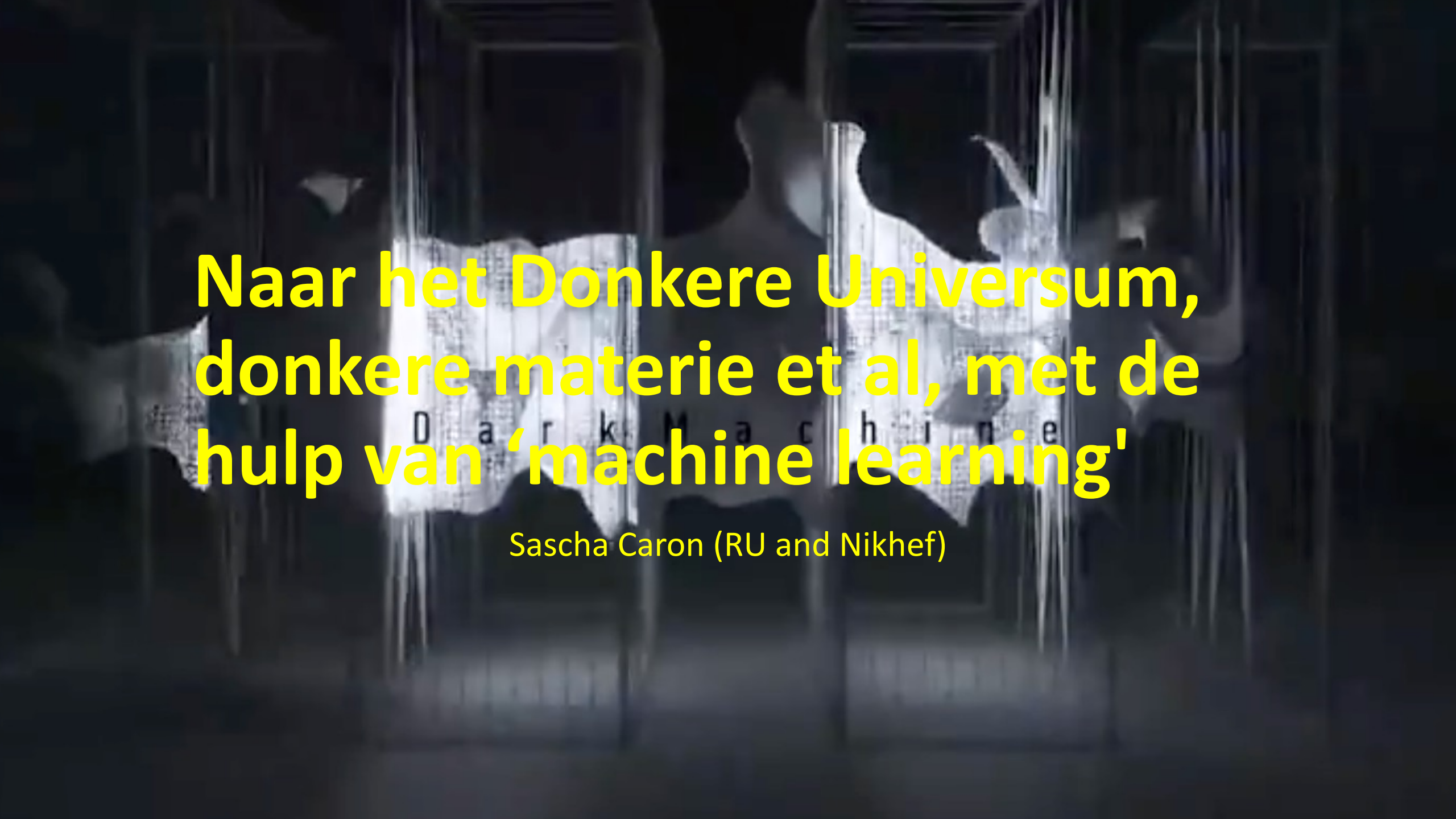


Naar het Donkere Universum, donkere materie et al, met de hulp van ‘machine learning'

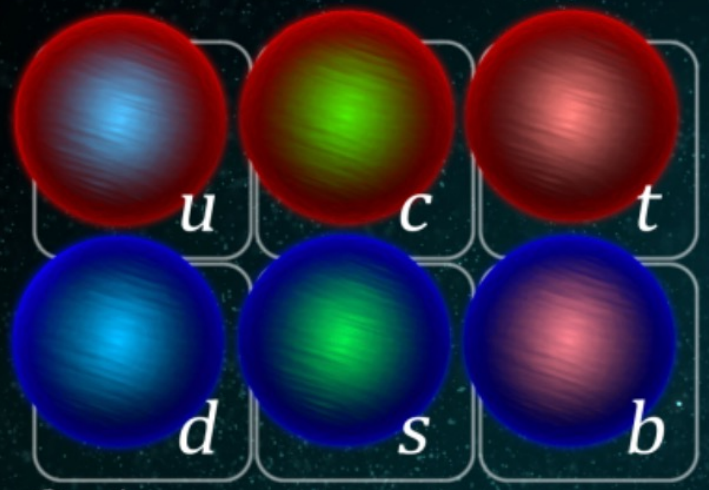
Sascha Caron (RU and Nikhef)

(Fundamenteel) Natuurkunde

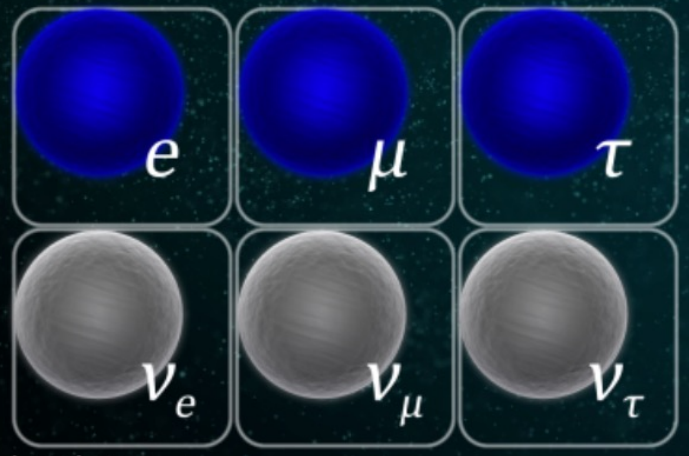
Het doel van de natuurkunde is om te begrijpen hoe de wereld/het universum zich gedraagt.

➔ Situatie in 2022 ?

Standard Model

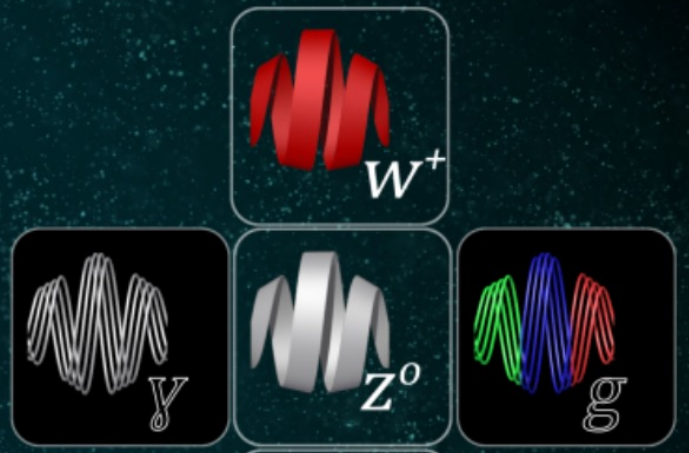


Quarks



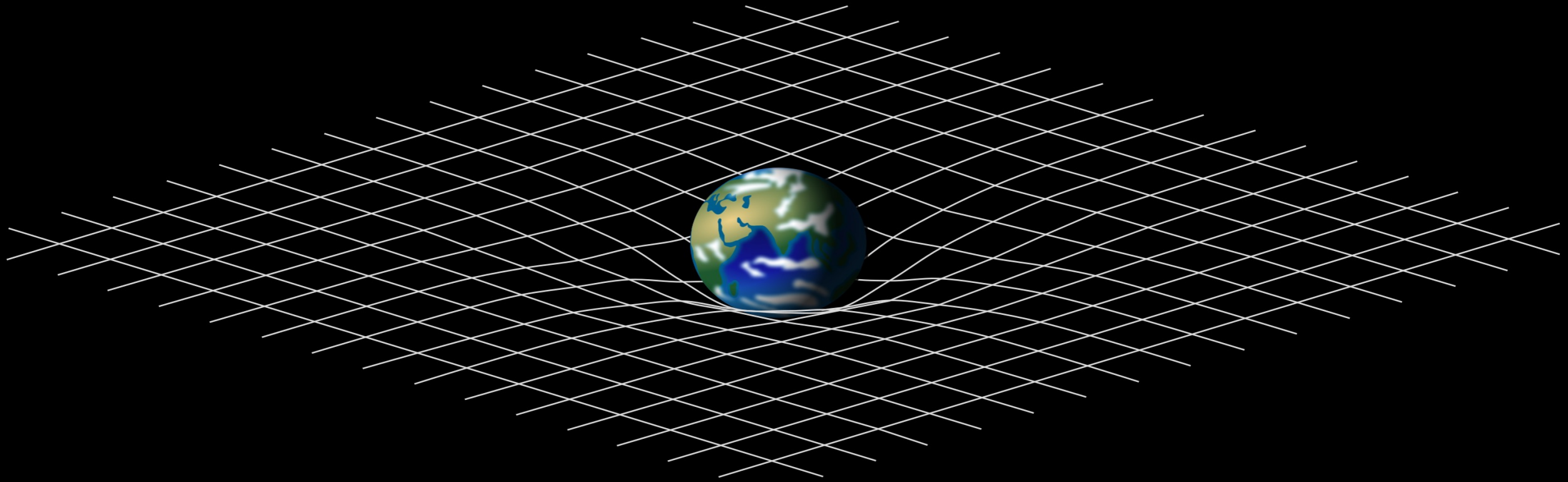
Leptons


Higgs boson

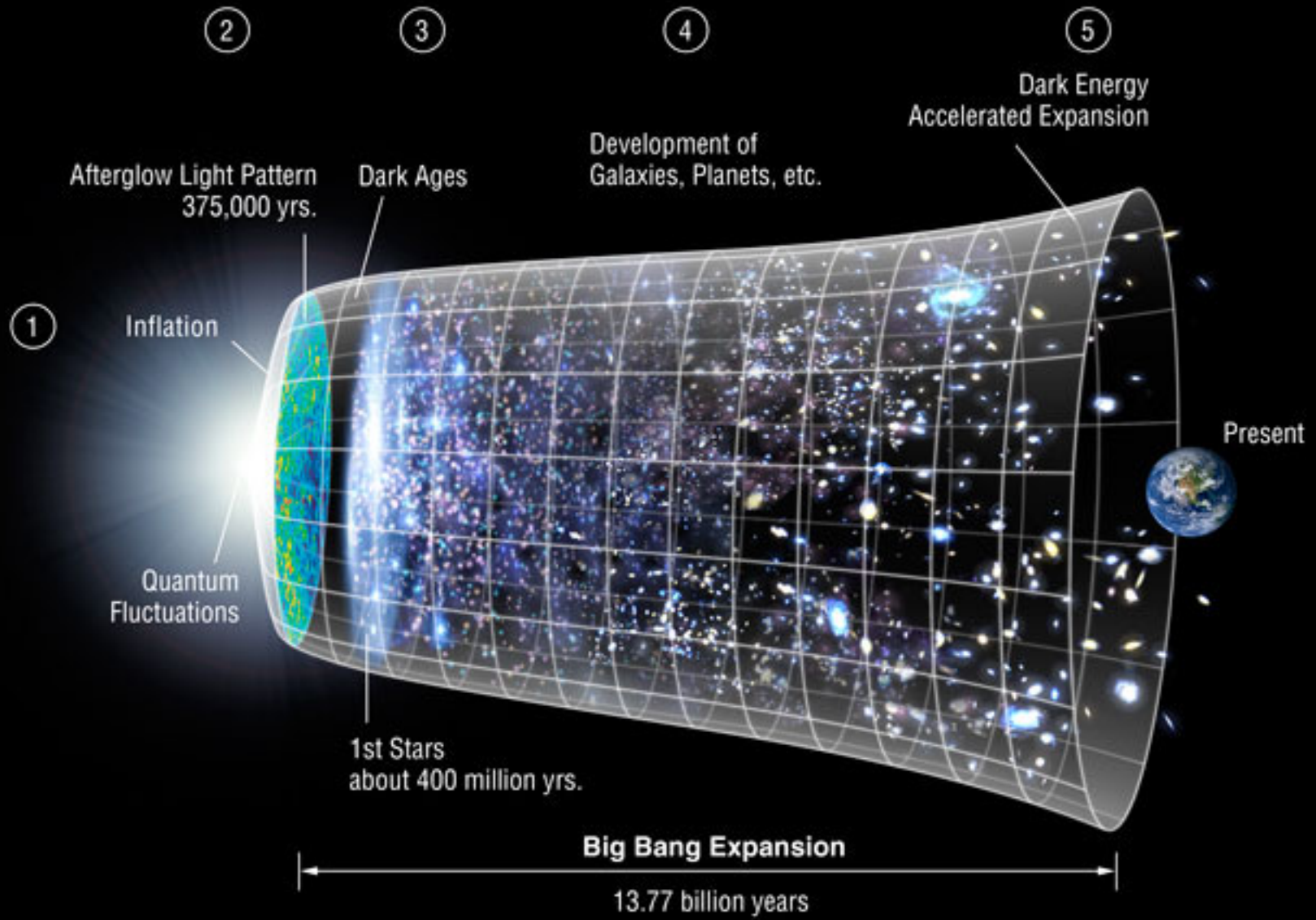


Forces

General Relativity



Cosmological model: Lambda Cold Dark Matter



We know that this is wrong

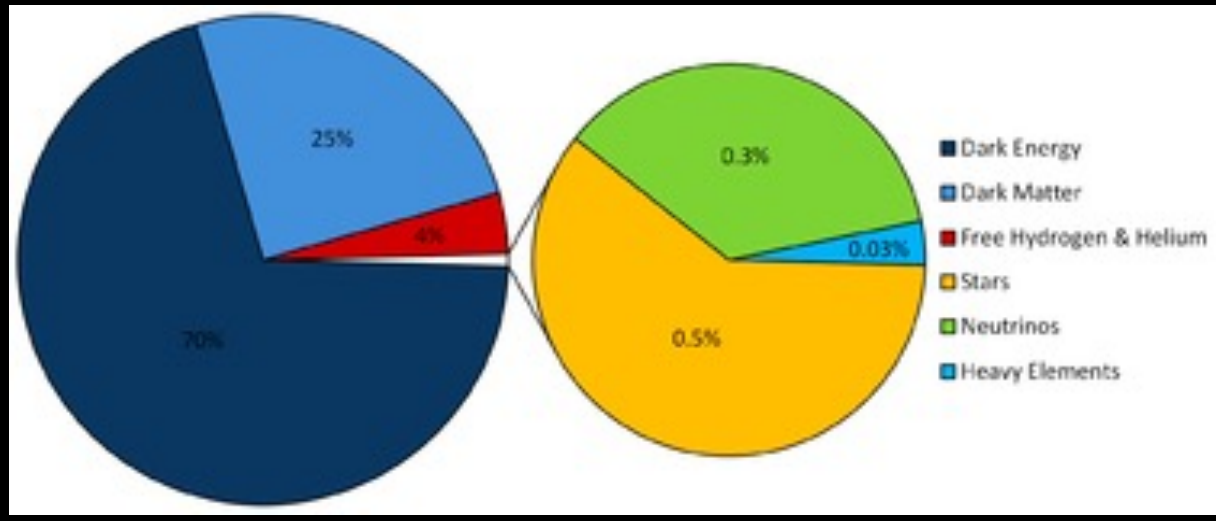
De Big Bang /Standard Model krachten hadden (bijna) gelijke hoeveelheden materie en antimaterie moeten creëren.

Waarom is er veel meer materie dan antimaterie in het universum?

Wat is donkere materie / donkere energie?

Wat doet de Algemene Relativiteitstheorie op kwantumniveau?

…



Natuurkunde

Het doel van de natuurkunde is om te begrijpen hoe de wereld/het universum zich gedraagt.

→Hoe kunnen we dit inzicht krijgen?

→Wij bouwen een ... nieuw/beter .. (wiskundig?) model van het heelal
→Onze taak: We willen het model bouwen dat gegevens/data beschrijft

Wat we meestal doen:

We doen voorspellingen met **ons/een model** en testen de voorspelling met gegevens?

Scientific discovery: The 5th paradigm ?

- **Eerste paradigma : Empirisch / Observatie**
- **2de paradigma : Theoretische modellen (analytisch oplosbaar?)**

Scientific discovery: The 5th paradigm ?

- **Eerste paradigma : Empirisch / Observatie**
- **2de paradigma : Theoretische modellen (analytisch oplosbaar?)**
- **3e paradigma (jaren 1970): Simulatie / Numerieke berekening**
-

Scientific discovery: The 5th paradigm ?

- **Eerste paradigma : Empirisch / Observatie**
- **2de paradigma : Theoretische modellen (analytisch oplosbaar?)**
- **3e paradigma (jaren 1970): Simulatie / Numerieke berekening**
- **2018 Het vierde paradigma: data-intensieve wetenschappelijke ontdekking (Jim Gray), data en machine learning ➔ maar ook hype**
- https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery

Scientific discovery: The 5th paradigm ?

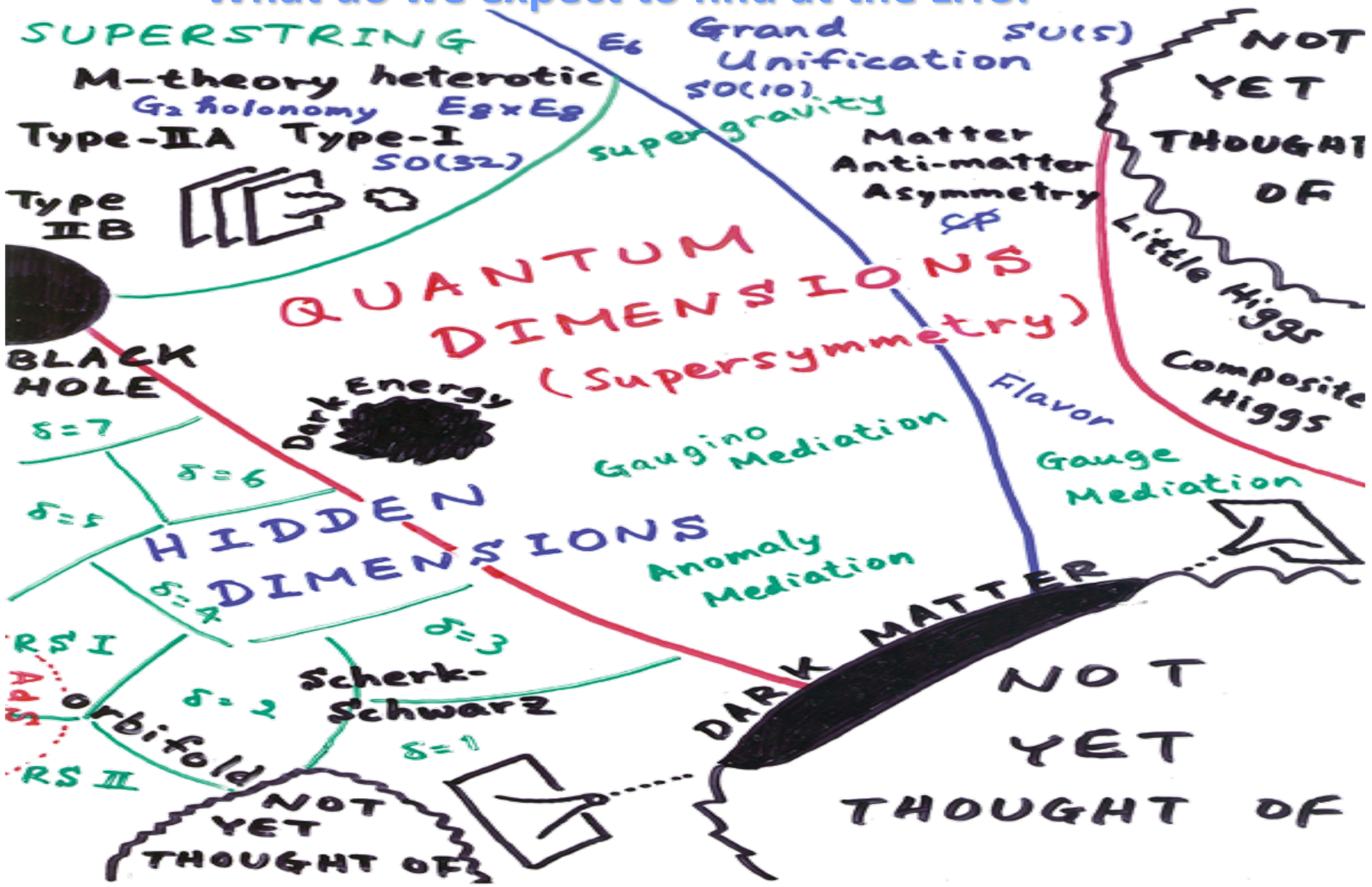
- **Eerste paradigma : Empirisch / Observatie**
- **2de paradigma : Theoretische modellen (analytisch oplosbaar?)**
- **3e paradigma (jaren 1970): Simulatie / Numerieke berekening**
- **2018 Het vierde paradigma: data-intensieve wetenschappelijke ontdekking (Jim Gray), data en machine learning ➔ maar ook hype**
- https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery
- **Juni 2022: AI4Science om het vijfde paradigma van wetenschappelijke ontdekking te versterken (Christopher Bishop, QFT-proefschrift), AI getraind op wetenschappelijke simulatoren (machine learning, kwantumfysica, computationele chemie, moleculaire biologie, vloeistofdynamica, software engineering en andere disciplines)**
- fth-paradigm-of-scientific-discovery/

Hoe de SM, Gravity, LambdaCDM uit te breiden?

Laten we eens kijken naar het Beyond the Standard Model landschap ...
(zwaartekrachtlandschap lijkt op elkaar...)

The situation in 2006

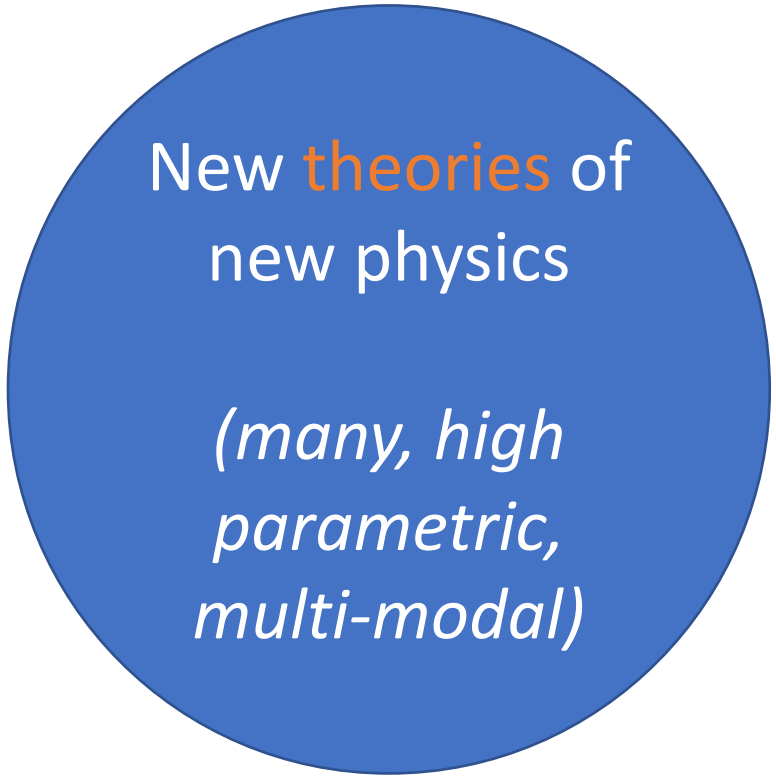
What do we expect to find at the LHC?

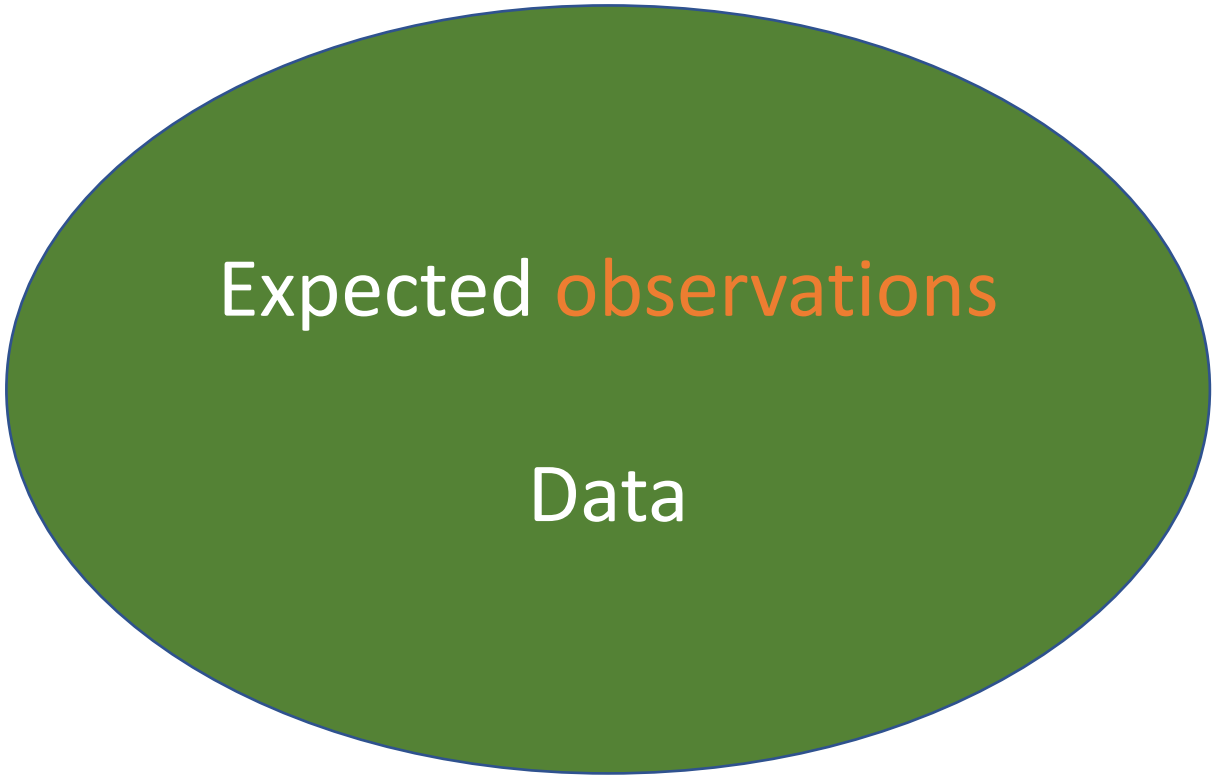


One physicist's schematic view of particle physics in the 21st century
(Courtesy of Hitoshi Murayama)

Each of those models have many parameters

Tegenwoordig zogenaamde "effectieve veldtheorieën" en-vogue

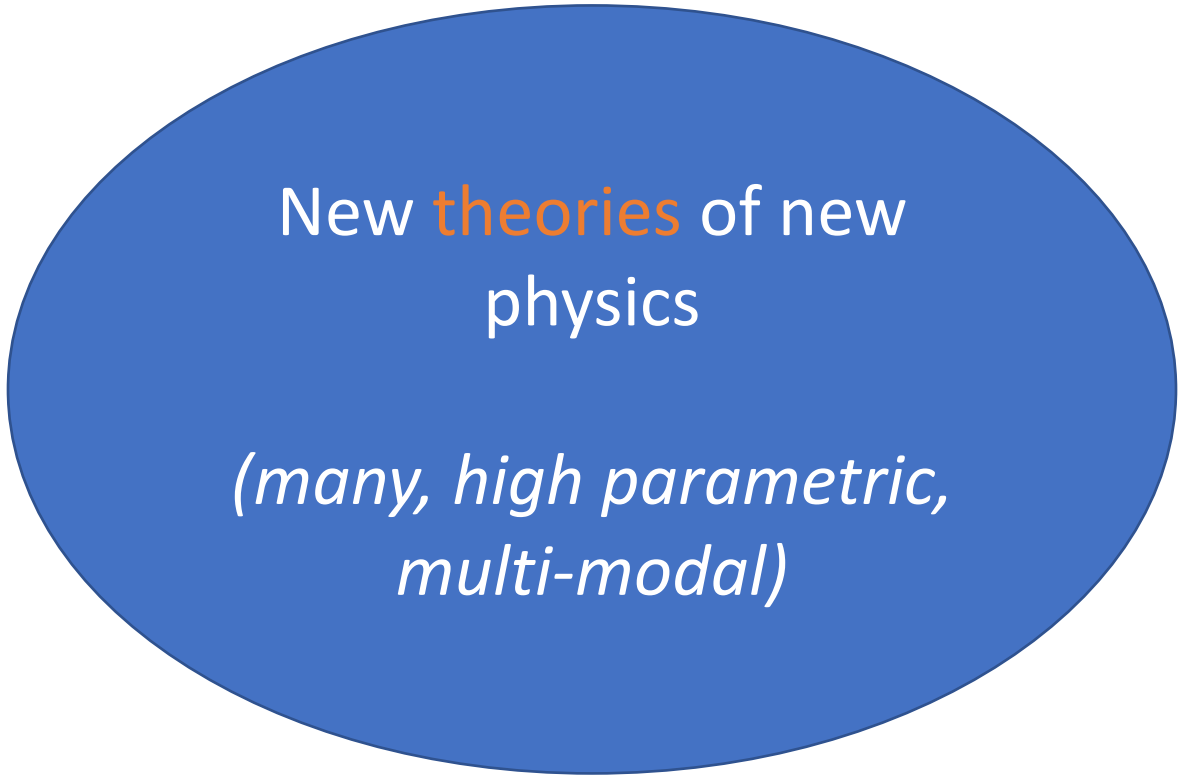




Each of those models have many parameters

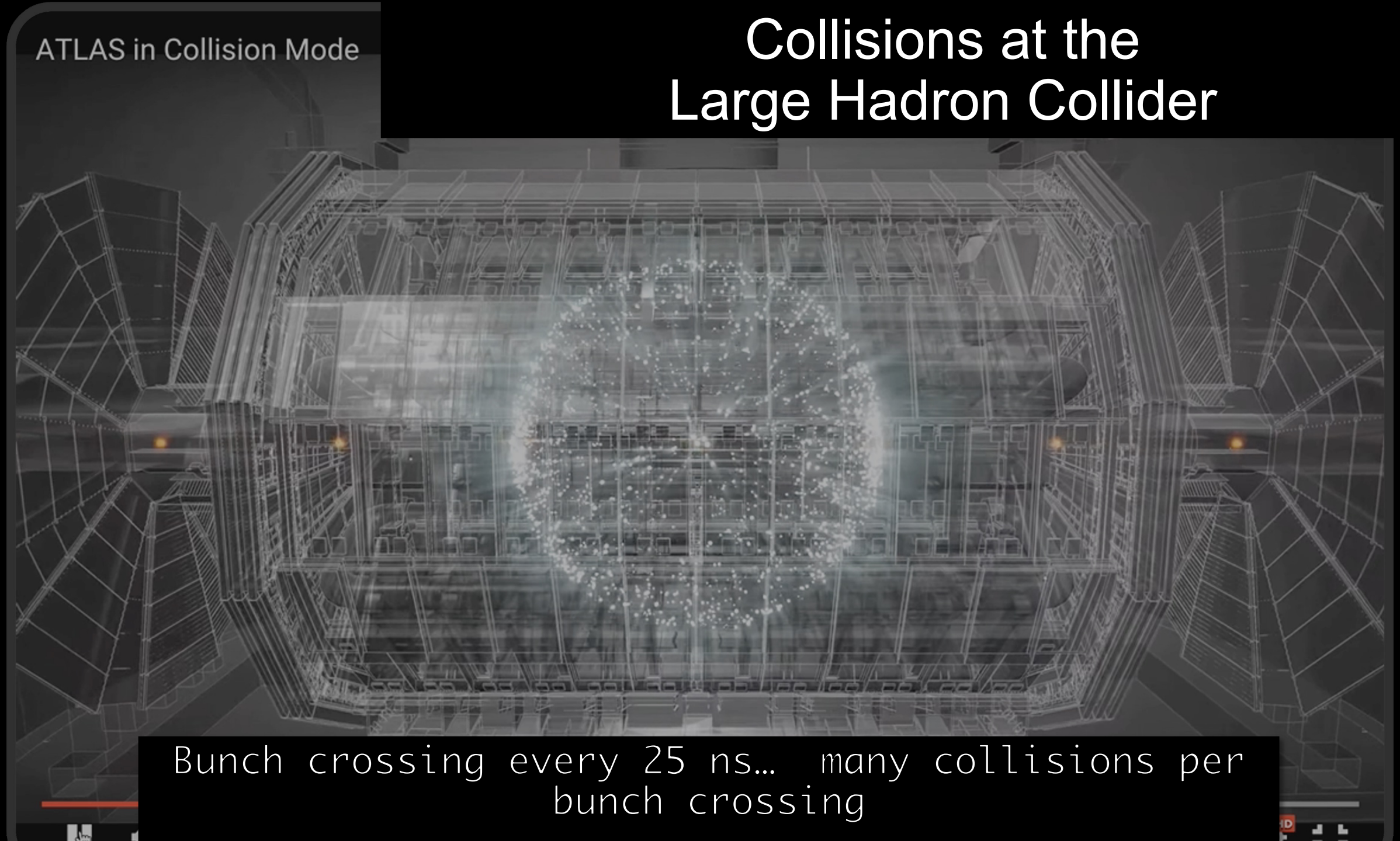
Tegenwoordig zogenaamde "effectieve veldtheorieën" en-vogue

→ 6000 vrije parameters....



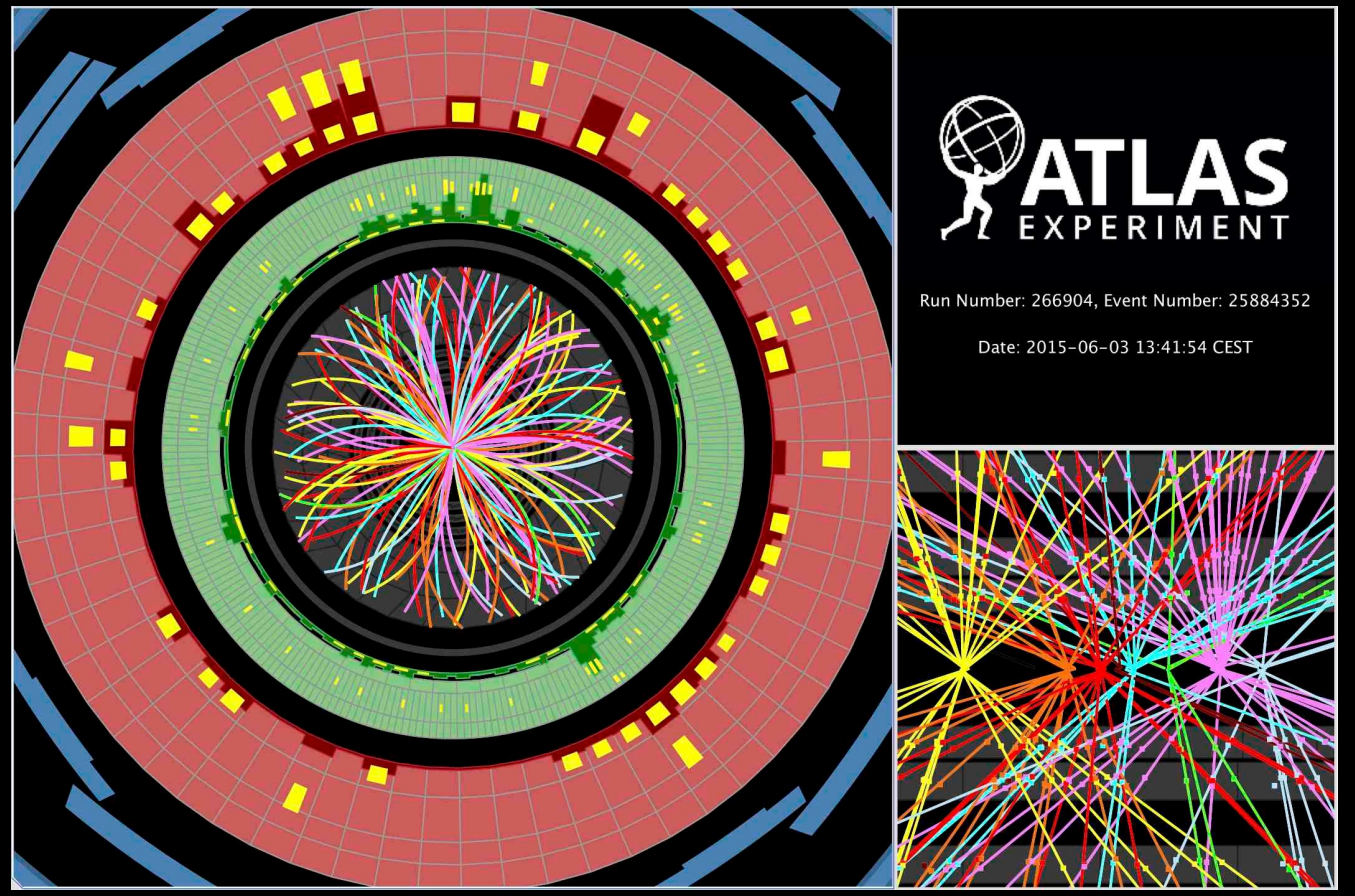
We kennen onze "verwachting" !!!
(Standaard model, GR, etc.)

Collisions at the Large Hadron Collider



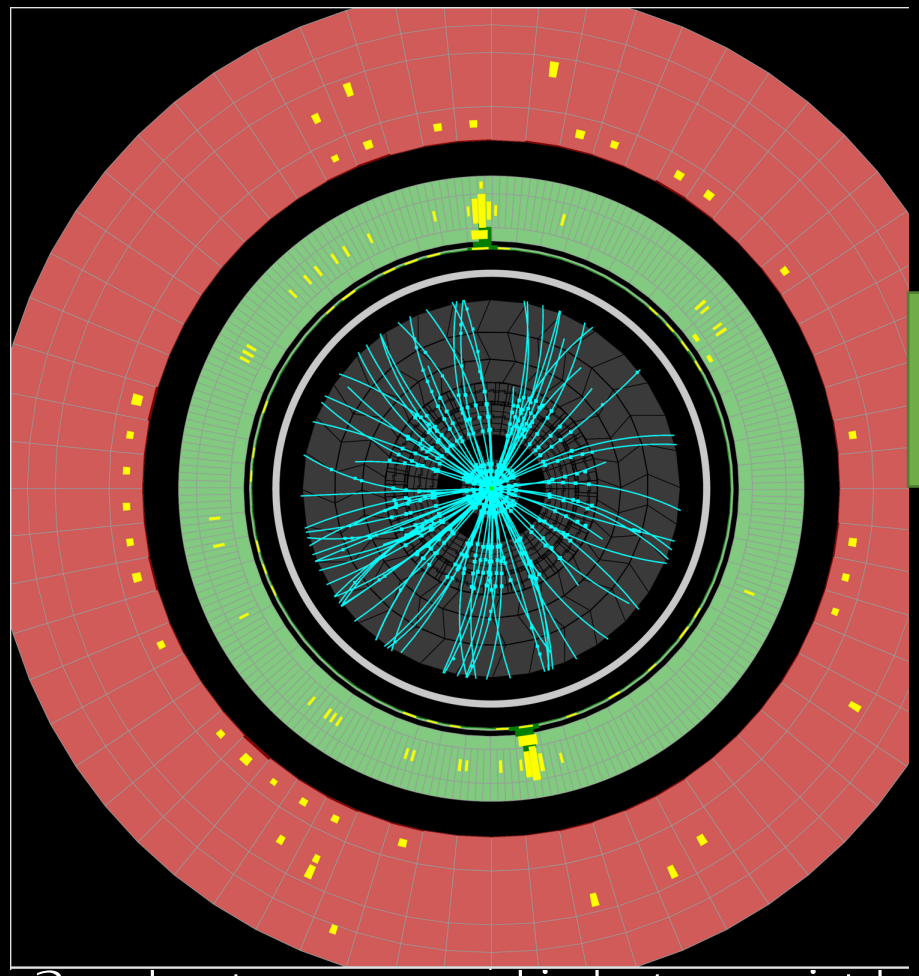
Bunch crossing every 25 ns… many collisions per bunch crossing

Most events look like this…



Event from LHC run-2

1 in >1000 billion events looks like this

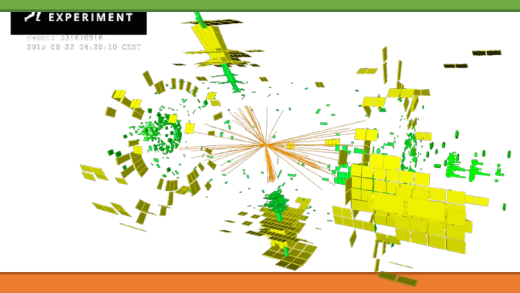


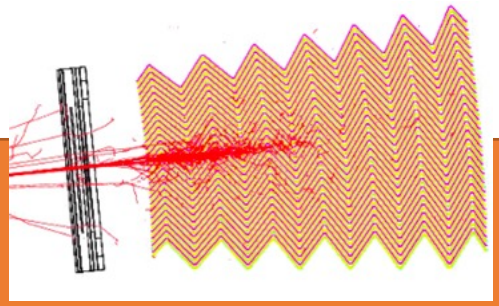
Mass of the Higgs is reconstructed with photon energies

Higgs to 2 photon candidate with mass of 125 GeV

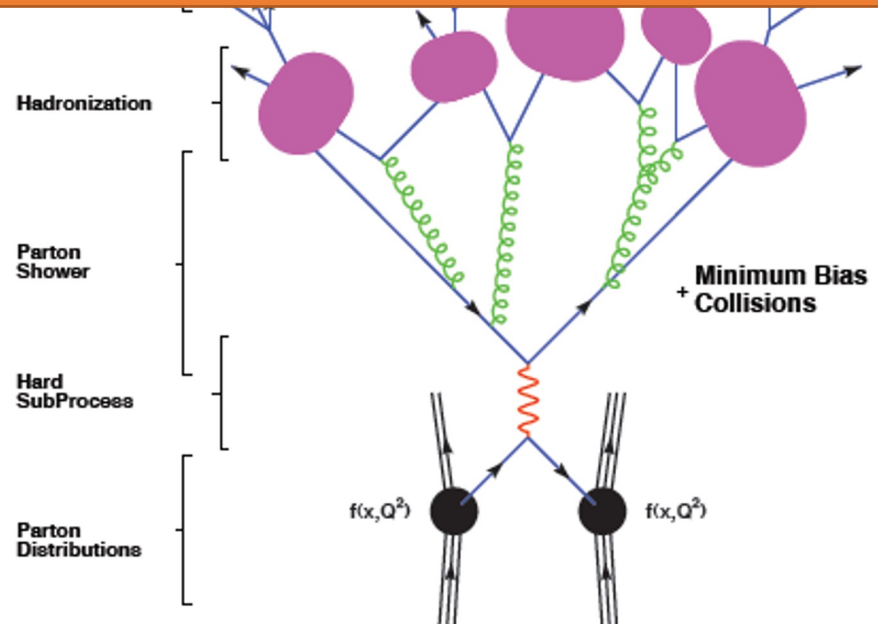
Simulation: Traditional

Energie en hoeken van gereconstrueerde deeltjes





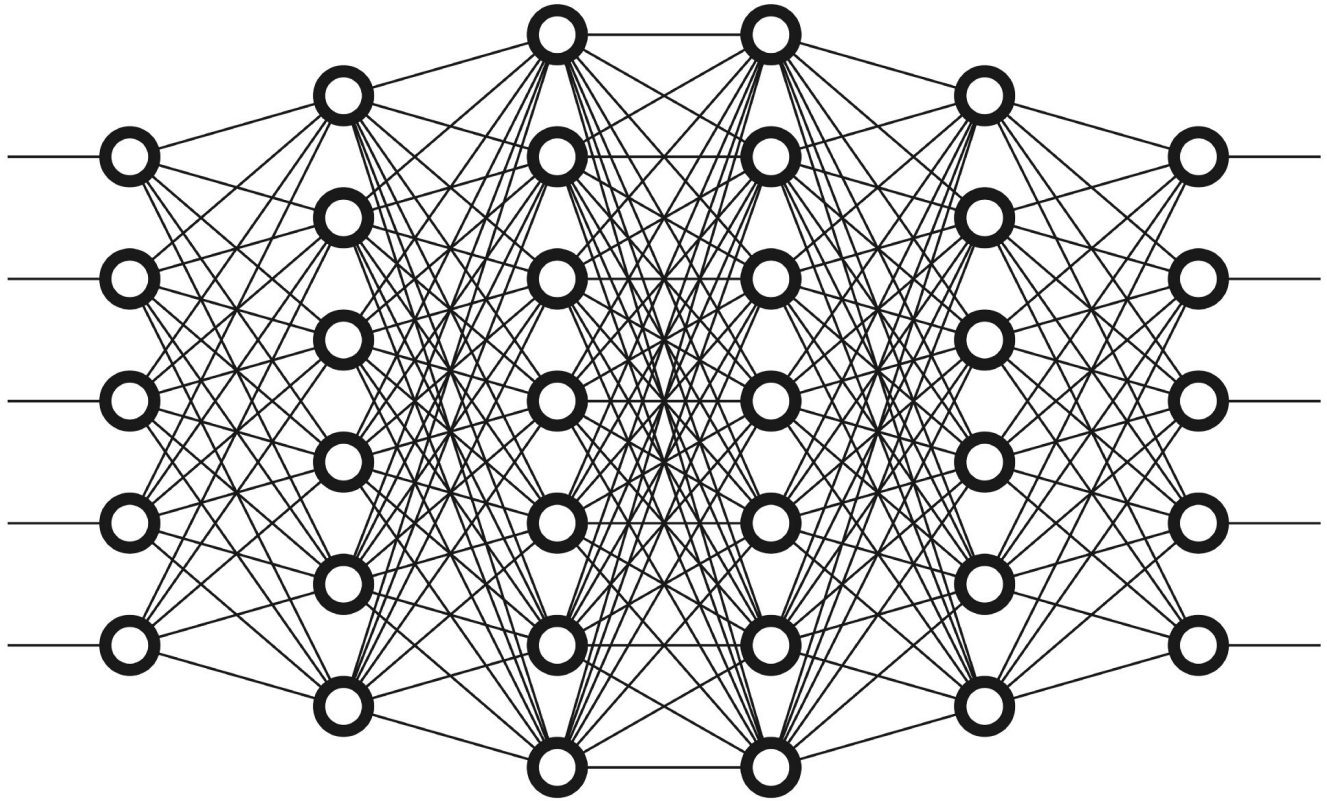
Detector Simulator



Invoer:
„random“ getallen

Simulation: Us

Energie en hoeken van gereconstrueerde deeltjes



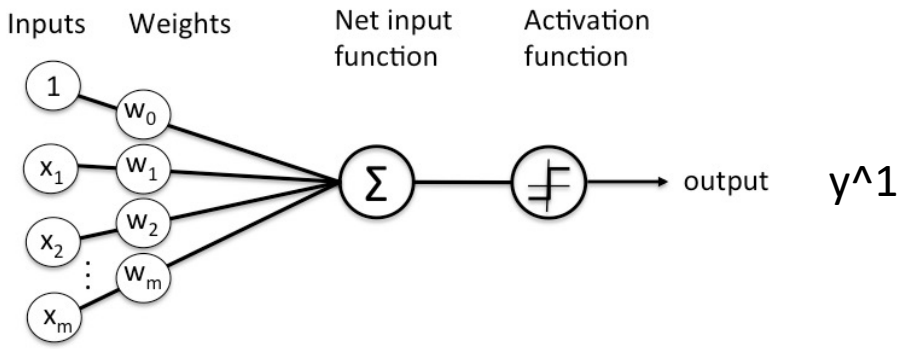
Invoer:
„random“ Willekeurige

Machine Learning on a few slides

The perceptron

Put (my) ZOOM pictures in here

$$y(\vec{x}, \mathbf{w}) = f(\sum_i w_i x_i)$$

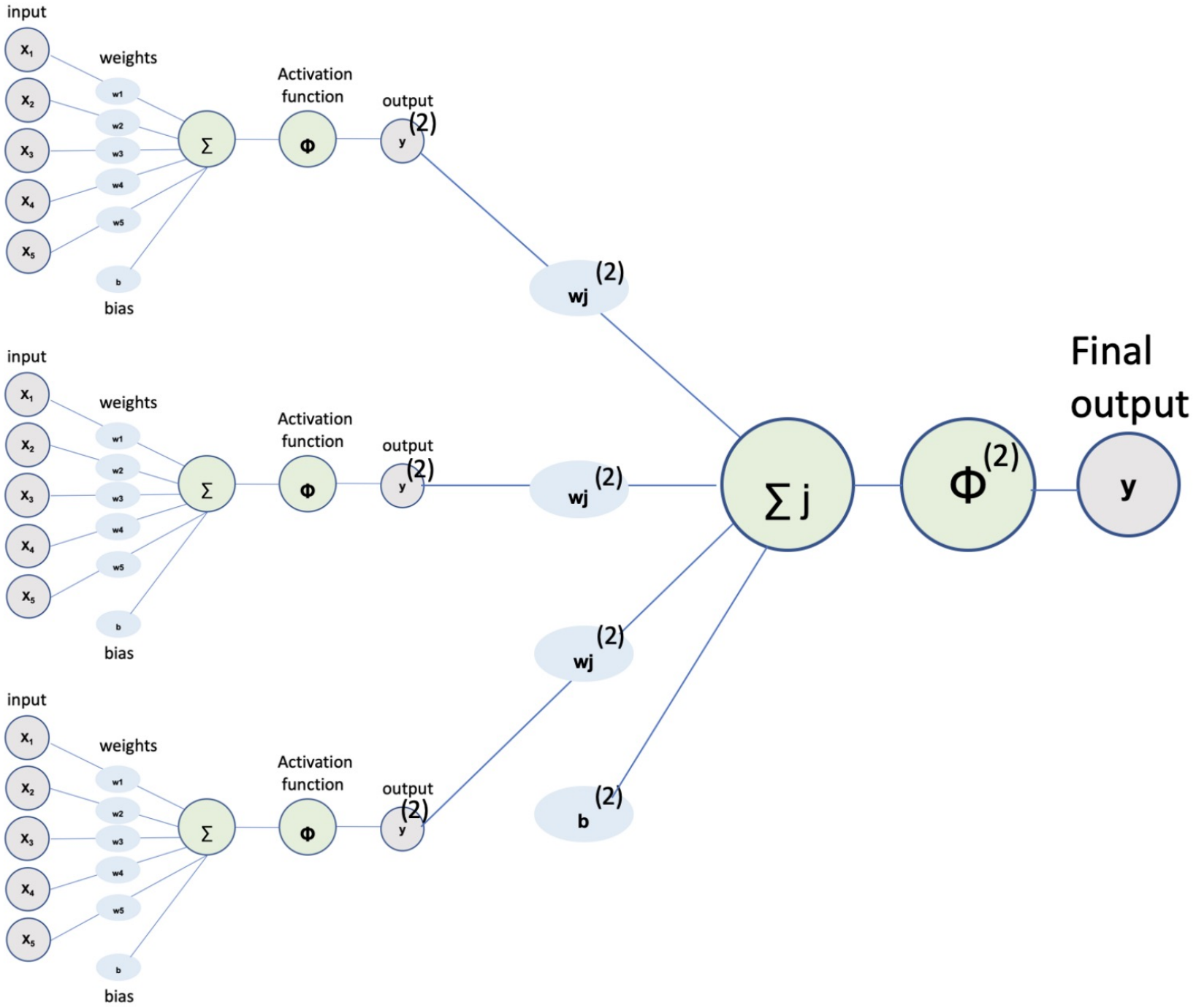


y^1

and f is a non-linear “activation function”
X is the input (numbers/pictures)
Y is the output

A multilayer perceptron

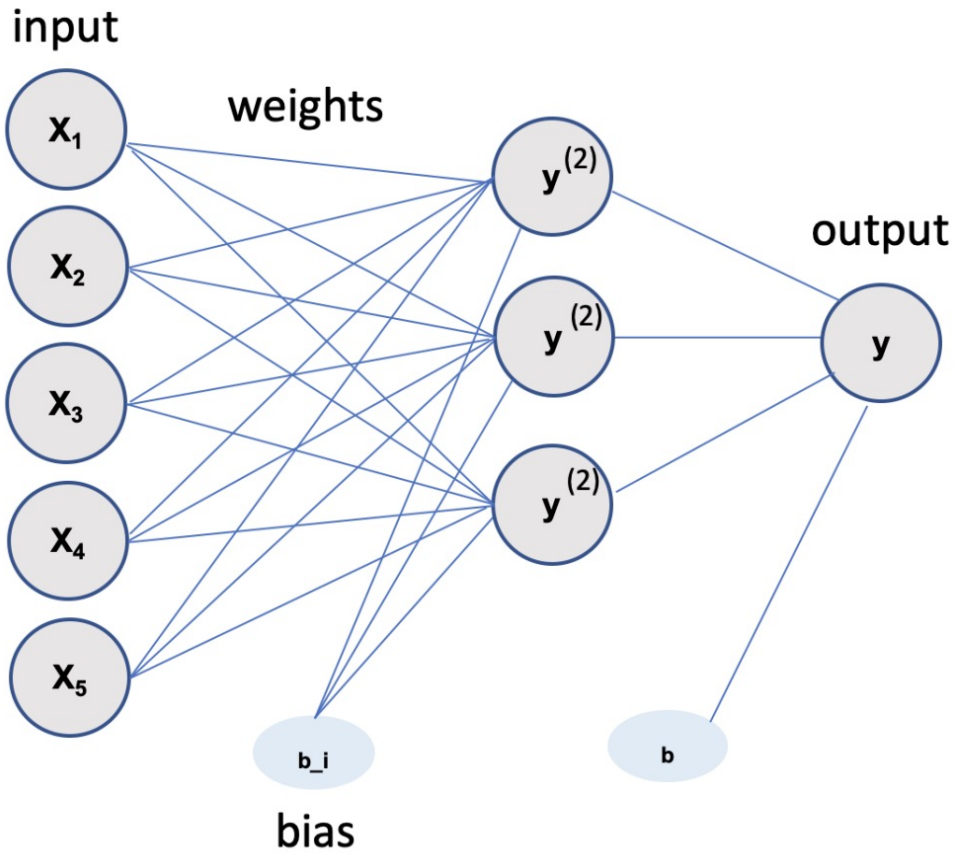
- Gewoon veel perceptrons op een rij en verbonden…



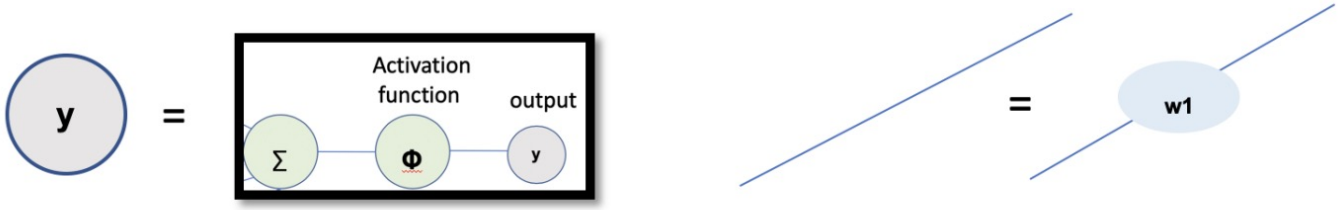
Put (my) ZOOM pictures in here

A simpler way of graphing this

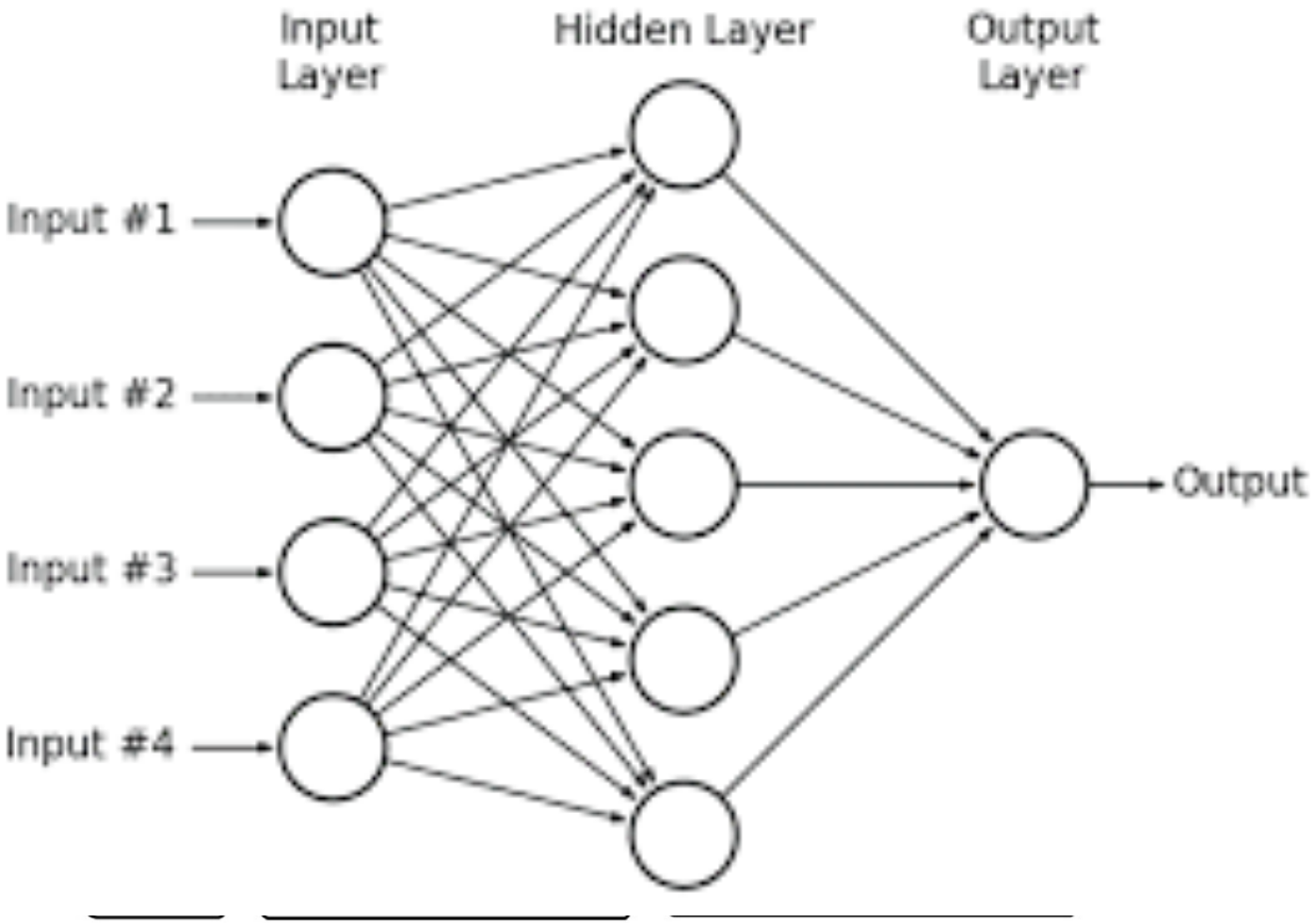
Put (my) ZOOM pictures in here



Put (my) ZOOM pictures in here



Put (my) ZOOM pictures in here



“Training” a Neural Network

Bepaal het doel, bijvoorbeeld de output y moet katten en honden classificeren

(y dicht bij 0 hond, y dicht bij 1 kat)

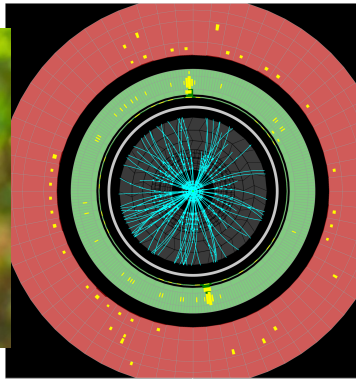
Bepaal de gewichten van dit "circuit / netwerk" dat deze uitvoer correct geeft voor de *meeste gegevens* ...

→ Veel gegevens nodig…

How kunnen wij met dit nieuwe fysica zoeken?

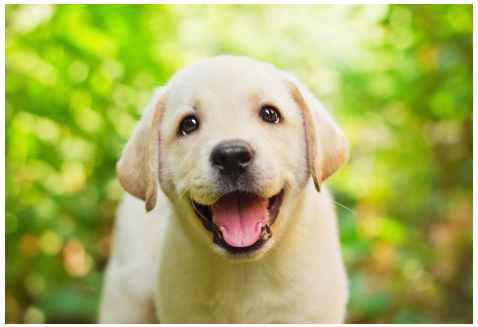
De bekende fysica (dieren) zijn katten….

Traditional approach “Model driven



1. Kies een model van nieuwe fysica (**honden** !)
2. Vereenvoudigen
3. Kies een waarschijnlijke (?) set parameters
4. Doe een voorspelling
5. Train ML classifier om de voorspelling te testen (bijv. signaal versus achtergrondclassificatie)
6. Hypothesetest met data|old model vs data|nieuw model op classifier output
7. Het modelparameterpunt uitsluiten ?
8. Ga naar 3 of 1

- Beste aanpak als het model + parameterset waar is
- => Voorspel het "juiste signaal"
- ➔ Werkte voor het Higgs deeltje !



- Beste aanpak als het model + parameterset waar is
- □ Voorspel het "juiste signaal"
- Slechte aanpak als het model + parameterset verkeerd is. Hoe erg ?

Idee: Breid deze aanpak uit

Wat kunnen we veranderen / verbeteren?

Higgs: Je hebt een nieuw speeltje, het is een playmobil kasteel met een grootte tussen de 1-100 cm. Kunt u het vinden ?

Vandaag: ik heb een nieuw speeltje voor je, ik zet het ergens in je kamer. De grootte is 0,1-100 cm. Kunt u het vinden ?

Idea: **Breid eenvoudige nieuwe op fysica gebaseerde modelgebaseerde aanpak uit**

What can we change / improve ?

Nog 3 routebeschrijvingen gevonden (zijn er meer?):

- Zoek systematisch in alle gegevens naar nieuwe fysica (brute kracht)
- Hyperklasse augmentatie: Train een ML-classificator op veel modellen van nieuwe fysica
- Anomaliedetectie: Train ML-classificator alleen op bekende natuurkunde

Brute force

- Het brute force algoritme probeert alle (vele) mogelijkheden uit totdat er geen oplossing wordt gevonden.
-

A strategy for a general search for new phenomena using data-derived signal regions and its application within the ATLAS experiment

Goal:

Strategy paper. Generalize previous attempts.

Define a “meta-algorithm“ for

automated / generic / unsupervisedLHC searches

Show with 2015 data that this is - in principle – possible

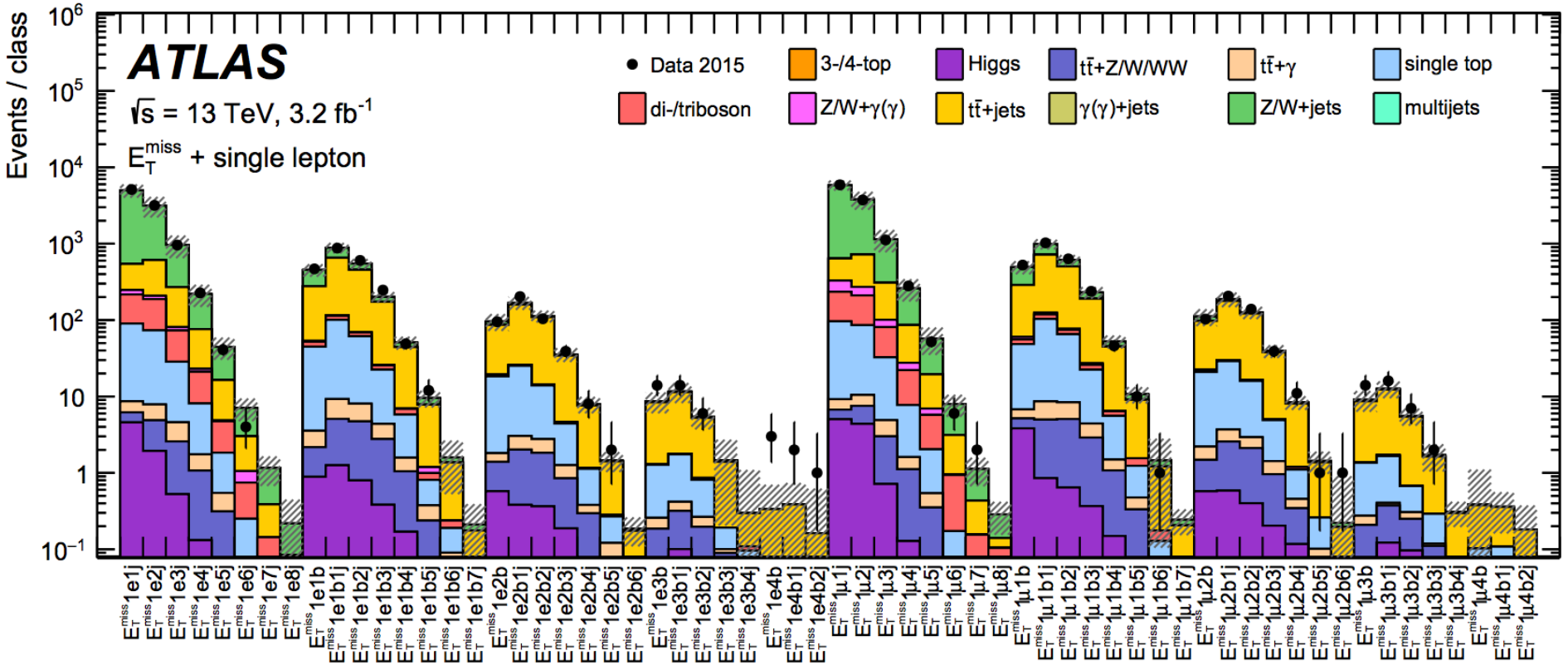
https://arxiv.org/pdf/1807.07447.pdf

A strategy for a general search for new phenomena using data-derived signal regions and its application within the ATLAS experiment

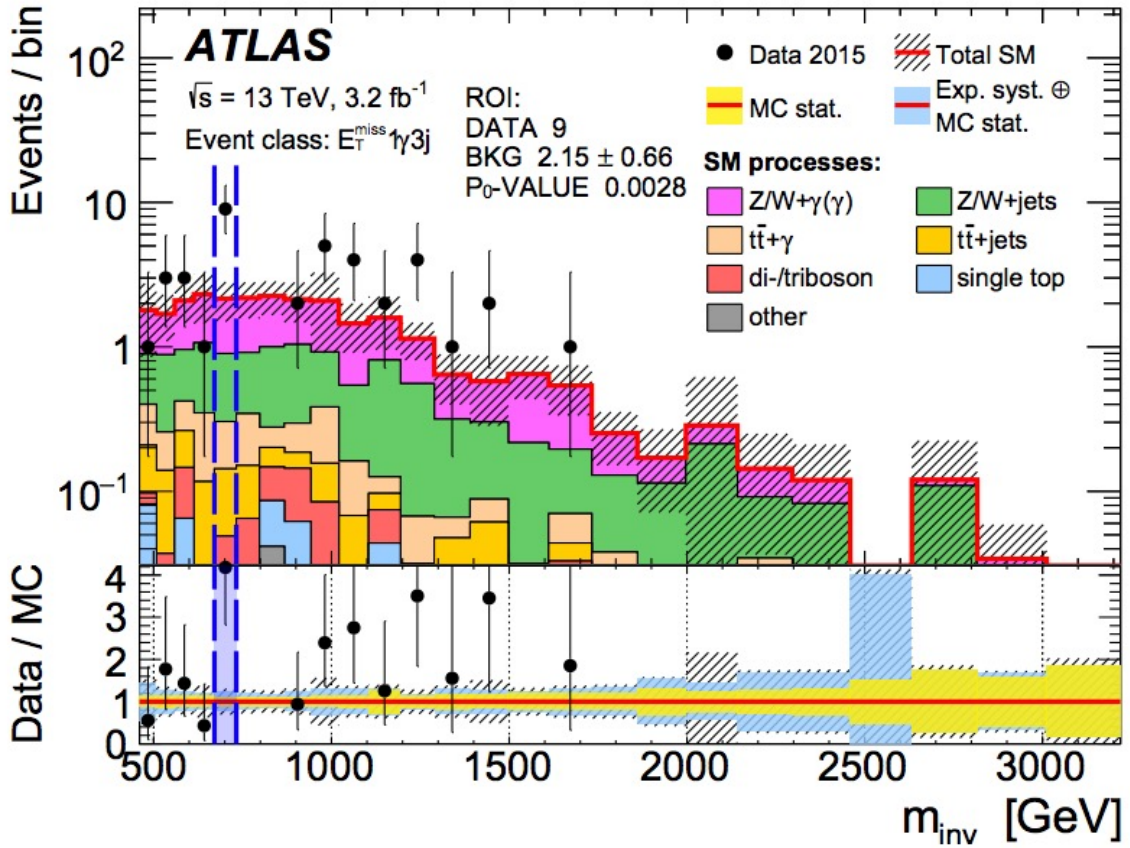
Define a 2-step approach:
First put available resources on generality
Then use available resources to test most interesting deviations…

1. General Search: Automatically testing a large set of signal regions
Observation of one or more significant deviations in some phase-space region(s)
➔ Trigger to perform dedicated and model-dependent analyses
where these **'data-derived' phase-space region(s) can be used as signal regions**

In ATLAS > 800 channels !
> 10^5 signal regions/hypothesis tests !



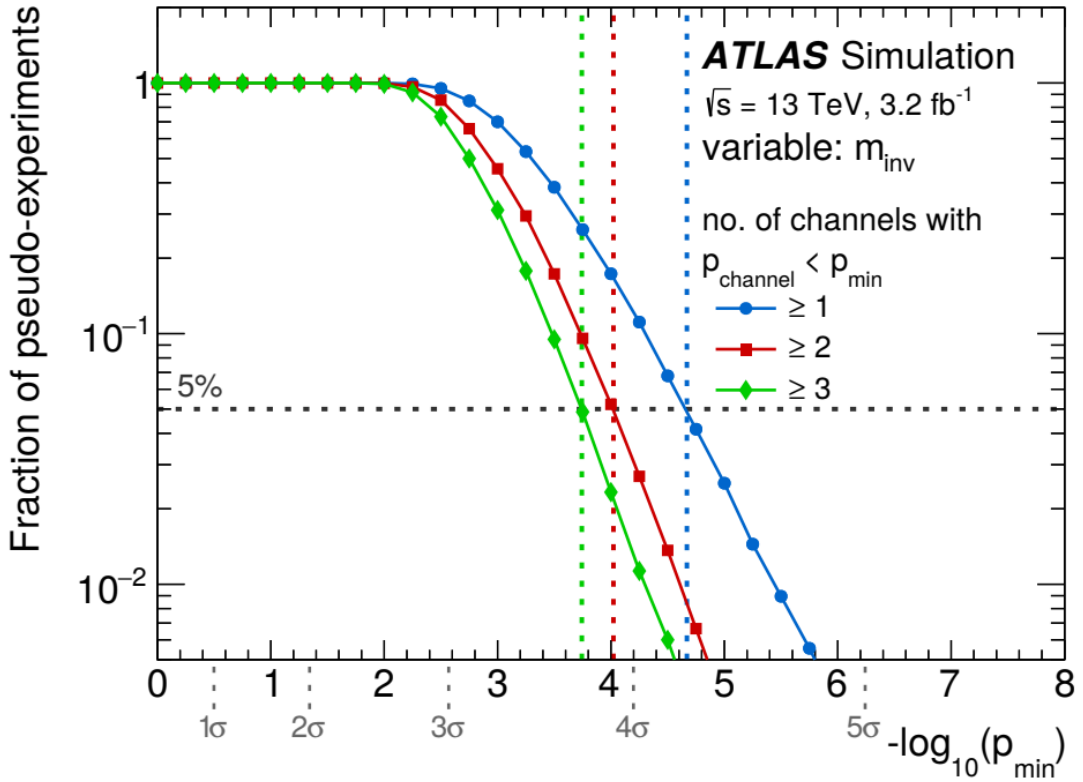
> 800 channels …. (plot shows a small selection)



> 30000 regions (hypothesis tests)

Determine p-value thresholds by asking how many toy datasets would give such a deviation

→ *A regions is interesting if you find channels with p-values more significant than in 95% of the toys*



A strategy for a general search for new phenomena using data-derived signal regions and its application within the ATLAS experiment

Define a 2-step approach:
First put available resources on generality
Then use available resources to test most interesting deviations…

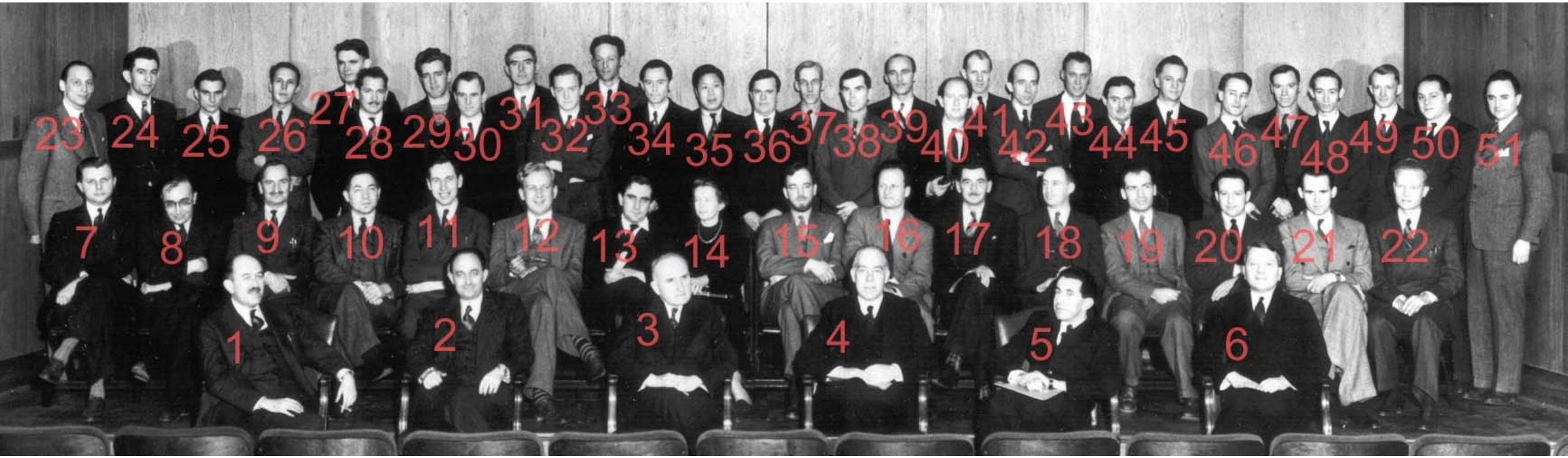
1. General Search: Automatically testing a large set of signal regions
Observation of one or more significant deviations in some phase-space region(s)
➔ Trigger to perform dedicated and model-dependent analyses
where these **'data-derived' phase-space region(s) can be used as signal regions**
2. Dedicated Search
- "Wave function collapsed" to test most interesting deviations with available resources
- On 2nd dataset (➔ Statistically independent, unbiased p-value !!)


Hyperclass of models

Zoeken via “Hyperclass: Mixture of theories”

Stel dat het model/de parameterset niet de juiste is, maar enige kennis bevat over het nieuwe fenomeen dat we in de gegevens verwachten.

Misschien moeten we de kennis van de theoriegemeenschap mengen.



https://home.gwu.edu/~kargaltsev/HEA/washington-conferences.html

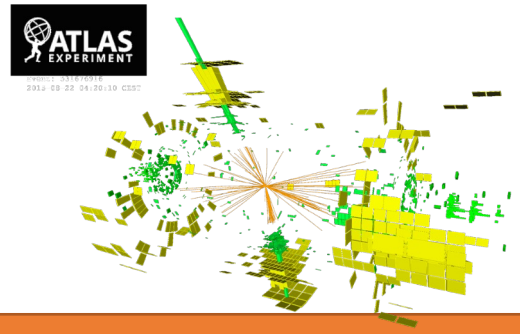
Our approach Model driven

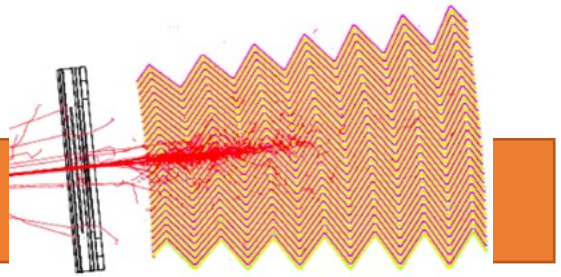
1. Pick many “model of new physics”
2. Pick many likely (?) sets of parameters!
3. Make many predictions
4. Mix them
5. Train a NN
6. Test the prediction
7. Exclude the model parameter point
8. Go to 3 or 1

Detectie van anomalieën: buiten distributie

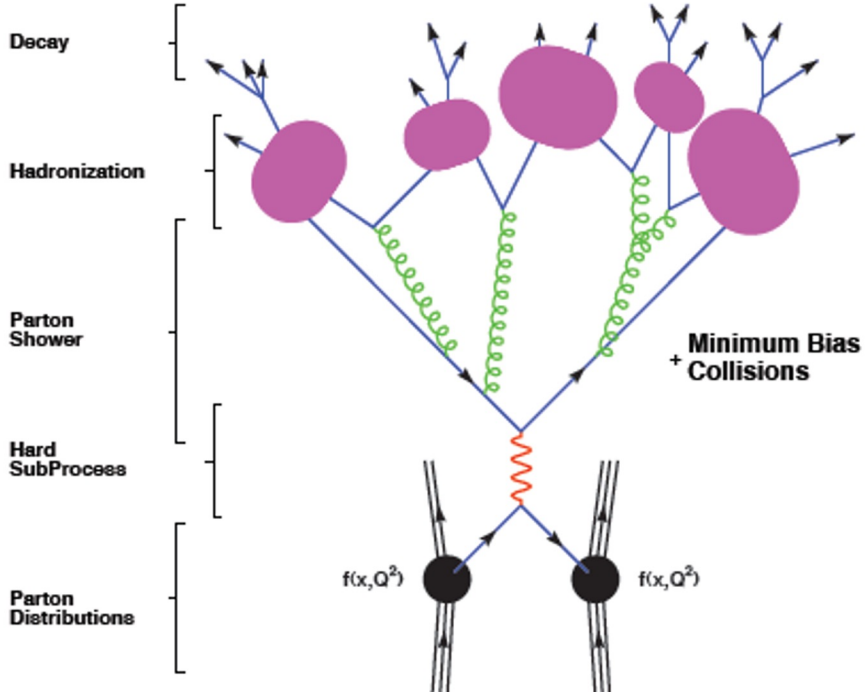
Is the data in the simulation ?

Energy and angles of reconstructed particles





Detector Simulator



Input:
Random numbers

Anomaly detection

1. Kies **geen** "nieuw natuurkundig model"
2. Leer het **achtergrondmodel / Standard Model**
3. Train ML classifier om de voorspelling te testen (is de achtergrond van de gebeurtenis of niet?)
4. Hypothesetest met data|backgroundmodel op classifieroutput
5. Het achtergrondmodel uitsluiten?

How kunnen wij met dit nieuwe fysica zoeken?

De bekende fysica (dieren) zijn katten…

Vraag niet meer: Is het een kat of een hond

Vraag: Is het een kat ?

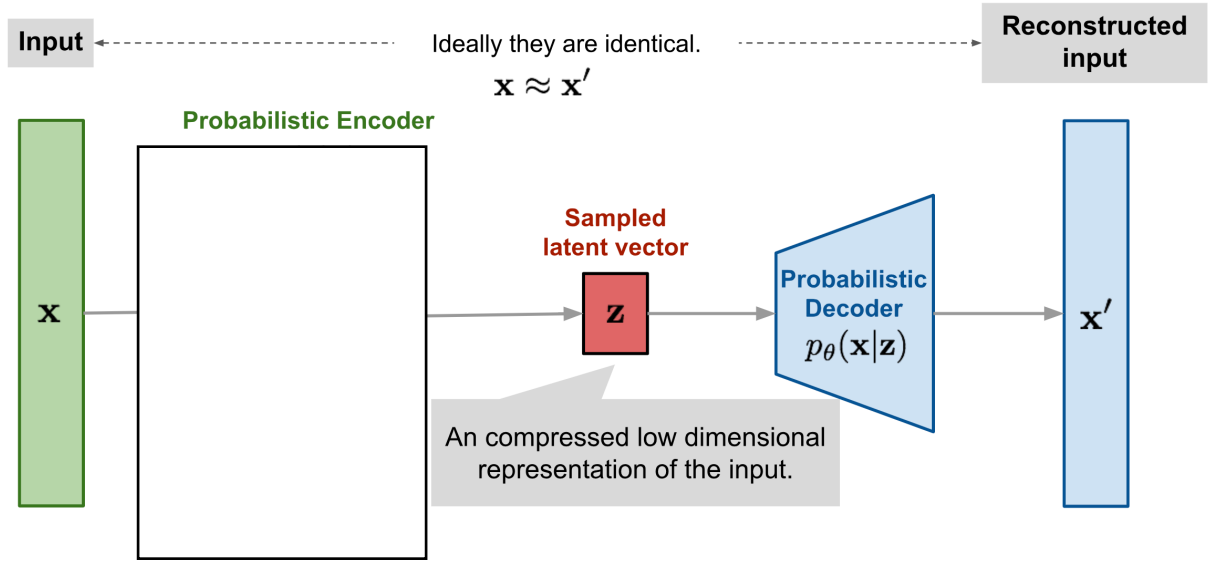
→ Moijelijker dan jullie denken…

In which variable should you search?
Need a variable to "flag" an outlier

How to define anomalies ? ML approaches

2018: The new standard approach

Various papers on arxiv now proposing this → Autoencoder



Then determine a distance between x and x' , e.g. MSE = (x-x')^2

But various other possibilities… needs comparison etc.

Rare and Different

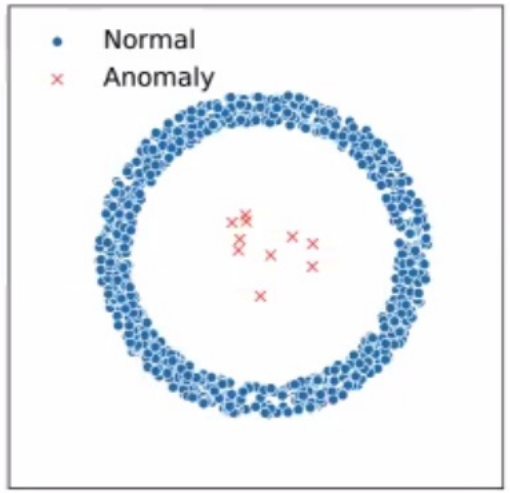
Idea:

Afwijkingen kunnen zeldzaam zijn, wat betekent dat deze gebeurtenissen een minderheid vormen in de normale gegevensset, of anders, wat betekent dat ze waarden hebben die zich niet in de gegevensset bevinden.

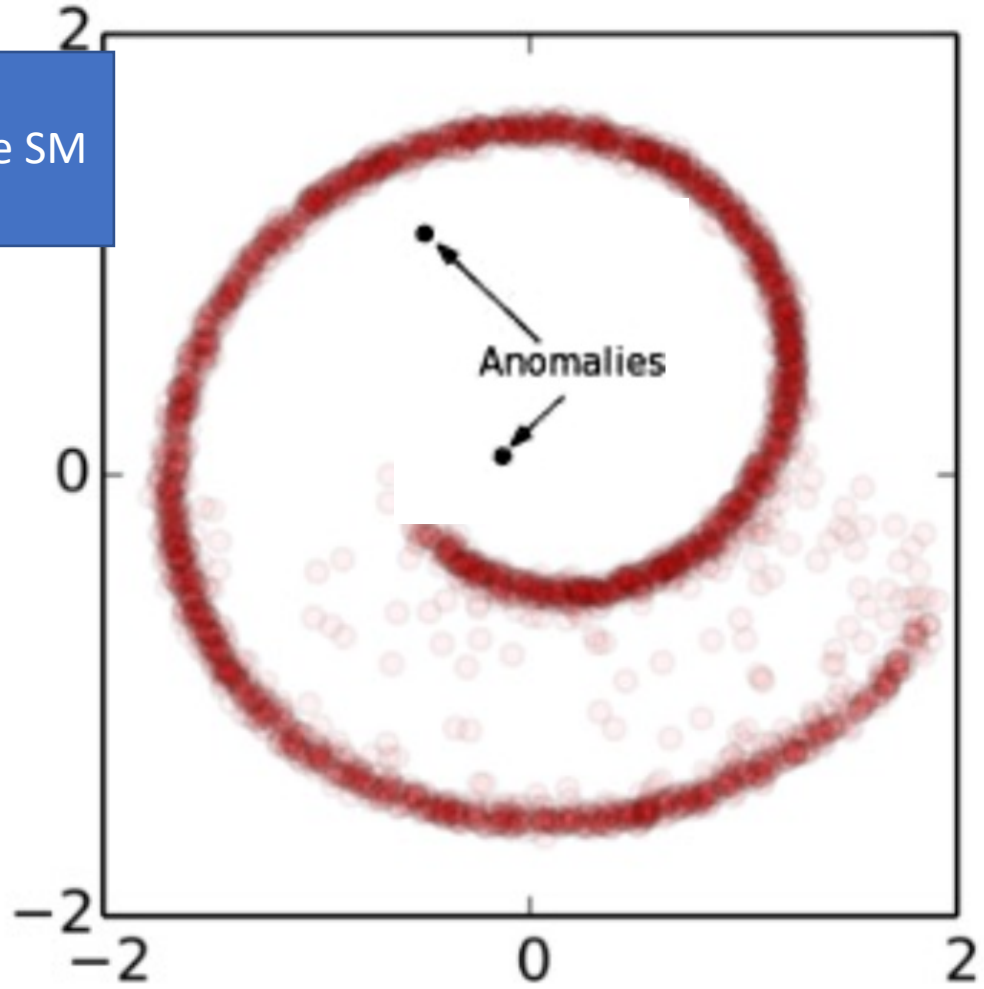
We kwantificeren en combineren deze twee eigenschappen/doelstellingen

https://arxiv.org/abs/2106.10164

Rare → Density estimation



Idea:
Signal region is region outside the SM /simulation



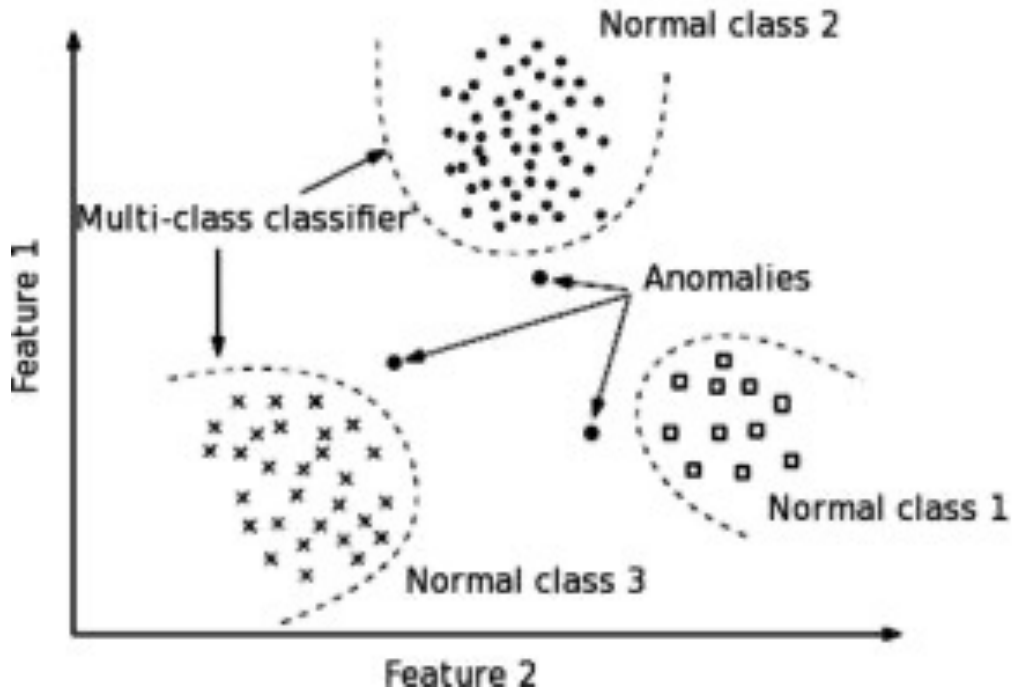
Series of paper on flow models from RU :

- Bob and Rob Verheyen 2021
- Luc Hendris, SC, Rob Verheyen, 2021
- Rob Verheyen : Surjective normalizing flows work even,

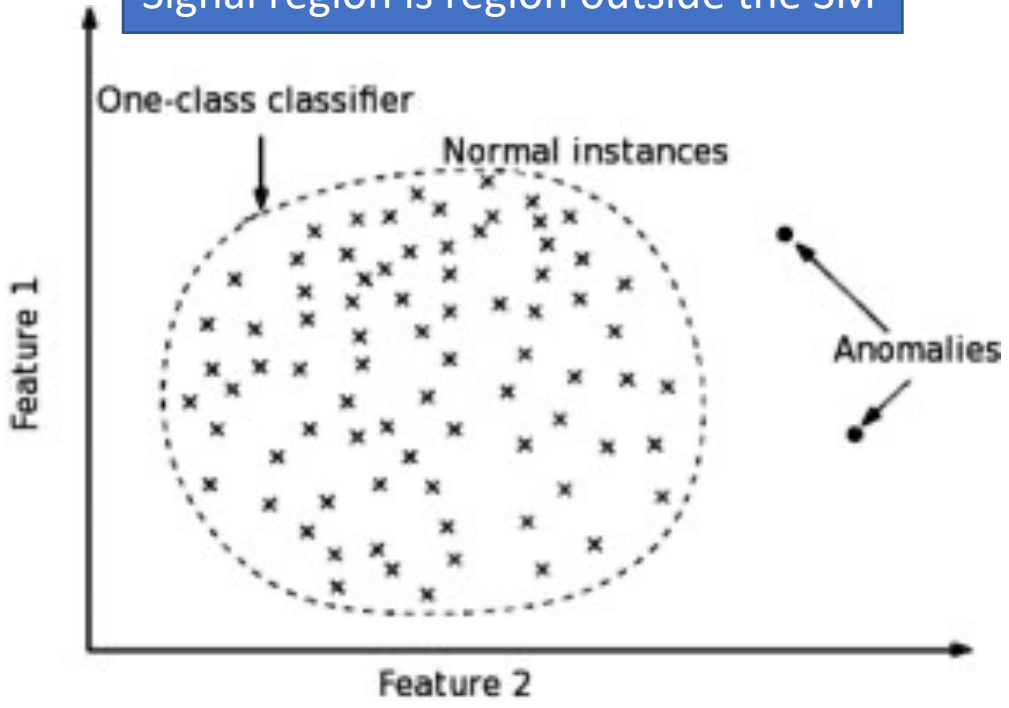
better as anomaly detectors…

→ https://inspirehep.net/literature/2077178

Different ? One class classification



Idea:
Signal region is region outside the SM



Different? Deep SVDD

Alternatively one could try to pass the events through a trained “filter” that only allows events to pass if they belong to the training data

Here: Deep SVDD

X → Network → 42

Anomaly score:

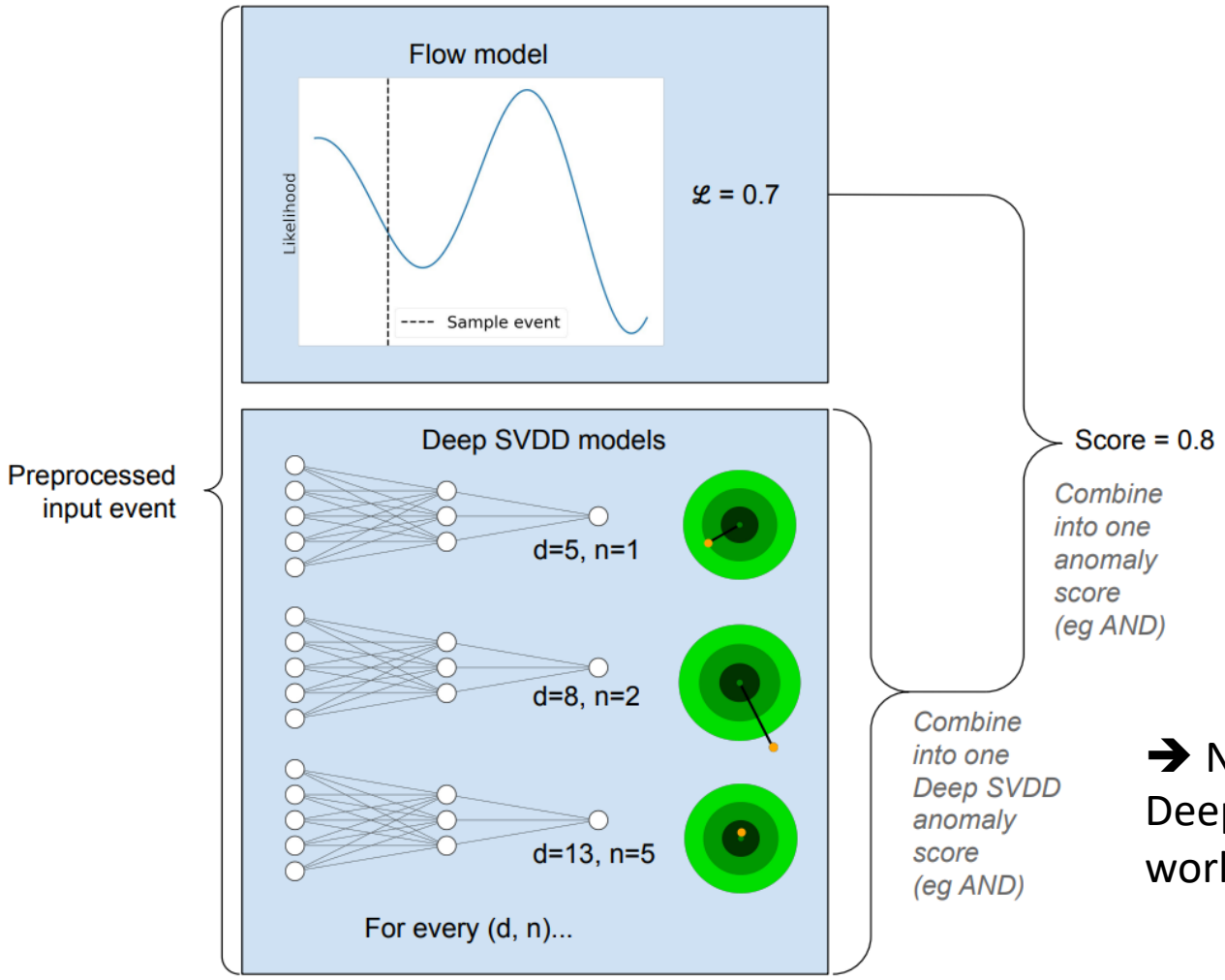
Difference from 42 !

The Deep SVDD network is similar to the encoder component of an autoencoder. The loss is defined as

$$s(x) = O_n^d - \text{Model}(x), \tag{3}$$

where the model maps the input x to the same tensor shape as the manifold O. In our case, O is a vector of identical scalar values, with the subscript n defining the scalar value and superscript d the number of elements in the vector. For example, O_3^4 identifies the vector $(3, 3, 3, 3)$. The optimisation of the Deep SVDD model is fundamentally very simple: it is a NN that receives some input x and transforms it to some output O_n^d.

Rare and Different



➔ Need an ensemble of Deep SVDDs to make it work

Luc Hendriks, Rob Verheyen, SC: https://arxiv.org/abs/2106.10164


Compare them all (besides brute force)

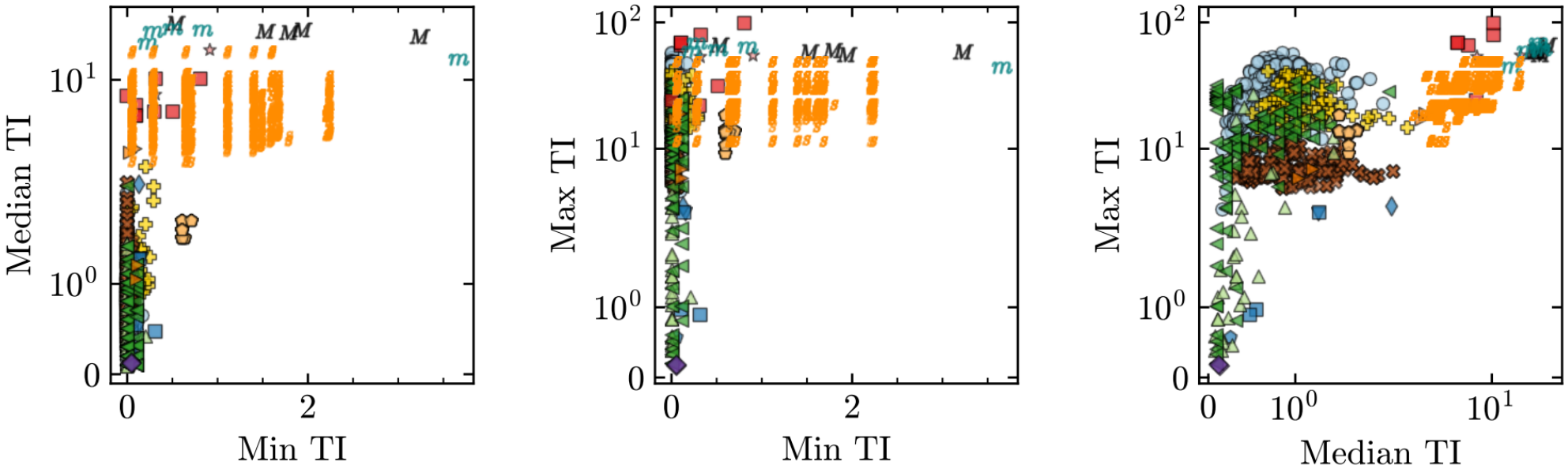
Compare them all

Compared:

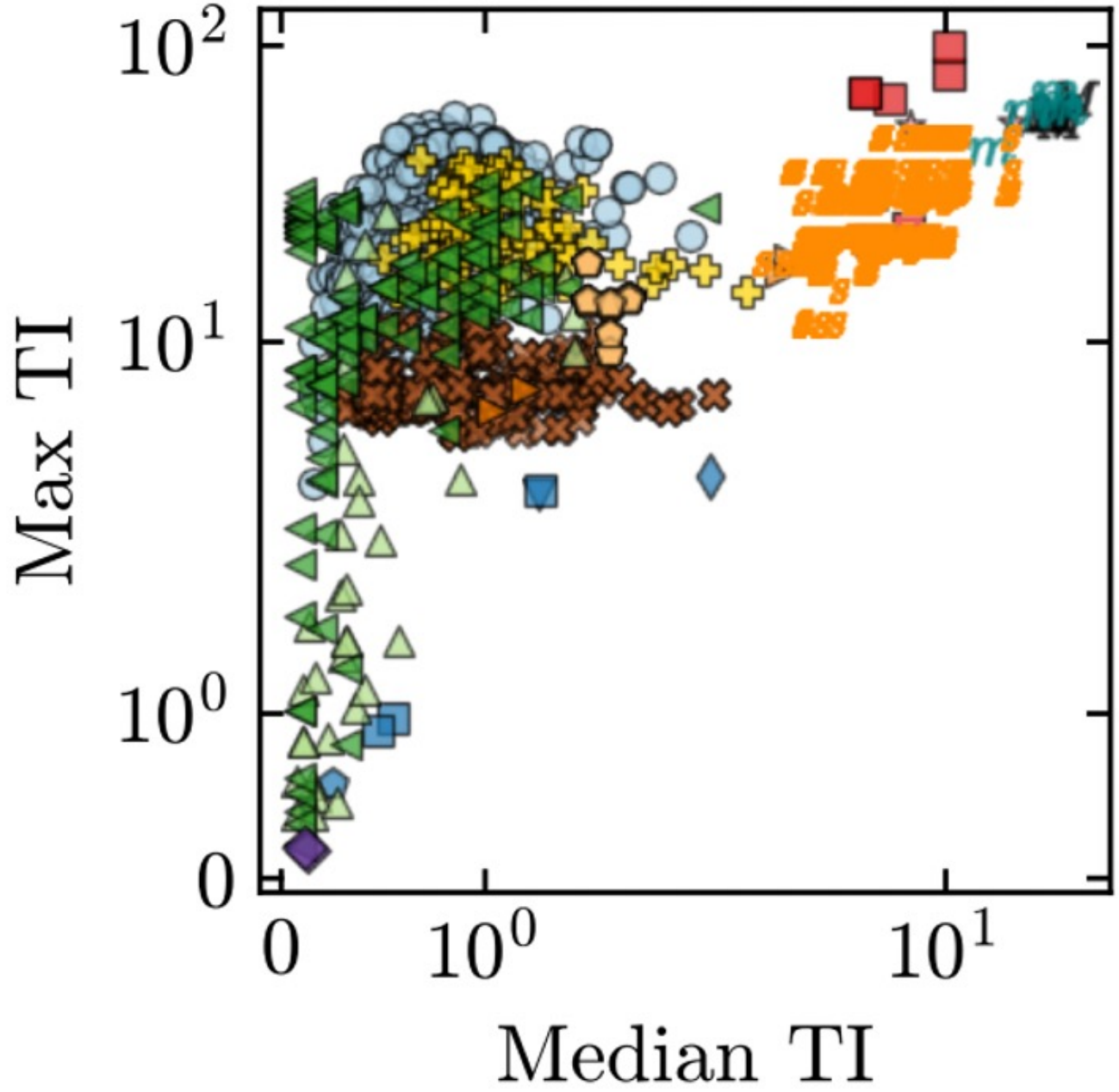
- Oude benaderingen (100s getraind op verschillende "single" signalen)
- Mix van theorie benadering
- Anomalie detectie benaderingen

Who wins?

Total Improvement for models over all signals on
Dark Machines Unsupervised Challenge Hackathon Data



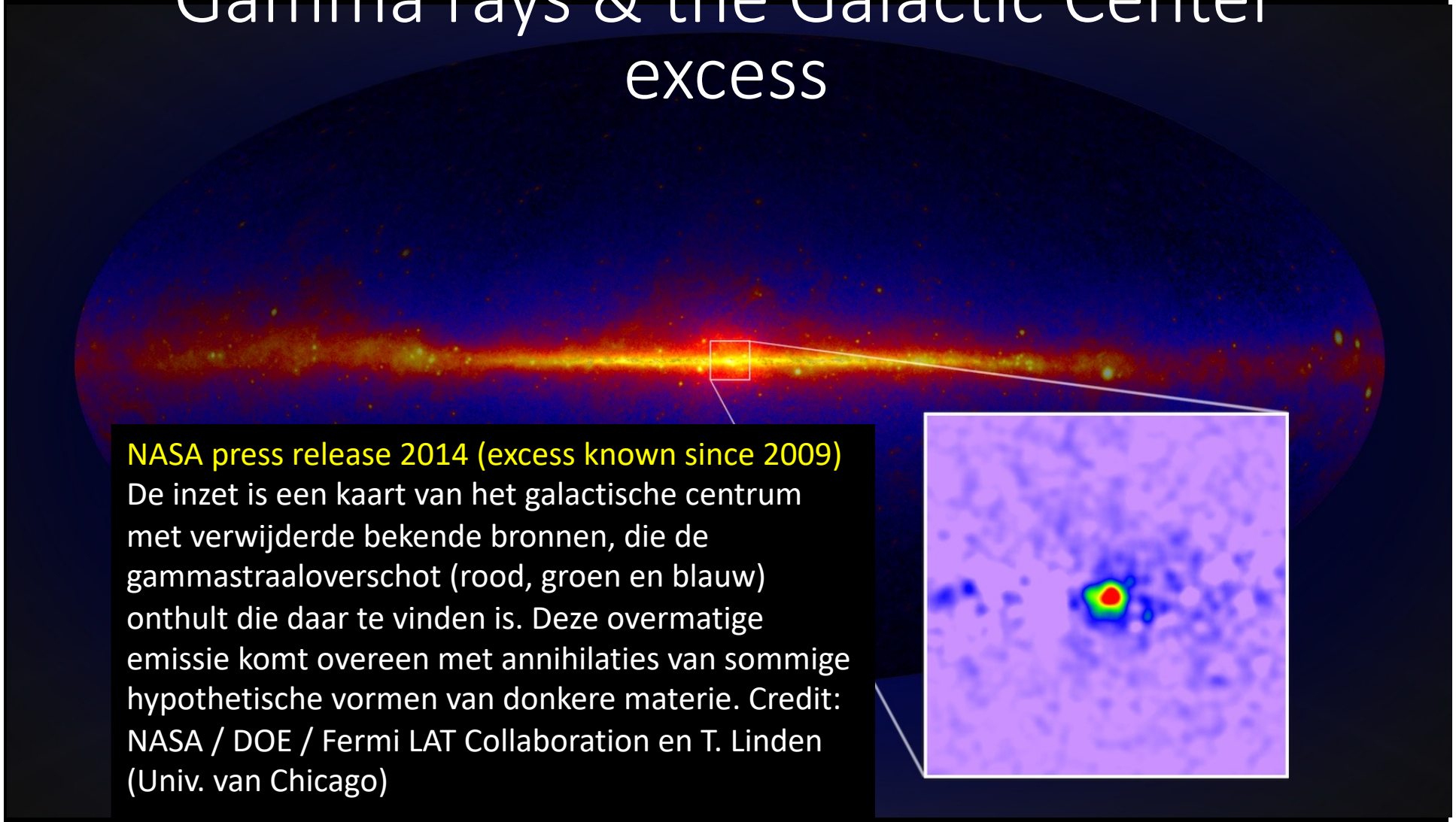
Latent Space	SNF	Flow	SimpleAE
ALAD	IAF	Combined	*M* Mixture of Theories
DAGMM	ConvF	Deep SVDD	*m* mini Mixture of Theories
ConvVAE	KDE	Deep Set	*S* Single Signal Approach
Planar	VAE	CNN(β)VAE	


10^2
10^1
10^0
0
Max TI
0
10^0
10^1
Median TI

Gamma rays: Galacitic Center and the reality gap

→ Zien we donkere materie in het centrum van ons melkwegstelsel?

Gamma rays & the Galactic Center excess



NASA press release 2014 (excess known since 2009)
De inzet is een kaart van het galactische centrum met verwijderde bekende bronnen, die de gammastraaloverschot (rood, groen en blauw) onthult die daar te vinden is. Deze overmatige emissie komt overeen met annihilaties van sommige hypothetische vormen van donkere materie. Credit: NASA / DOE / Fermi LAT Collaboration en T. Linden (Univ. van Chicago)

Official paper in 2015

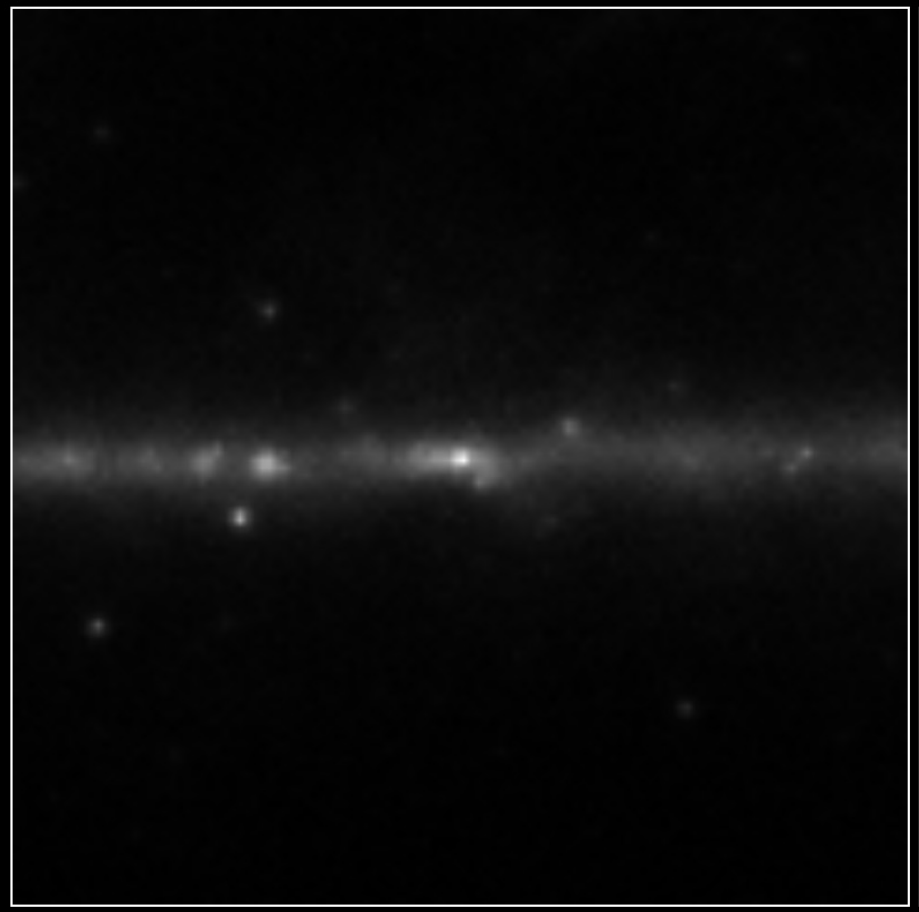
Fermi-LAT Observations of High-Energy Gamma-Ray Emission Toward the Galactic Center
Fermi-LAT Collaboration (M. Ajello (Clemson U.) *et al.*). Nov 9, 2015. 29 pp.
e-Print: **arXiv:1511.02938 [astro-ph.HE]** | PDF

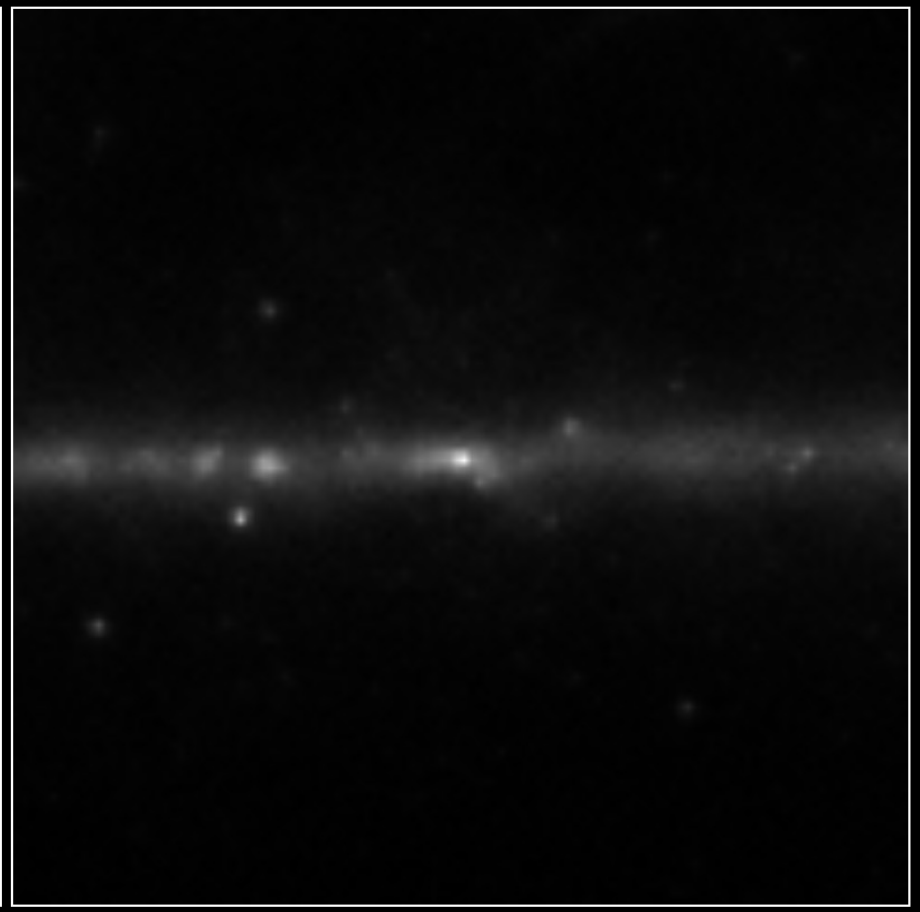
Guess the fraction of point sources

www.mydarkmachine.org

What is this fraction?

This is 0.5





Your prediction:

Invert image:

Guess

Simulatie met parameters ➔ Foto’s

Eerste idee: Train Conv Network voor

Fotos → Parameters

Werkt dit ? Werkt het beter dan conventionele methoden? Waarom?

Our 2017 convolutional network

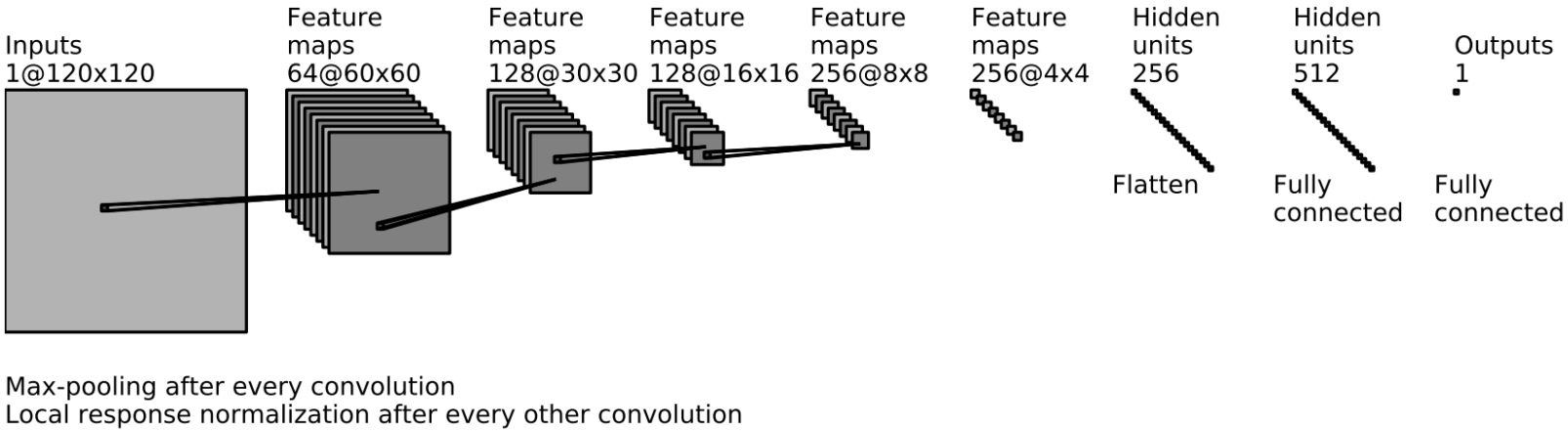


Figure 6: Visualization of the convolutional neural network. The network consists of an input layer, 5 convolutional + pooling layers, 2 fully connected layers and finally an output layer.

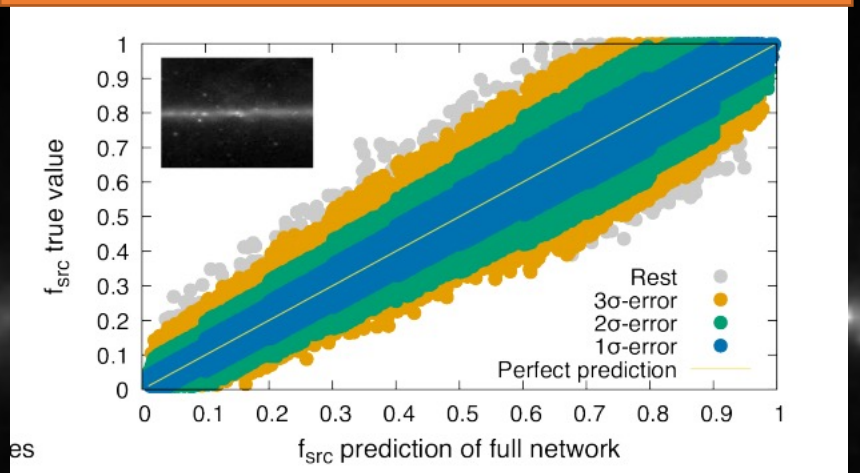
Can a NN determine the number of unresolved point sources relative to isotropic radiation ?

•Published in: *JCAP* 05 (2018) 058, e-Print: 1708.06706 [astro-ph.HE]

What is this fraction?

This is 0.5

Network can generalize over randomness



(b) Prediction of the full network versus true values.

Your prediction:

Invert image:

Truth: 0.052

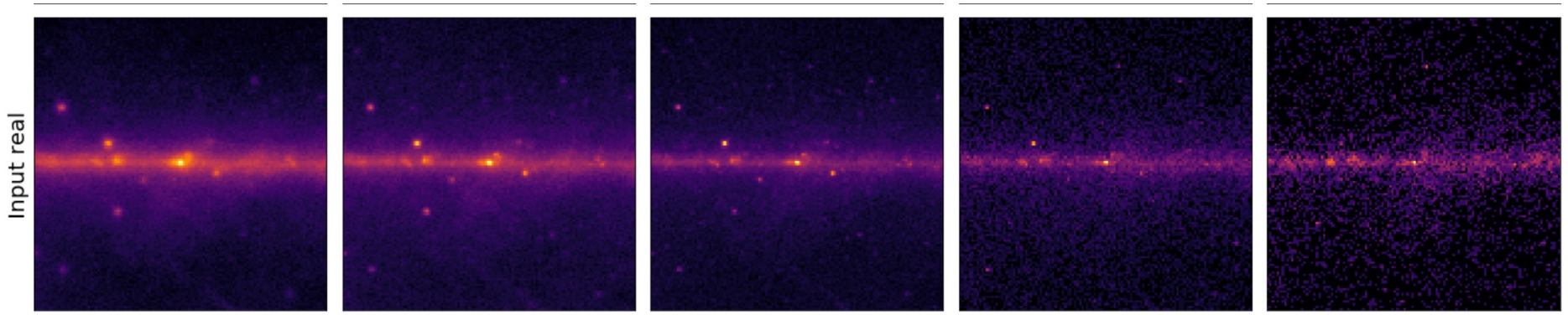
Network: 0.1230

Your guess: 0.5

Who is better? The network

Interpretation here is frequentists and relies on the model to be correct (uncertainties from toy experiments, no p-value yet)

Today: More wavelenghts→ Bayesian determination of 25 parameters

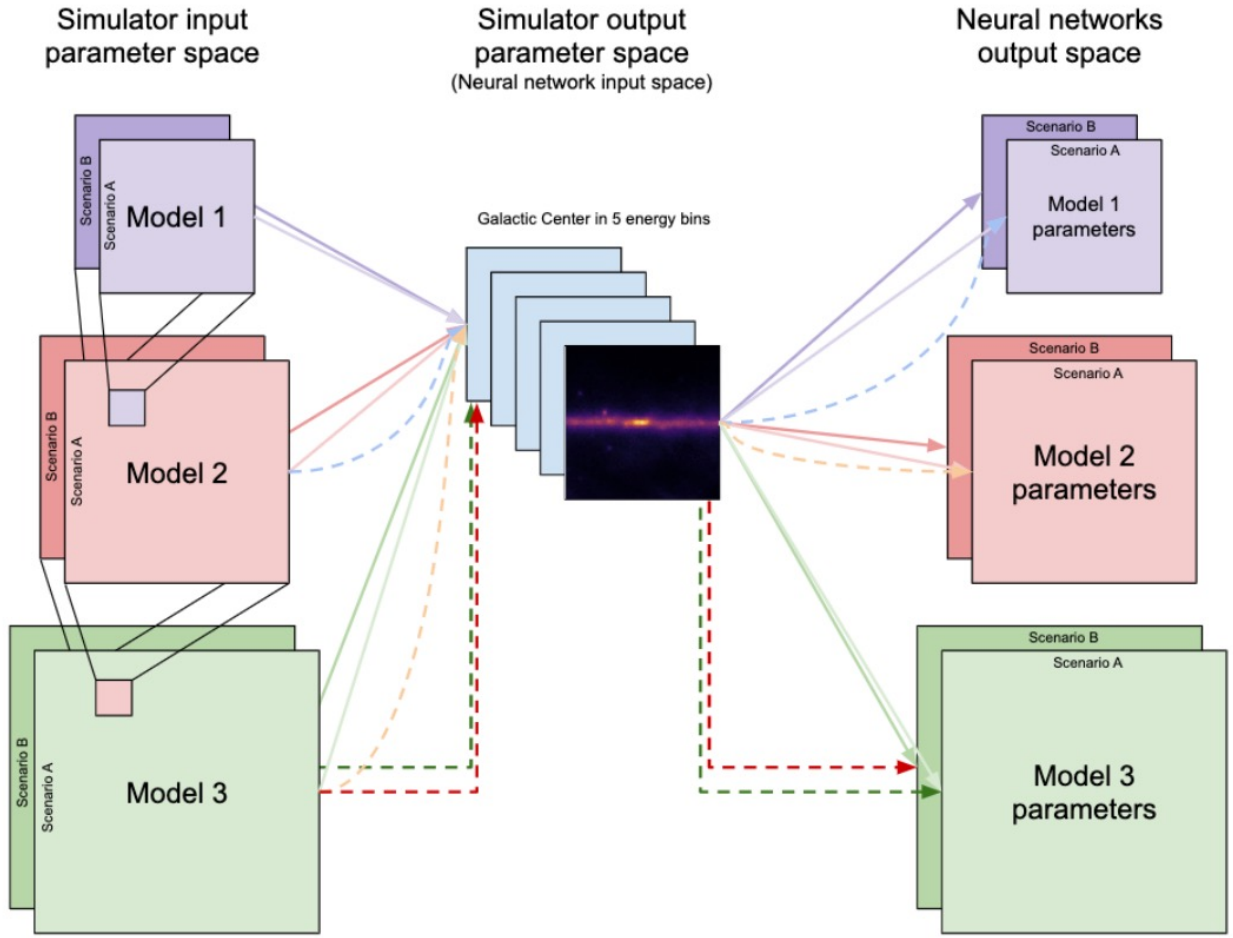


LET NETWORK OUTPUT all parameters AND all uncertainties

New paper after 5 years !

RU Internal / in Fermi-LAT review:

Idee: Test complexere simulaties, leer de beste simulaties uit data, neem voor het eerst alle onzekerheden op (ook "out of simulation")



Kunnen we de puntbronnen ook direct bepalen?

Ja, ander project …

Astronomy and gamma rays: autosourceID

→ **Automatische identificatie van astronomische objecten**

Automatic ID of astrophysical objects: AutosourceID, slides by Fiorenzo Stoppa

IDEA: FASTER / REALTIME ID OF ASTRONOMIC SOURCES

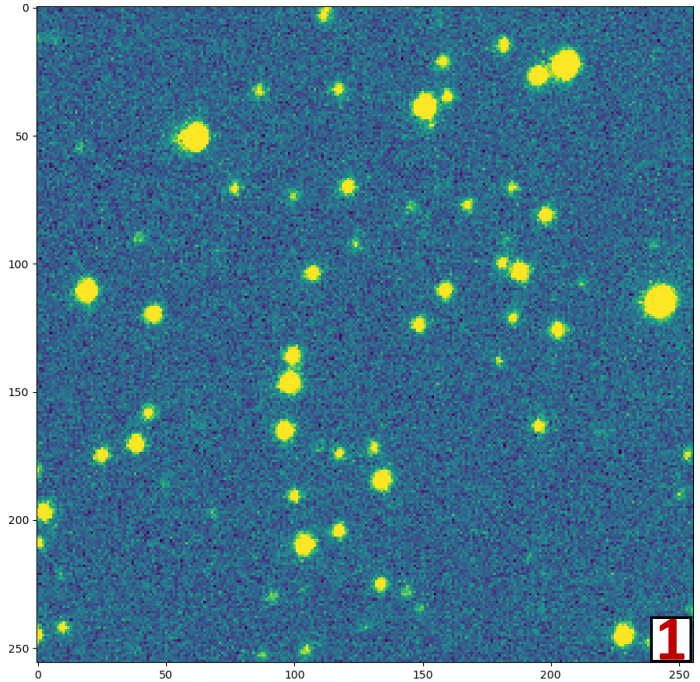
Full field (10.5k x 10.5k pixels) is 3.7 seconds for AutosourceID and 120 for SExtractor.

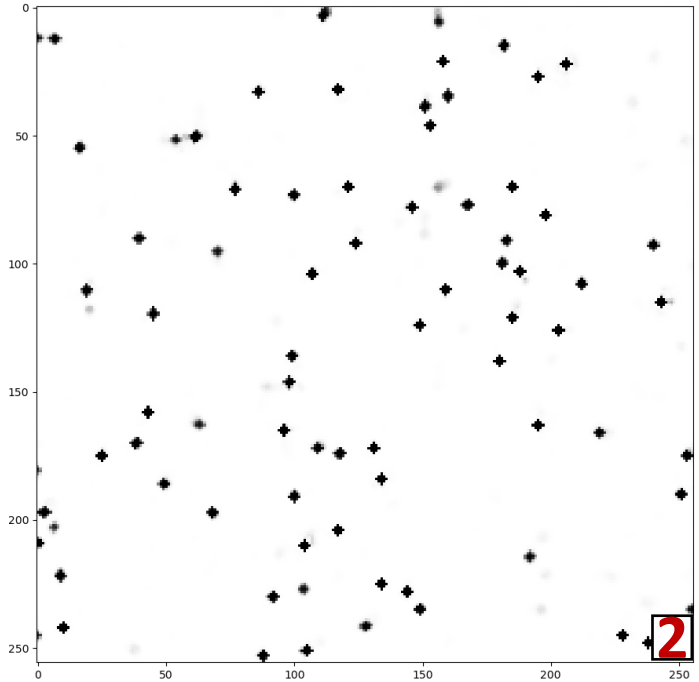
Train de ML op de simulatie en/of de astronoom.

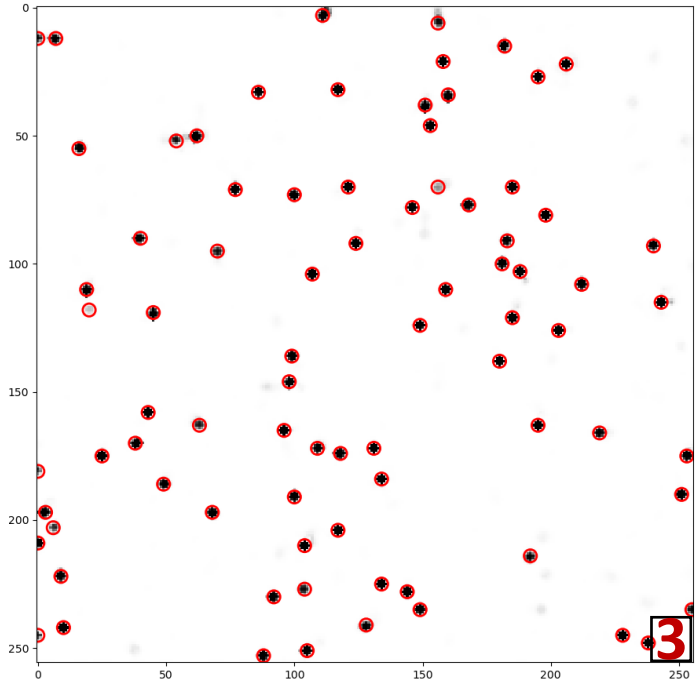
1 Input optical image

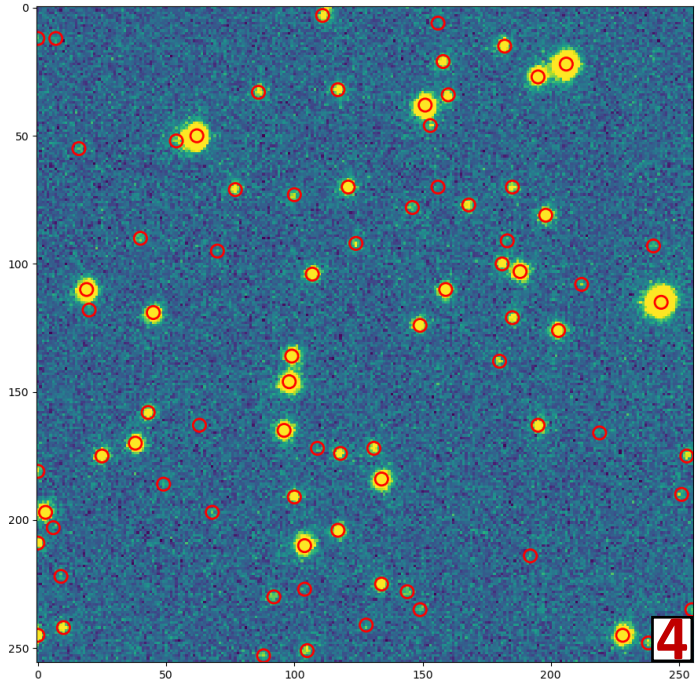
2 Predicted mask

3 Laplacian of Gaussian Results

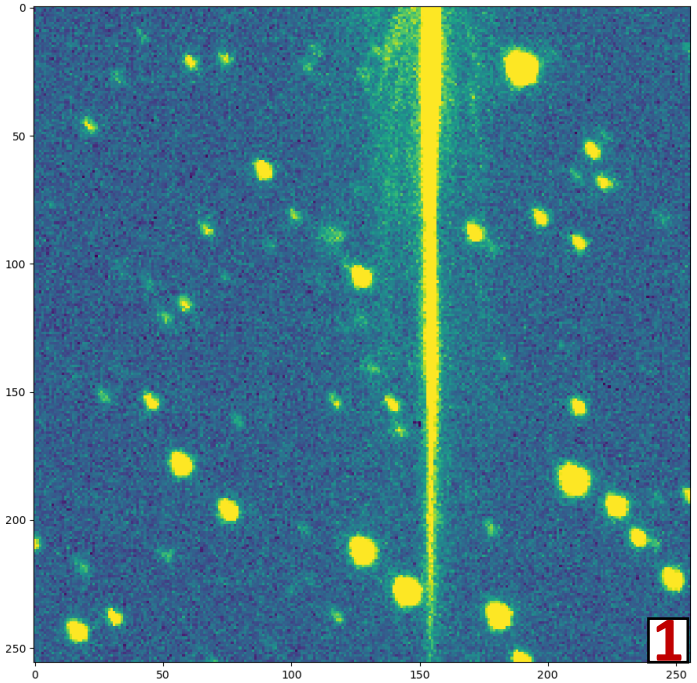


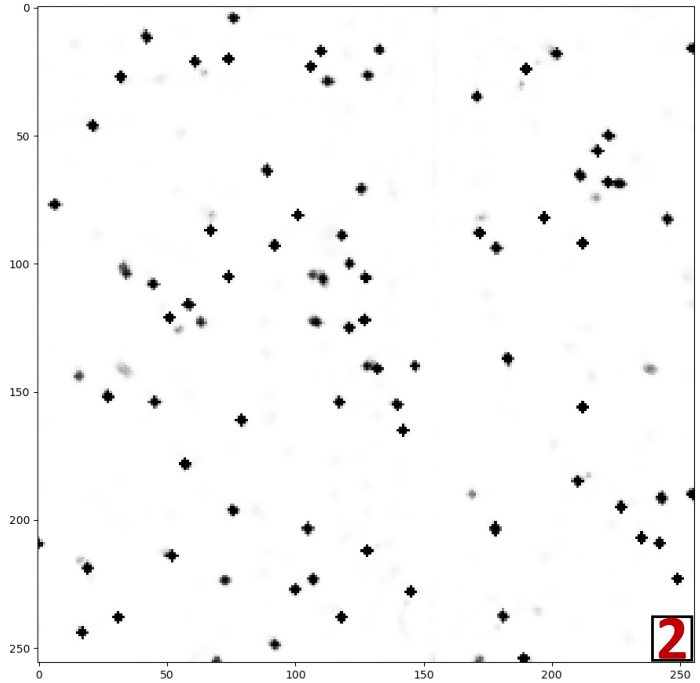


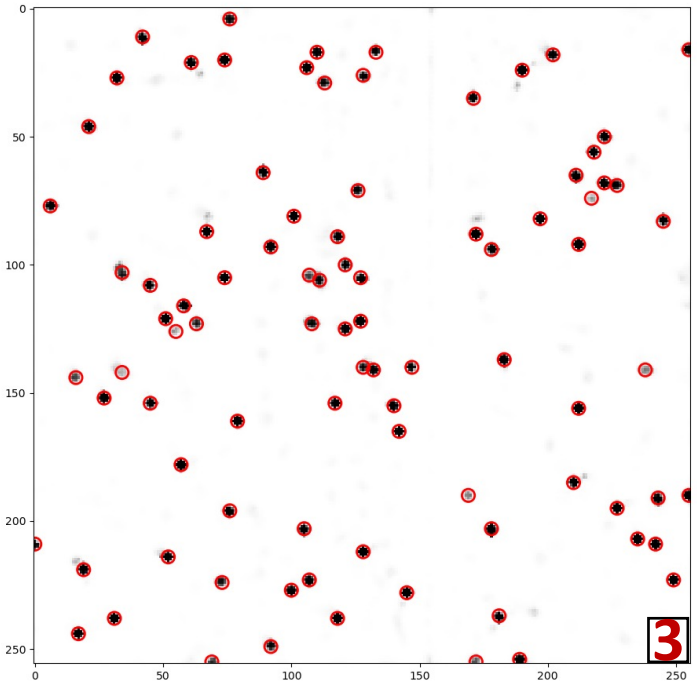


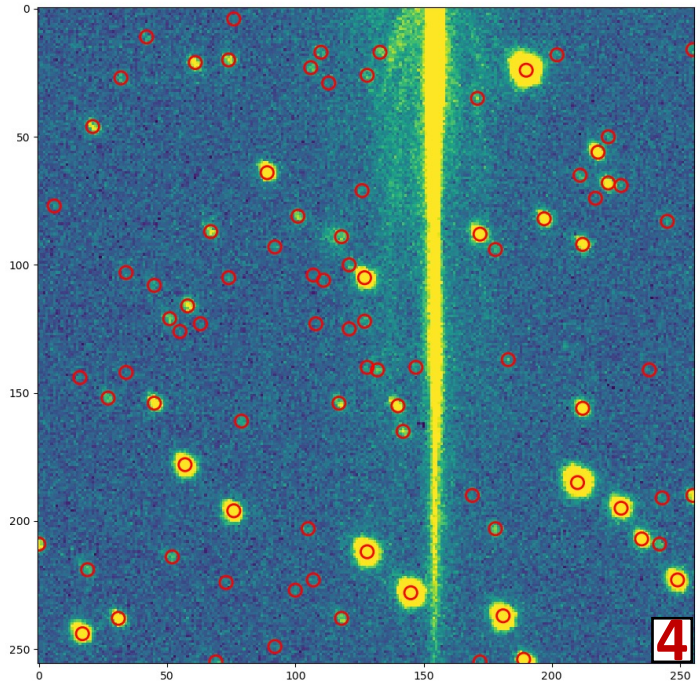


1 Input optical
2 image
3 Predicted mask
4 Laplacian of Gaussian Results





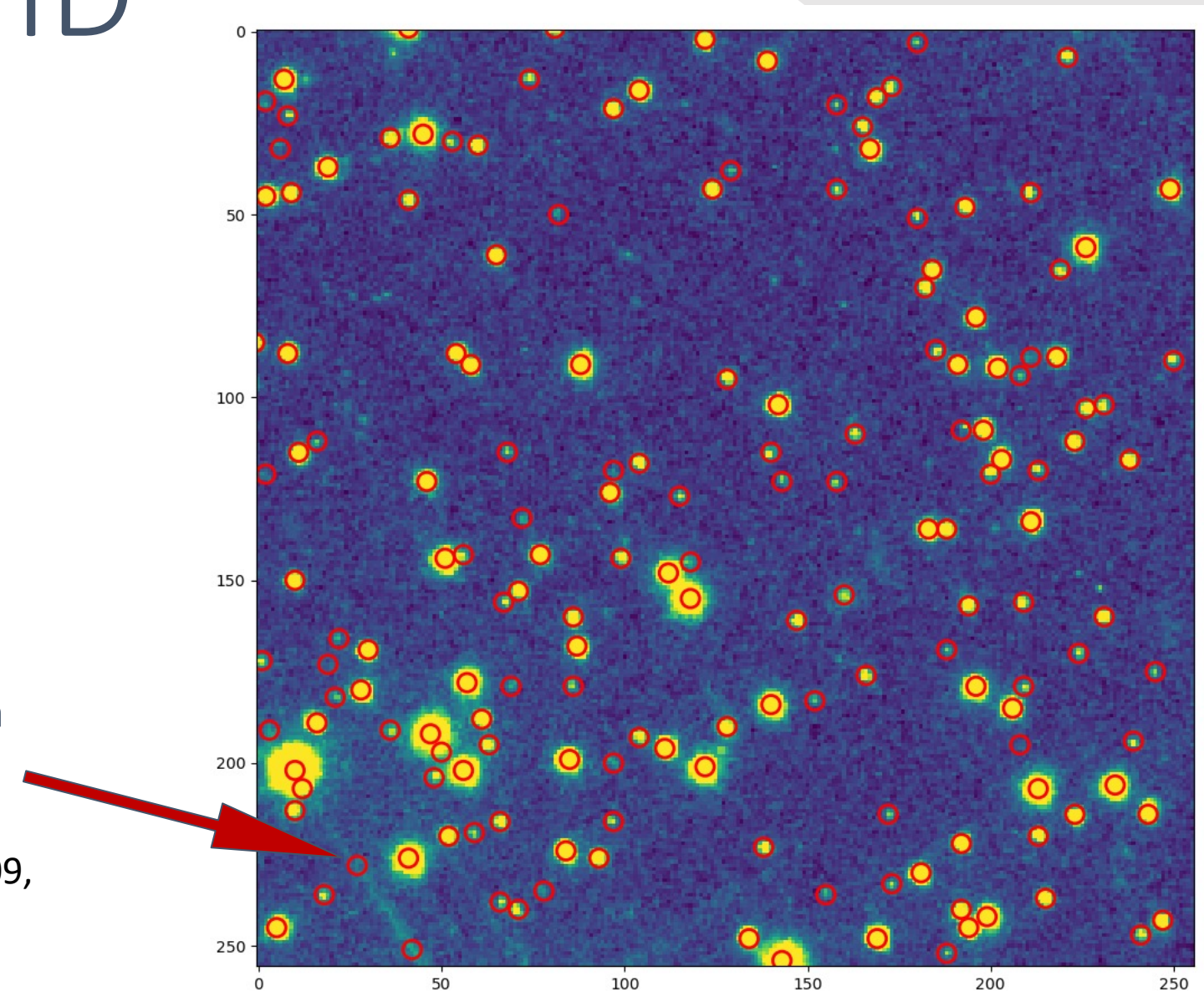




Hubble HD images

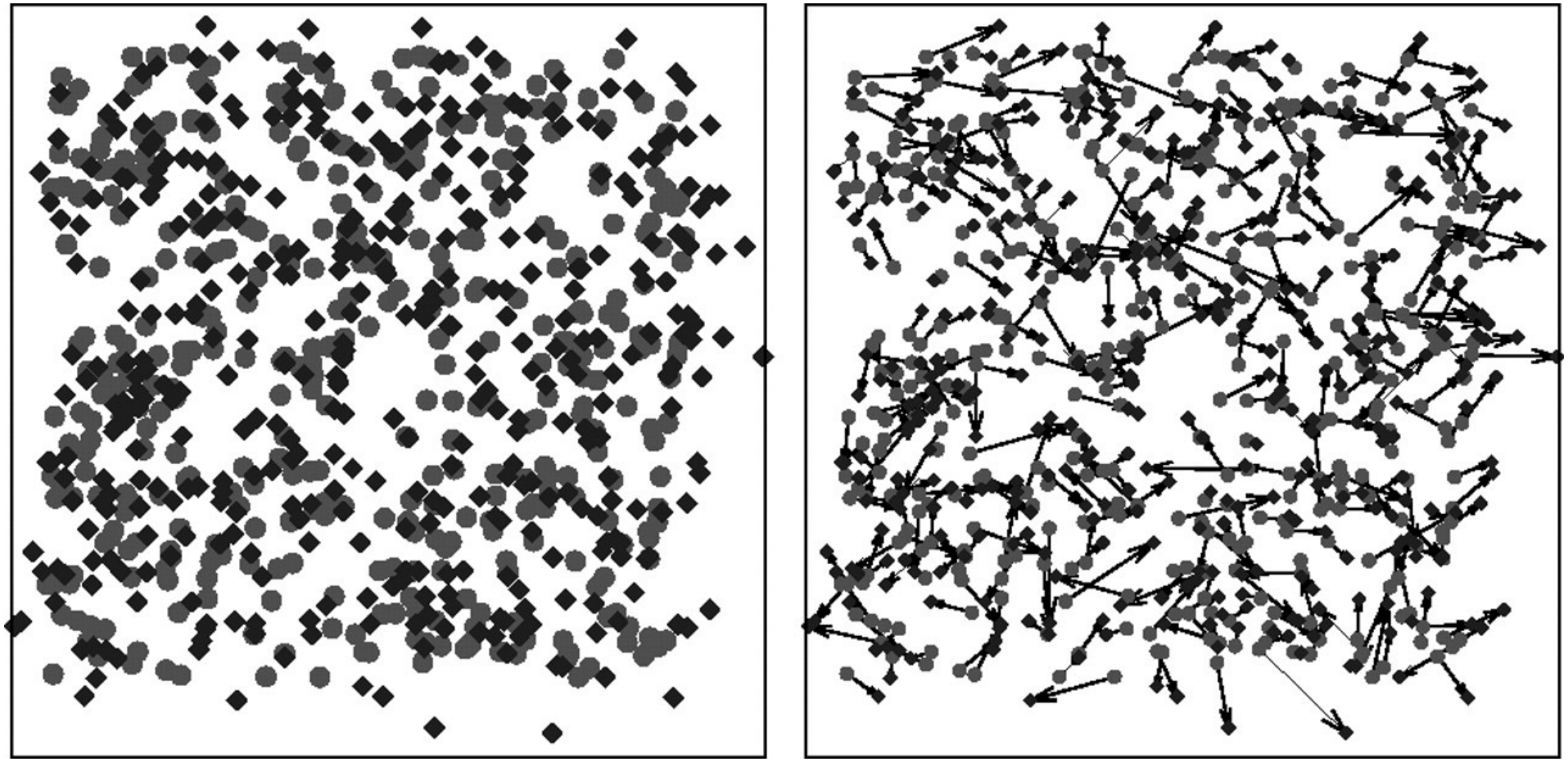
Most of the visible sources correctly localized.

Small problems with a diffraction spike



•*Astron.Astrophys.* 662 (2022) A109,
With Astro department
(mainly Fiorenzo Stoppa)

ATLAS: tracking → inference at 40 MHz ?

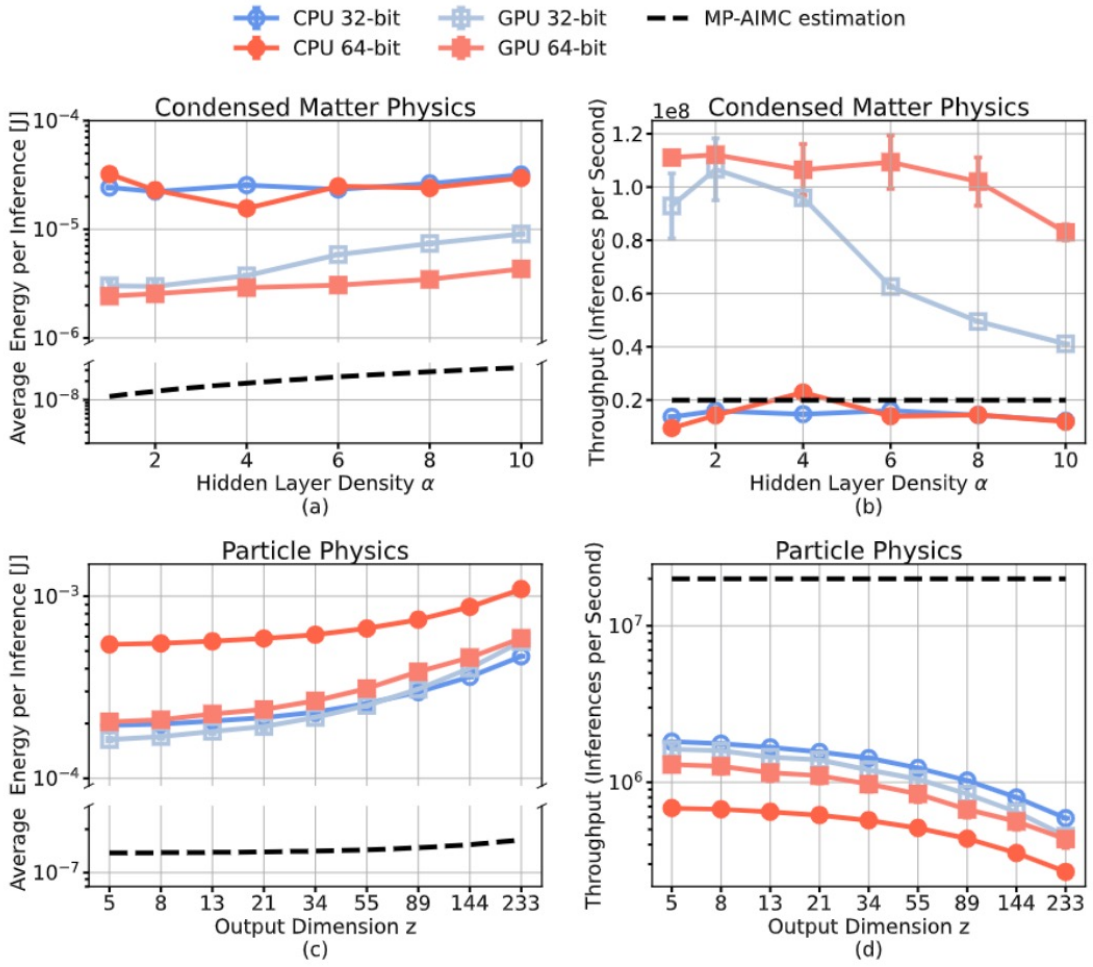


Idea : Train u-nets to go from **(almost) ALL pixels → (almost) ALL tracks** (including ALL uncertainties) in one step
Do this on dedicated hardware accelerators (GPUs, FPGAs, neuromorphic, future quantum ?) → CODE?

Neuromorphic computing

→ The next step: Print wetenschappelijke neurale netwerken op computerchips

LHC etc. : Neuromorphic Computing on AIMC architecture with IBM and IMM



How fast can neuromorphic chips process scientific data?
→ ATLAS trigger
How much energy do they consume ?

(also compare to quantum hardware, maybe enourmous gain!)

Benchmarking energy consumption and latency for neuromorphic computing in condensed matter and particle physics

Dominique J. Kösters,[1,2,3] Bryan A. Kortman,[1,4] Irem Boybat,[3] Elena Ferro,[3,5] Sagar Dolas,[6] Roberto Ruiz de Austri,[7] Johan Kwisthout,[8] Hans Hilgenkamp,[1,9] Theo Rasing,[10] Heike Riel,[3] Abu Sebastian,[3] Sascha Caron,[4,11] and Johan H. Mentink[10]

[1)] *University of Twente, Faculty of Science and Technology, P.O. Box 217 7500 AE, Enschede, The Netherlands*
[2)] *Radboud University, Institute for Molecules and Materials, Heyendaalseweg 135, 6525 AJ, Nijmegen, The Netherlands*
[3)] *IBM Research Europe - Zürich, 8803 Rüschlikon, Säumerstrasse 4, Switzerland*
[4)] *Nikhef, P.O.Box 41882 1098 XG Amsterdam, The Netherlands*
[5)] *Eidgenössische Technische Hochschule Zürich, Department of Information Technology and Electrical Engineering, Gloriastrasse 35, 8092 Zürich, Switzerland*
[6)] *SURF Cooperation, Innovation Team, Moreelespark 48, 3511 EP, Utrecht, The Netherlands*
[7)] *University of Valencia-CSIC, Instituto de Física Corpuscular, Parc Científic UV, c/ Catedrático José Beltrán 2, E-46980 Paterna, Spain*
[8)] *Radboud University, Donders Institute for Brain, Cognition and Behaviour, P.O. Box 9104 6500 HE, Nijmegen, The Netherlands*

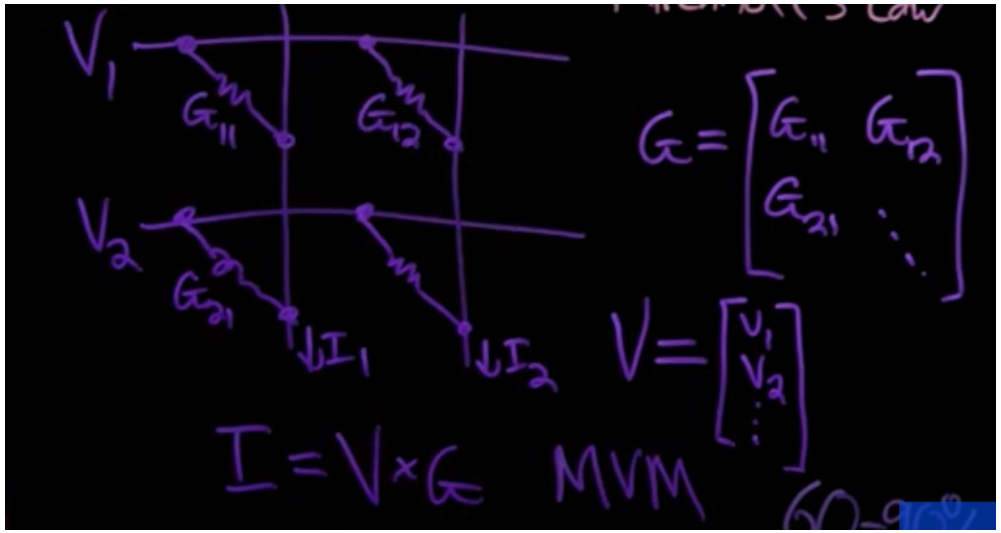
Answering referee comments

→ published in a few weeks

Neuromorphic computing

I = current
V= voltage
G= resistances

NN =



(source IBM video)

Various approaches ranging from classifical FPGA, ASICs to In memory computing (previous slide), spiking NN on chips (Inter Loihi) or even photonic !
Etc.

→ Need dedicated study, will likely become highly important for compuational science
→ Main topic of our NWA proposal “datascope” (was nextgraspp before)
→ RU could become a leader here ?

Samenvatting

- Nieuwe fysica is onbekend…
-
- Nieuwe methoden om effective naar nieuwe fysica te zoeken