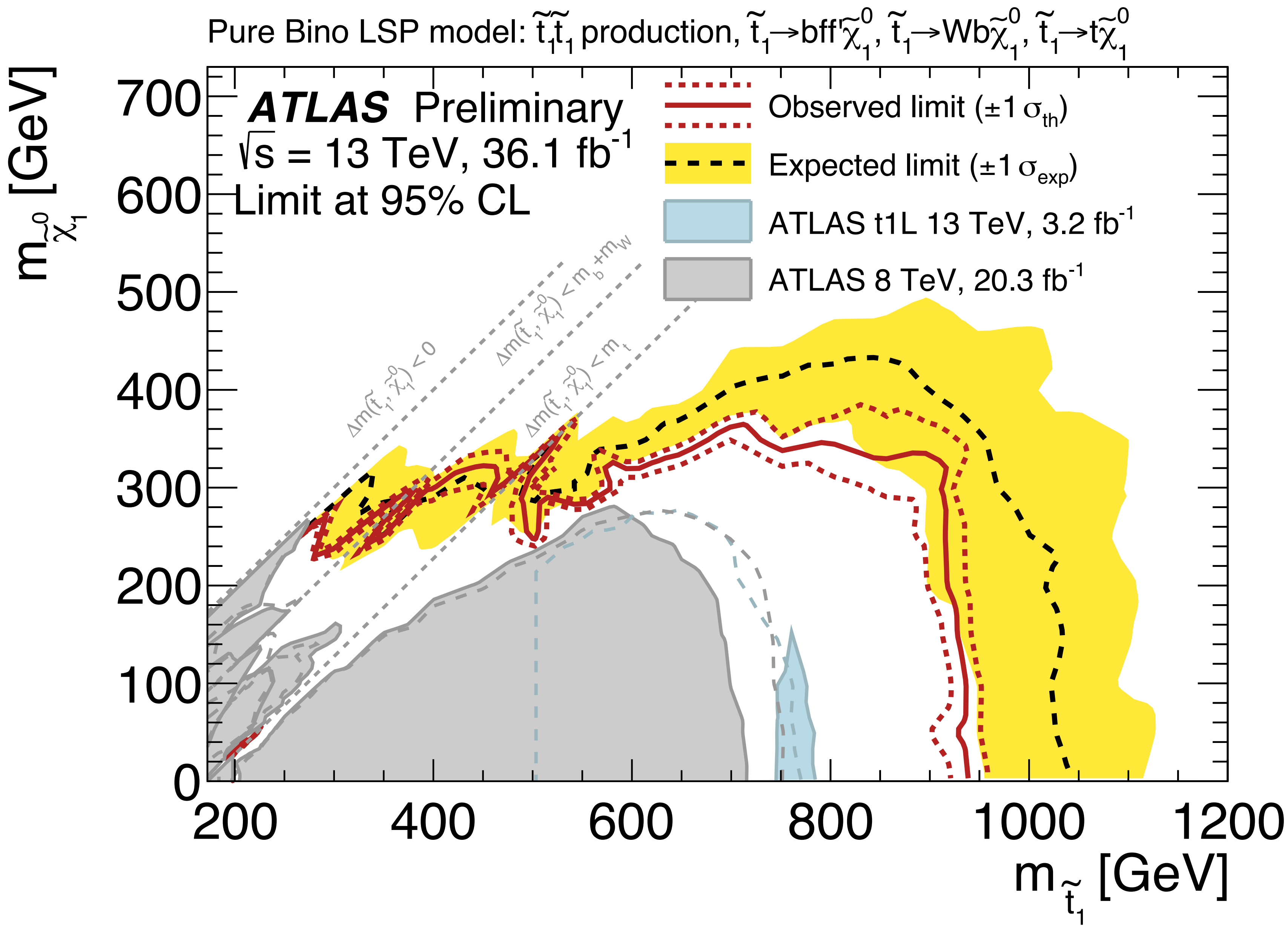
→ Analysis performs **3 independent BDTs** along the diagonal line



Searching for supersymmetry with ML techniques

arxiv:1711.11520

- After likelihood fit **no** significant **excesses** are observed compared to SM expectation
- Exclusion limits are derived for model of top squark pair production
- Large improvement of the expected limit using BDT compared to the previous analysis



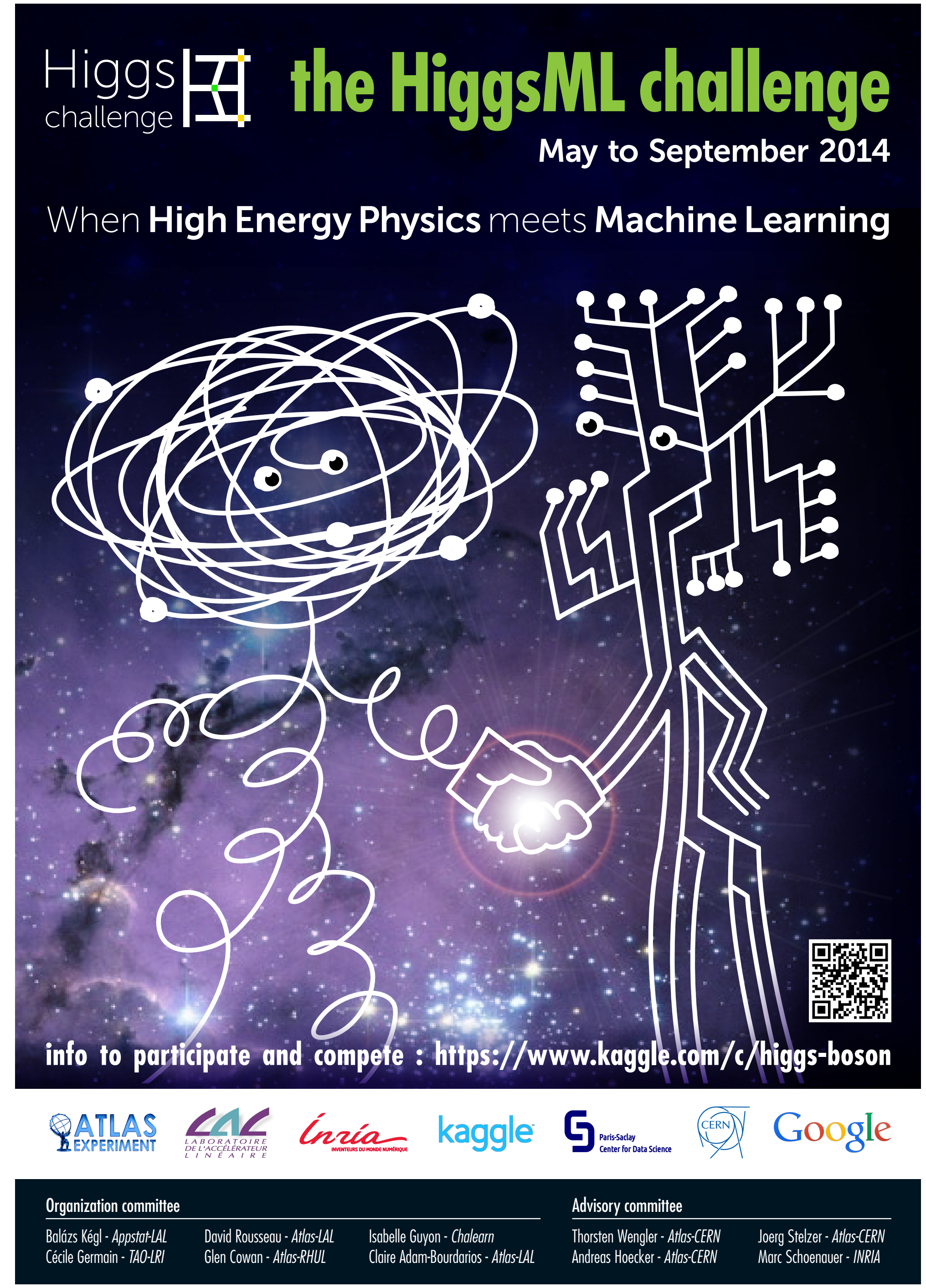
The HiggsML challenge

Public competition organised by ATLAS in 2014 (https://higgsml.lal.in2p3.fr)

Goal was to separate ATLAS simulated $H \rightarrow \tau\tau$ events from **background**

After 4 weeks almost 200 teams had beaten the in-house benchmark

In total 1785 teams or individuals participated in the competition



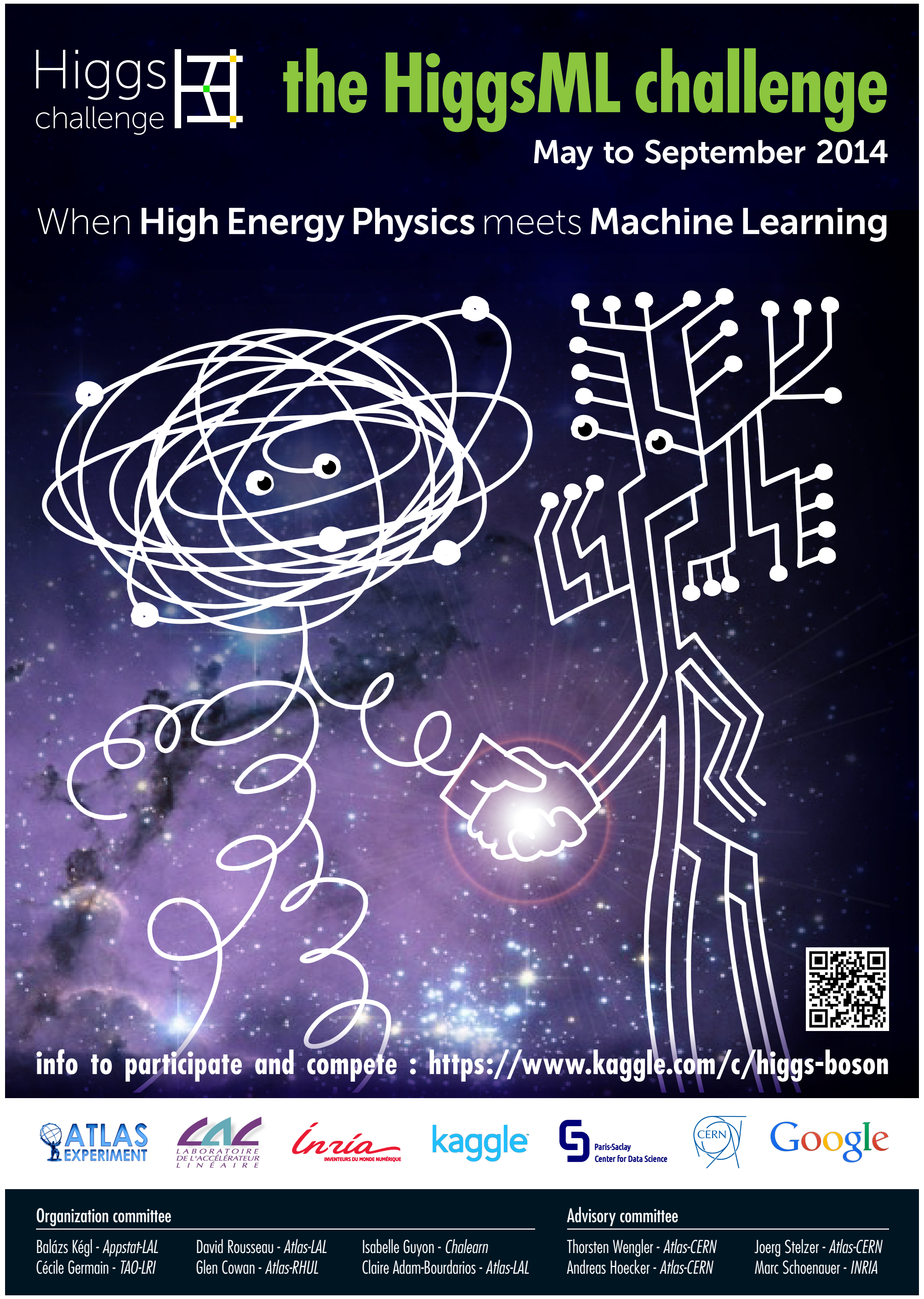
The HiggsML challenge

The **winner** performed an algorithm using the **average** of **70 DNNs** with 35 inputs, 3 hidden layers of 600 nodes each, and 2 outputs

This is a classifier with more than **70 million fitted parameters!**

Another award was given to the team that submitted a model **potentially most useful** to the collaboration

The winners' software framework is commonly known as ***XGBoost***



Summary

Massive amounts of data are processed and analysed by the ATLAS collaboration

Machine learning techniques attract more and more attention at the experiment

Several fields of applications exploit the benefits of advanced learning algorithms:

- Particle reconstruction and identification
- Separation of new signatures from standard model background

ML applications outside HEP care less about systematics — In HEP those effects are essential!

Intensive **optimisation** and **validation studies** are necessary

Understanding what the algorithm learns is vital!