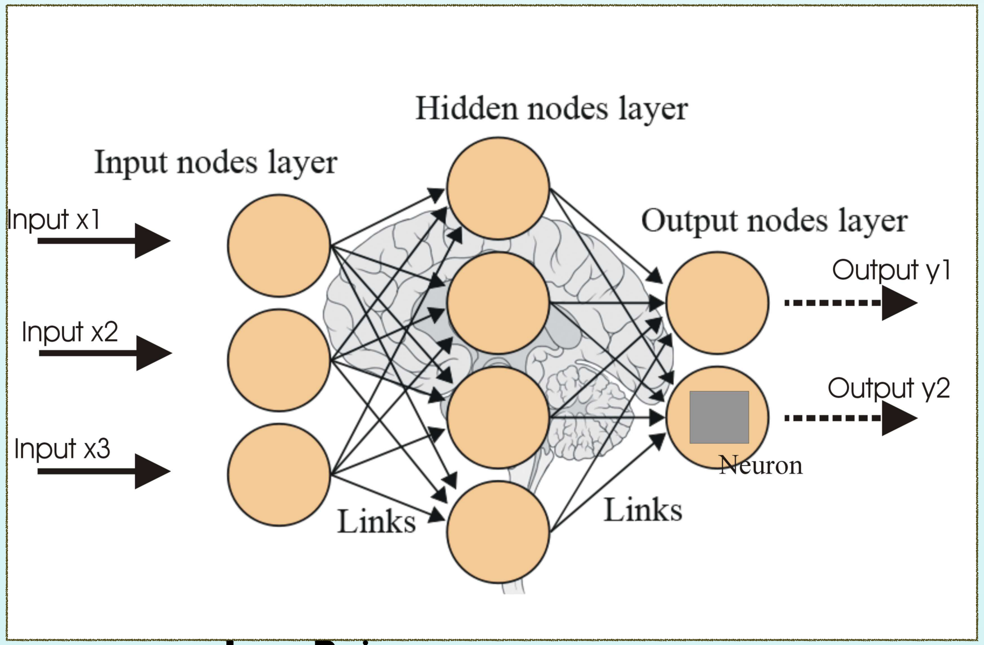
Machine Learning for High-Energy Physics

Juan Rojo
VU Amsterdam & Theory Group, Nikhef

Master's Lunch
Master in Physics and Astronomy VU+UvA
Amsterdam Science Park, 12/10/2017

Machine Learning and Artificial Neural Networks

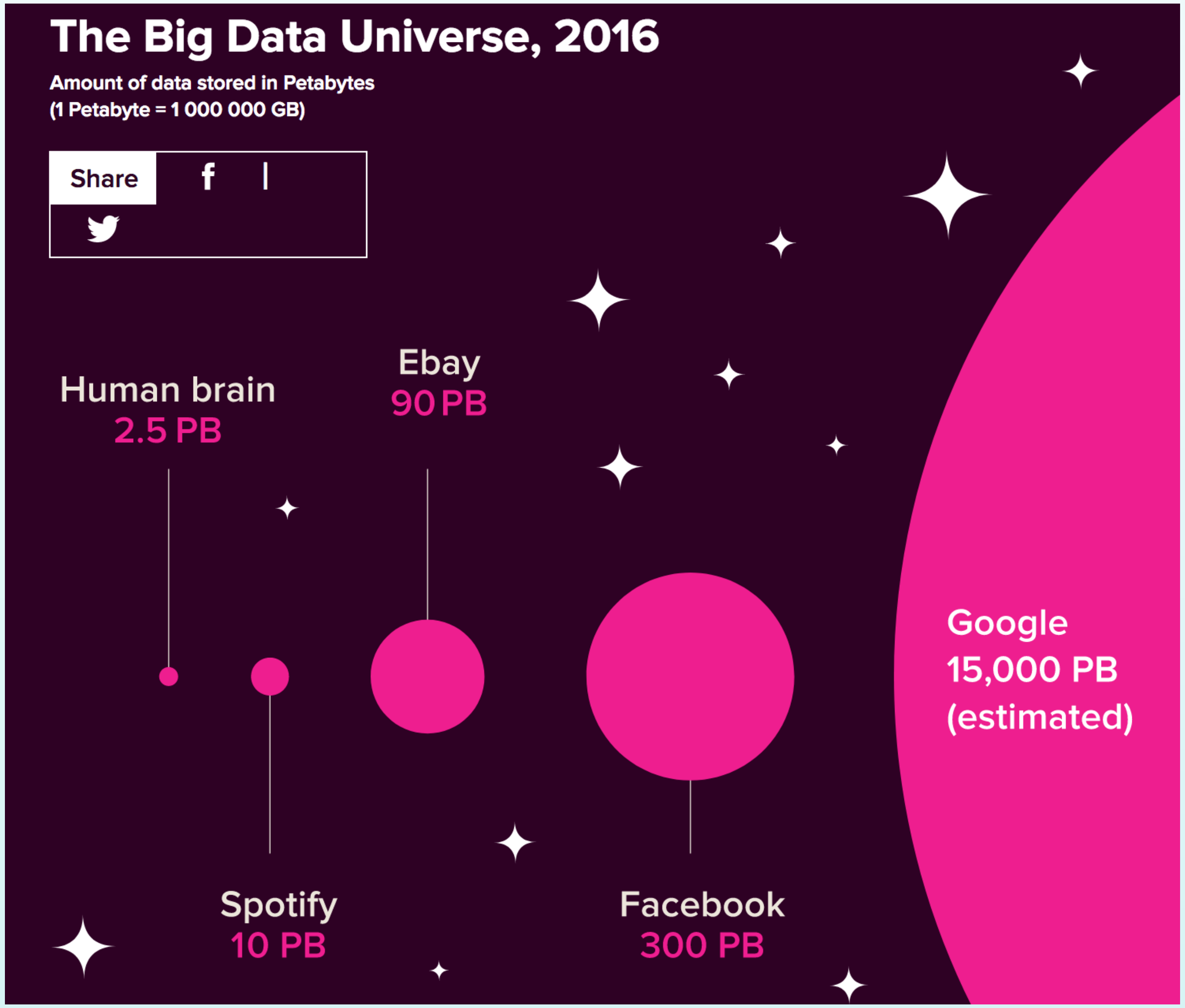


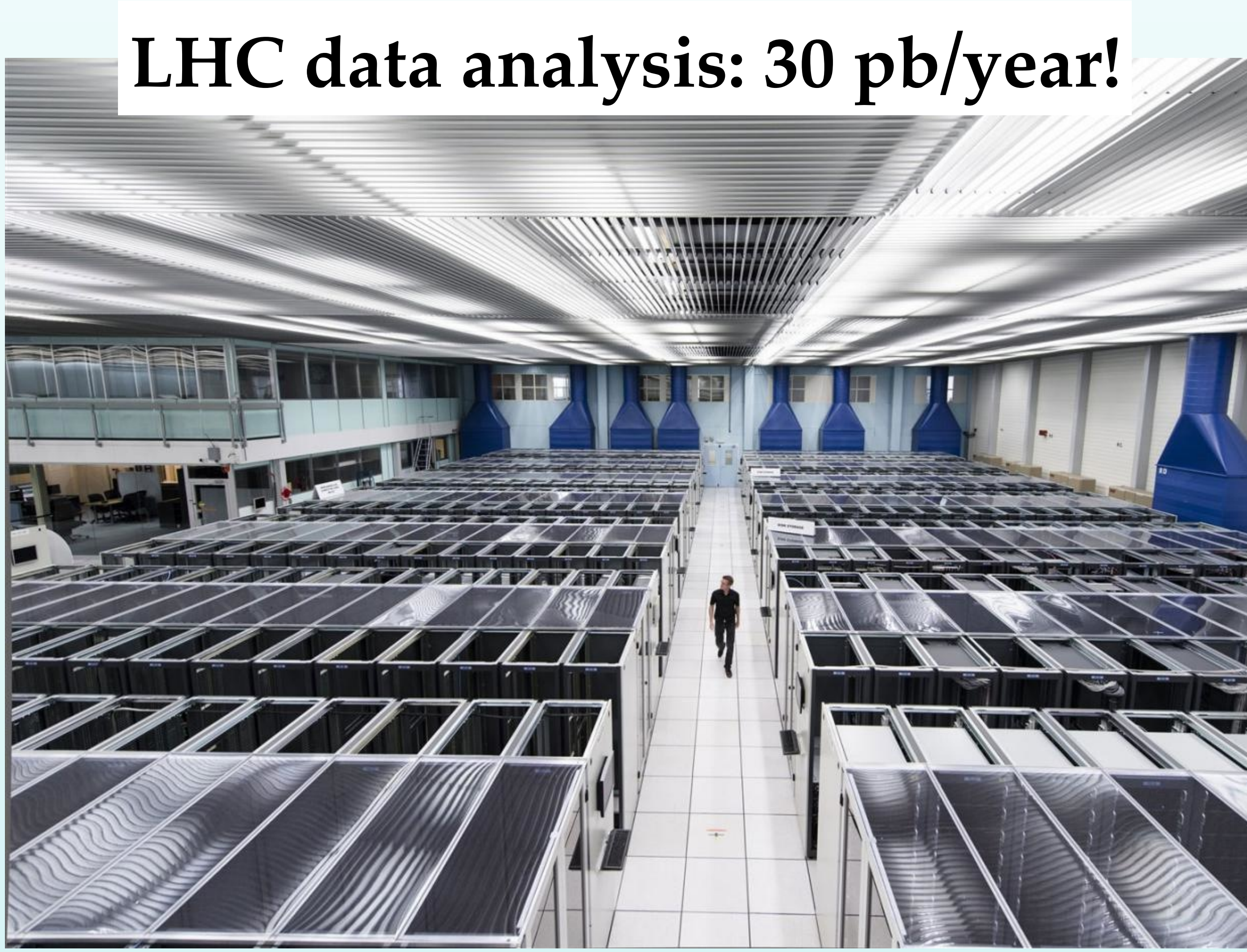
What is machine learning?

THE ROYAL SOCIETY

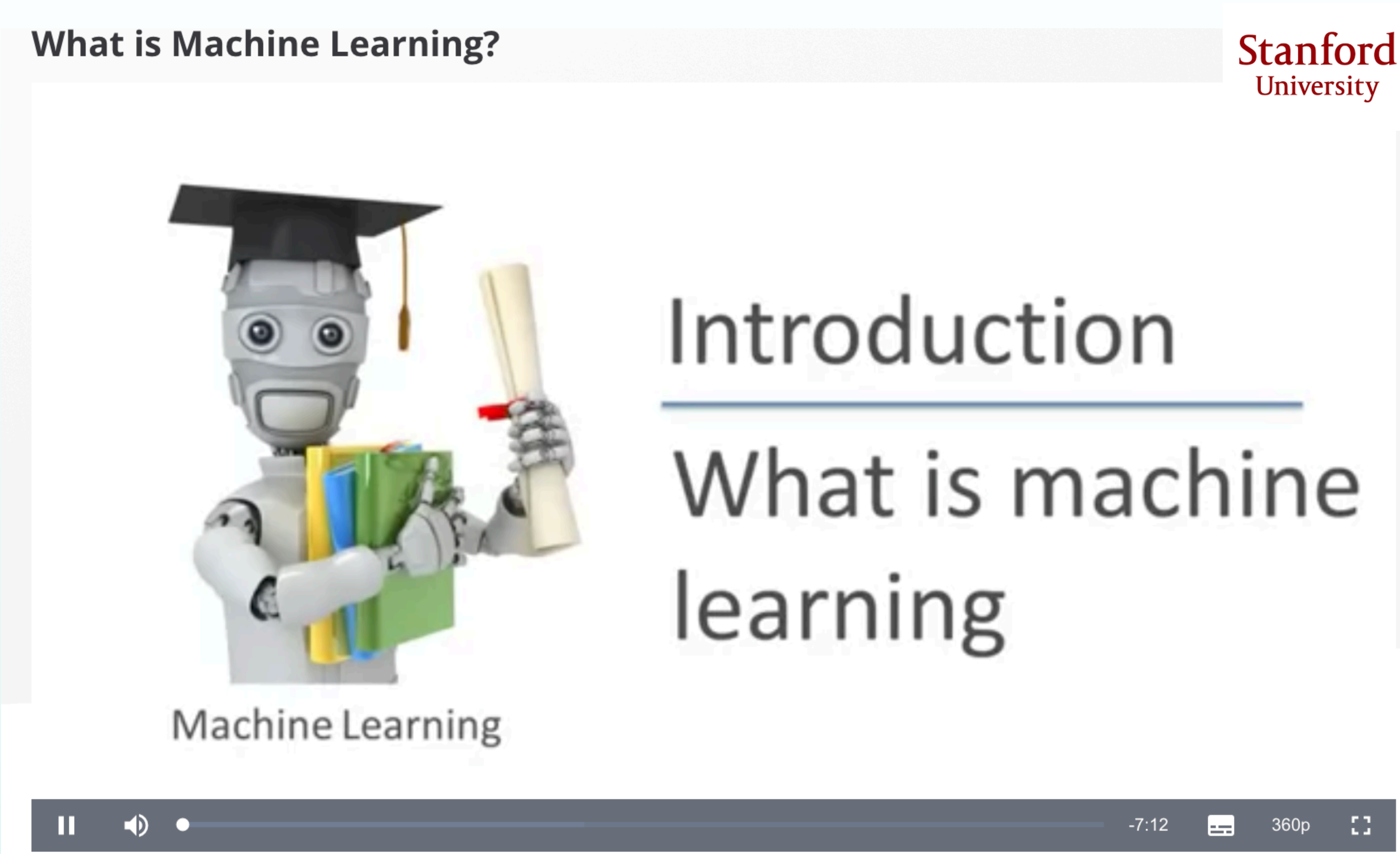
Machine Learning at the LHC

- By **Machine Learning** we denote those families of computer algorithms that **learn how to excel on a task** based on a **large sample of examples**, rather than on some a priori fixed rules
- ML algorithms are nowadays ubiquitous, from **driverless cars** to **Amazon's purchase suggestions**, to **automated medical imaging recognition** to beating the words best players at Go and chess
- ML tools rely on the **efficient exploitation of immense datasets**. And the **LHC** has a lot of data!



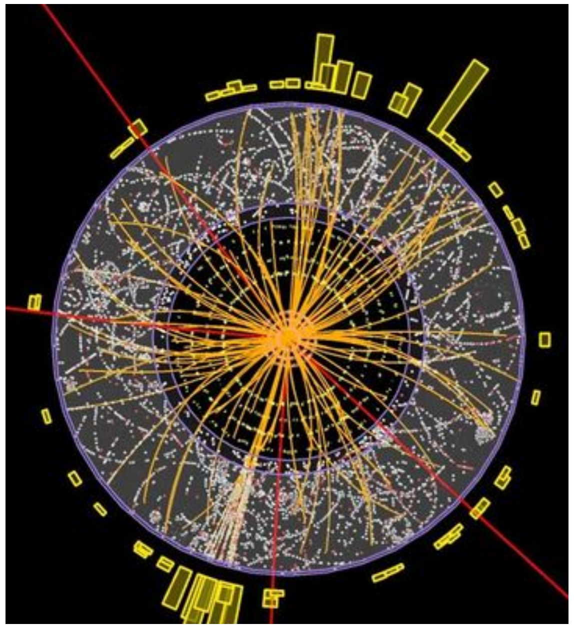


What is machine learning?

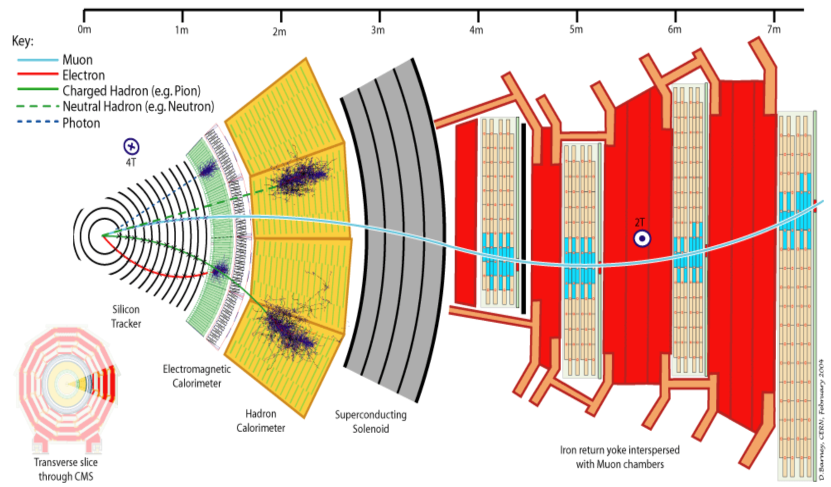


https://www.coursera.org/learn/machine-learning/lecture/Ujm7v/what-is-machine-learning

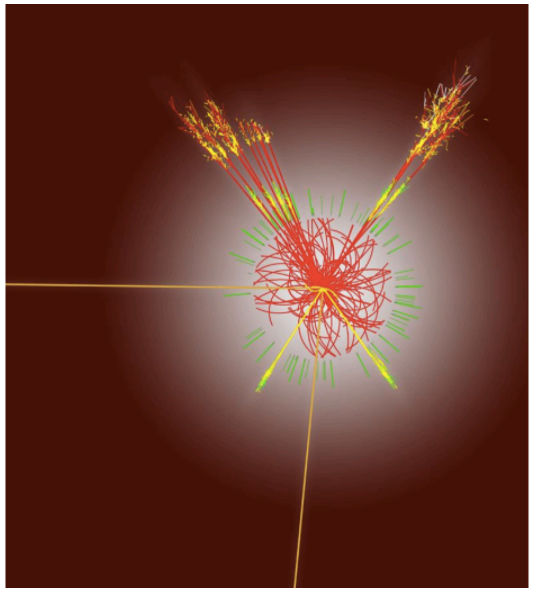
Machine Learning tools are everywhere!



Deep Kalman RNNs

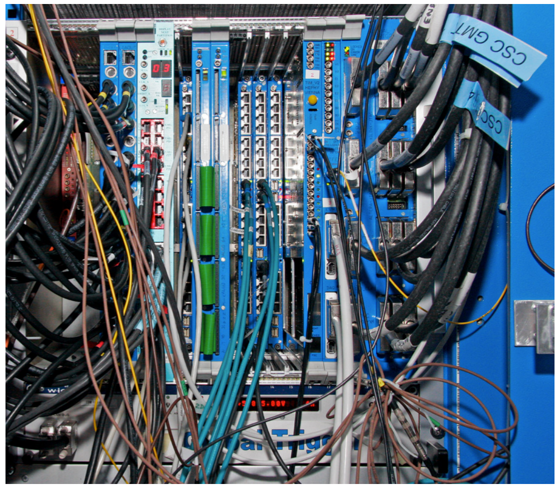


Generative Models, Adversarial Networks

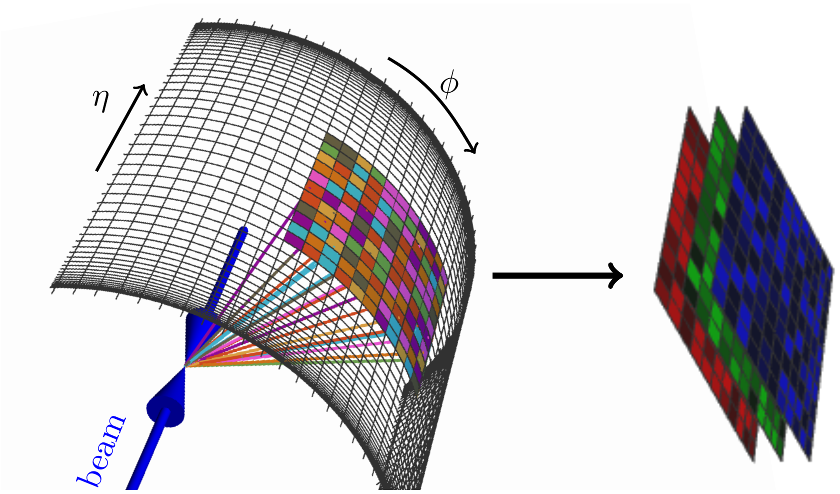


FCN, Recurrent, LSTM NN

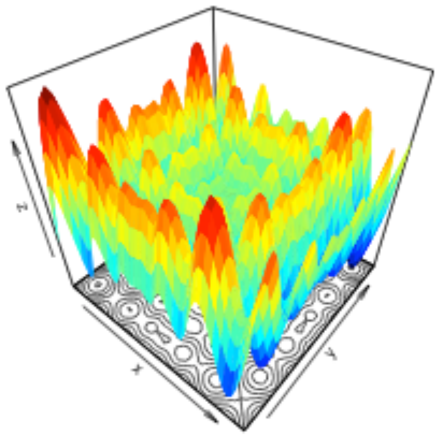
S. Glayzer



Deep ML +FPGA



Convolutional DNN



Multiobjective Regression

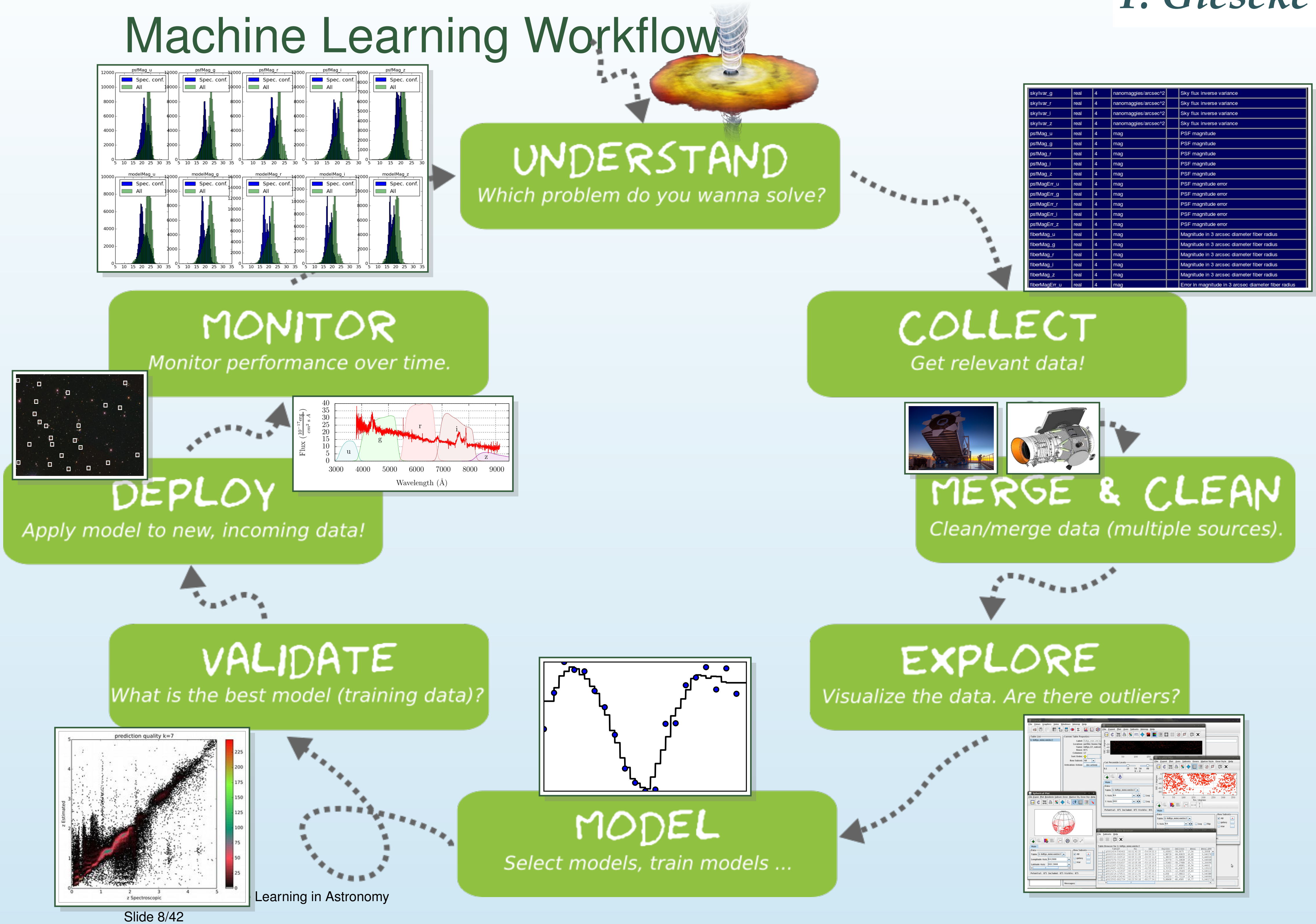
)6/19/2017

15

For many crucial applications, ML tools not just one option, but **the only option**

ML cheat sheet

F. Gieseke

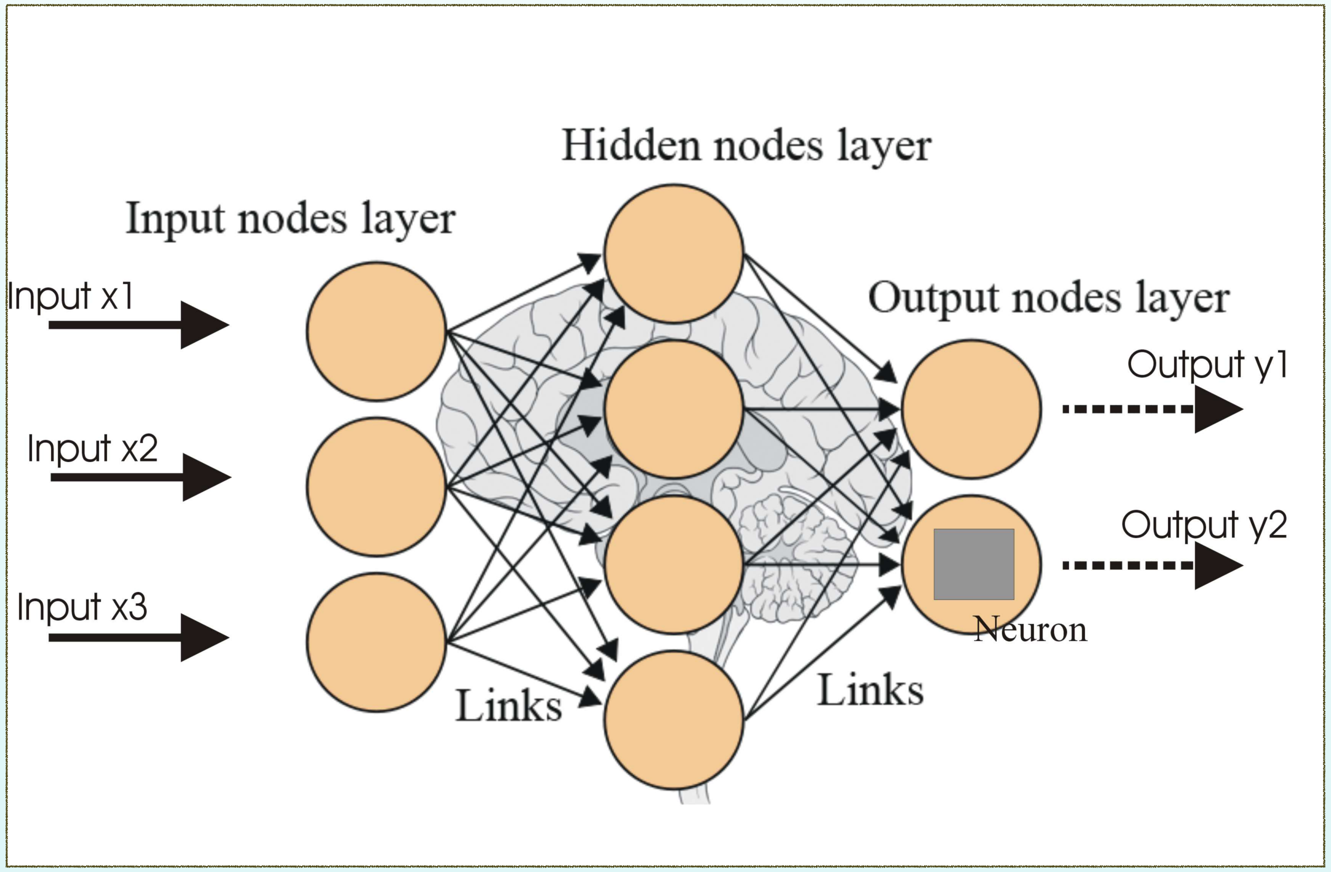


Endless possibilities - but also many non-trivial hurdles to overcome

Artificial Neural Networks

Inspired by **biological brain models**, **Artificial Neural Networks** are **mathematical algorithms** widely used in a wide range of applications, from **HEP** to **targeted marketing** and **finance forecasting**

From Biological to Artificial Neural Networks



Artificial neural networks aim to excel where domains as their **evolution-driven counterparts outperforms traditional algorithms in tasks such as pattern recognition, forecasting, classification**, …

Neural Networks Demystified [Part 1: Data and Architecture]

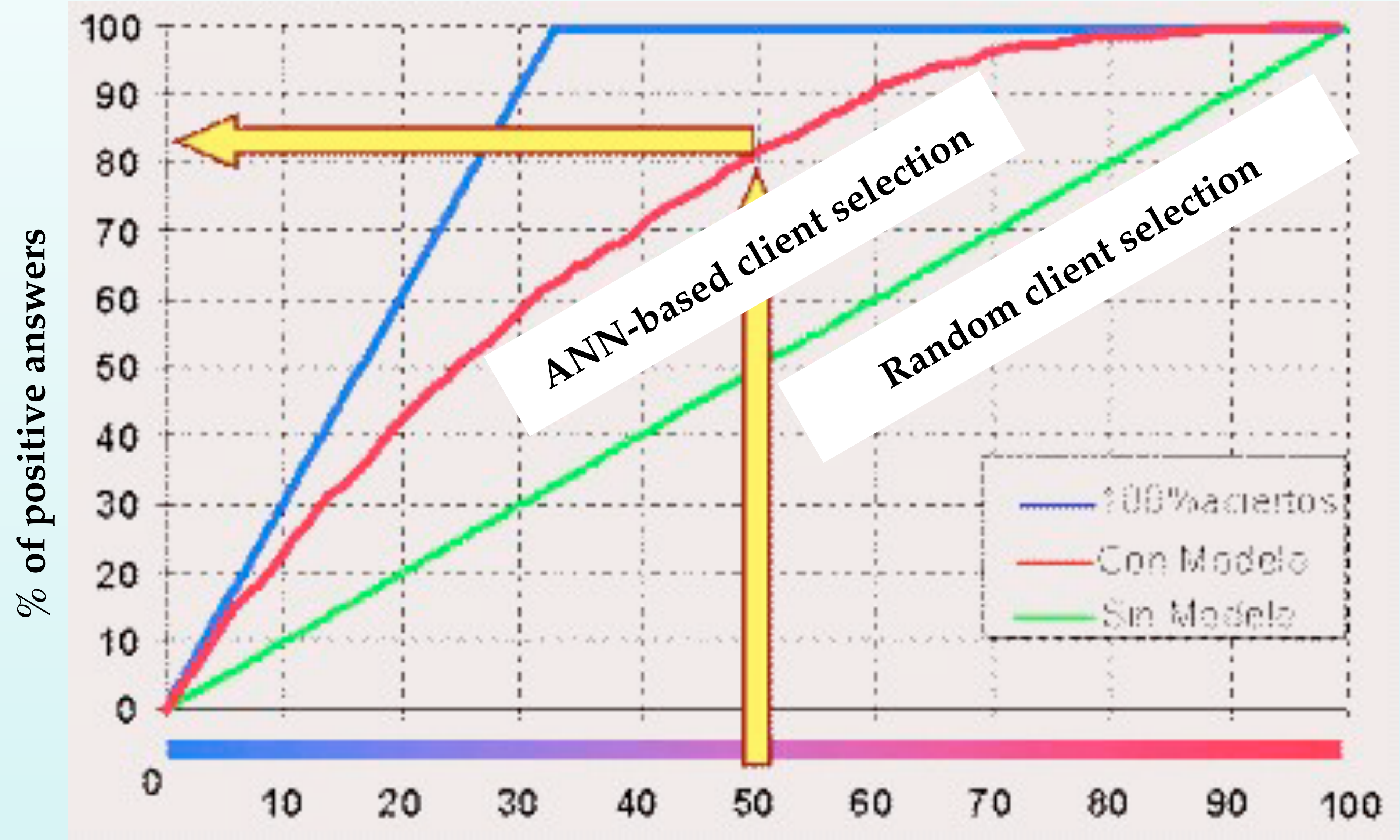
https://www.youtube.com/watch?v=bxe2T-V8XRs

ANNs - a marketing example

A bank wants to offer a new credit card to their clients. Two possible strategies:

- **Contact all customers**: slow and costly
- Contact **5%** of the customers, **train a ANN with their input** (gender, income, loans) and **their output** (yes/no) and use the information to c**ontact only clients likely to accept the product**

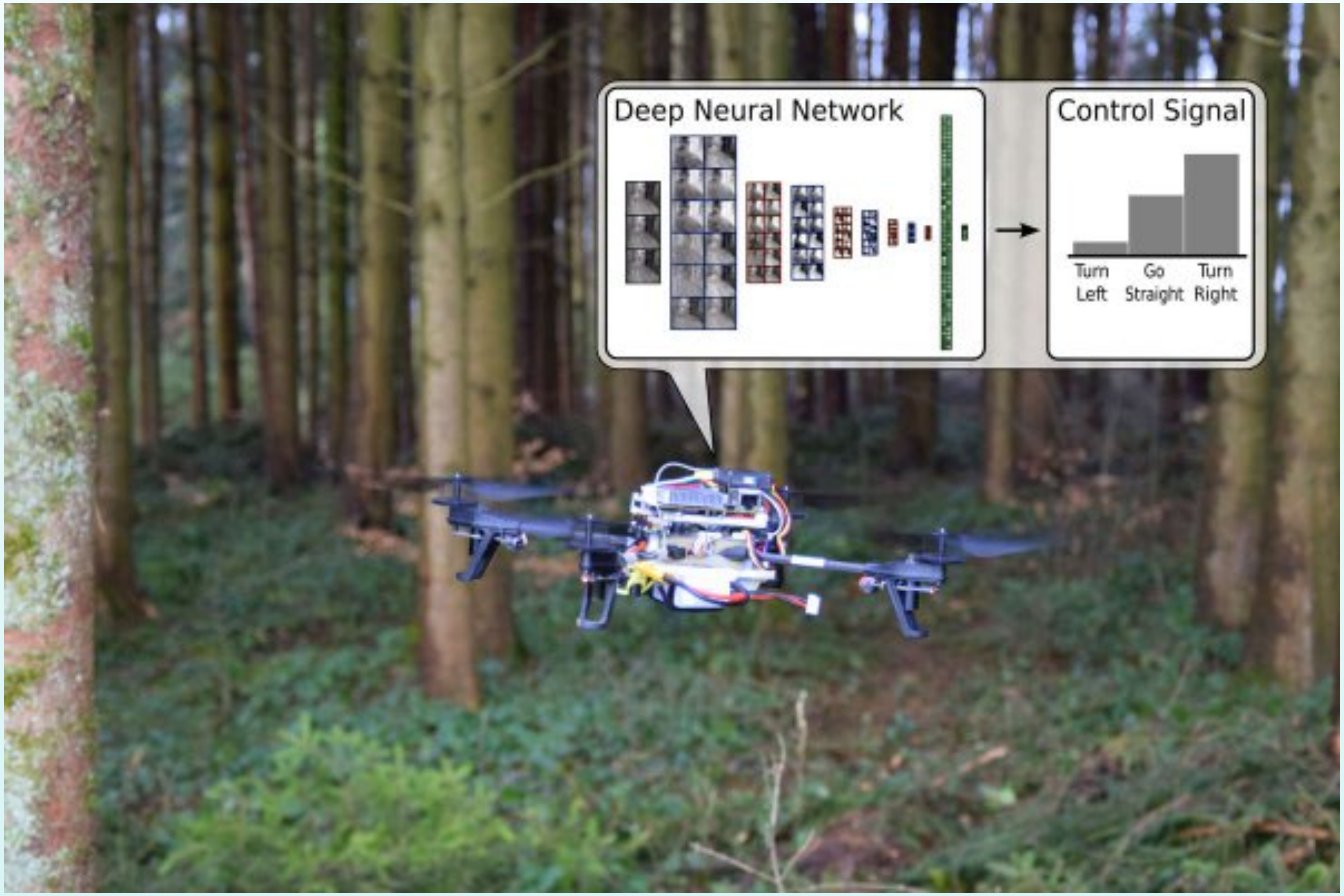
Cost-effective method to improve marketing performance!



% of customers contacted

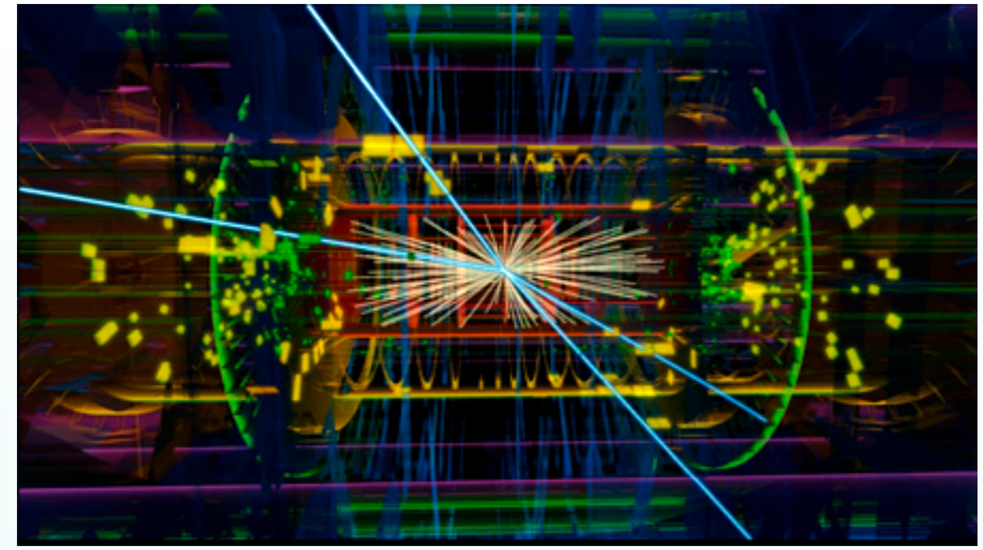
ANNs and pattern recognition

- ANNs can enable an **autonomous vision-control drone** to recognize and follow forest trails
- Image classifier operates directly on **pixel-level image intensities**
- If a trail is visible, the **software steers the drone** in the corresponding direction

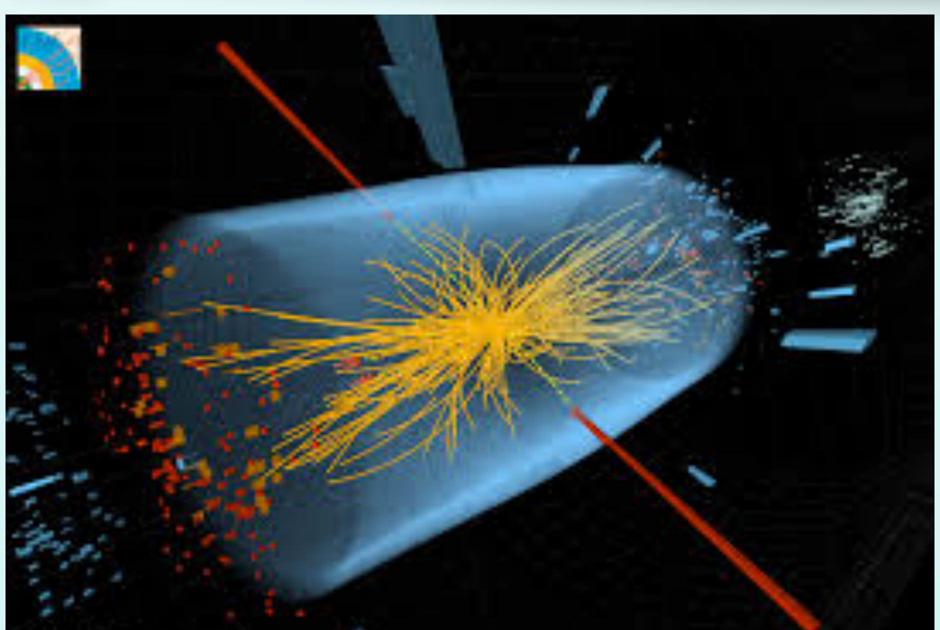


Giusti et al, IEEE Robotics and Automation Letters, 2016

Similar algorithms at work in self-driving cars!



Exploring the high-energy frontier at the Large Hadron Collider



Outstanding questions in Particle Physics

The Higgs boson

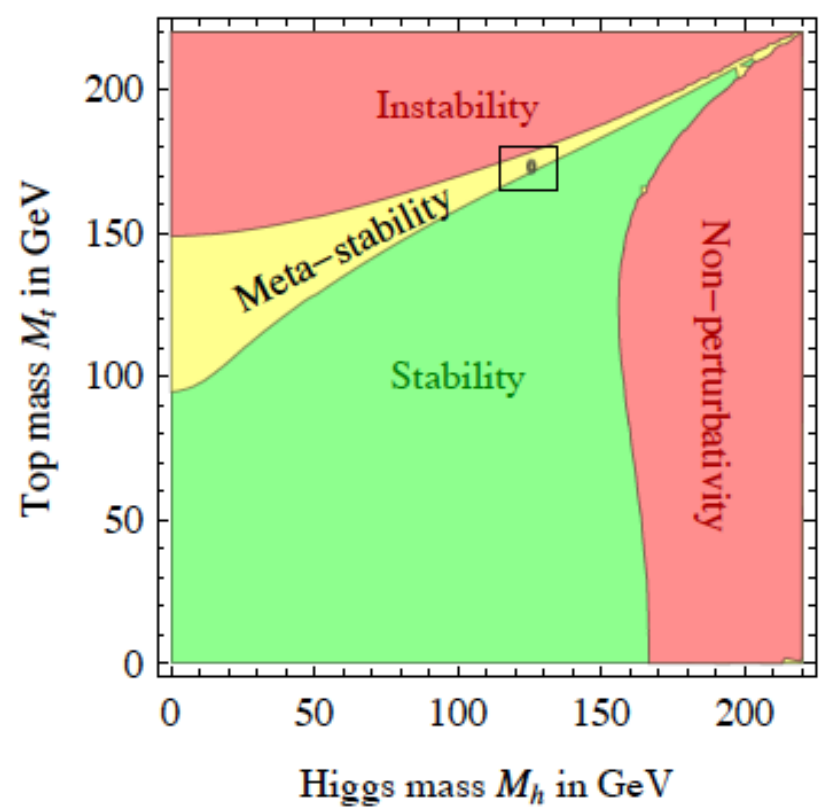
- ☑ Huge gap, **10^{17}**, between **Higgs and Plank scales**
- ☑ **Elementary or composite**? Additional Higgs bosons?
- ☑ Coupling to **Dark Matter**? Role in cosmological phase transitions?
- ☑ Is the **vacuum state of the Universe** stable?

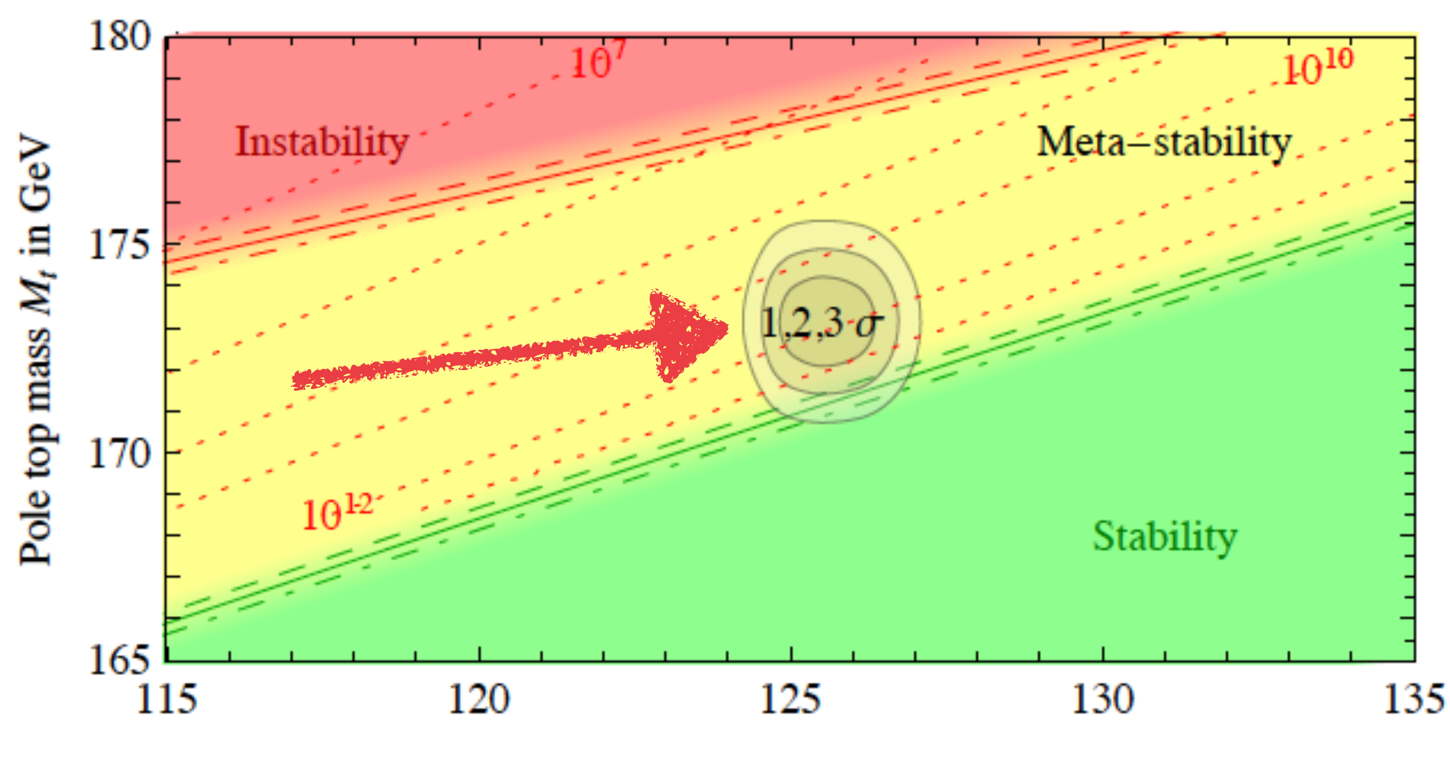
10^{17} GeV (Planck scale)

With radiative corrections, the *natural* value of the Higgs mass is Planck scale

125 GeV (Higgs mass)

1 GeV (Proton mass)





Degrassi et al 12

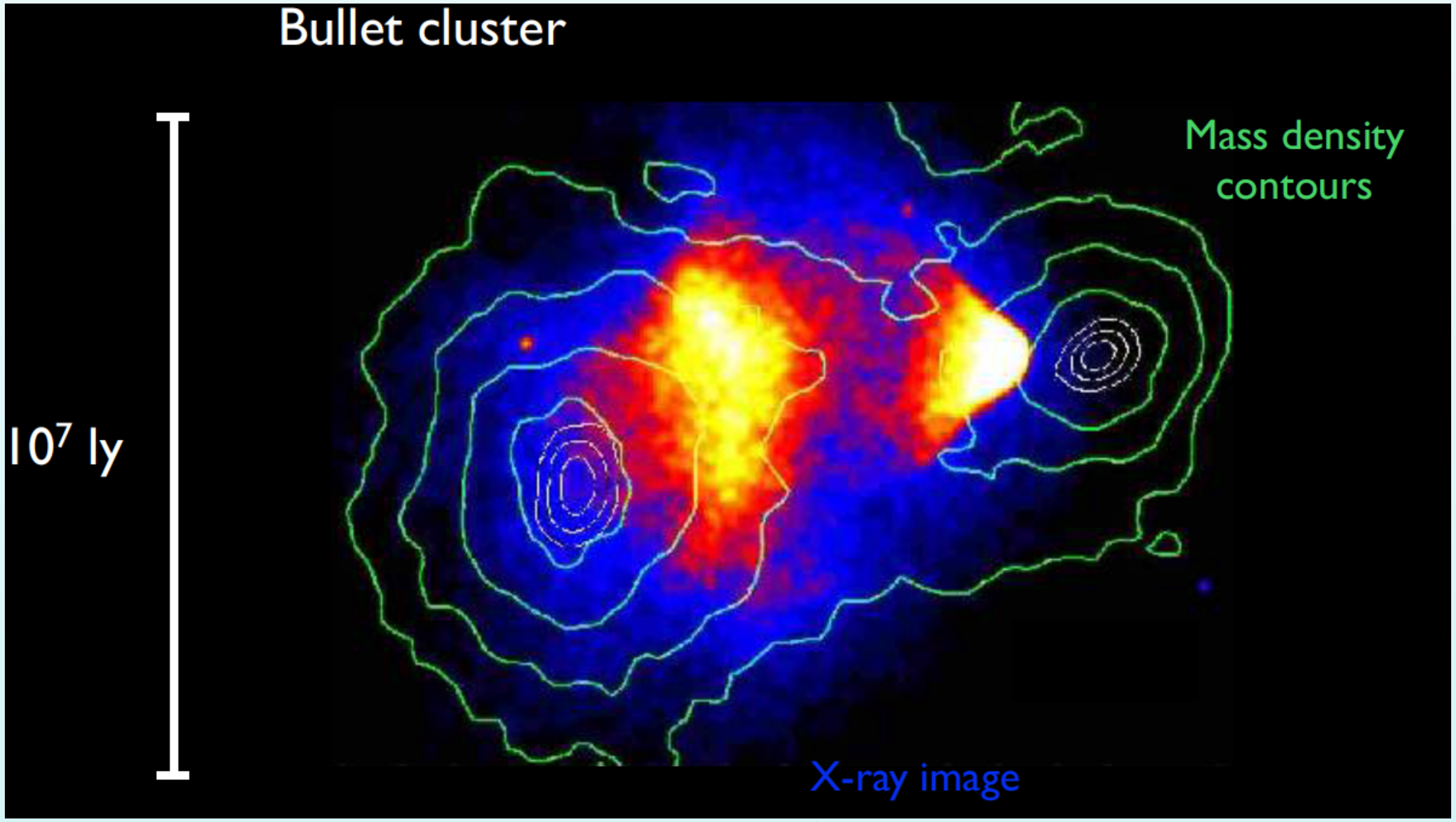
Outstanding questions in Particle Physics

The Higgs boson

- Huge gap, **10^{17}**, between **Higgs and Plank scales**
- **Elementary or composite**? Additional Higgs bosons?
- Coupling to **Dark Matter**? Role in cosmological phase transitions?
- Is the **vacuum state of the Universe** stable?

Dark Matter

- **Weakly interacting massive particles**? Sterile neutrinos? **Extremely light particles** (axions)?
- **Interactions** with Standard Model particles?
- What is the **structure of the Dark Sector**? Is Dark Matter self-interacting?



Outstanding questions in Particle Physics

The Higgs boson

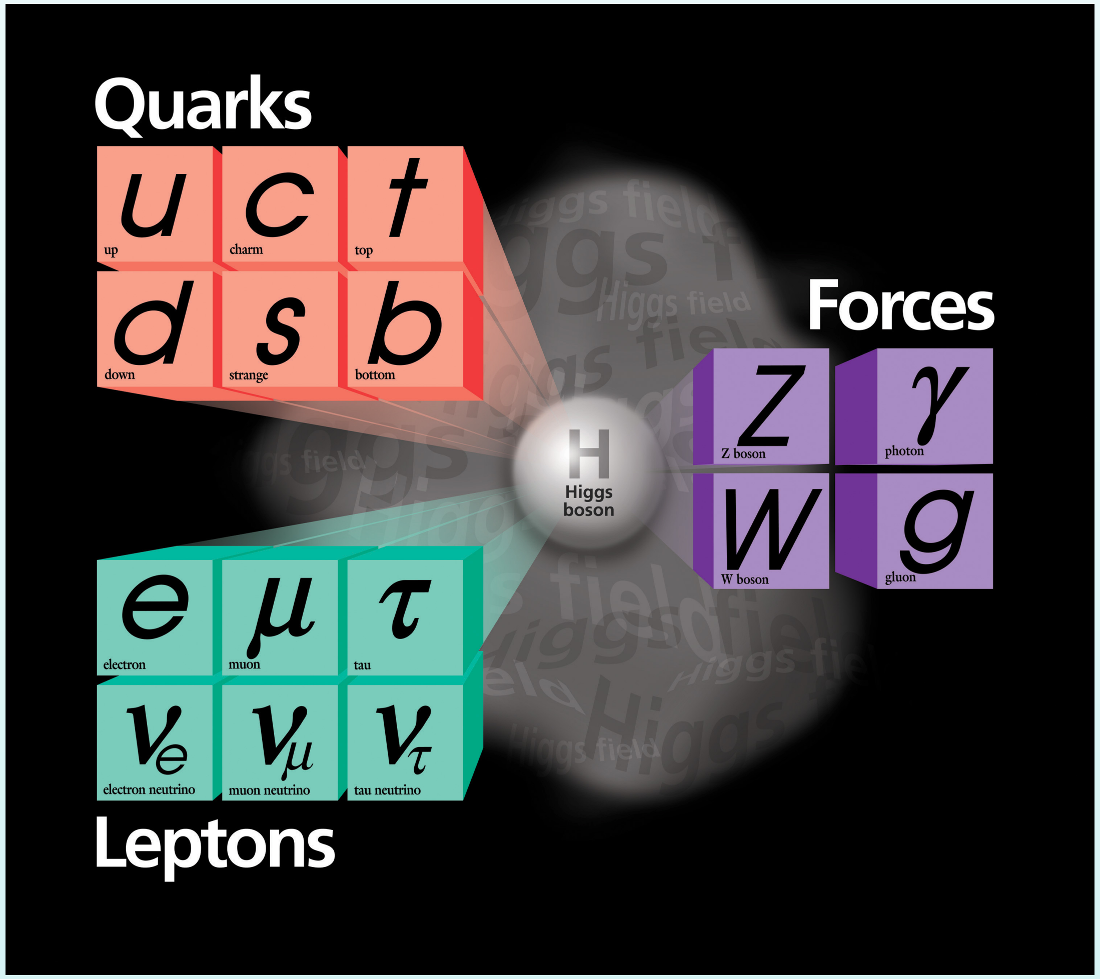
- Huge gap, **10^{17}**, between **Higgs and Plank scales**
- **Elementary or composite**? Additional Higgs bosons?
- Coupling to **Dark Matter**? Role in cosmological phase transitions?
- Is the **vacuum state of the Universe** stable?

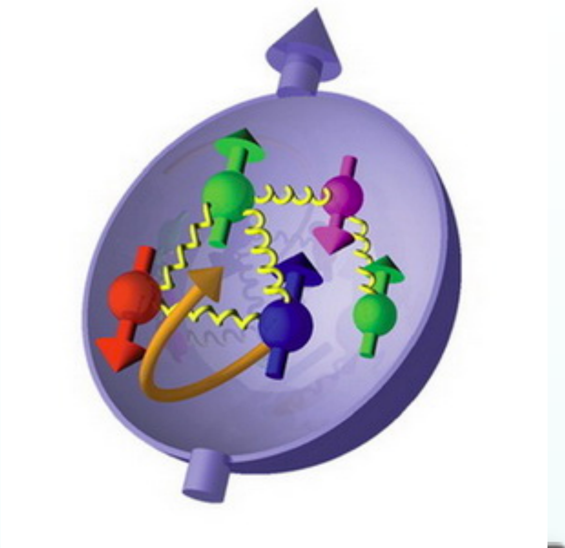
Quarks and leptons

- **Why three families?** Can we explain **masses and mixings**?
- Origin of **Matter-Antimatter asymmetry** in the Universe?
- Are **neutrinos Majorana or Dirac**? CP violation in the lepton sector?

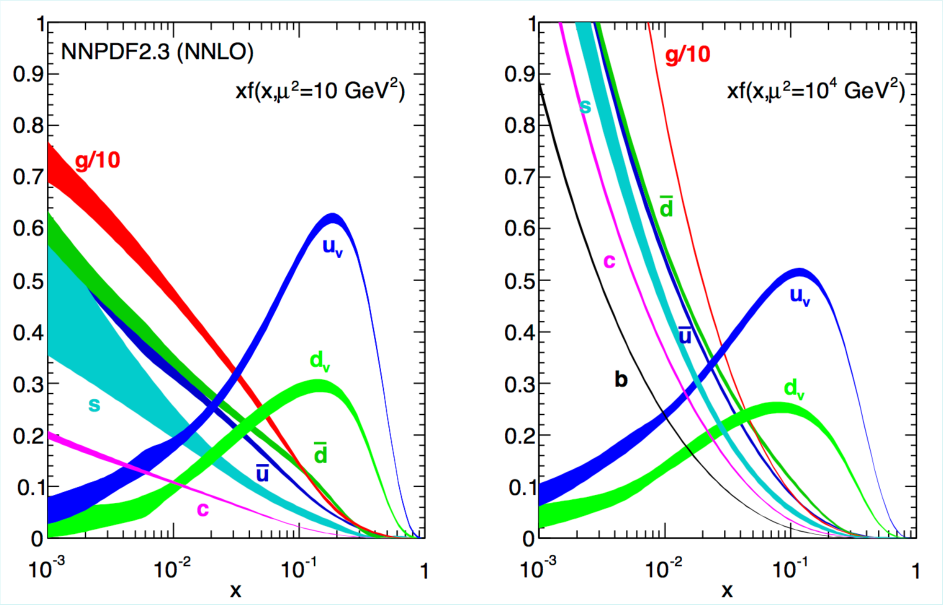
Dark Matter

- **Weakly interacting massive particles**? Sterile neutrinos? **Extremely light particles** (axions)?
- **Interactions** with Standard Model particles?
- What is the **structure of the Dark Sector**? Is Dark Matter self-interacting?



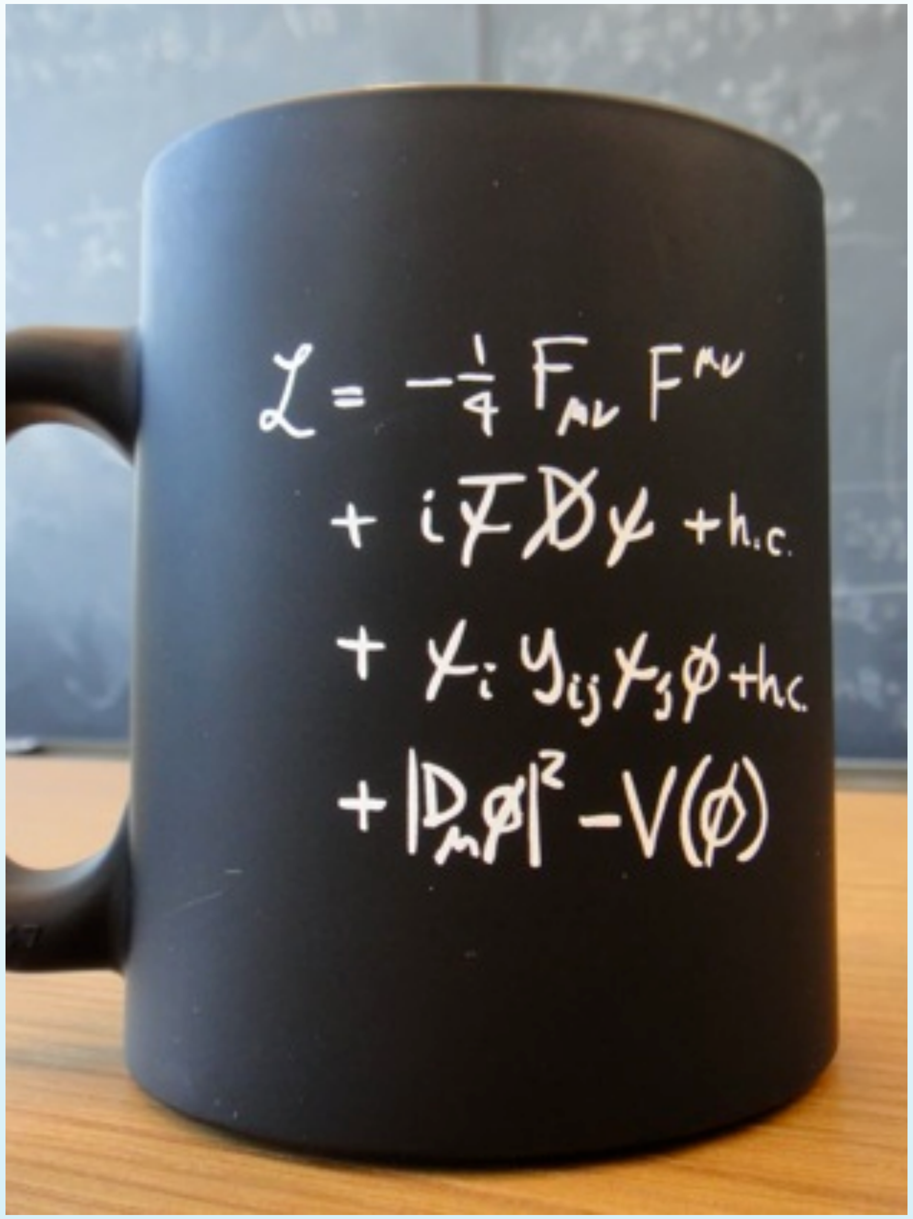


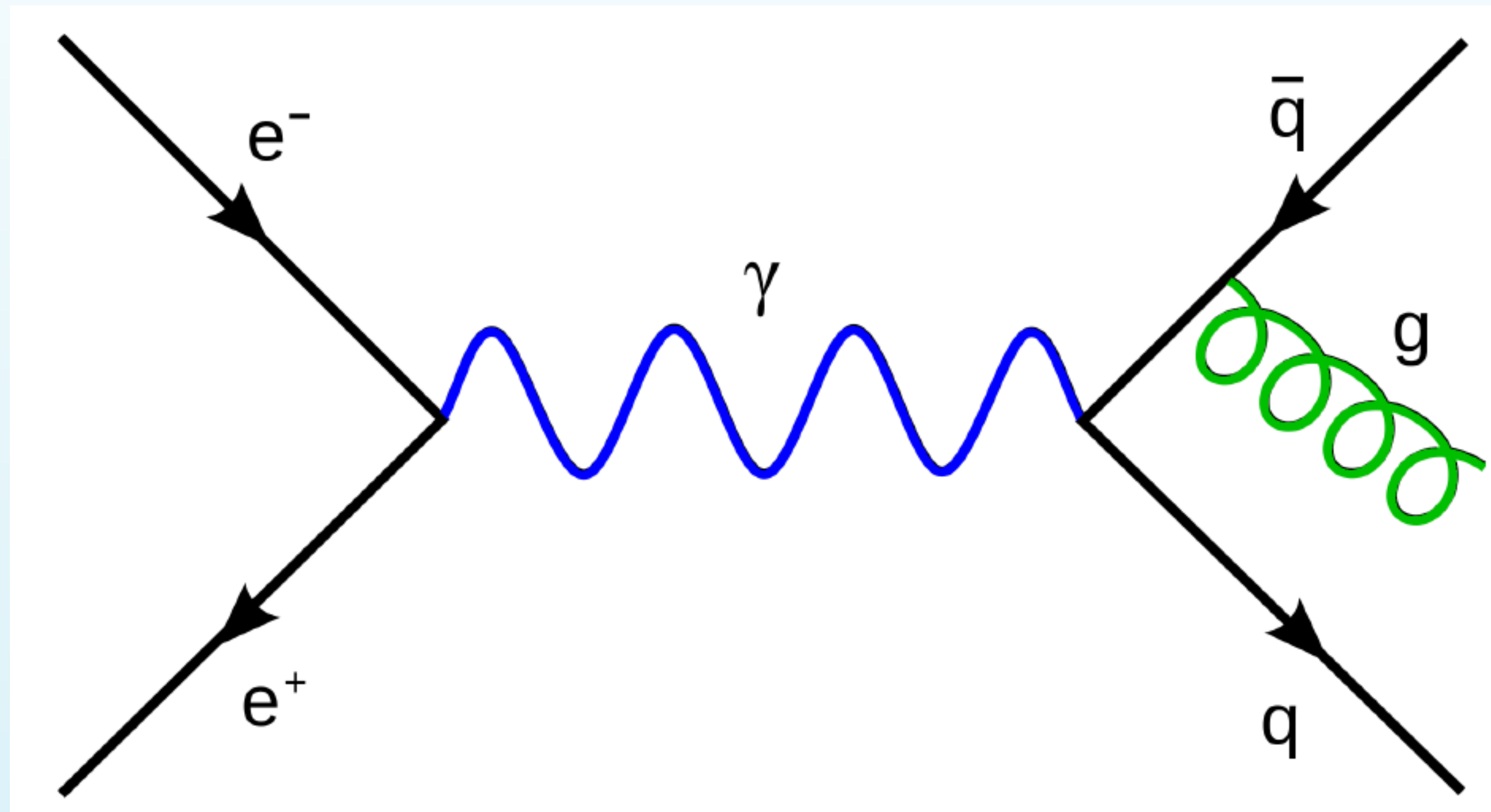
The inner life of protons : Parton Distribution Functions



Lepton vs Hadron Colliders

In high-energy **lepton colliders**, such as the **Large Electron-Positron Collider** (LEP) at CERN, the collisions involve **elementary particles** without substructure

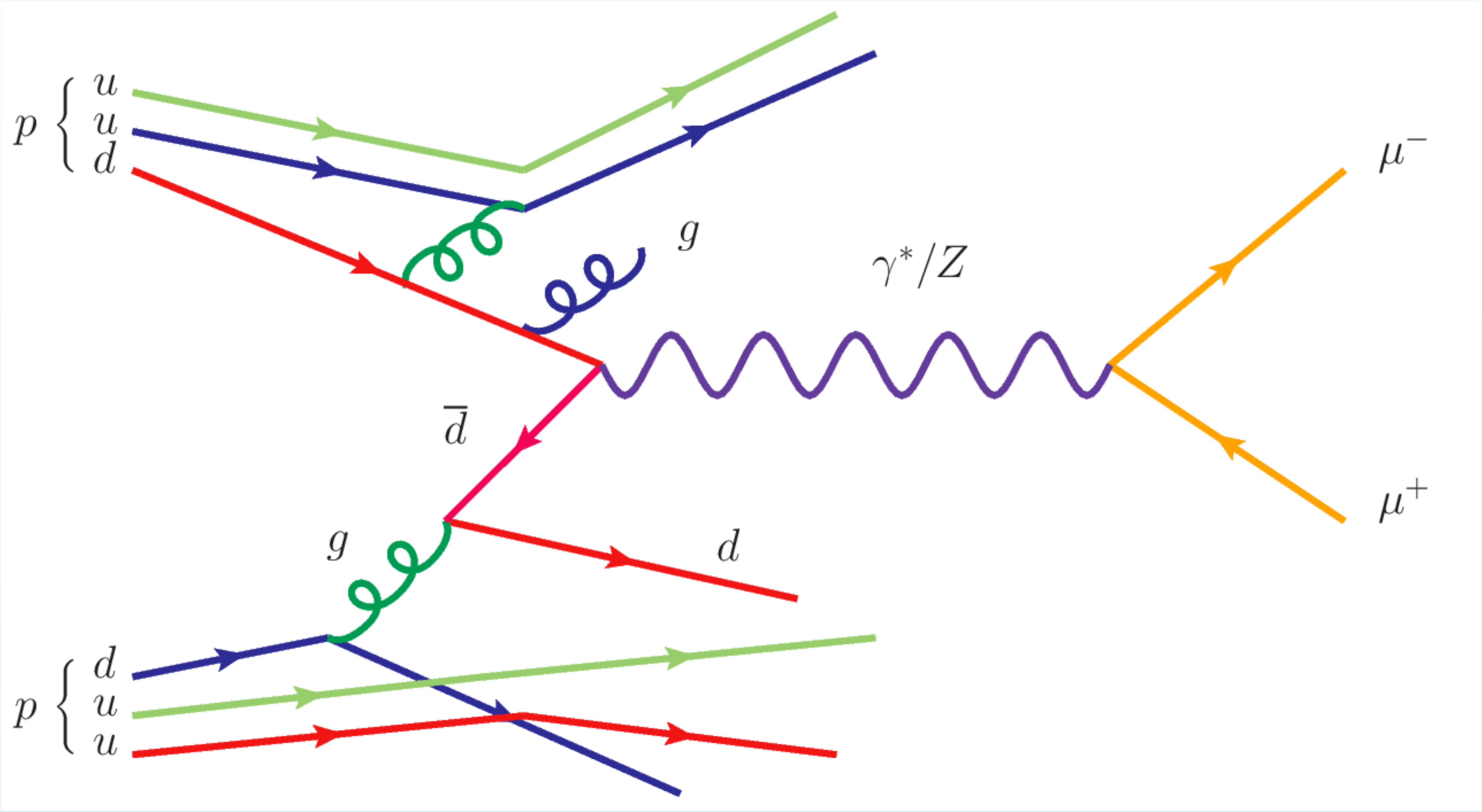




Cross-sections in lepton colliders can be computed in perturbation theory using the **Feynman rules of the Standard Model Lagrangian**

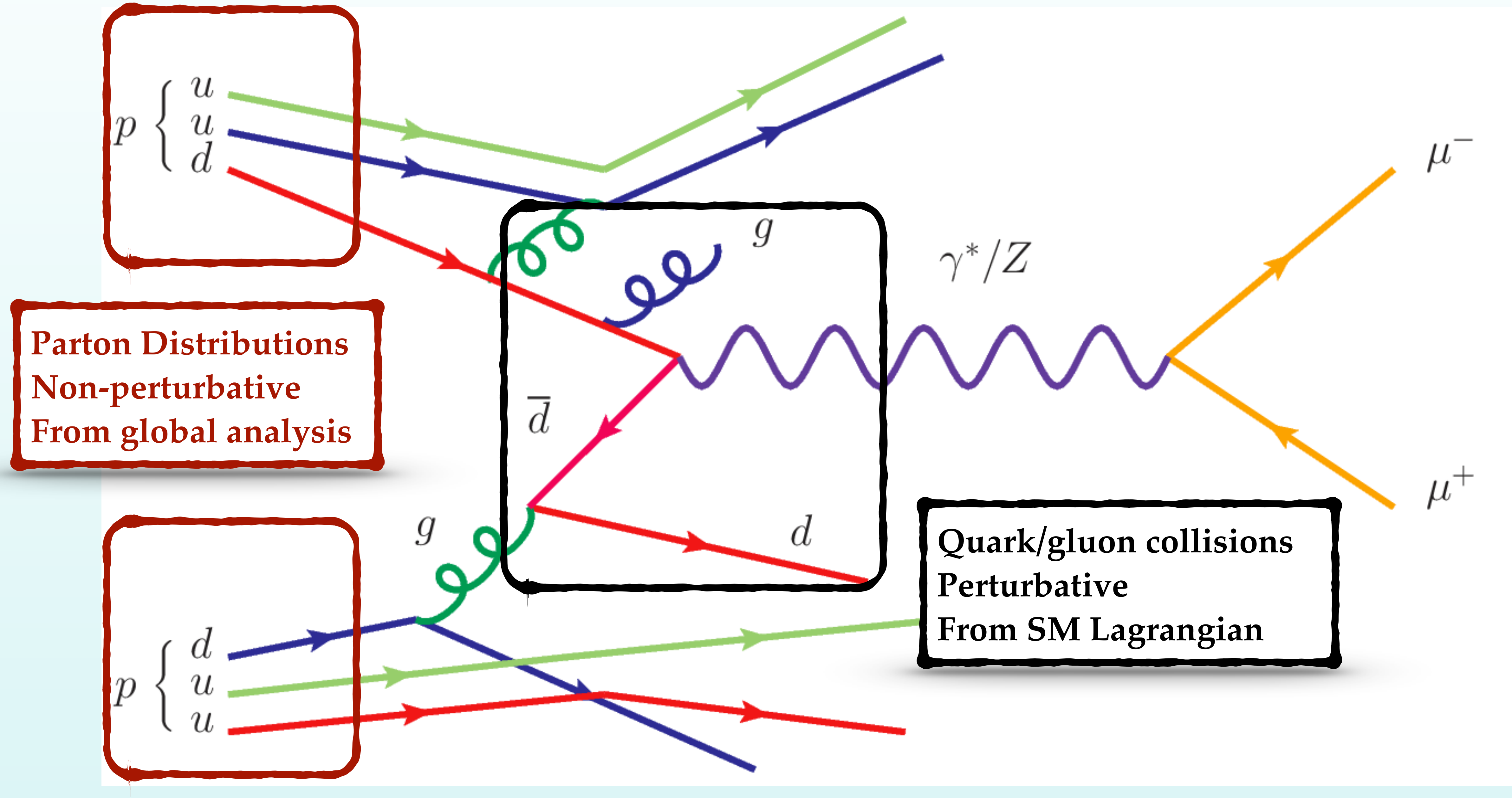
Lepton vs Hadron Colliders

In high-energy **hadron colliders**, such as the LHC, the collisions involve **composite particles** (protons) with internal structure (quarks and gluons)



Anatomy of a proton-proton collision

In high-energy **hadron colliders**, such as the LHC, the collisions involve **composite particles** (protons) with internal structure (quarks and gluons)



Calculations of c**ross-sections** in hadron collisions require the combination of **perturbative, quark/gluon-initiated processes**, and **non-perturbative, parton distributions**, information

Parton Distributions

The distribution of energy that **quarks and gluons carry inside the proton** is quantified by the **Parton Distribution Functions (PDFs)**

$$g(x, Q)$$

Q: **Energy of the quark/gluon collision**
Inverse of the resolution length

g(x,Q): **Probability** of **finding a gluon inside a proton**, carrying a fraction *x* of the proton momentum, when probed with energy *Q*

x: **Fraction of the proton's momentum**

Parton Distributions

The distribution of energy that **quarks and gluons carry inside the proton** is quantified by the **Parton Distribution Functions (PDFs)**

$$g(x, Q)$$

Q: **Energy of the quark/gluon collision**
Inverse of the resolution length

g(x,Q): **Probability** of **finding a gluon inside a proton**, carrying a fraction *x* of the proton momentum, when probed with energy *Q*

x: **Fraction of the proton's momentum**

PDFs are determined by **non-perturbative QCD dynamics:** cannot be computed from first principles, and need to be **extracted from experimental data** with a **global analysis**

Parton Distributions

The distribution of energy that **quarks and gluons carry inside the proton** is quantified by the **Parton Distribution Functions (PDFs)**

$$g(x, Q)$$

Q: **Energy of the quark/gluon collision**
Inverse of the resolution length

g(x,Q): **Probability** of **finding a gluon inside a proton**, carrying a fraction x of the proton momentum, when probed with energy Q

x: **Fraction of the proton's momentum**

PDFs are determined by **non-perturbative QCD dynamics:** cannot be computed from first principles, and need to be **extracted from experimental data** with a **global analysis**

- **Energy conservation** $$\int_0^1 dx \left(g(x,Q) + \sum_q q(x,Q) \right) = 1$$

- **Dependence with quark/gluon collision energy Q** determined in perturbation theory

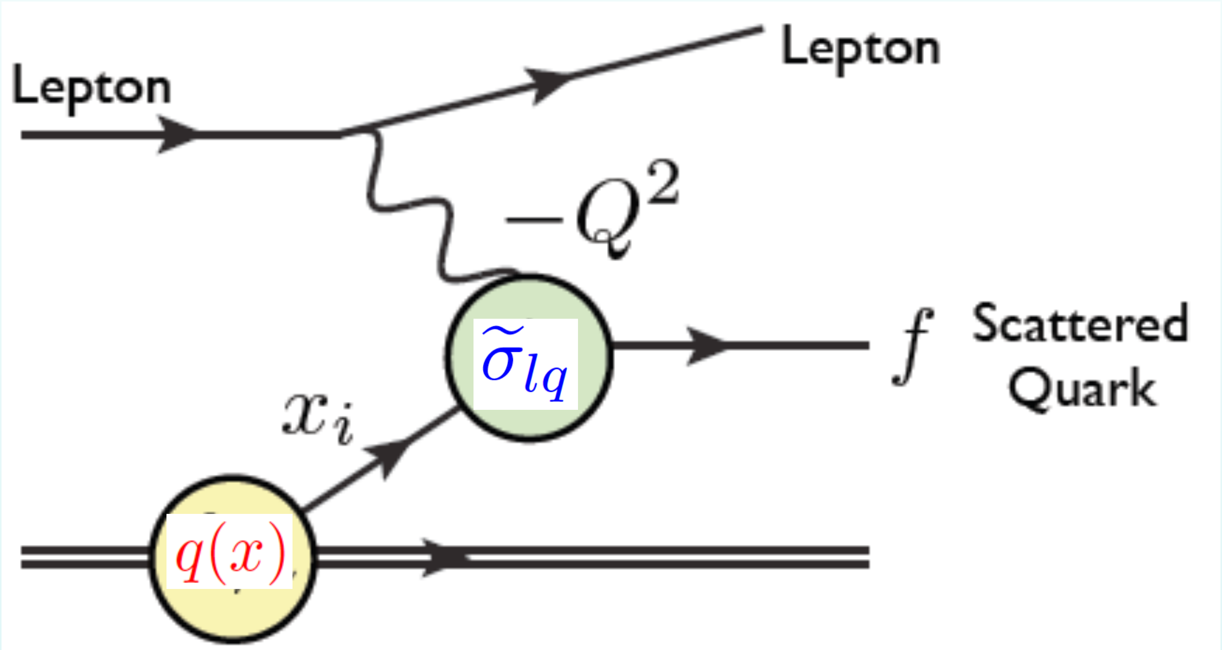
$$\frac{\partial g(x,Q)}{\partial \ln Q} = P_g\left(\alpha_s\right) \otimes g(x,Q) + P_q\left(\alpha_s\right) \otimes q(x,Q)$$

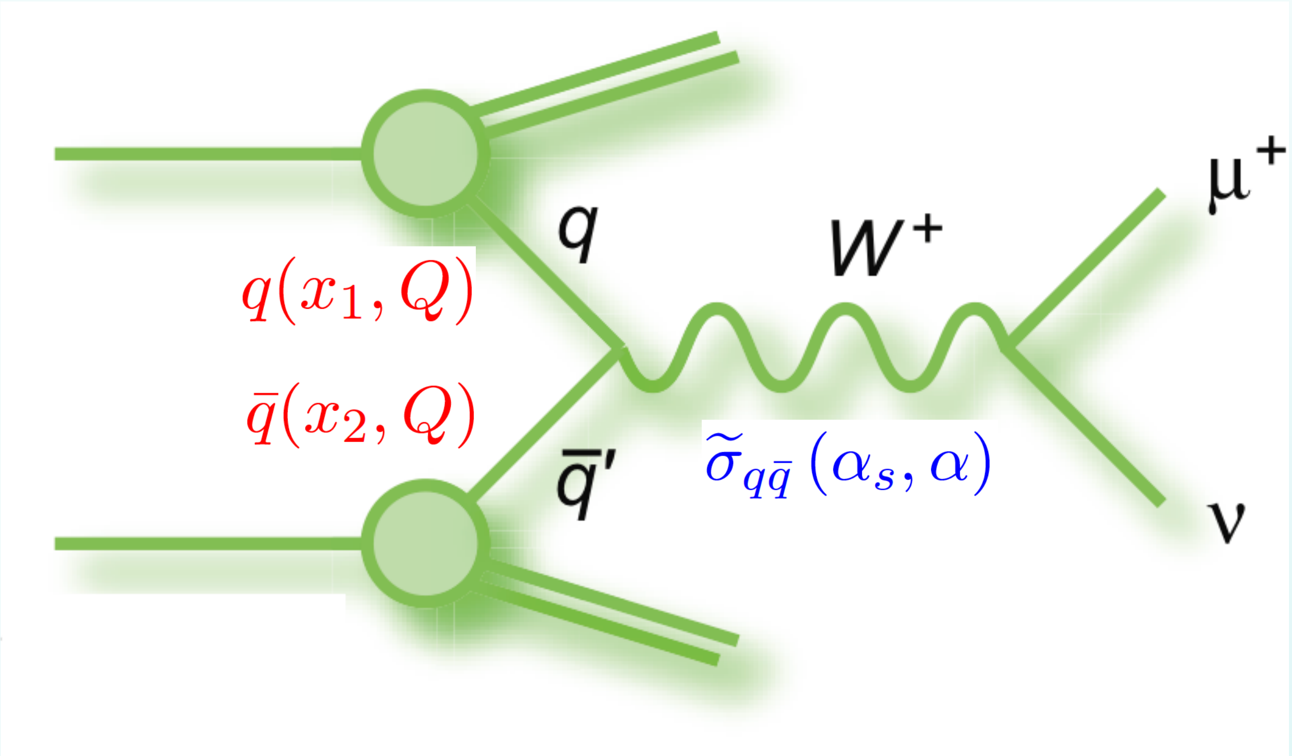
The Factorization Theorem

The **QCD Factorization Theorem** guarantees **PDF universality: extract them from a subset of** process and use them to provide pure predictions for new processes

$$\sigma_{lp} \simeq \widetilde{\sigma}_{lq}\left(\alpha_s, \alpha\right) \otimes q(x, Q)$$

$$\sigma_{pp} \simeq \widetilde{\sigma}_{q\bar{q}}\left(\alpha_s, \alpha\right) \otimes q(x_1, Q) \otimes \bar{q}(x_2, Q)$$





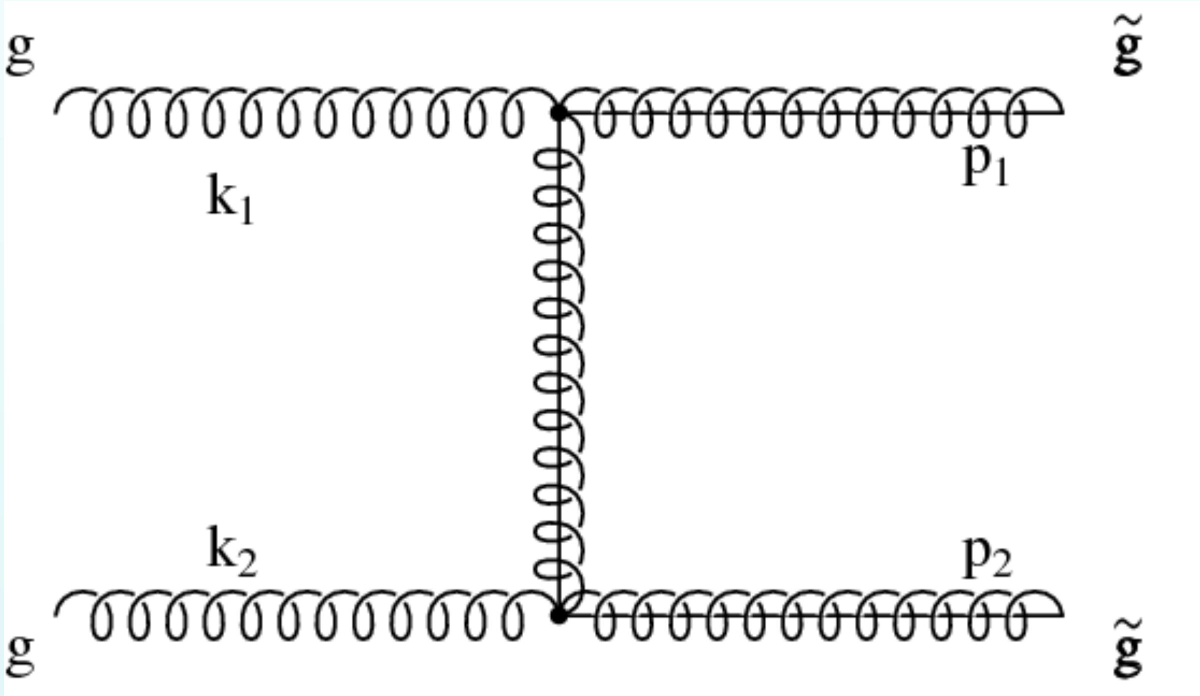
Determine PDFs in lepton-proton collisions ….

And use them to compute cross-sections in proton-proton collisions at the LHC

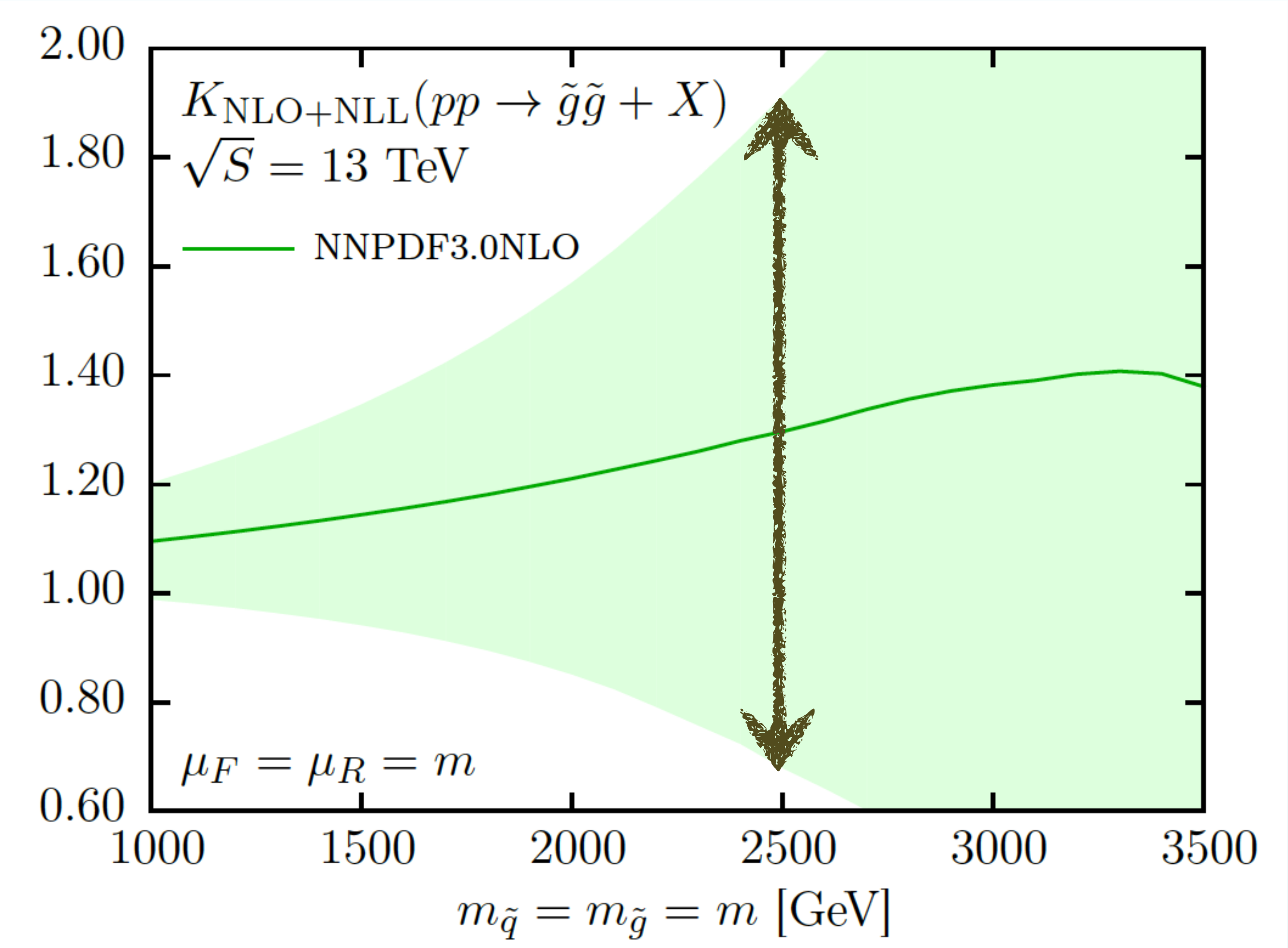
Beyond BSM discovery

PDF uncertainties in the production of **New Physics heavy resonances** can be al large as **100%!**

Crucial *i.e.* in searches for ***supersymmetry*** and any BSM scenario that predicts new heavy particles within the reach of the LHC



Gluino pair production at the LHC



Beenakker, Borchensky, Kramer, Kulesza, Laenen, Marzani, Rojo 15

*Unless we **improve PDF uncertainties**, even if we discover New Physics, it will be extremely **difficult to characterise the underlying BSM scenario***

ANNs as universal unbiased interpolants

ANNs provide **universal unbiased interpolants** to parametrize the non-perturbative dynamics that determines the **size and shape of the PDFs** from experimental data

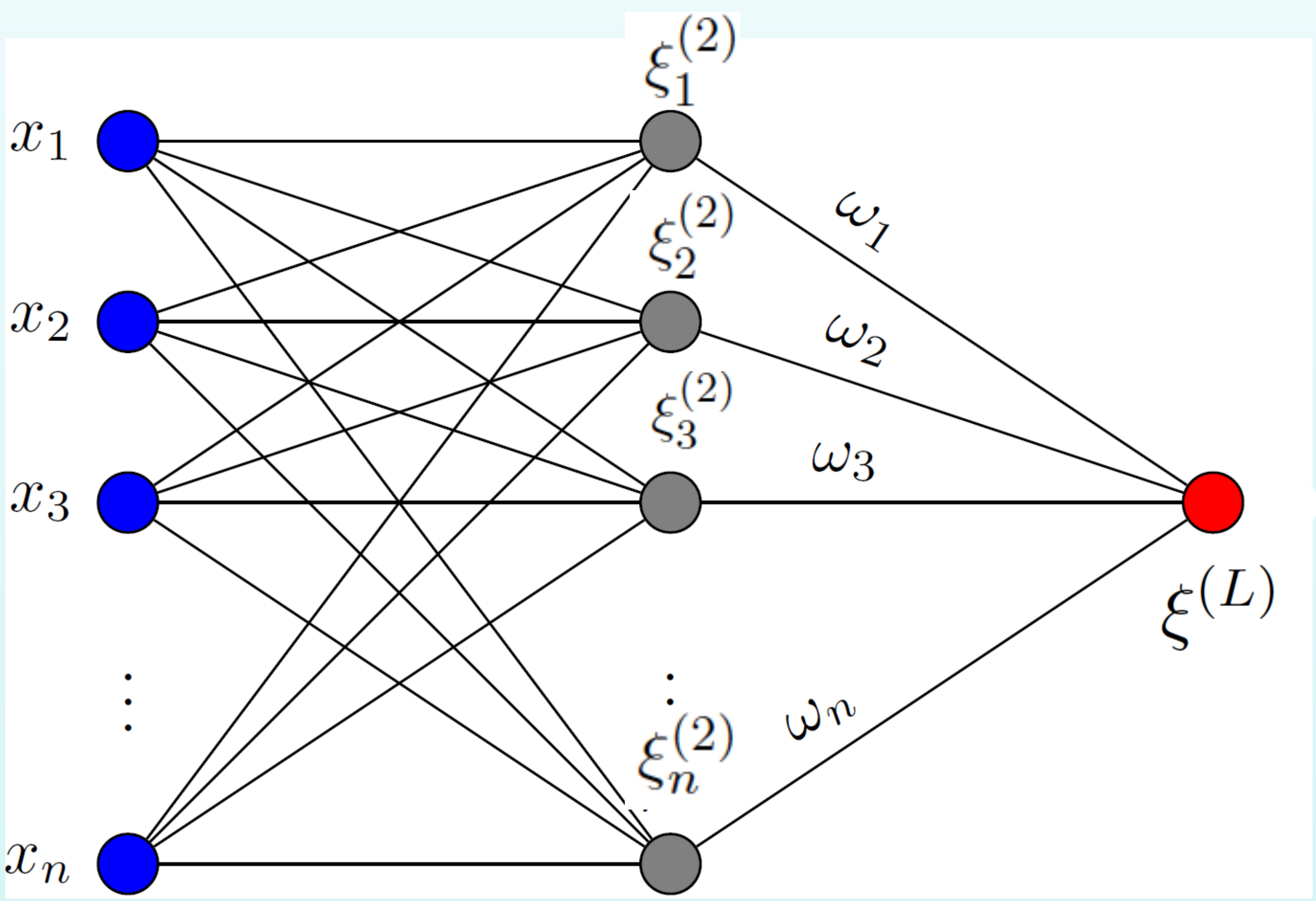
not from QCD!

Traditional approach

$$g(x,Q_0) = A_g(1-x)^{a_g}x^{-b_g}\left(1+c_g\sqrt{s}+d_gx+\ldots\right)$$

NNPDF approach

$$g(x,Q_0) = A_g\mathrm{ANN}_g(x)$$



$$\mathrm{ANN}_g(x) = \xi^{(L)} = \mathcal{F}\left[\xi^{(1)}, \{\omega_{ij}^{(l)}\}, \{\theta_i^{(l)}\}\right]$$

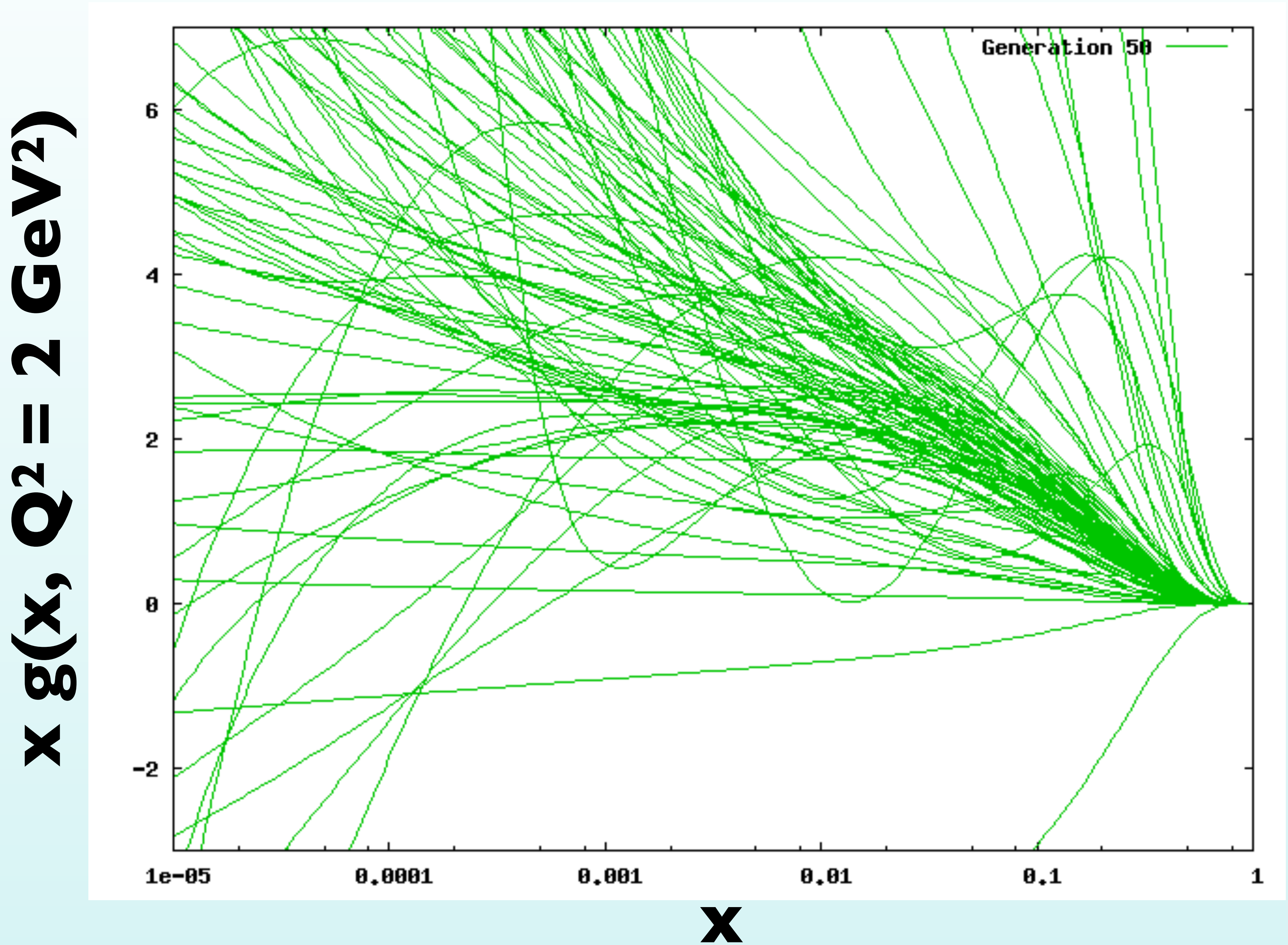
$$\xi_i^{(l)} = g\left(\sum_{j=1}^{n_{l-1}} \omega_{ij}^{(l-1)}\xi_j^{(l-1)} - \theta_i^{(l)}\right)$$

- ANNs eliminate **theory bias** introduced in PDF fits from choice of *ad-hoc* functional forms
- NNPDF fits used **O(400) free parameters**, to be compared with O(10-20) in traditional PDFs. Results stable if **O(4000) parameters used!**

PDF Replica Neural Network Learning

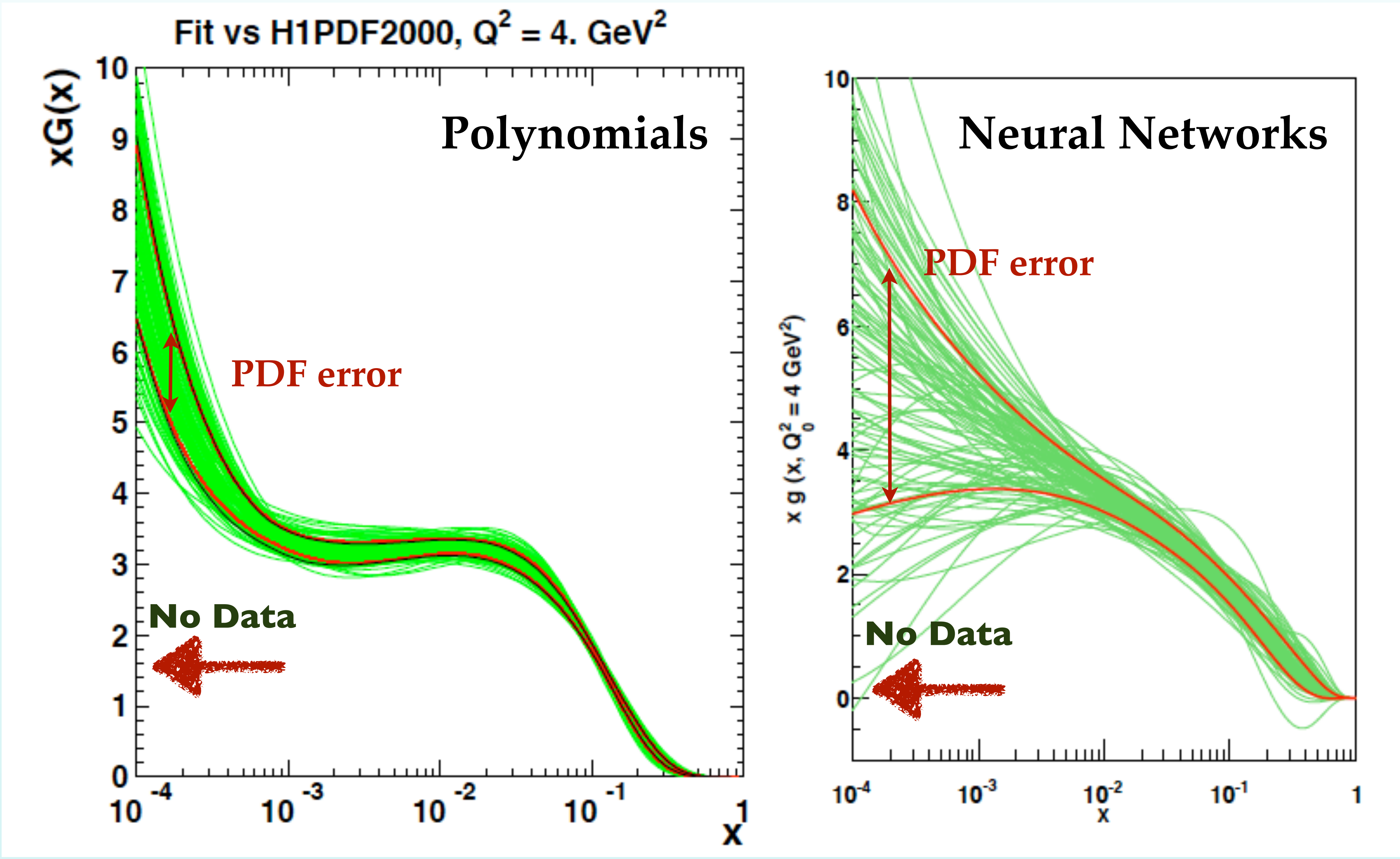
The minimisation of the **data vs theory χ^2** is performed using **Genetic Algorithms**

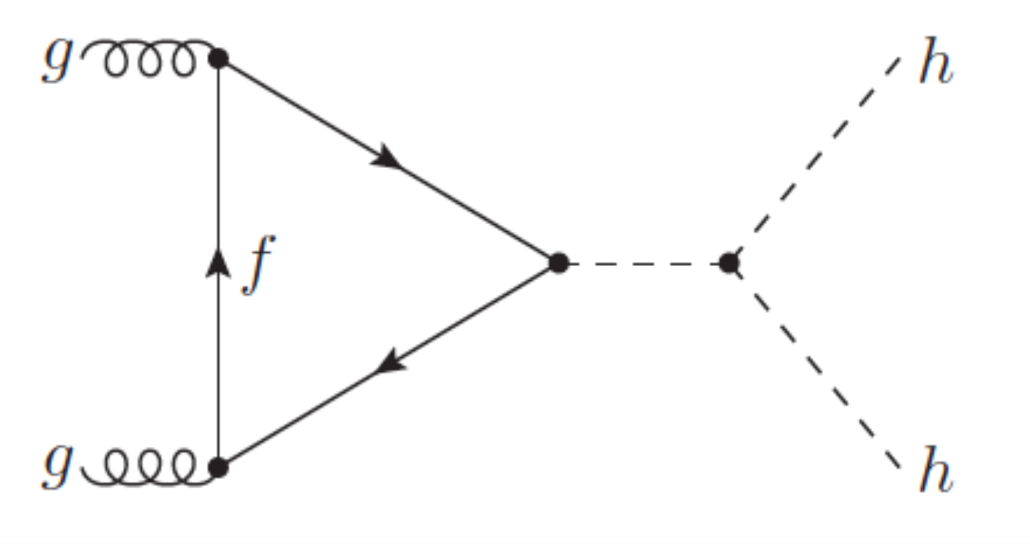
Each **green curve** corresponds to a **gluon PDF Monte Carlo** replica



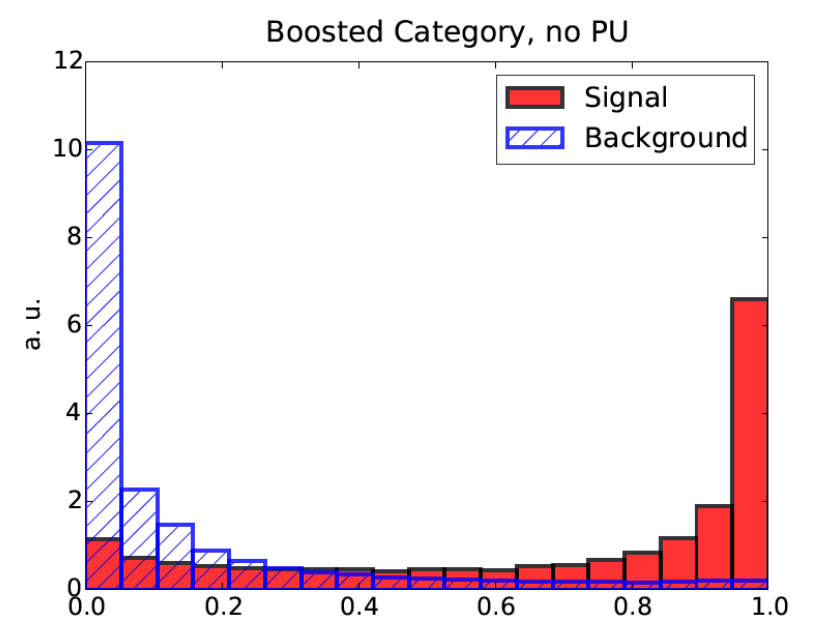
Artificial Neural Networks vs Polynomials

- Compare a **benchmark PDF analysis** where **the same dataset** is fitted with **Artificial Neural Networks** and with **standard polynomials,** other settings identical)
- **ANNs** avoid biasing the PDFs, **faithful extrapolation at small-x** (very few data, thus error blow up)





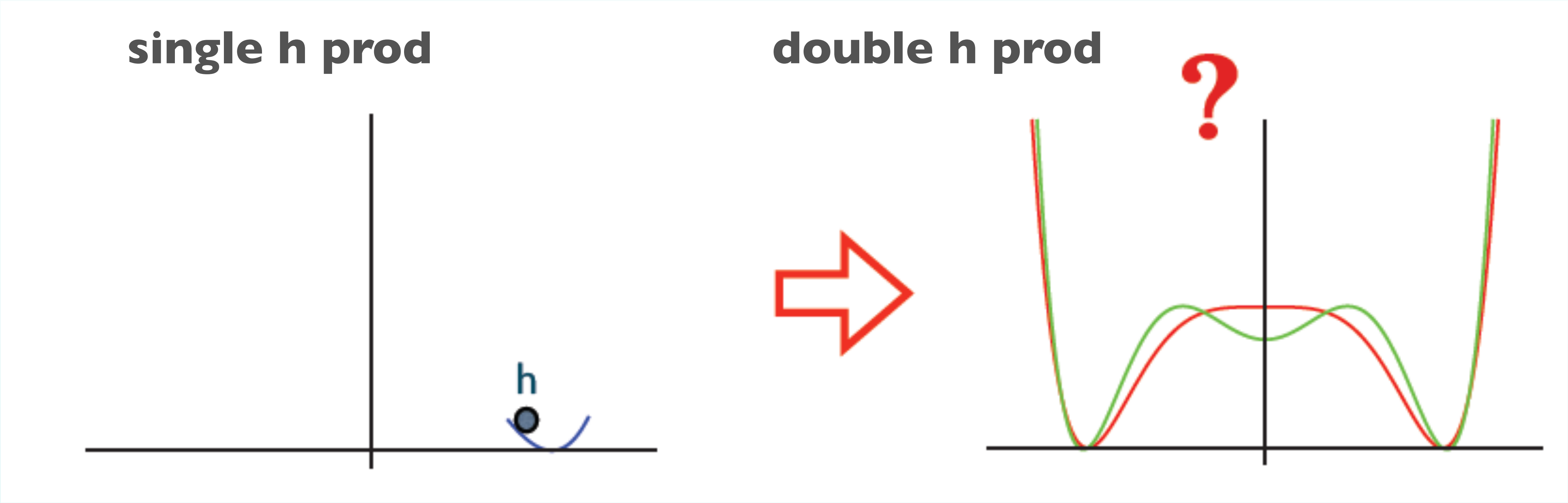
Unravelling the Higgs Self-Coupling



Berh, Bortolletto, Frost, Hartland, Issever, Rojo 15
Bishara, Contino, Rojo 16

Probing Electroweak Symmetry breaking

- **Current measurements** (couplings in single Higgs production) probe **Higgs potential close to minimum**
- Double Higgs production essential to **reconstruct the full Higgs potential** and clarify EWSB mechanism
- Higgs SM potential is *ad-hoc*: not fixed by the SM symmetries, **many other EWSB mechanisms conceivable**



Higgs mechanism

$$V(h) = m_h^2 h^\dagger h + \frac{1}{2}\lambda(h^\dagger h)^2$$

Coleman-Weinberg mechanism

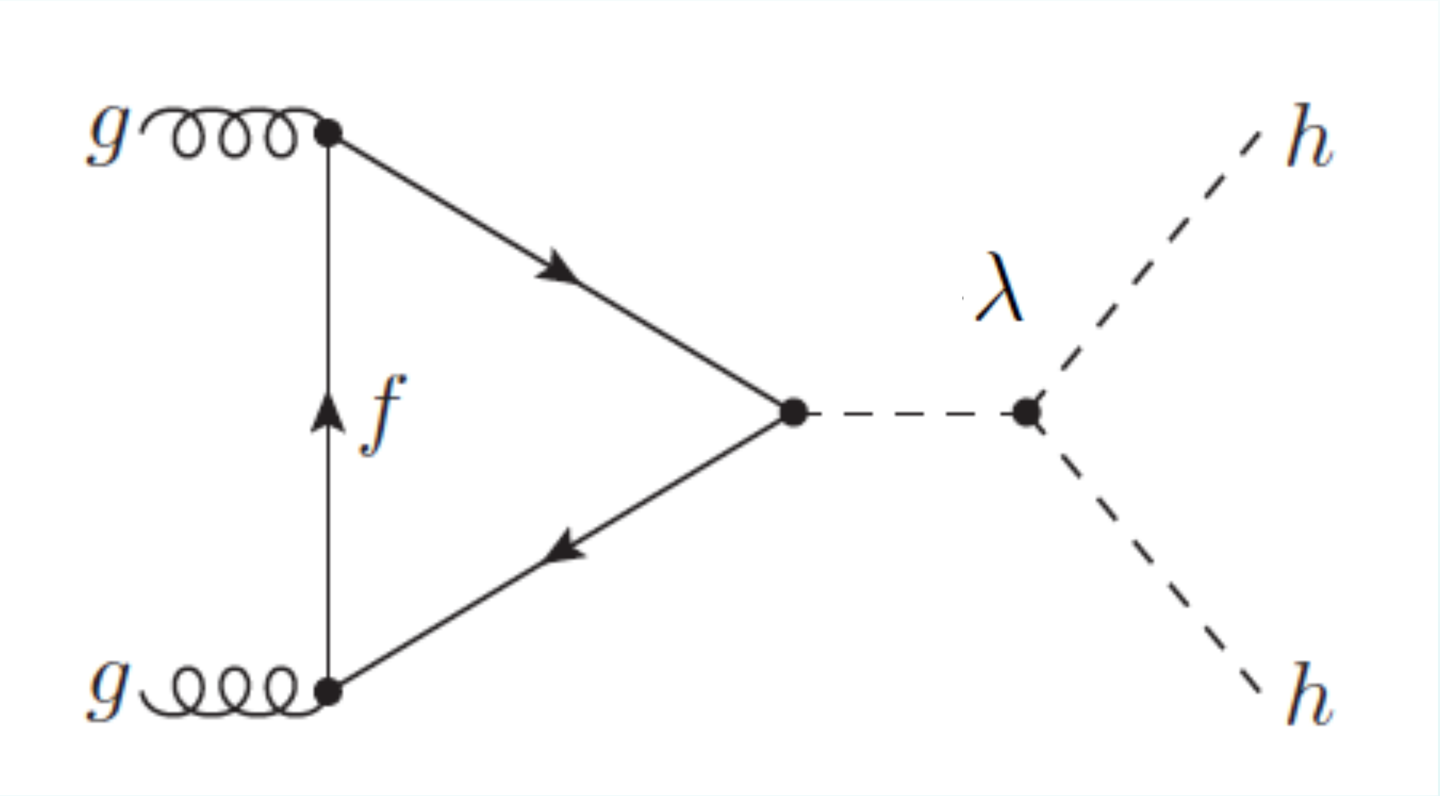
$$V(h) \to \frac{1}{2}\lambda(h^\dagger h)^2 \log\left[\frac{(h^\dagger h)}{m^2}\right]$$

Each possibility associated to **completely different EWSB mechanism**, with crucial implications for the **hierarchy problem**, the structure of quantum field theory, and **New Physics at the EW scale**

Arkani-Hamed, Han, Mangano, Wang, arxiv:1511.06495

Probing Electroweak Symmetry breaking

- **Current measurements** (couplings in single Higgs production) probe **Higgs potential close to minimum**
- Double Higgs production essential to **reconstruct the full Higgs potential** and clarify EWSB mechanism
- The Higgs potential is *ad-hoc*: **many other EWSB mechanisms conceivable**



Higgs mechanism

$$V(h) = m_h^2 h^\dagger h + \frac{1}{2}\lambda(h^\dagger h)^2$$

Coleman-Weinberg mechanism

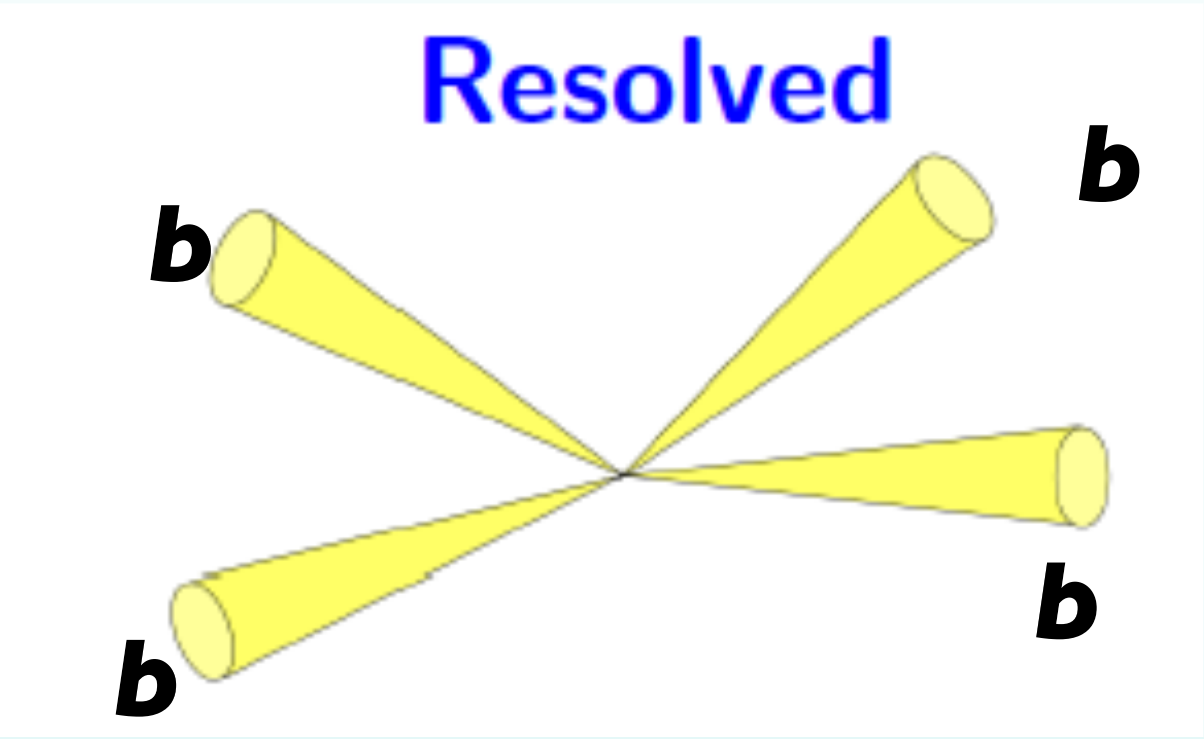
$$V(h) \to \frac{1}{2}\lambda(h^\dagger h)^2 \log\left[\frac{(h^\dagger h)}{m^2}\right]$$

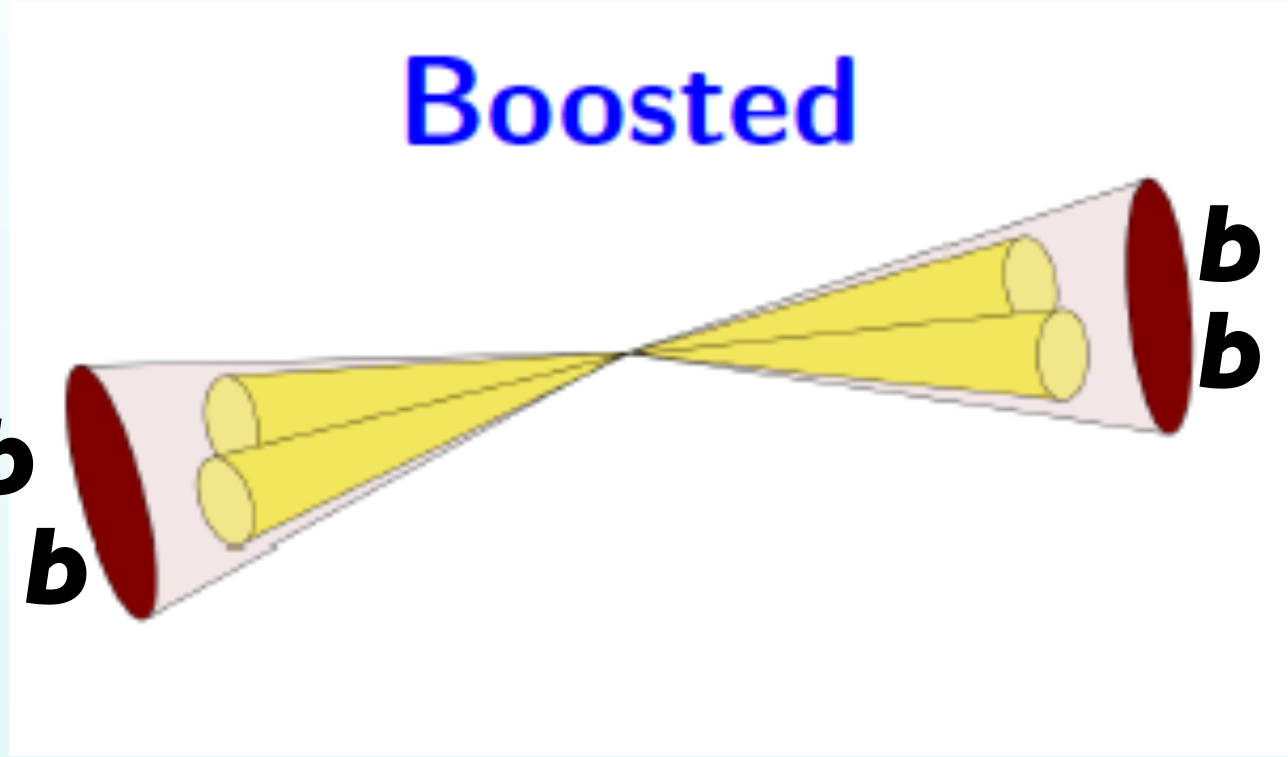
Each possibility associated to **completely different EWSB mechanism**, with crucial implications for the **hierarchy problem**, the structure of quantum field theory, and **New Physics at the EW scale**

Arkani-Hamed, Han, Mangano, Wang, arxiv:1511.06495

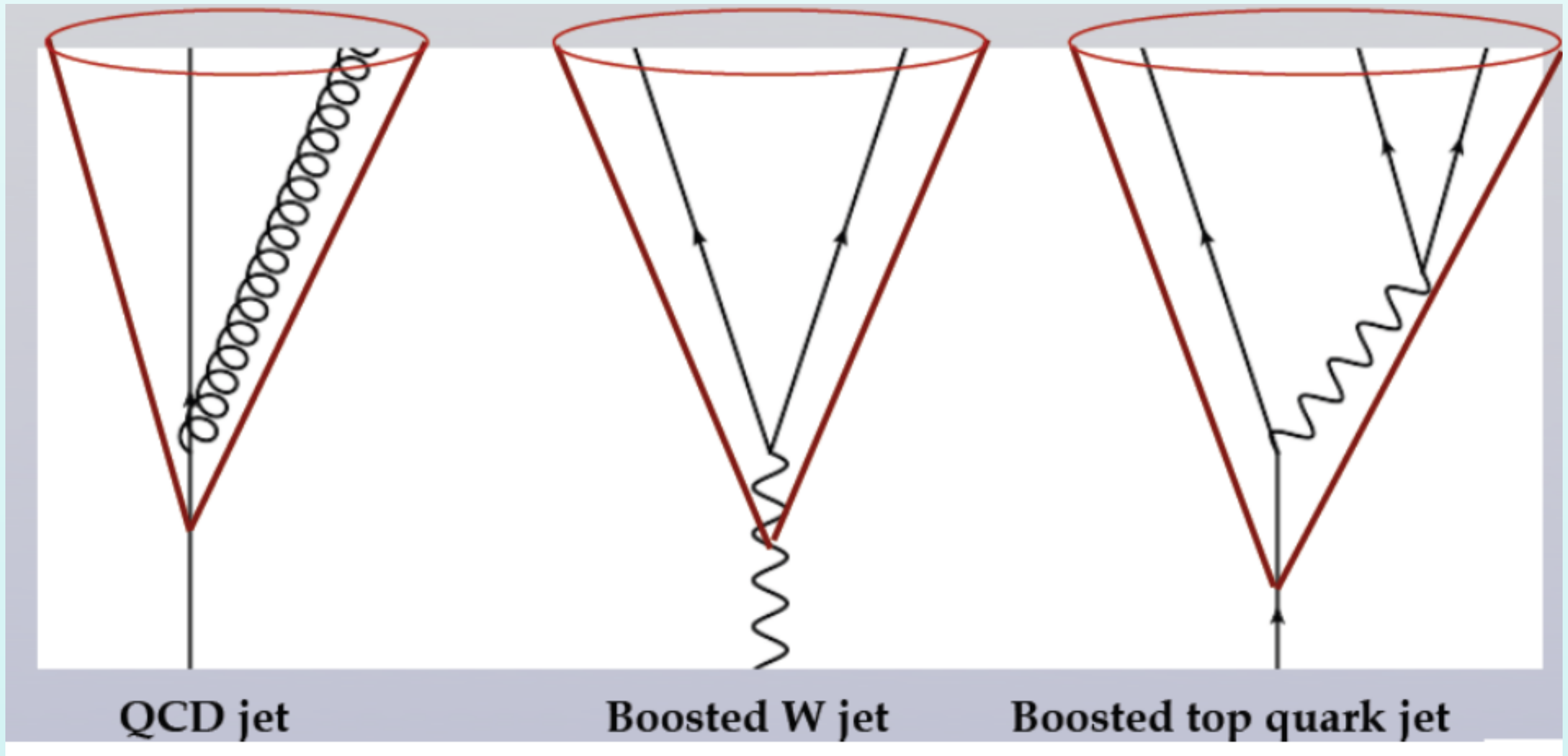
hh->bbbb: selection strategy

- Exploit **4b final state**: highest signal yields, but **overwhelming QCD background** (by orders of magnitude!)
- Carefully chosen selection strategies ensure that **all relevant event topologies can be reconstructed**

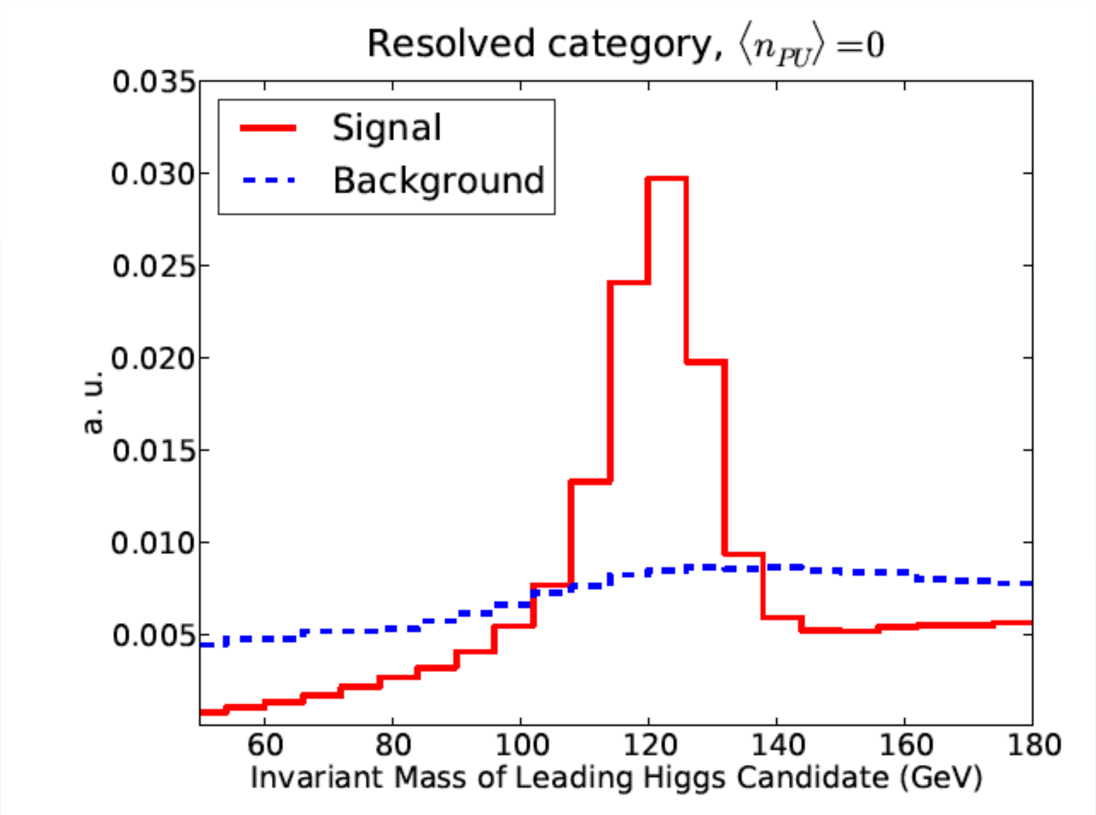


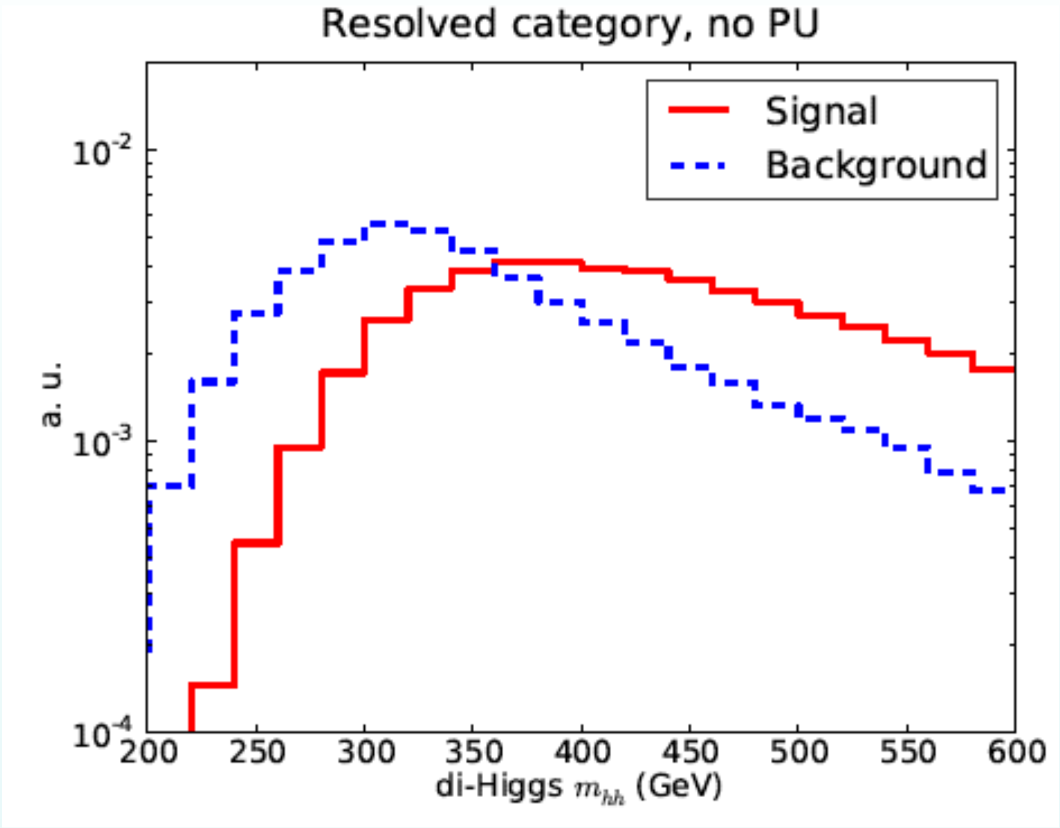


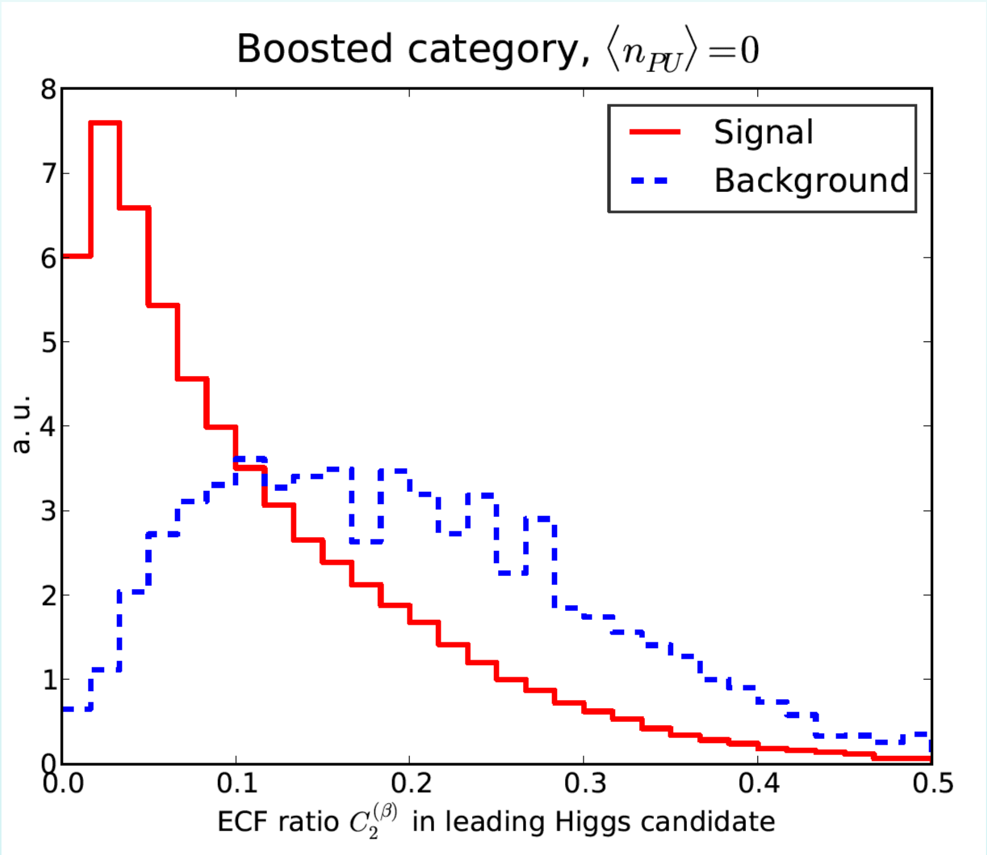
Recent progress in **jet substructure** techniques important to reduced QCD background in the **boosted regime**

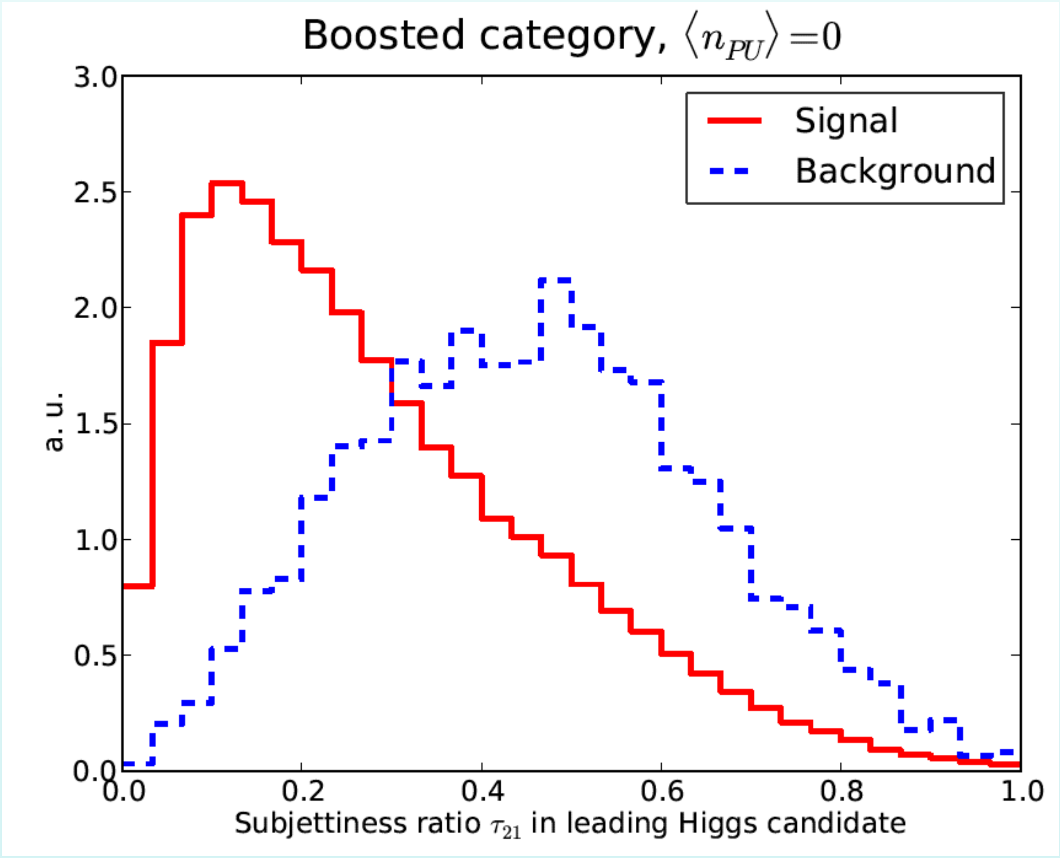


di-Higgs kinematic distributions

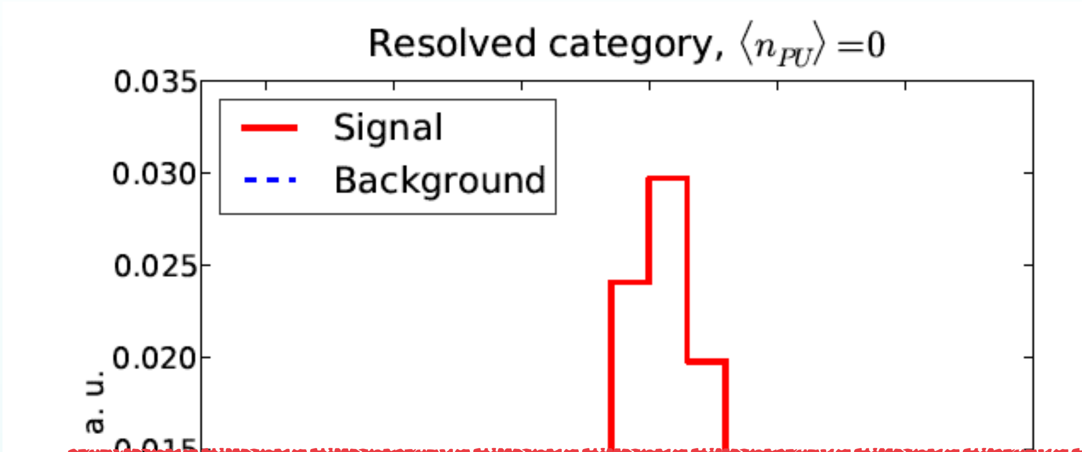


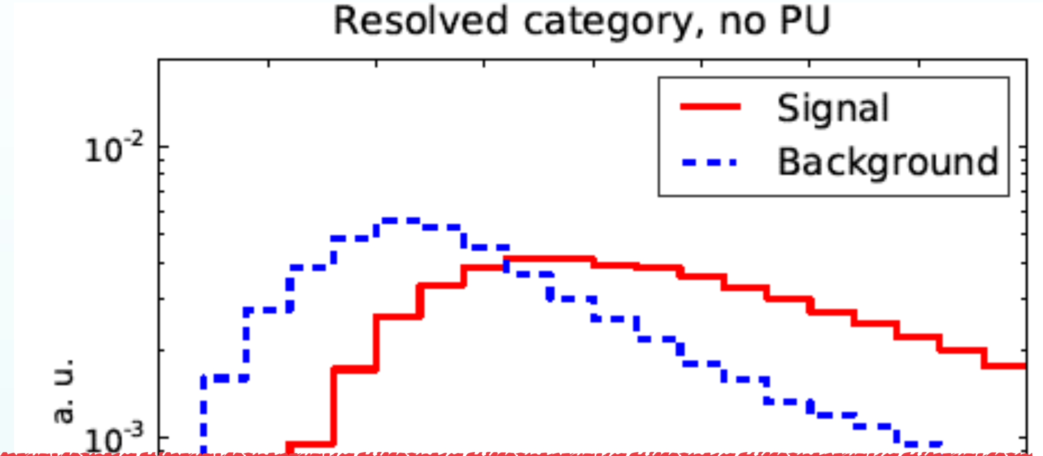






di-Higgs kinematic distributions

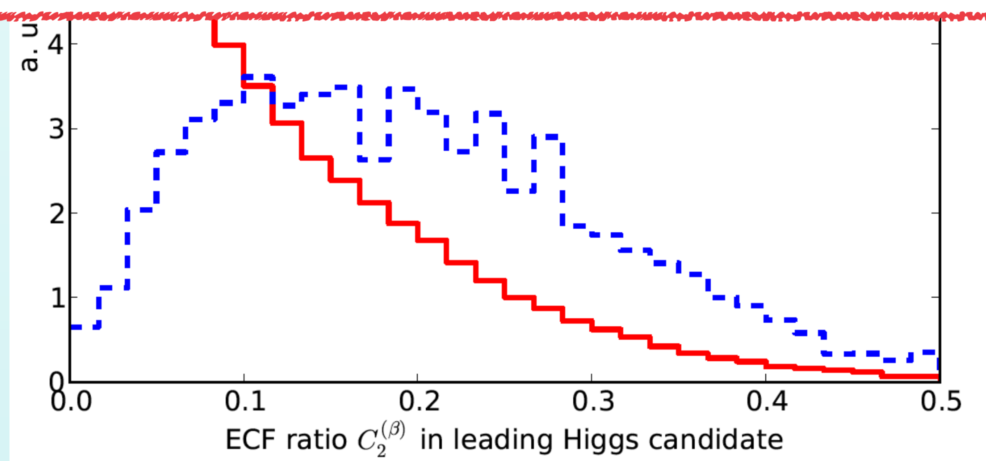


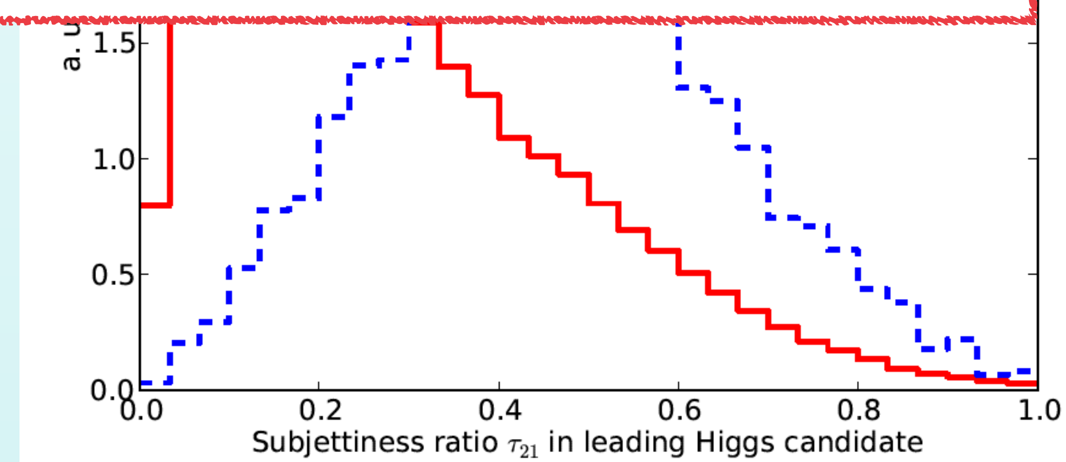


Many kinematic variables can be used to **disentangle signal and background**

How do we select which ones to use? And the optical cuts? And the cross-correlations among variables?

We don't need to! Use **ML methods to identify automatically** the combination of kinematical variables with the highest discrimination power!





Multivariate techniques

Input

Higgs pT

Higgs m

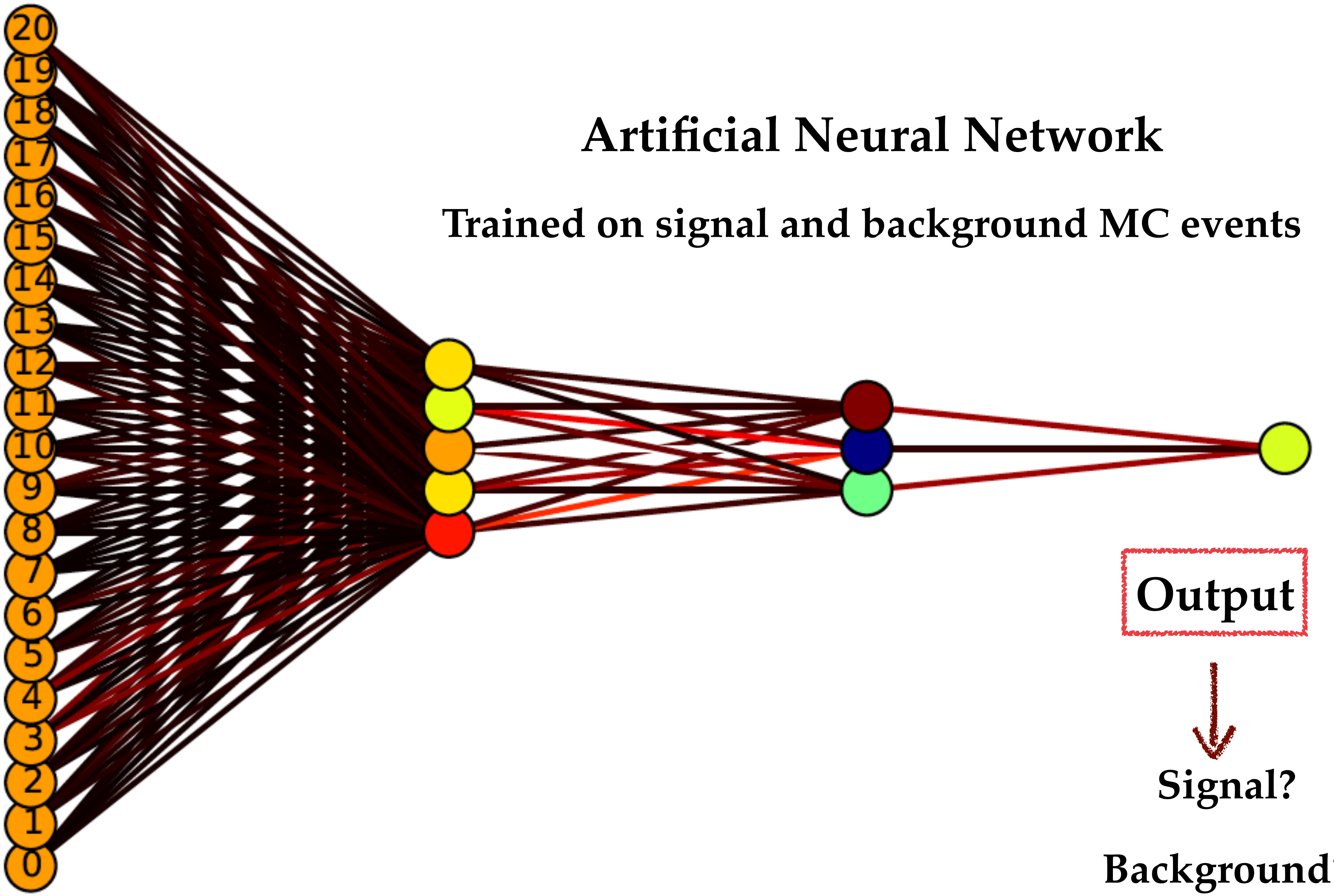
di-Higgs m

ECF

τ_{12}

Subjet pT

.....



Artificial Neural Network

Trained on signal and background MC events

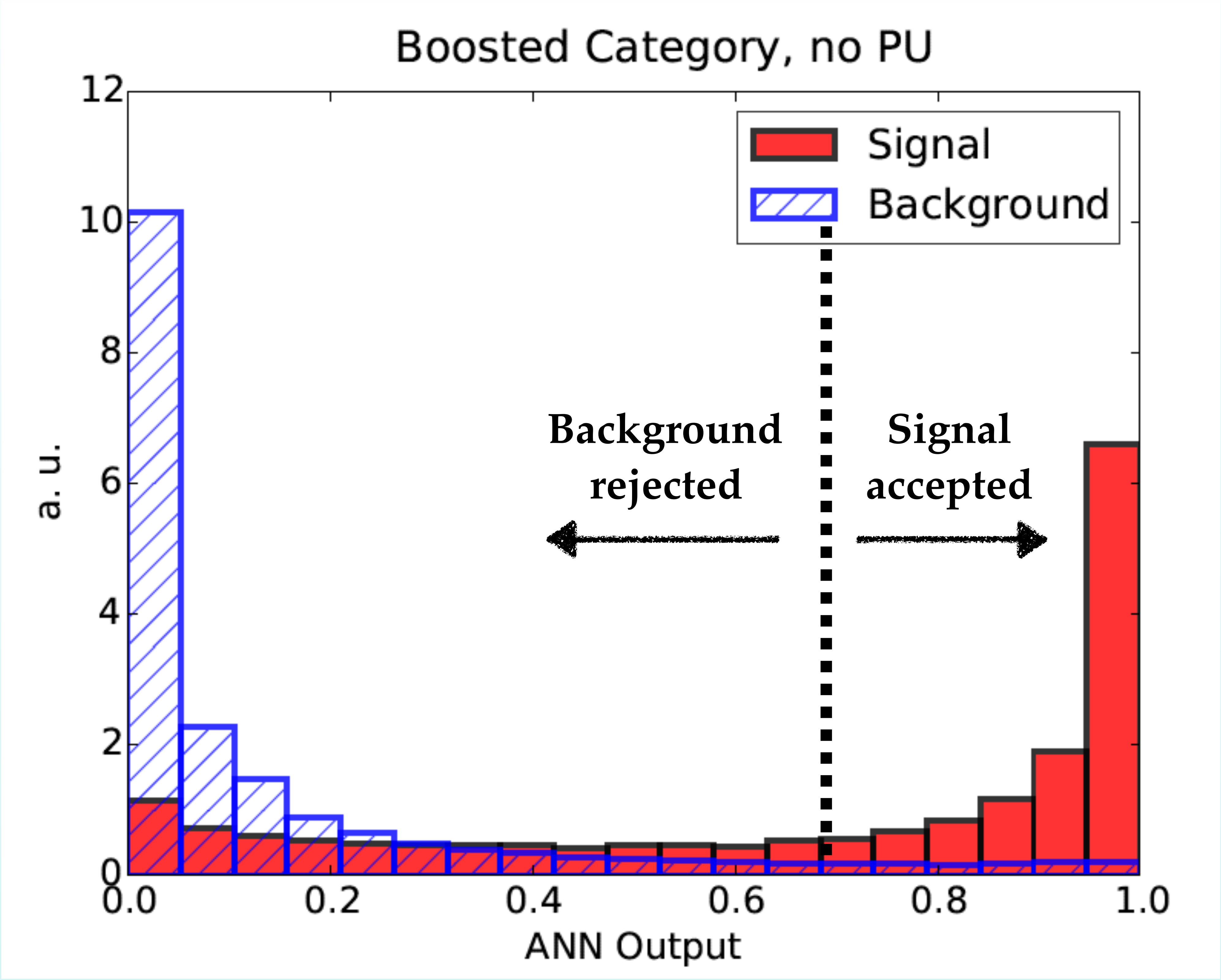
Output

Signal?

Background?

Multivariate techniques

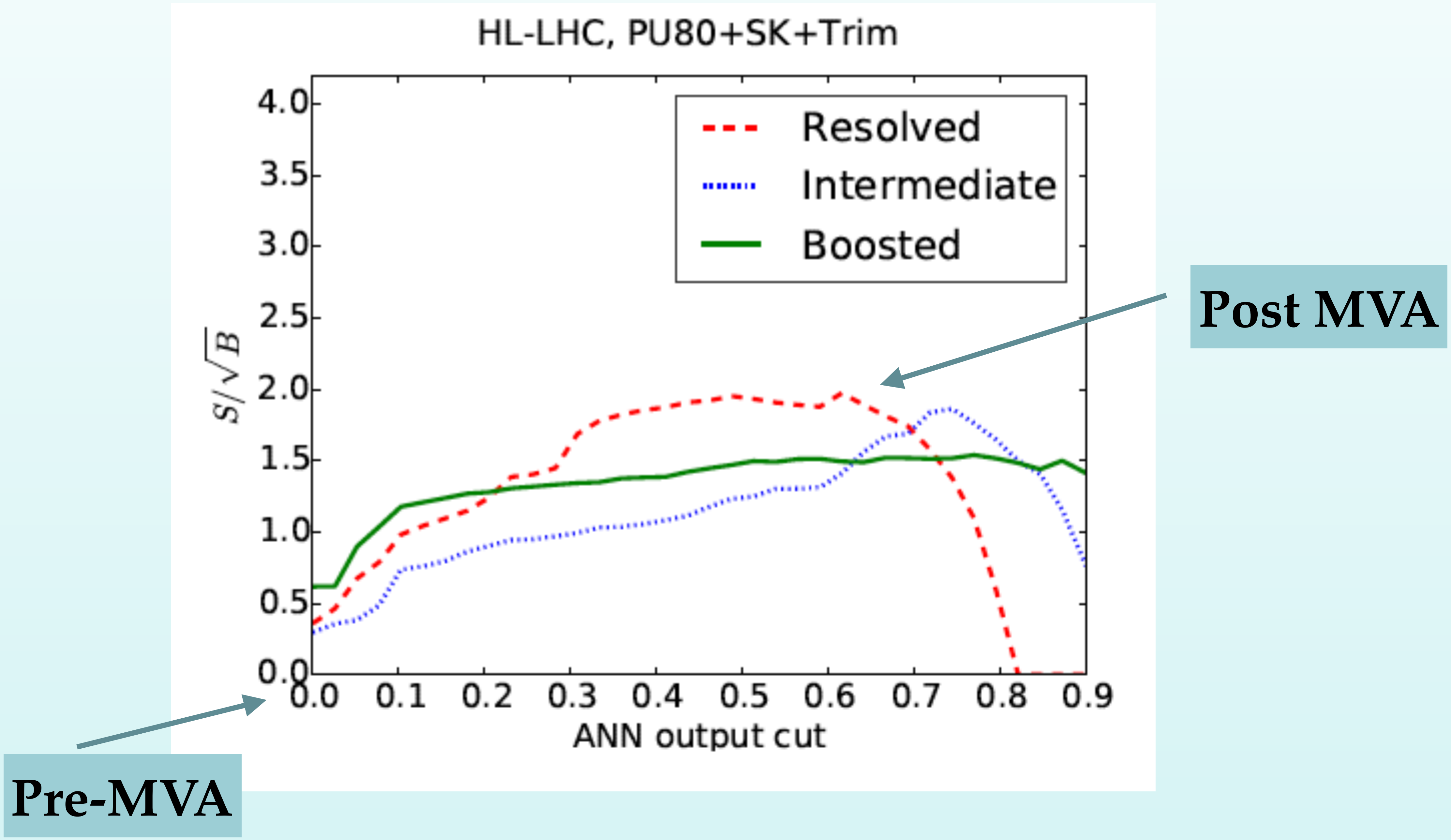
Combining information from all kinematic variables in MVA: **excellent signal/background discrimination**



Discovering Higgs self-interactions

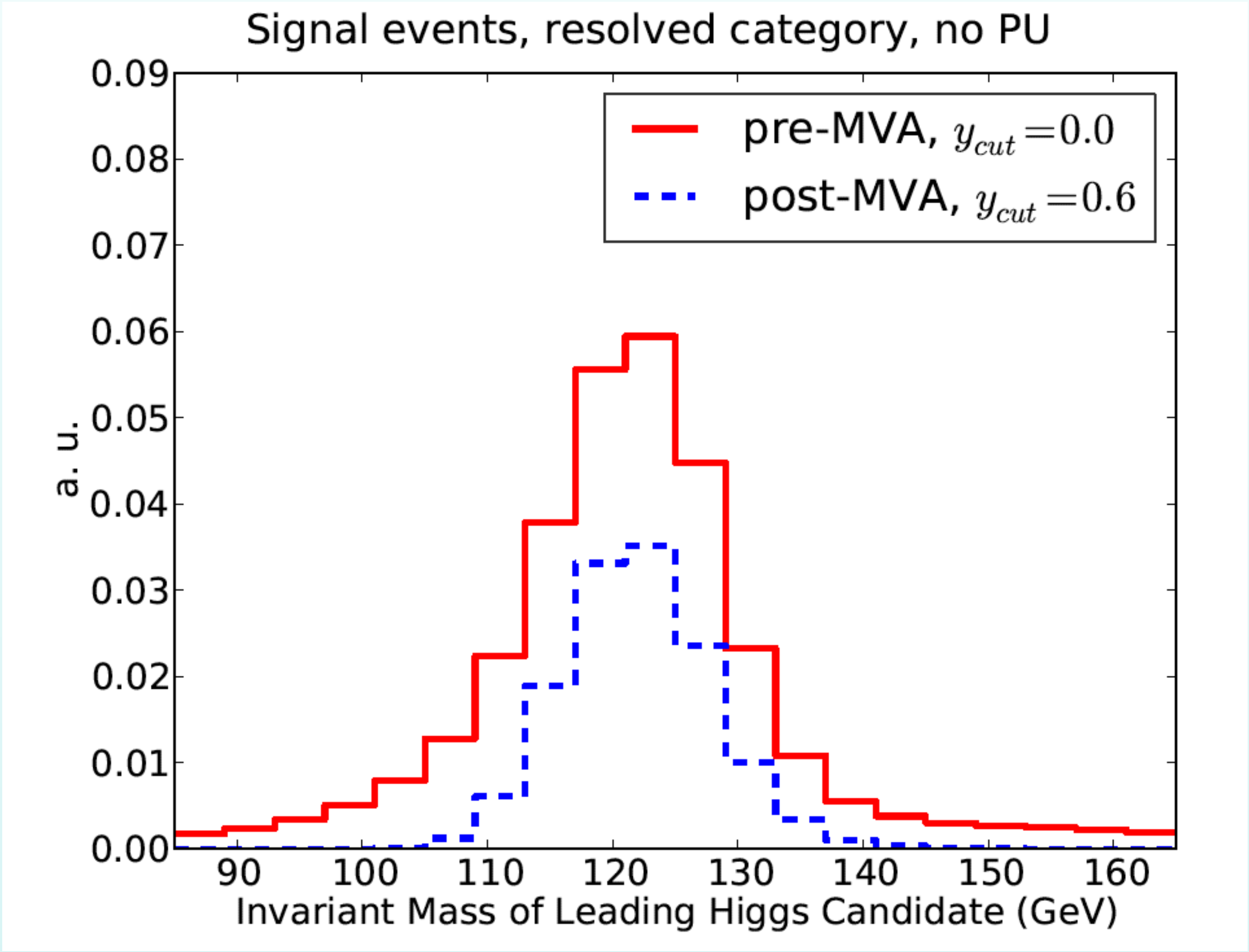
ML techniques allow to **substantially improve the signal significance** for this process **observe Higgs pair production in the 4b final state** at the HL-LHC. Observation (maybe discovery) within reach!

$$\left(\frac{S}{\sqrt{B_{4b}}}\right)_{\rm tot} \simeq 4.7\ (1.5)\,, \quad \mathcal{L} = 3000\ (300)\ {\rm fb}^{-1}$$



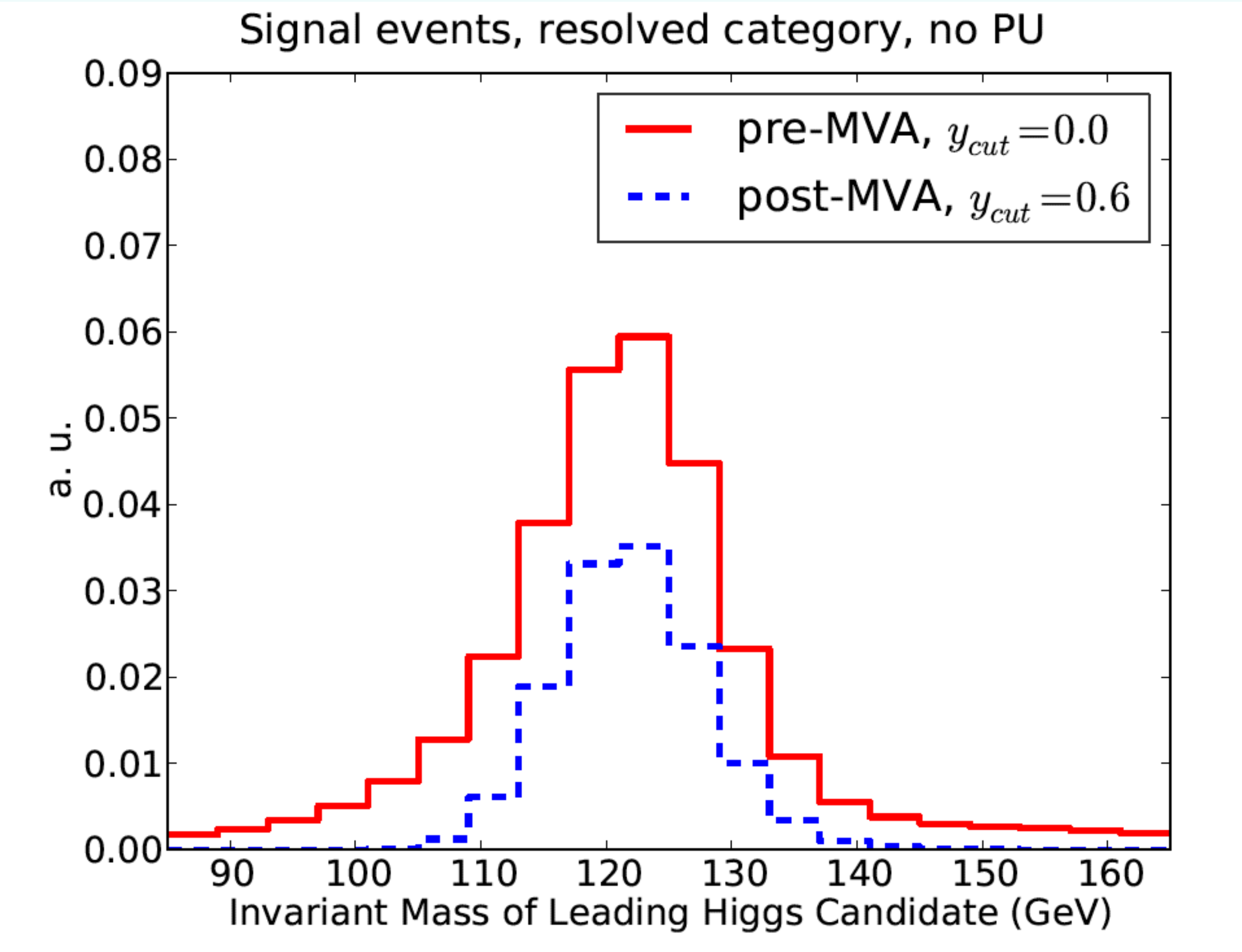
Opening the Black Box

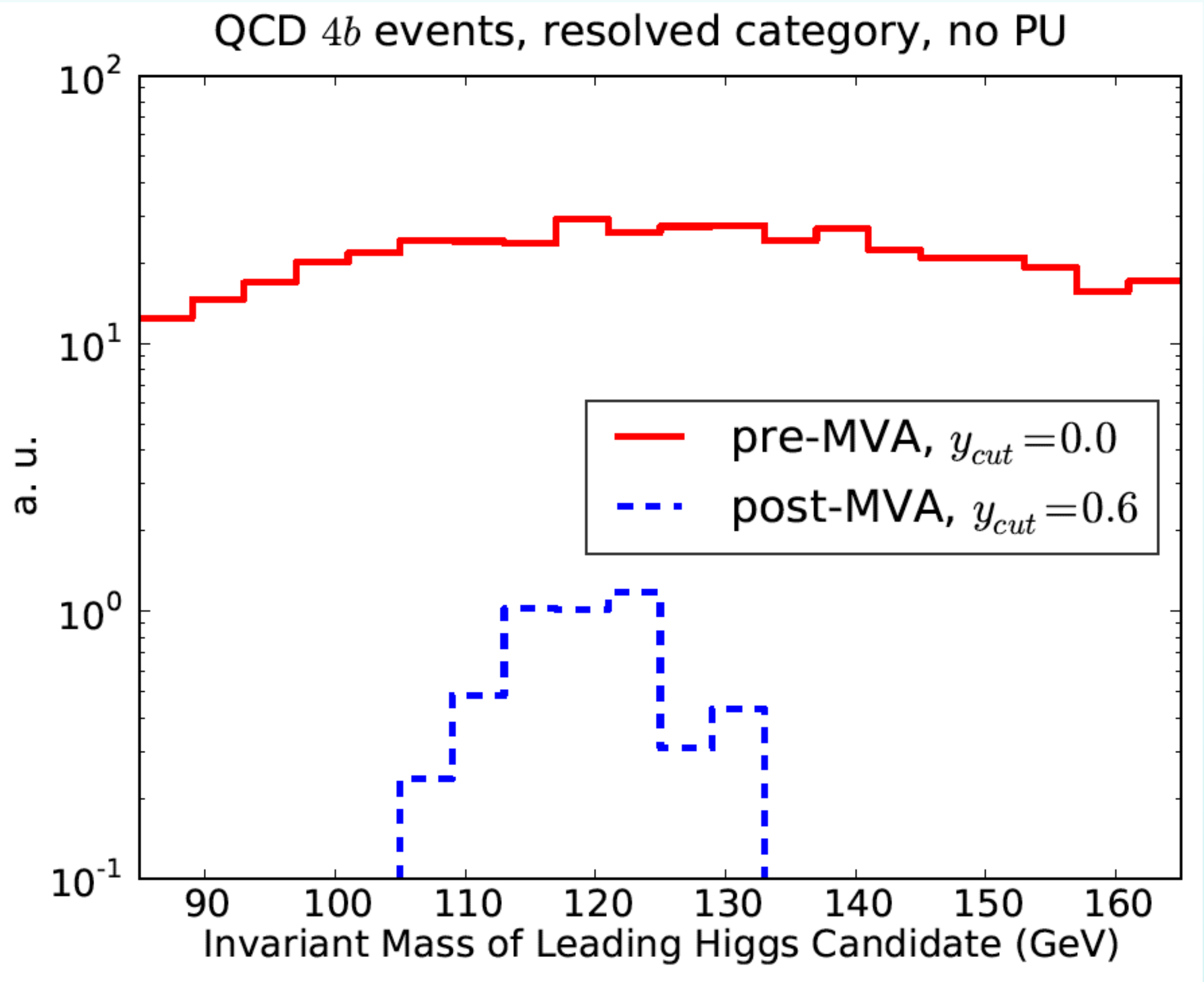
- ANNs are sometimes criticised by being **black boxes**, with little understanding of what happens inside them
- But ANNs are simply a **set of combined kinematical cuts**, nothing mysterious in them!
- Kin distributions **after and before the ANN cut** allow determining the **effective kinematic cuts** being optimised by the MVA, which would allow a cut-based analysis



Opening the Black Box

- ANNs are sometimes criticised by being **black boxes**, with little understanding of what happens inside them
- But ANNs are simply a **set of combined kinematical cuts**, nothing mysterious in them!
- Kin distributions **after and before the ANN cut** allow determining the **effective kinematic cuts** being optimised by the MVA, which would allow a cut-based analysis





The MVA sculpts a Higgs peak in the QCD background!

Juan Rojo

Wrapping up: more cool ML applications!

https://www.youtube.com/watch?v=Bui3DWs02h4

ANNs and LHC phenomenology

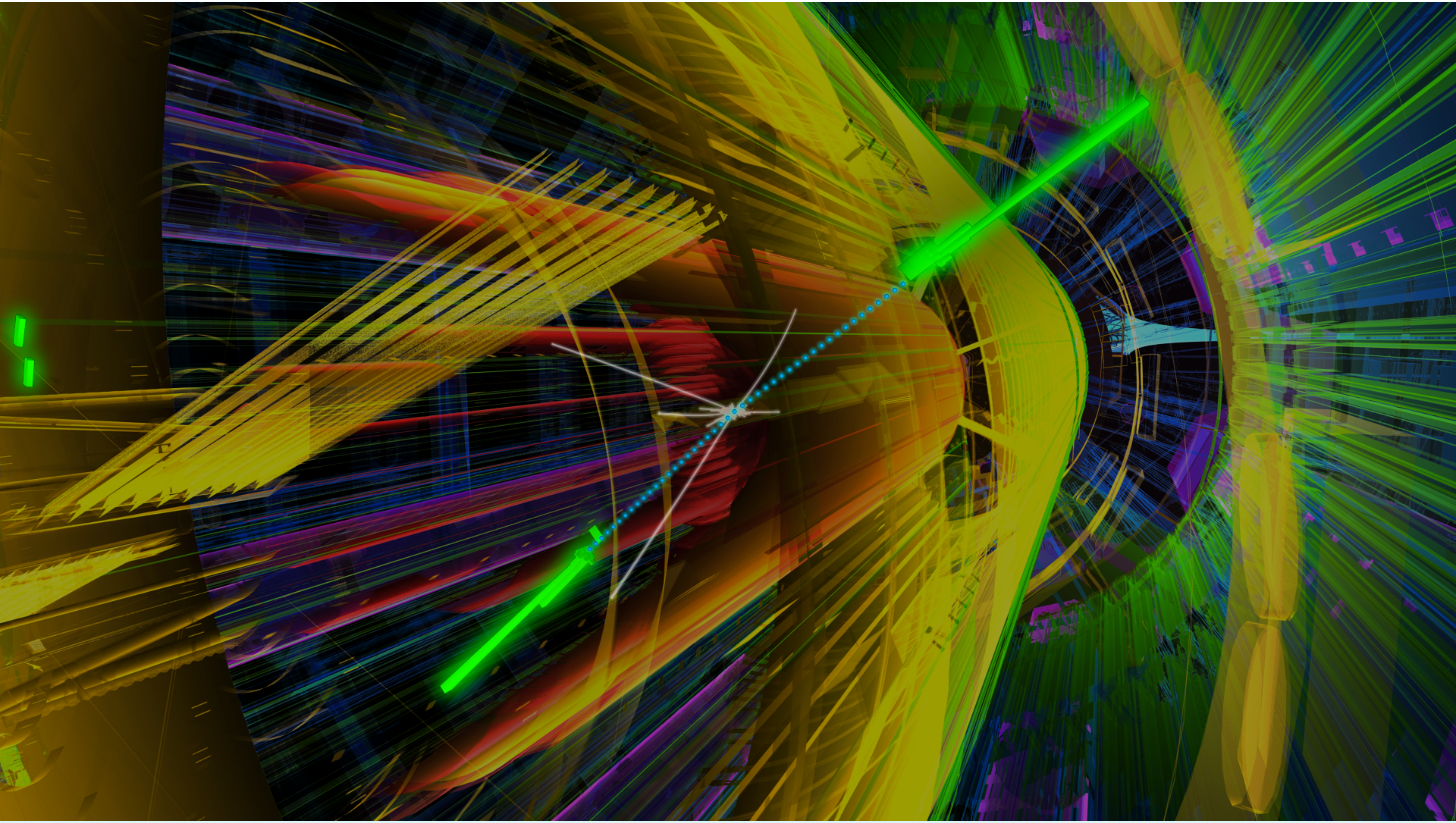
- **Machine Learning algorithms** are already **transforming our world**, from the way we move, shop and heal ourselves, to our understanding of what makes us unique as humans

- In the context of **LHC data analysis and interpretation, ML tools are ubiquitous**, from event selection deep in the detector chain (triggering) to bottom-quark tagging and automated BSM models classification (and exclusion)

- Artificial Neural Networks can be used as **universal unbiased interpolators in global analysis of the proton structure**, with implications from BSM heavy particle production to ultra-high energy neutrino astrophysics

- ANNs can also be used as **classifiers (discriminators) between signal and background** in very busy collision environments, improving LHC physics prospects *i.e.* for **Higgs pair production**

Fascinating times ahead at the high-energy frontier!



And stay tuned for news from the LHC!

Fascinating times ahead at the high-energy frontier!

And stay tuned for news from the LHC!