Artificial Intelligence in High Energy Physics II

First, how did we see the Higgs Boson ?

David Rousseau

IJCLab-Orsay

rousseau@lal.in2p3.fr

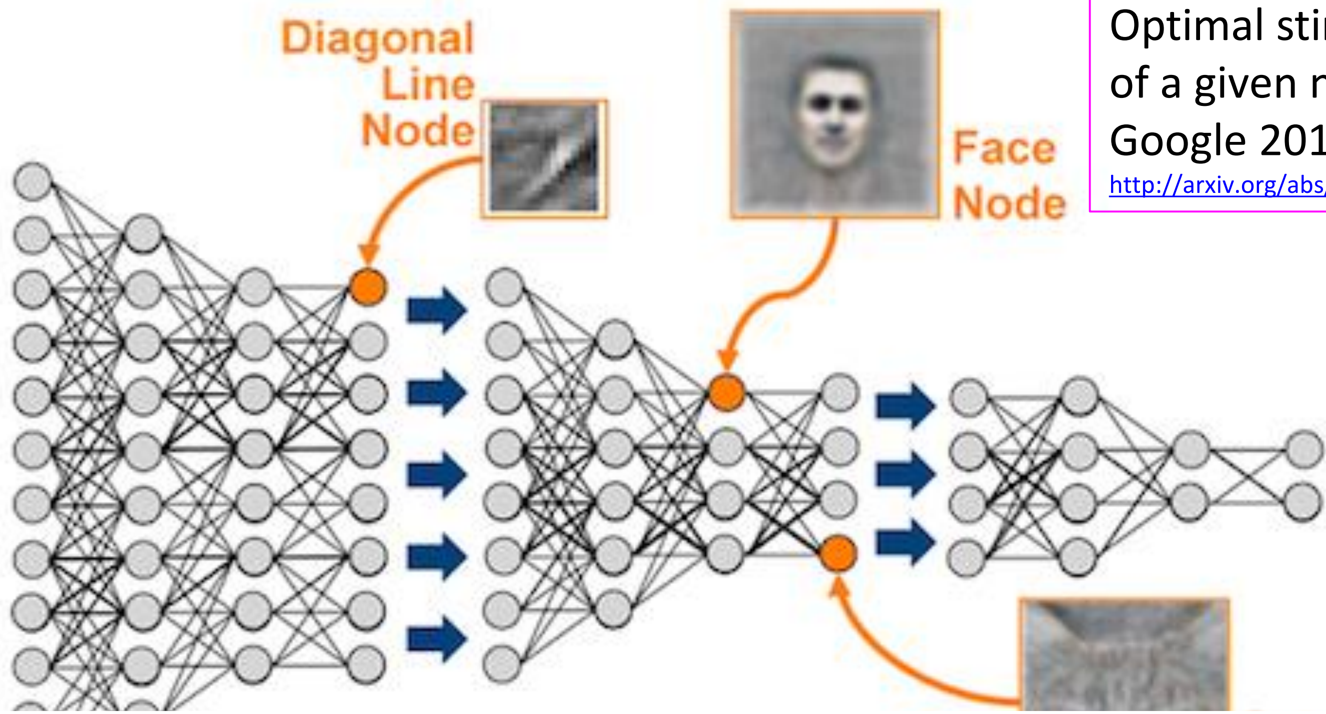
Rencontres d'été de physique

de l'infiniment grand à l'infiniment petit

Juillet 2022

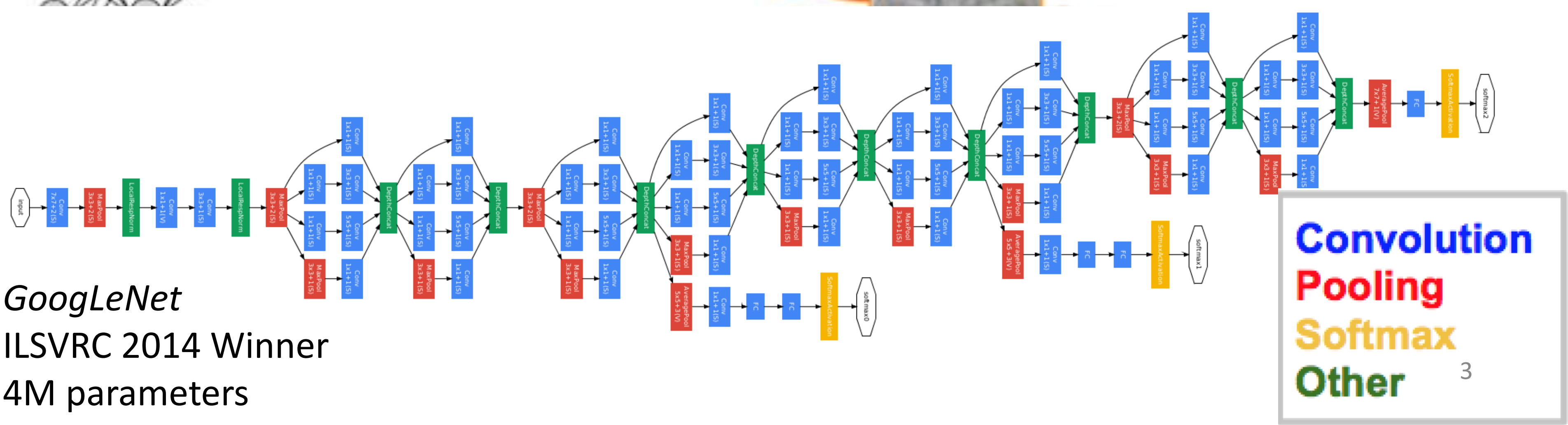
Apprentissage profond

Apprentissage profond



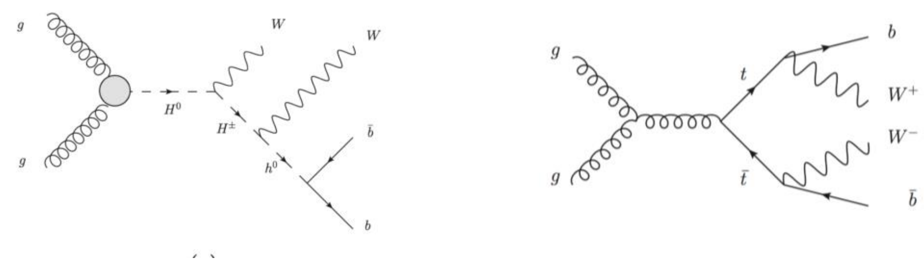
Optimal stimulus
of a given neuron
Google 2012
http://arxiv.org/abs/1112.6209

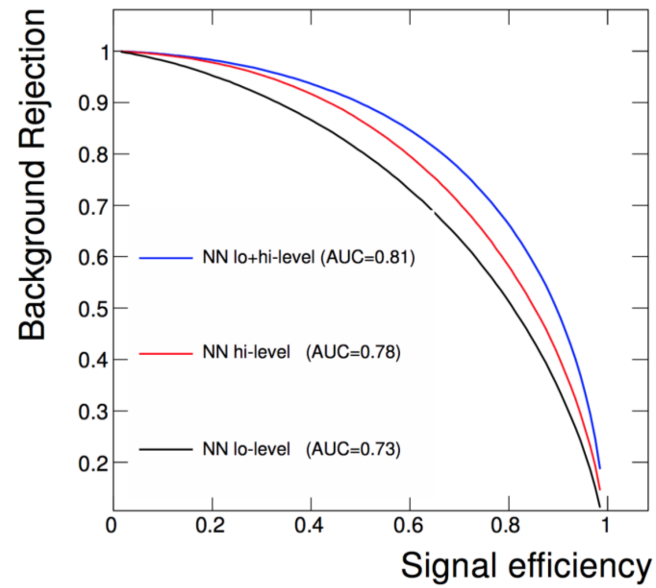
˜1950
1990 chiffres manuscrits
2000 deep learning explosion



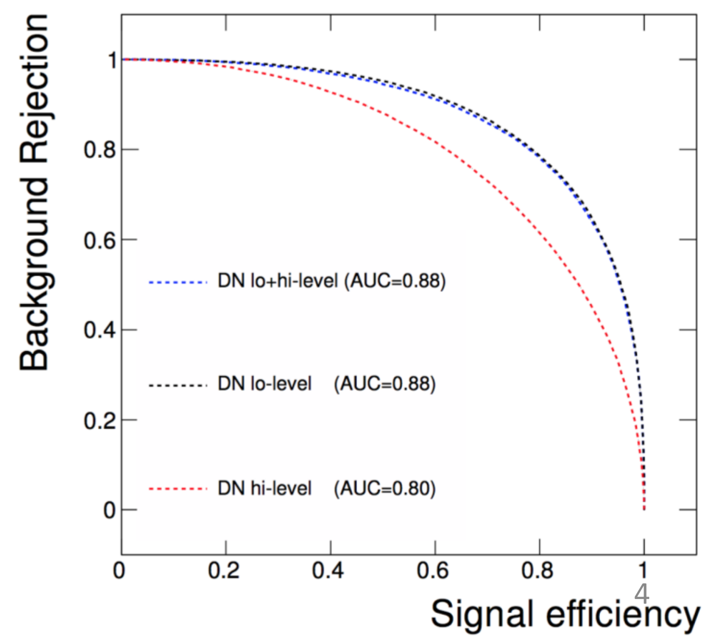
GoogLeNet
ILSVRC 2014 Winner
4M parameters

[1402.4735](1402.4735) Baldi, Sadowski, Whiteson



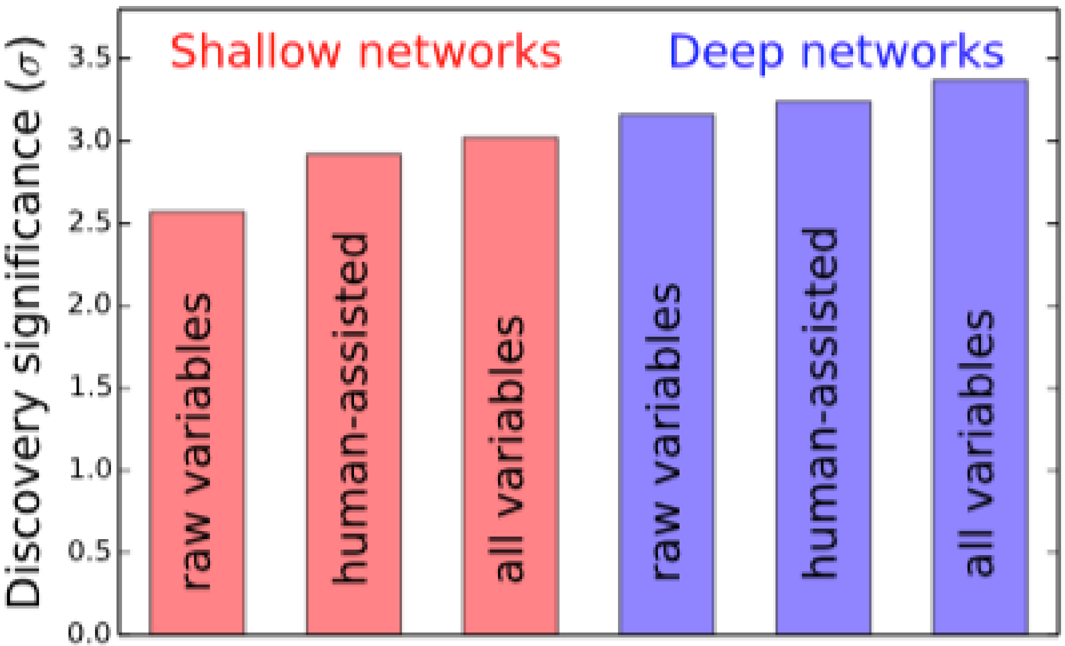


- Recherche d'un Higgs rare : H^0➔WWbb vs tt➔WWbb
- Simulation simplifiée d'événements LHC
- Variables de bas niveau
  - Quadri-vecteur particules
- Variables de haut niveau
  - Masses invariantes, angles etc…
- ➔DNN (Deep Neural Network) marche mieux que réseau simple
- ➔DNN n'a pas besoin de variables de haut niveau
- ➔le DNN apprend la physique ????



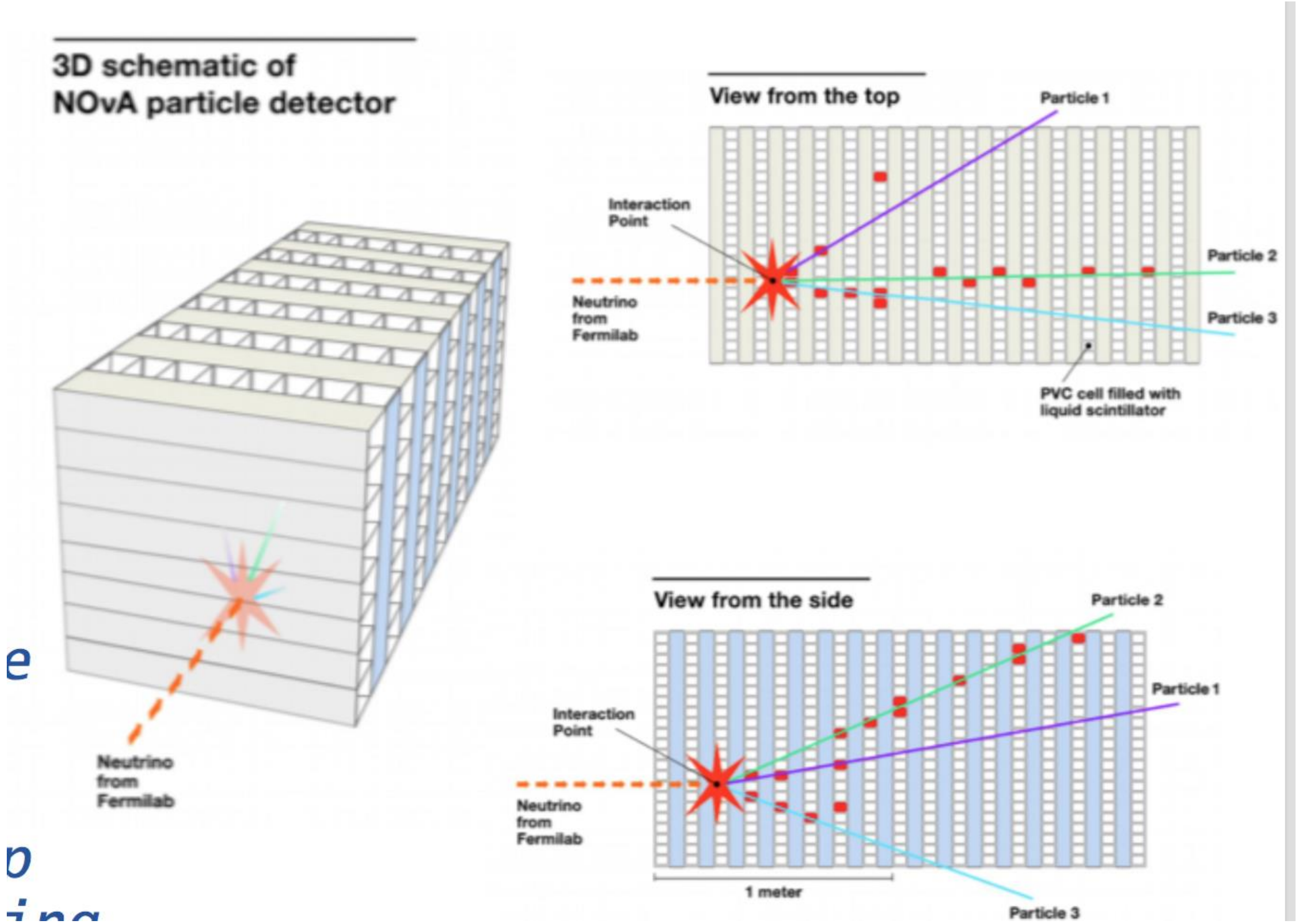
[1410.3469](1410.3469) Baldi Sadowski Whiteson

- Analyse H tautau : H➔tautau vs Z➔tautau
  - Bas niveau
  - Haut niveau



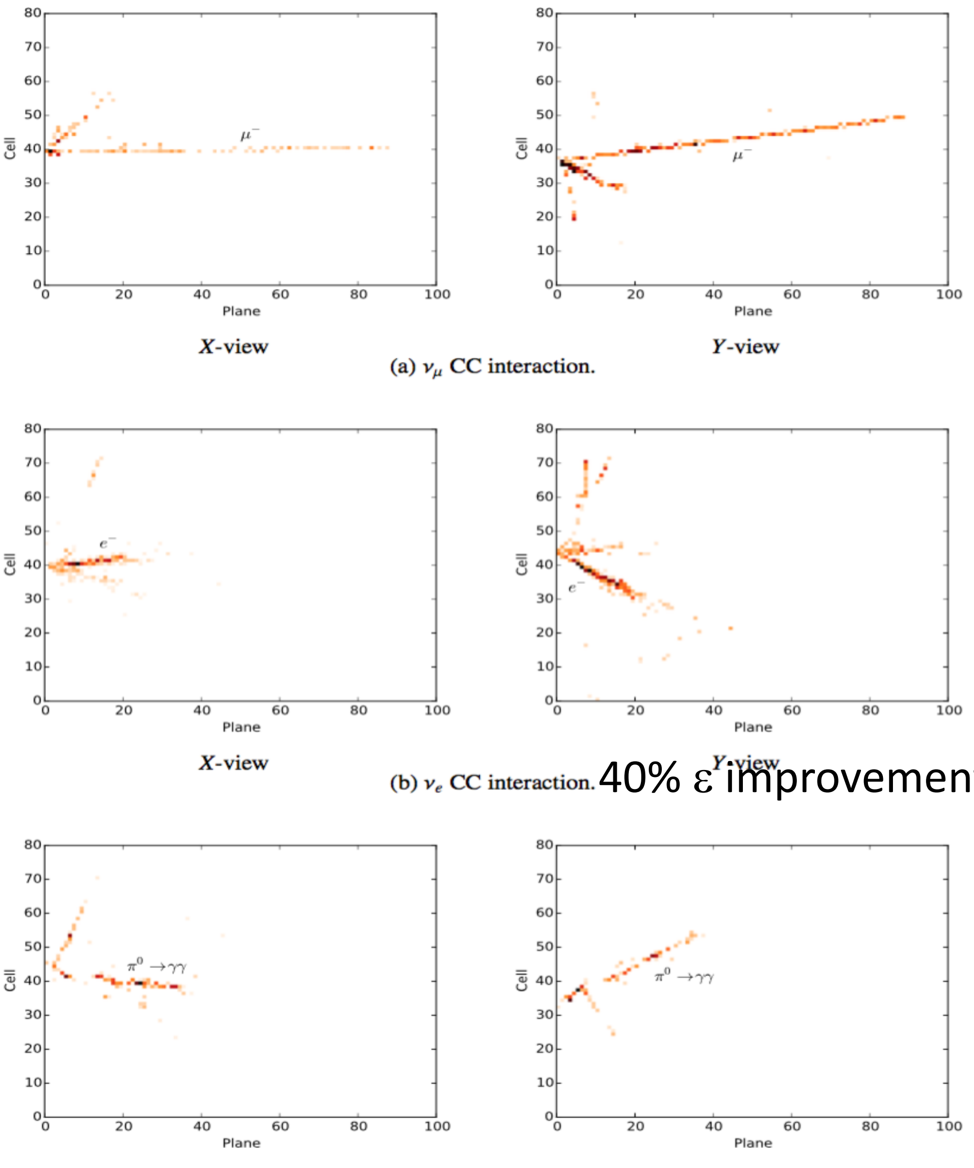
- DNN marche mieux que réseau simple
- Mais cette fois il a besoin de variables de haut niveau !
- Ca se complique…
- Pour l'instant, pas de résultat de physique utilisant DNN sauf ...

Un succès du Deep Learning : NOVA



NOVA (2)

arXiv 1604.01444 Aurisano et al



(a) ν_μ CC interaction.

(b) ν_e CC interaction.

(c) NC interaction.

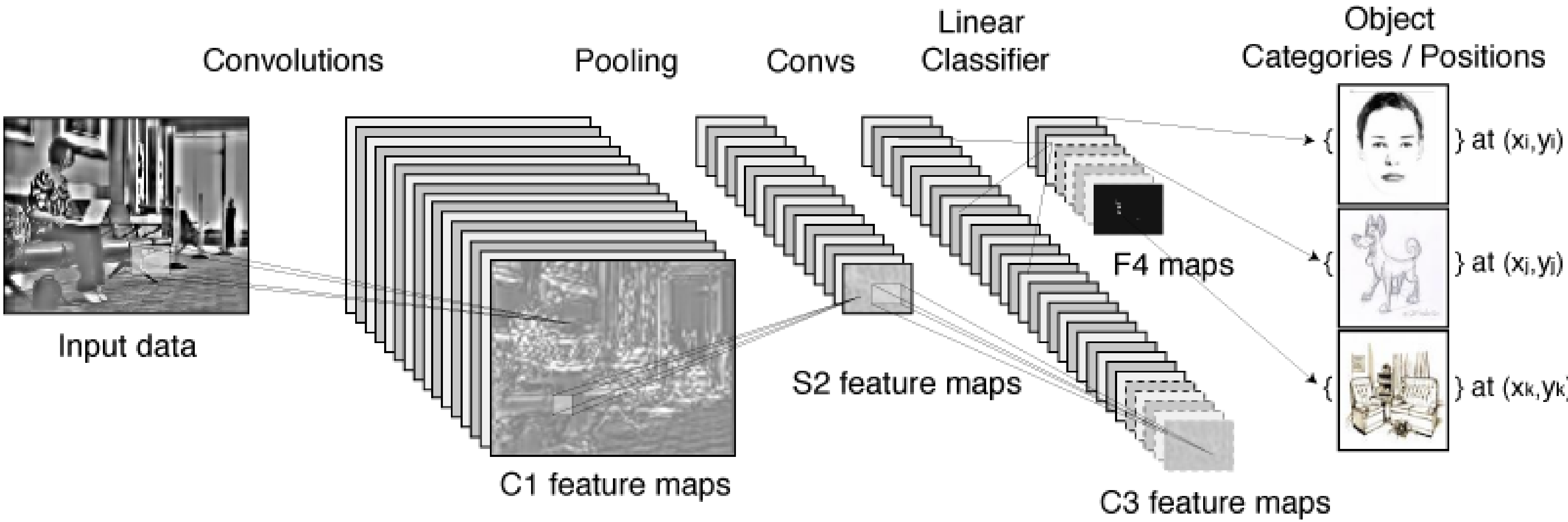
40% ε improvement

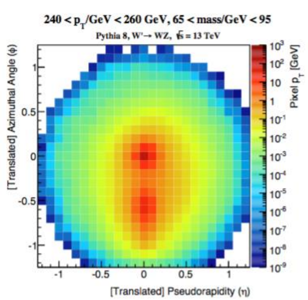
Neutrino interaction classification

Using Convolutionnal Neural Network (GoogleNet)

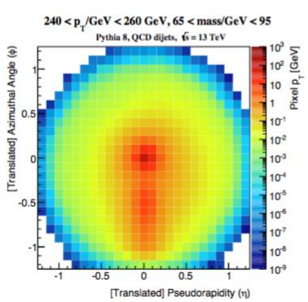
Actually used in physics results 1703.03328 and 1706.04592

Réseau de Neurones Convolutifs (CNN)





Higgs Lourd



Bruit de fond

‘Godfathers of AI’ honored with Turing Award, the Nobel Prize of computing

3

Yoshua Bengio, Geoffrey Hinton, and Yann LeCun laid the foundations for modern AI

By James Vincent | Mar 27, 2019, 6:02am EDT

2019

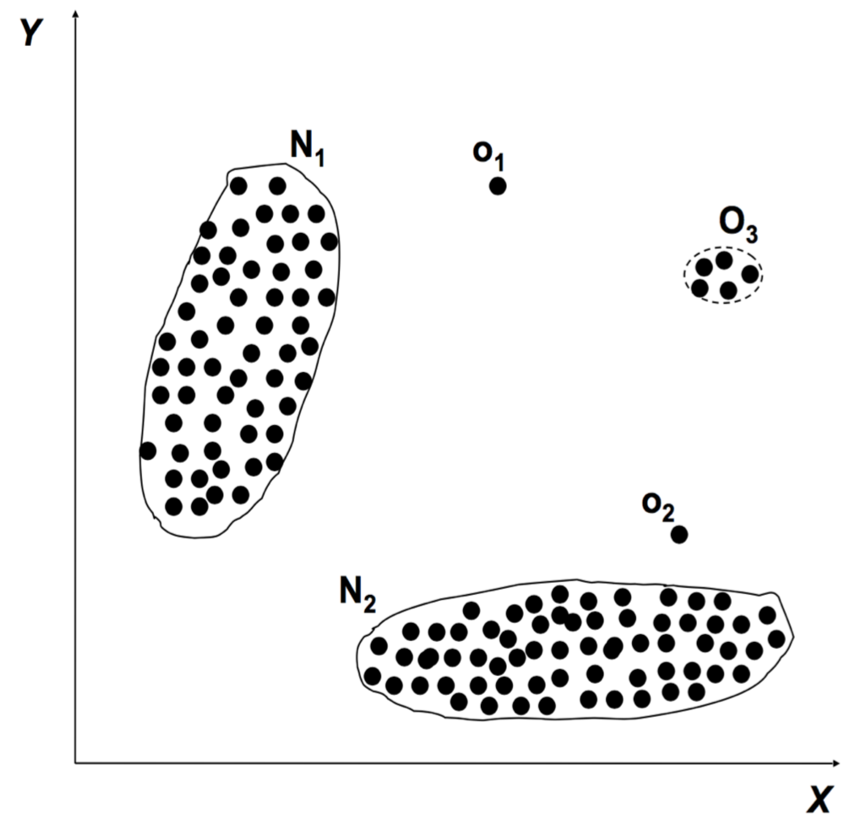
SHARE

Détection d’anomalie

Détection d’anomalie

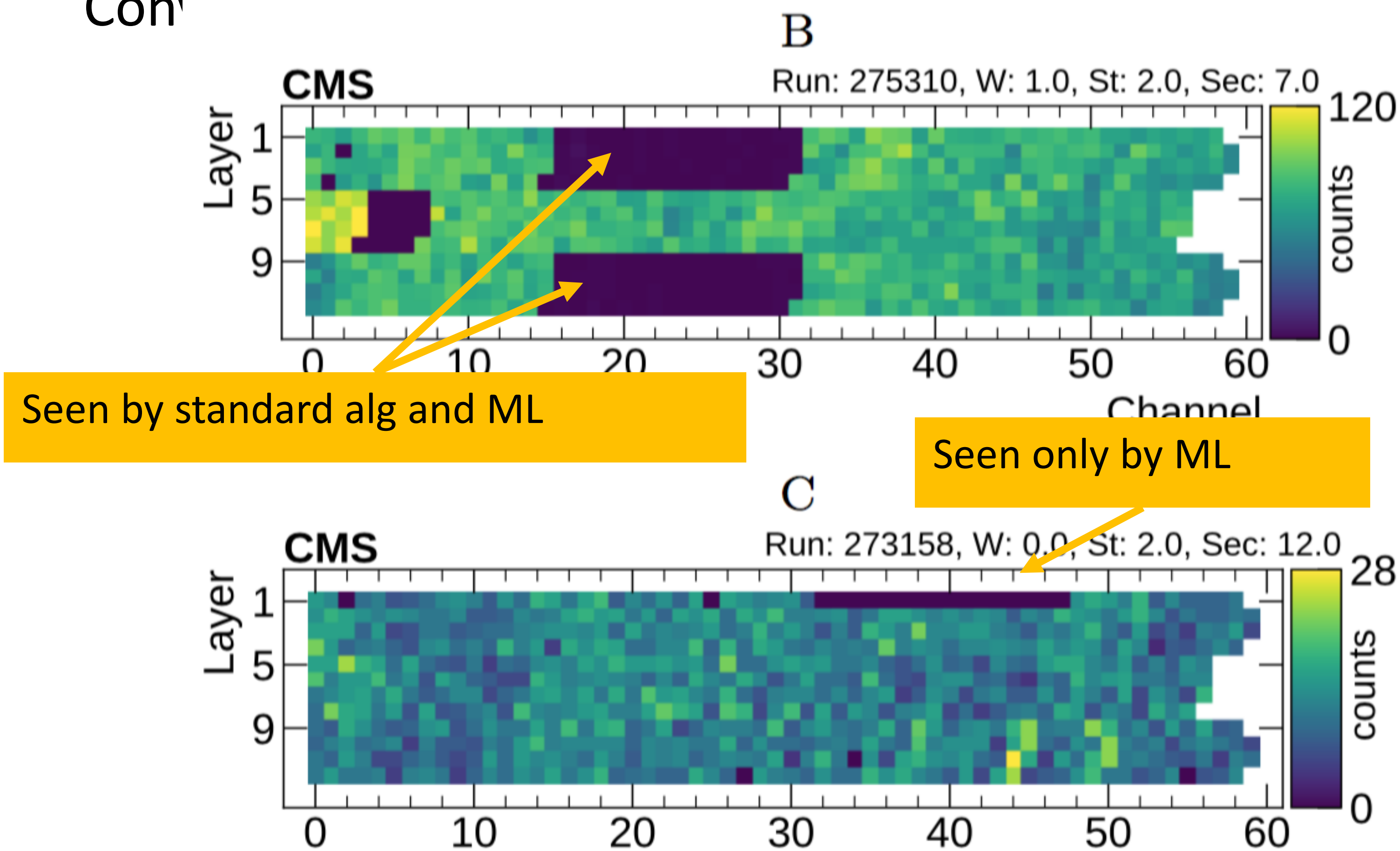
- Comment détecter les anomalies “O”utliers par rapport aux “N”ucleus
- Méthode supervisée : on a un modèle pour N et O
- Méthode non-supervisée : pas de modèles
- Méthode semi-supervisée : modèle pour N, pas pour O

- Applications : détection de spam, détection de pannes



Application au monitoring

- Example application CMS muon chamber monitoring (with Con



Demo on real data.

Application à la physique

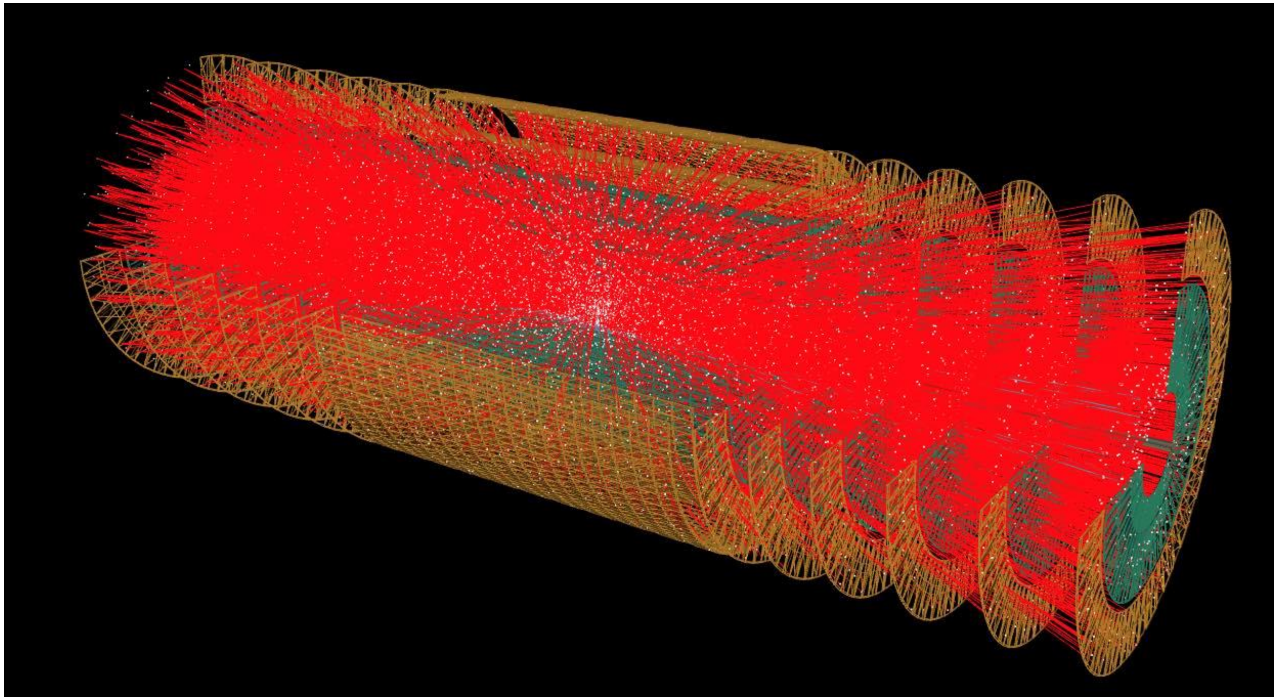
Application à la physique

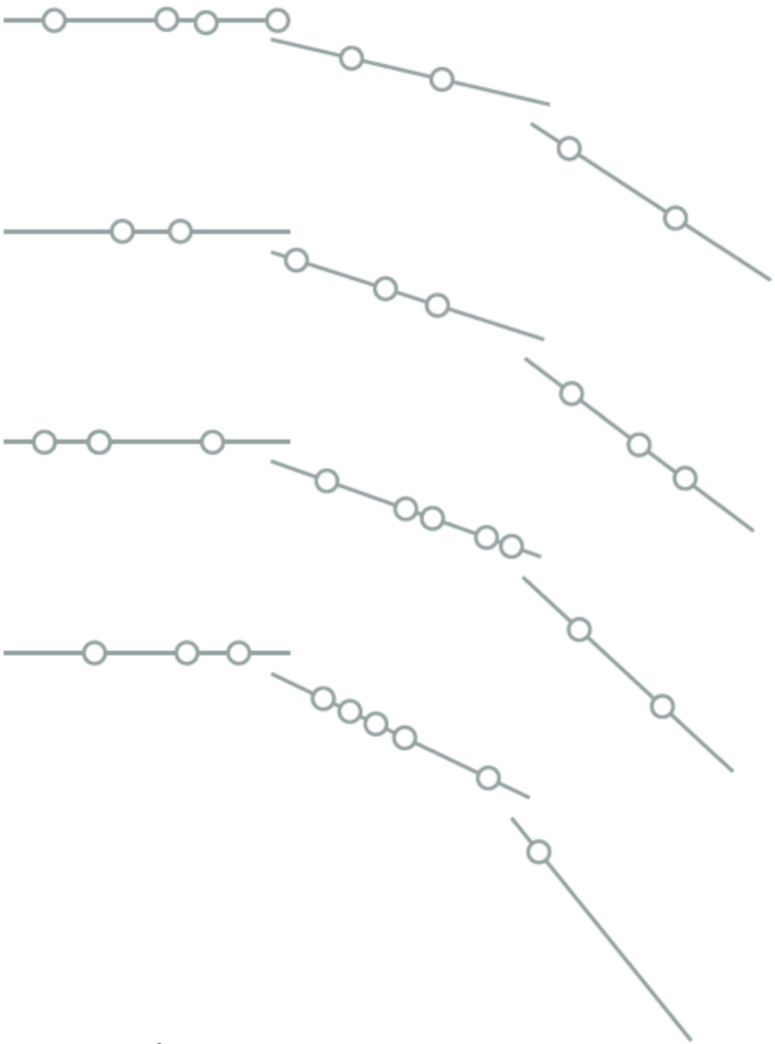
Tracking challenge

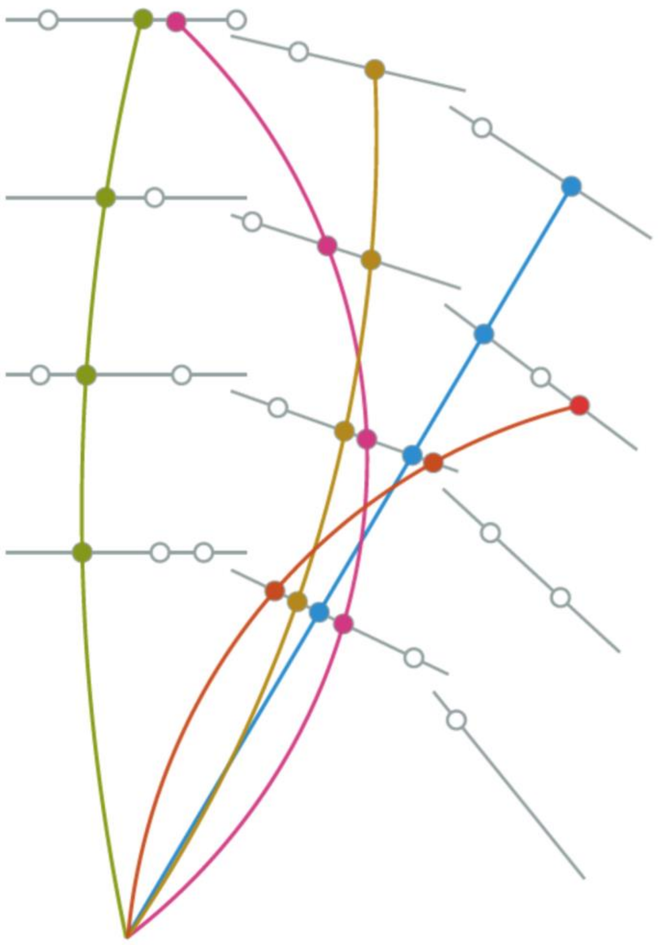
https://sites.google.com/site/trackmlparticle/
https://www.kaggle.com/c/trackml-particle-identification
https://twitter.com/trackmllhc

L’enjeu

- LHC en 2025
- Augmentation du nombre de collisions de proton
- ➔superposition d’événements ~200 (contre ~50 maintenant)
- ➔événements tres complexe
- En particulier pour la trajectographie : relier les points
- Algorithmes actuels ~10 fois trop lents







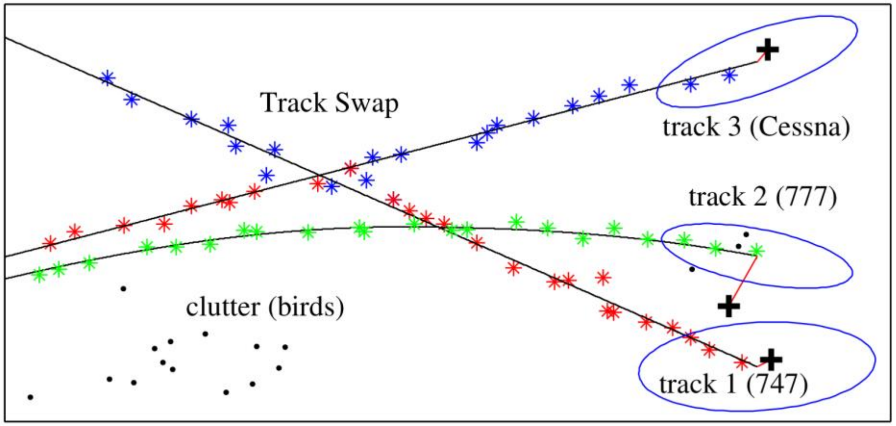
Mais :

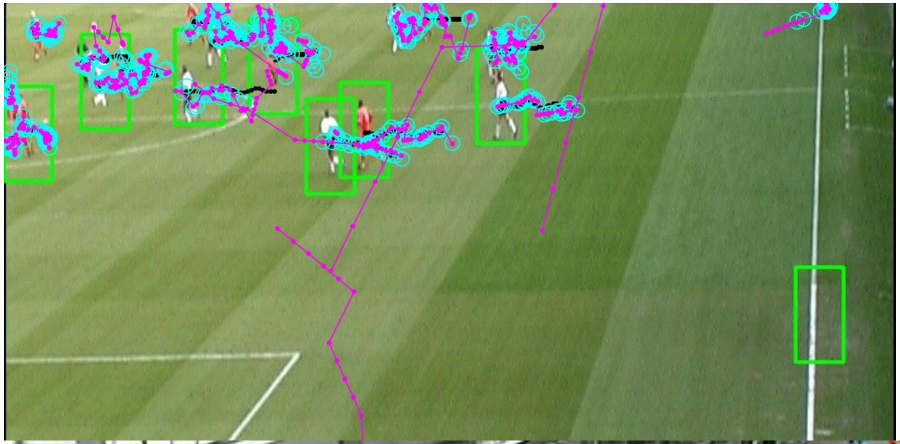
- 3 dimensions
- 10'000 traces x 10 points

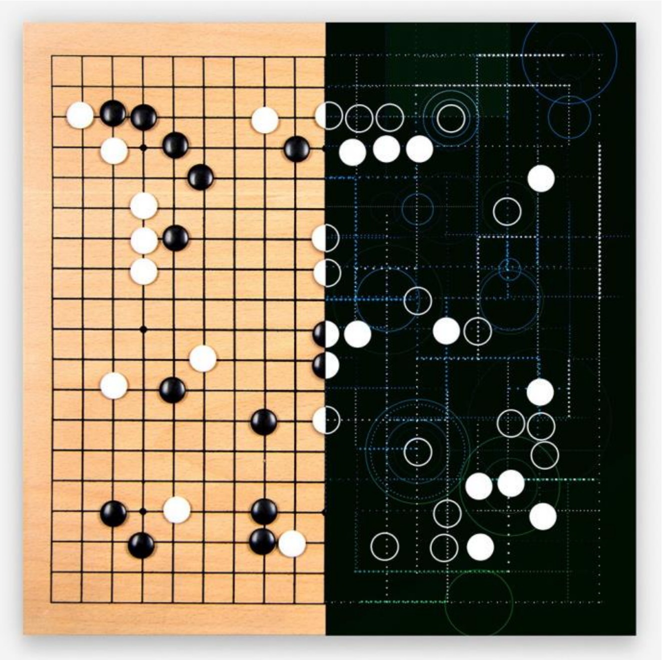
Pourquoi c'est difficile ?

- 100'000 points à grouper en 10'000 traces de 10 points
  - ➔ $\sim 10^{450'000}$ combinaisons
  - ⇒la "force brutale" n'a (vraiment) aucune chance
- Precision des points : ~50μm sur un volume ~40 m^3
  - ➔ $3\ 10^{14}$ voxels!
  - projection 2D ➔ $2\ 10^9$ pixels !
  - ⇒ algorithme de traitement d'image n'ont (vraiment) aucune chance
- Pas un problème classique

Trajectographie (Tracking)

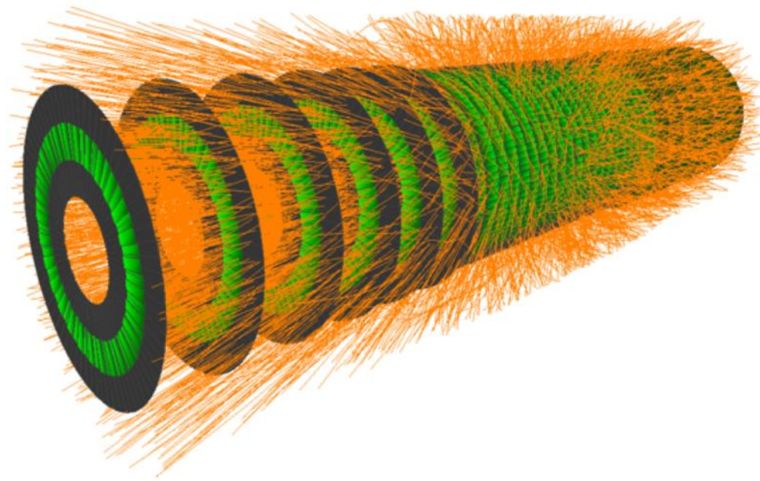






Tracking Machine Learning challenge

https://twitter.com/trackmllhc



Idée : mettre les points (x,y,z) sur le web

demander aux participants de relier les points

Objectif : obtenir de nouveaux algorithmes

- Accuracy phase sur Kaggle, uniquement la précision : mai-aout 2018
- Throughput phase (compétition officielle NeurIPS 2018) prévision + vitesse : octobre 2018-mars 2019

Dataset

- Hit file

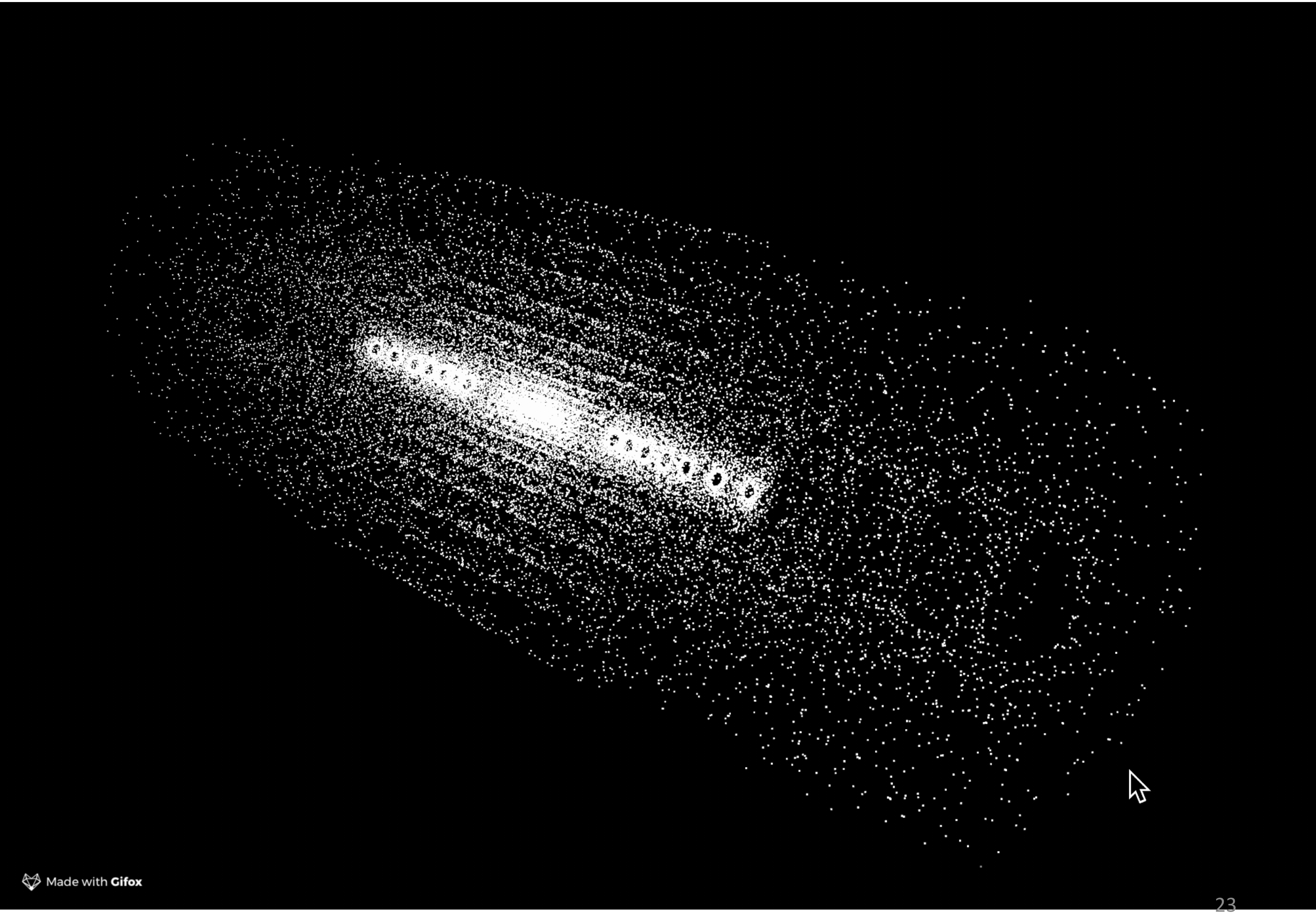
(measured position mm)

	hit_id	x	y	z	volume_id	layer_id	module_id
0	1	-64.409897	-7.163700	-1502.5	7	2	1
1	2	-55.336102	0.635342	-1502.5	7	2	1
2	3	-83.830498	-1.143010	-1502.5	7	2	1
3	4	-96.109100	-8.241030	-1502.5	7	2	1
4	5	-62.673599	-9.371200	-1502.5	7	2	1

- Truth file

(true position mm   particle momentum GeV)

	hit_id	particle_id	tx	ty	tz	tpx	tpy	tpz	weight
0	1	0	-64.411598	-7.164120	-1502.5	250710.000000	-149908.000000	-956385.00000	0.000000
1	2	22525763437723648	-55.338501	0.630805	-1502.5	-0.570605	0.028390	-15.49220	0.000010
2	3	0	-83.828003	-1.145580	-1502.5	626295.000000	-169767.000000	-760877.00000	0.000000
3	4	297237712845406208	-96.122902	-8.230360	-1502.5	-0.225235	-0.050968	-3.70232	0.000008
4	5	418835796137607168	-62.659401	-9.375040	-1502.5	-0.281806	-0.023487	-6.57318	0.000009

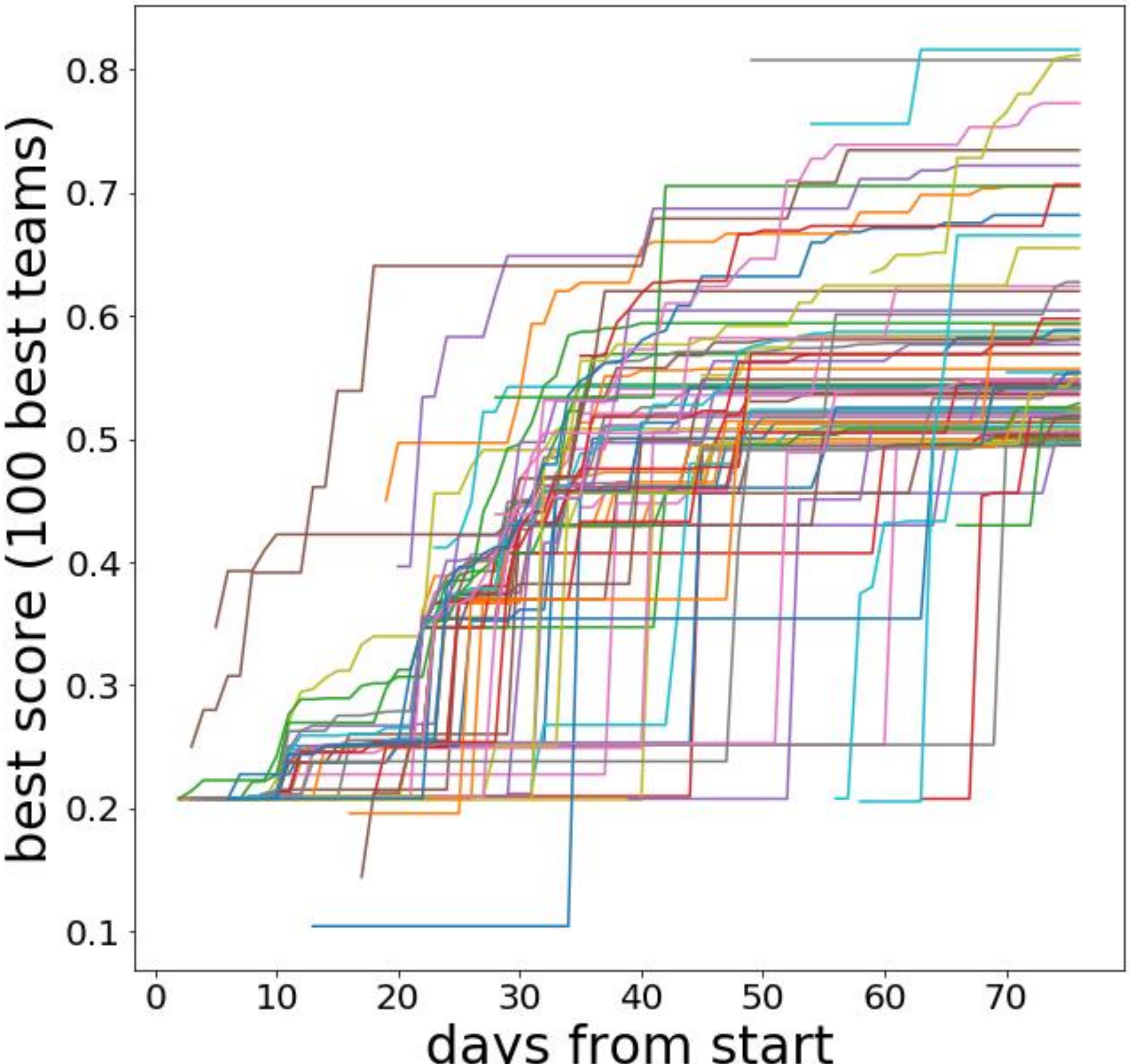

Made with Gifox

Visualisation spin-off

- Visit at CERN Tobias Isenberg visualisation scientist at LRI-Orsay with PhD student Xiyao Wang
- Will use TrackML dataset to experiment with visualisation/interaction with Microsoft' Hololens

Leaderboard

#	△1w	Team Name	Kernel	Team Members	Score	Entries	Last
1	—	Edwin Steiner			0.8161	2	13d
2	▲ 2	demelian			0.8118	16	9h
3	▼ 1	outrunner			0.8070	1	1mo
4	▼ 1	yuval r			0.7726	19	5h
5	—	Mickey			0.7345	10	18d
6	—	Zidmie			0.7220	11	10d
7	▲ 3	On est les champions !!!!!			0.7065	10	2d
8	▼ 1	icecuber			0.7054	3	1mo
9	▼ 1	Vicens Gaitan			0.7049	13	6d
10	▼ 1	Félicitations à la France			0.6818	42	6d
11	—	Victor Nedel'ko			0.6653	4	10d
12	—	John Sweeney			0.6551	16	5d
13	▲ 3	trian2018			0.6274	23	21h
14	▼ 1	Seb B			0.6240	20	15d
15	▼ 1	Grzegorz Sionkowski			0.6200	22	25d
16	▼ 1	Andrea Lonza			0.6043	7	1mo



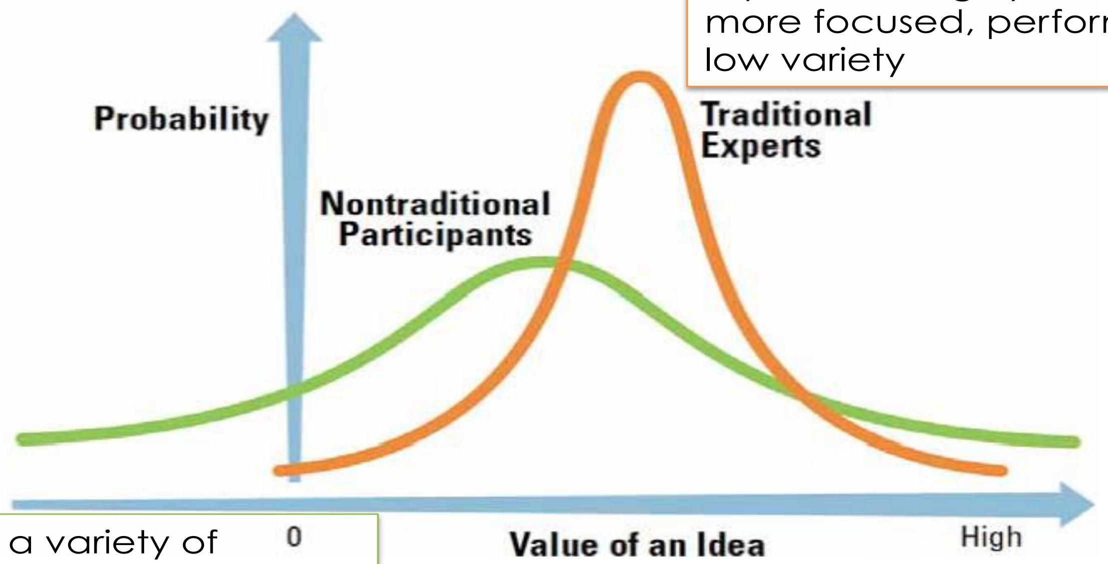
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
best score (100 best teams)
0
10
20
30
40
50
60
70
days from start

Why challenges work ?

MOTIVATION OF ORGANIZING CONTESTS: EXTREME VALUE

Courtesy : Lakhani 2014

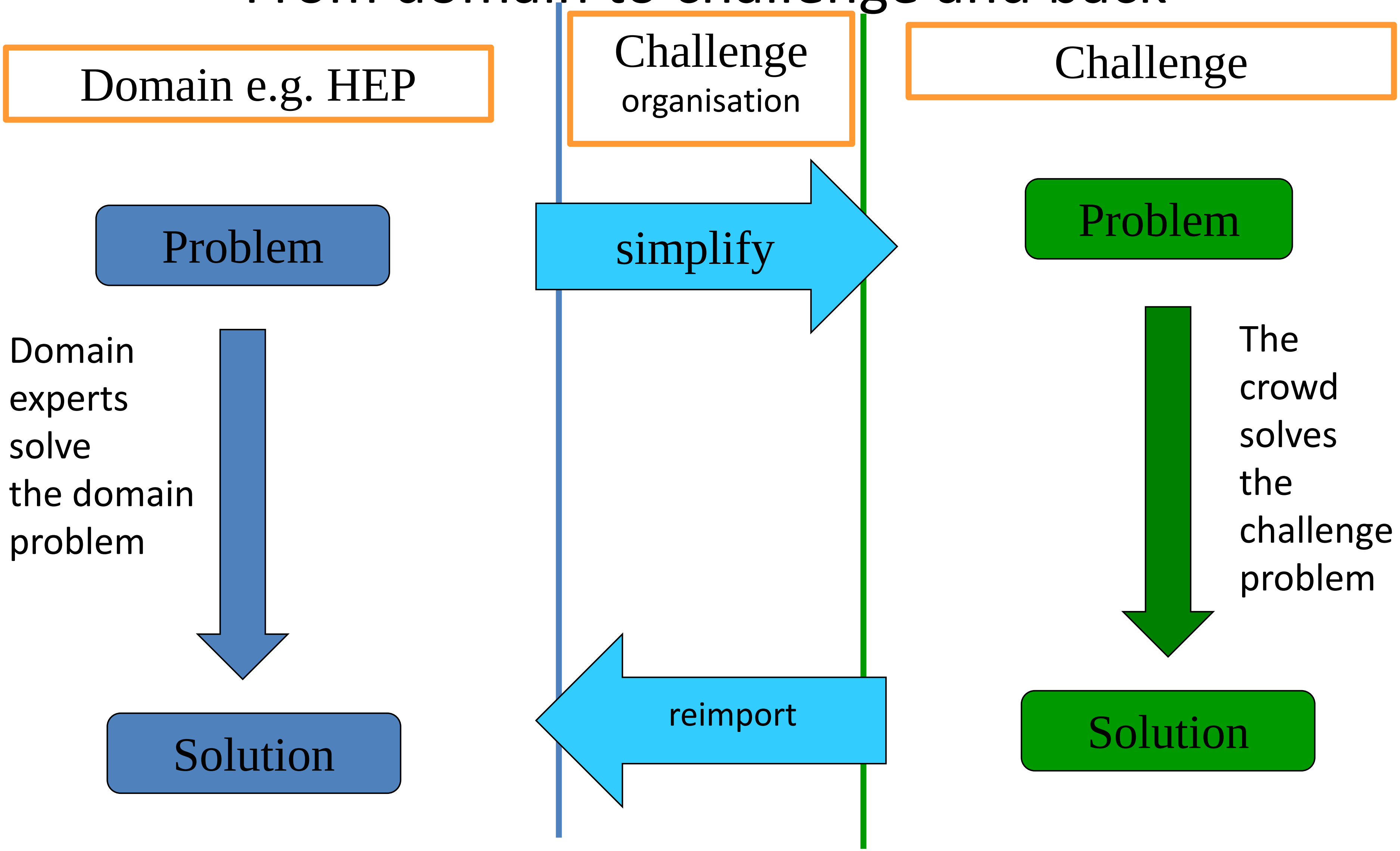
Experts are highly skilled, trained - > more focused, performed solution, low variety



OI is suitable for a variety of nonconventional surprising ideas that are « far » from traditional expertise - > high volatility

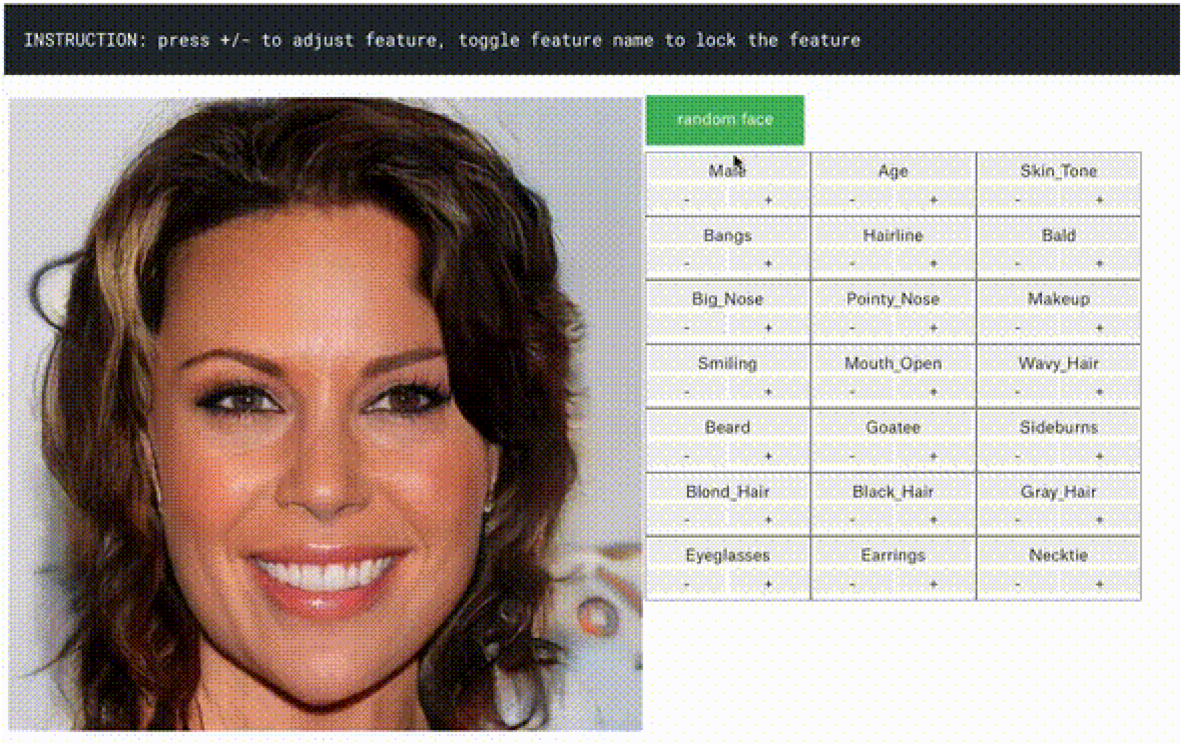
Not just ML, but a general trend:
Open Innovation

From domain to challenge and back

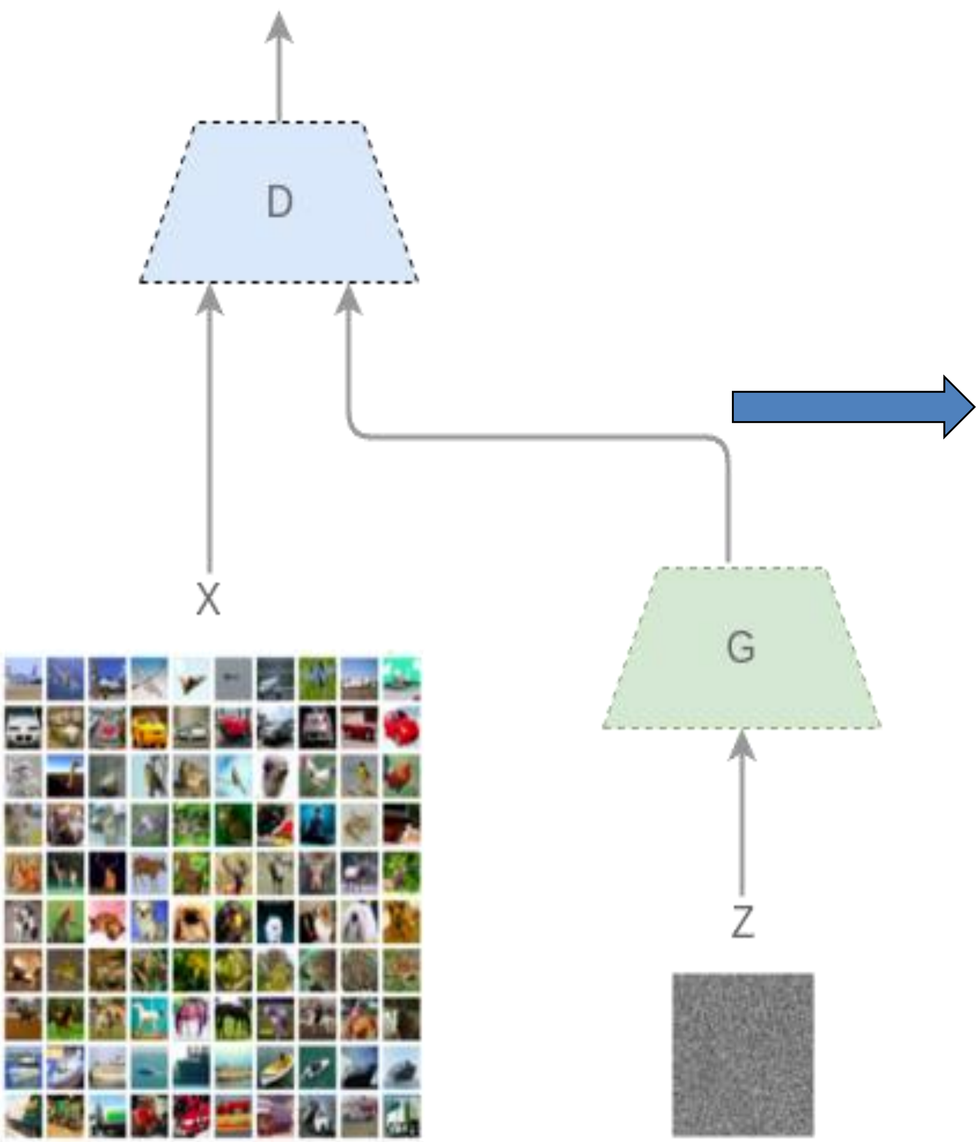


Generative Adversarial Network

Generative model

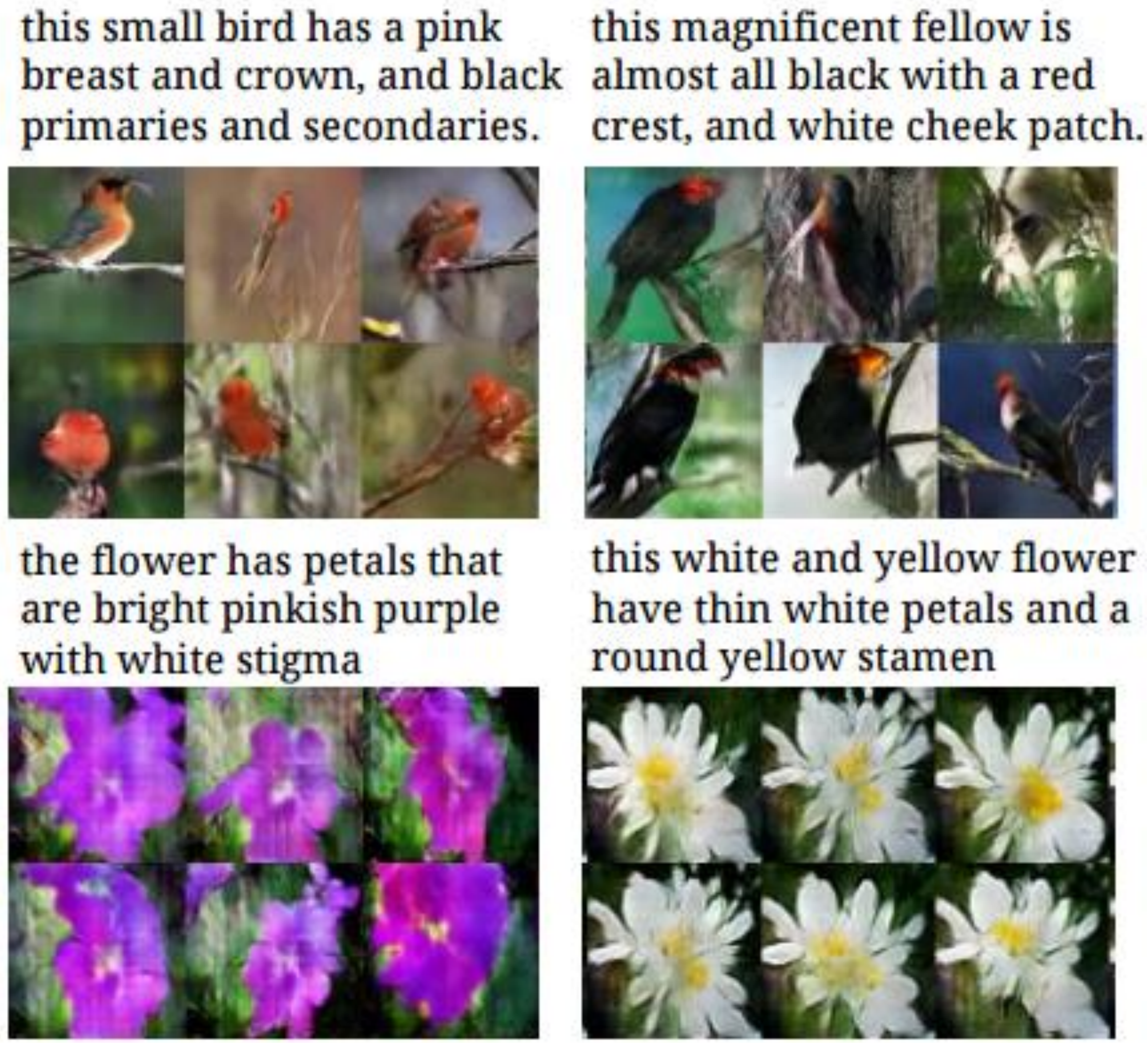


Réseau Génératif Adversaire (Generative Adversarial Network GAN)



Condition GAN

Text to image

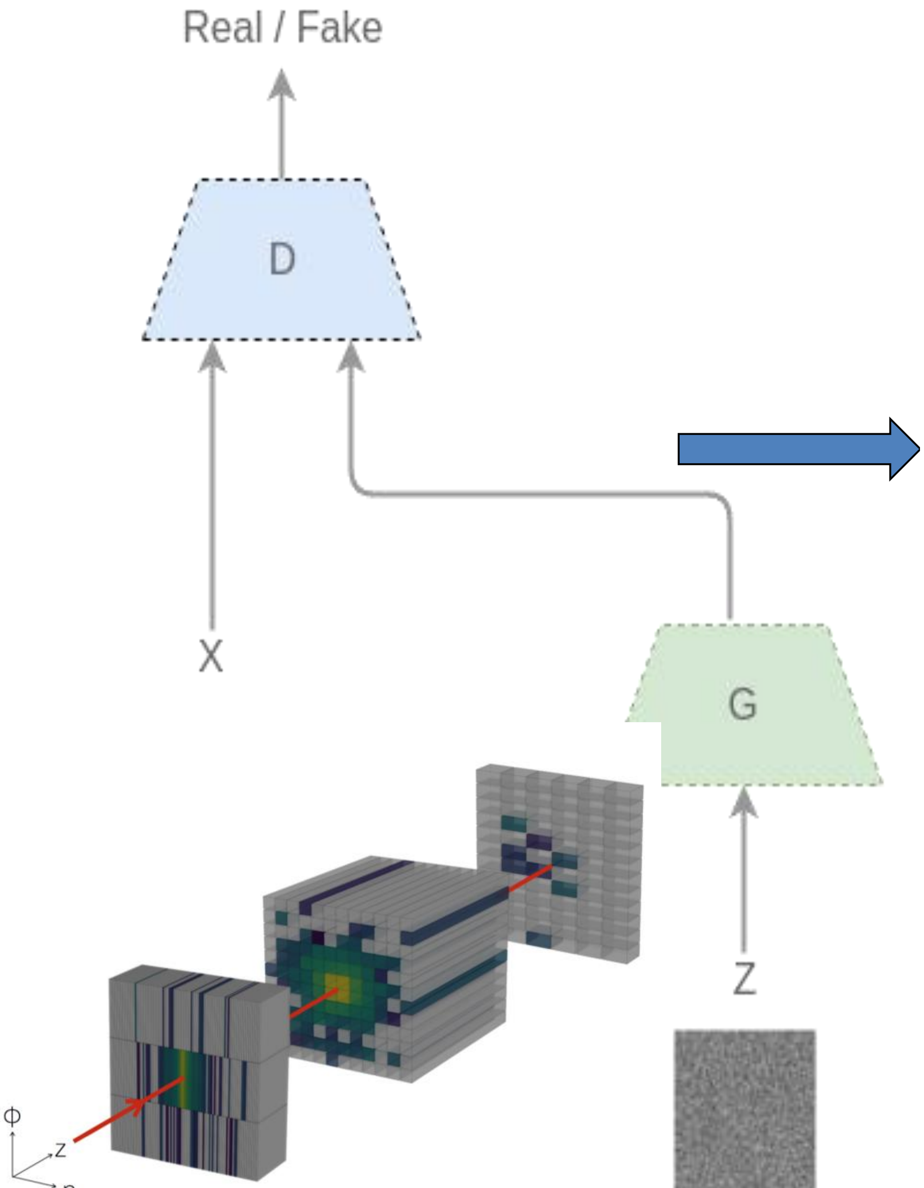


Text to image

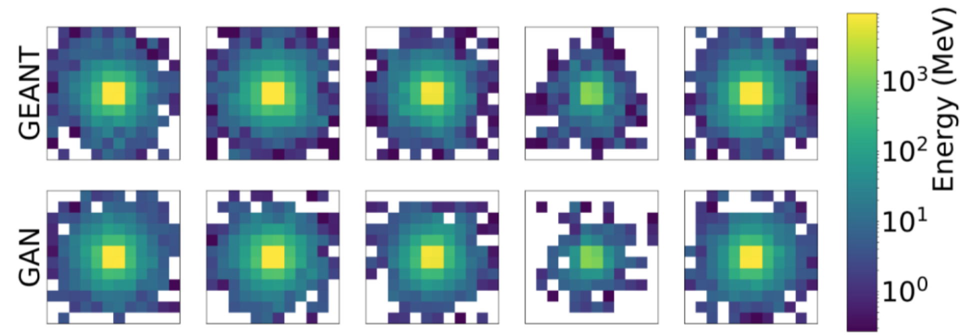
Dall.e

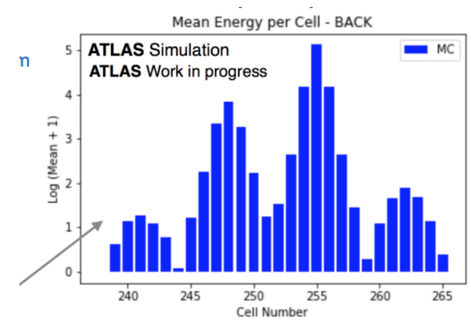
2022 « Large Foundation Model »

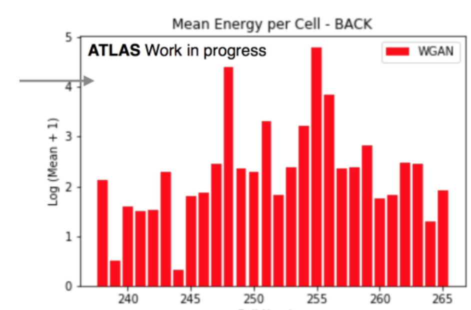
GAN pour la simulation de particules



Gain en temps de calcul x1000







Attention aux surprises

- https://twitter.com/goodfellow_ian/status/937406530743287808

La méthode scientifique

Publication scientifique

Phys.Rev.Lett. 114 (2015)191803

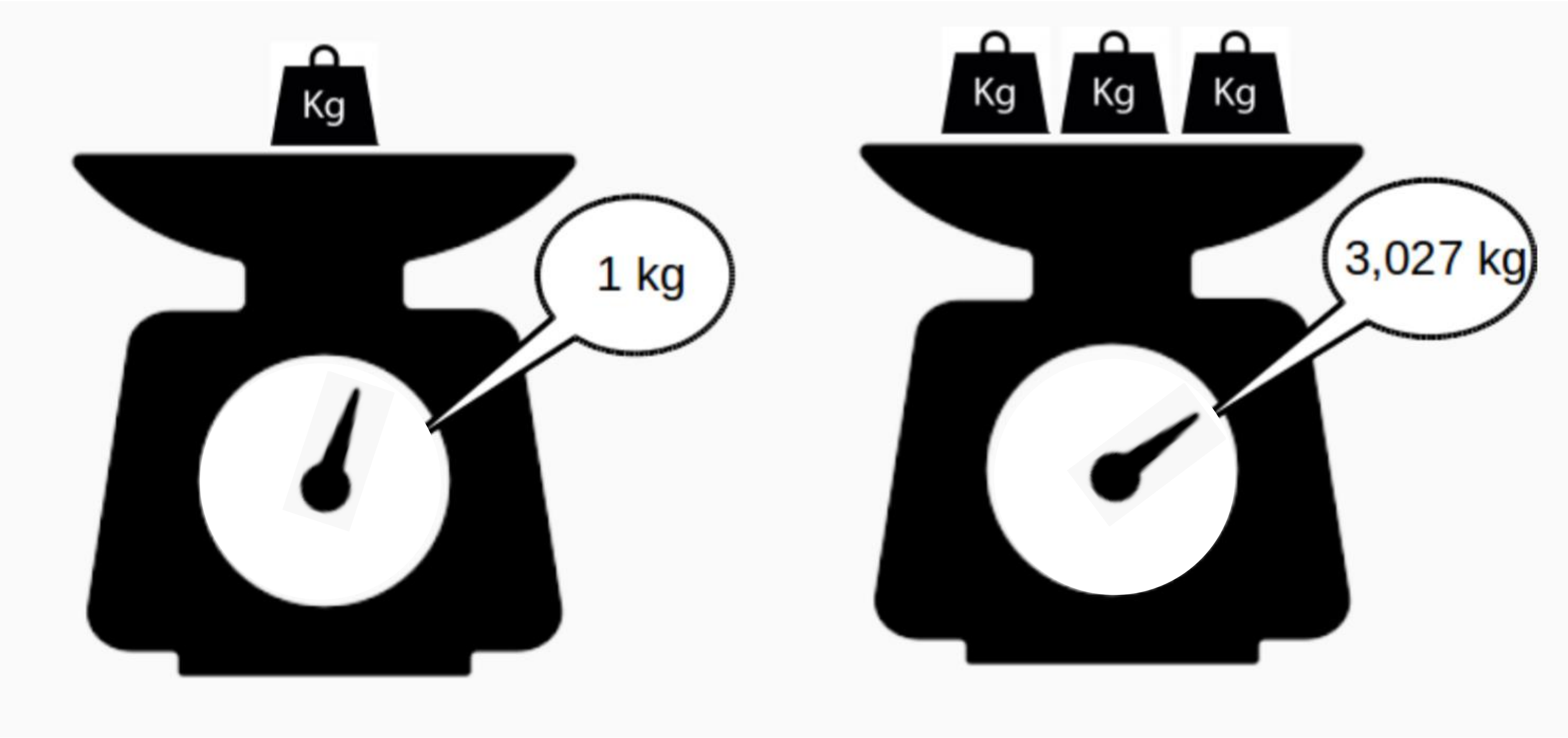
Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments

(ATLAS Collaboration)†
(CMS Collaboration)‡
(Received 25 March 2015; published 14 May 2015)

A measurement of the Higgs boson mass is presented based on the combined data samples of the ATLAS and CMS experiments at the CERN LHC in the $H \to \gamma\gamma$ and $H \to ZZ \to 4\ell$ decay channels. The results are obtained from a simultaneous fit to the reconstructed invariant mass peaks in the two channels and for the two experiments. The measured masses from the individual channels and the two experiments are found to be consistent among themselves. The combined measured mass of the Higgs boson is $m_H = 125.09 \pm 0.21\,(\mathrm{stat}) \pm 0.11\,(\mathrm{syst})$ GeV.

Incertitudes systématiques :
tout ce dont on n'est pas sur…
Et si ? Et si ? ... Les inconnues connues et inconnues jusqu'à ce que nos collègues soient convaincus

Biais expérimental



Pres de 500 termes de systématiatiques

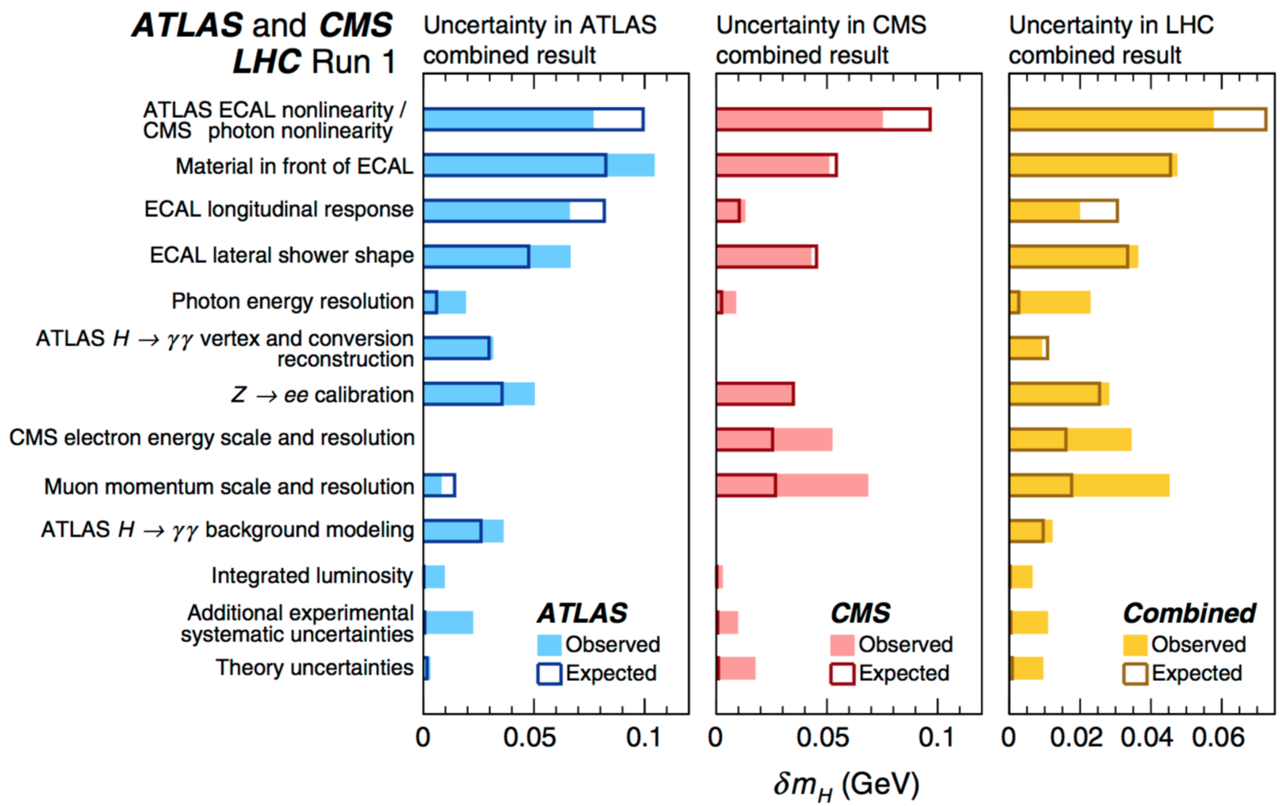


FIG. 3 (color online). The impacts δm_H (see text) of the nuisance parameter groups in Table I on the ATLAS (left), CMS (center), and combined (right) mass measurement uncertainty. The observed (expected) results are shown by the solid (empty) bars.

>>>89-90: Where is this measurement documented in this note?
**That has now been finished and published, and references will be updated in the notes
JHEP 09 (2014) 079 arXiv:1407.5532

>>>116: Doesn't "3 GeV" contradict your previous statement that we cover a pT range "from 1 GeV..."
**No, those muons with pT~1 GeV and p~3 GeV are at high eta. But this is all irrelevant, as we only use muons with pT>4 GeV.

>>>116: I'm not sure what you mean when you say that they "require" the combined reconstruction. I would just say "Muons used in this analysis are reconstructed using a statistical combination of an MS track and ID track."
**Only the muons that have a successful statistical combination of an MS track and ID track are used in this analysis, the text is updated.

>>>128: what is fully efficient for signal candidates? It's not clear what the word "which" refers to.
**This related to the vertex quality criteria applied to the fit.

>>>142: Please mention the total number of di-muon candidates here.
**This is added to the text, 7.8M for 2011, and ~65M for 2012.

>>>171: could you provide more information about how the psi is assigned to a primary vertex? I'm surprised by your statement that "few" events contain multiple vertices; I thought pileup was a significant issue at 7 TeV.
** Not really. The only relevance the primary vertex has is to measure Lxy, which is measured in the transverse direction only, and hence is not changing much from one collision vertex to another. However, the determination of the primary vertex position depends on whether the two muon tracks were used in its fit or not, hence we need to know which vertex was it. But at 7 TeV there was not much ambiguity, the vertex which jpsi came from was almost always the main primary one. Studies from the 2011 Jpsi Phi analysis -- where vertex choice may have been an issued-- showed that there was no impact in the few cases of an incorrect choice of vertex.

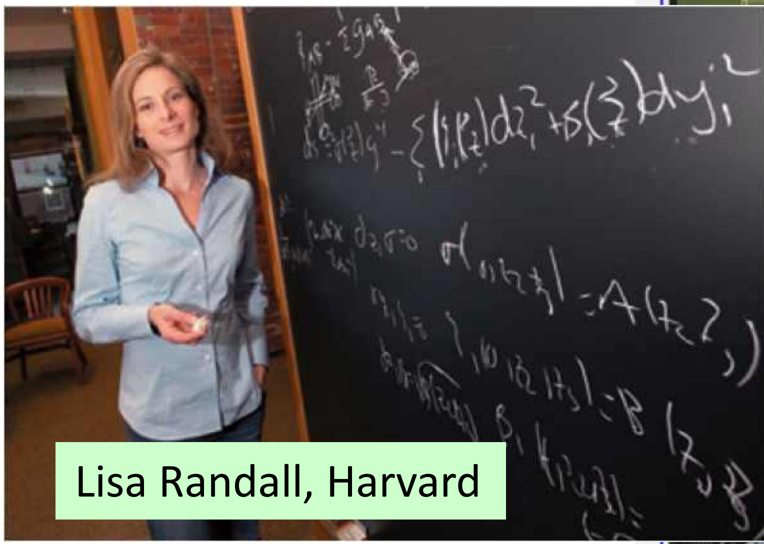
>>>Fig 1: Do you understand the eta dependence of this plot? What is the z-axis -- number of events per bin? The bins are extremely small -- what size bins did you choose?
** Yes. This is the scatter plot of dimuon candidates. The x-axis is the absolute rapidity "y" of the dimuon candidate, with the structure roughly reflecting the (smeared) geometry of the muon chambers, with dips near/around the cracks and edges.
The z-axis is just the candidate yield, the bin-width in |y| is 5e-3 and in pT is 320 MeV

>>>228: The phi* definition is unclear to me, in particular what the "psi production" is. Do you mean the psi momentum vector?
**phi* is the angle between the psi production plane (defined by psi momentum and colliding proton

ue

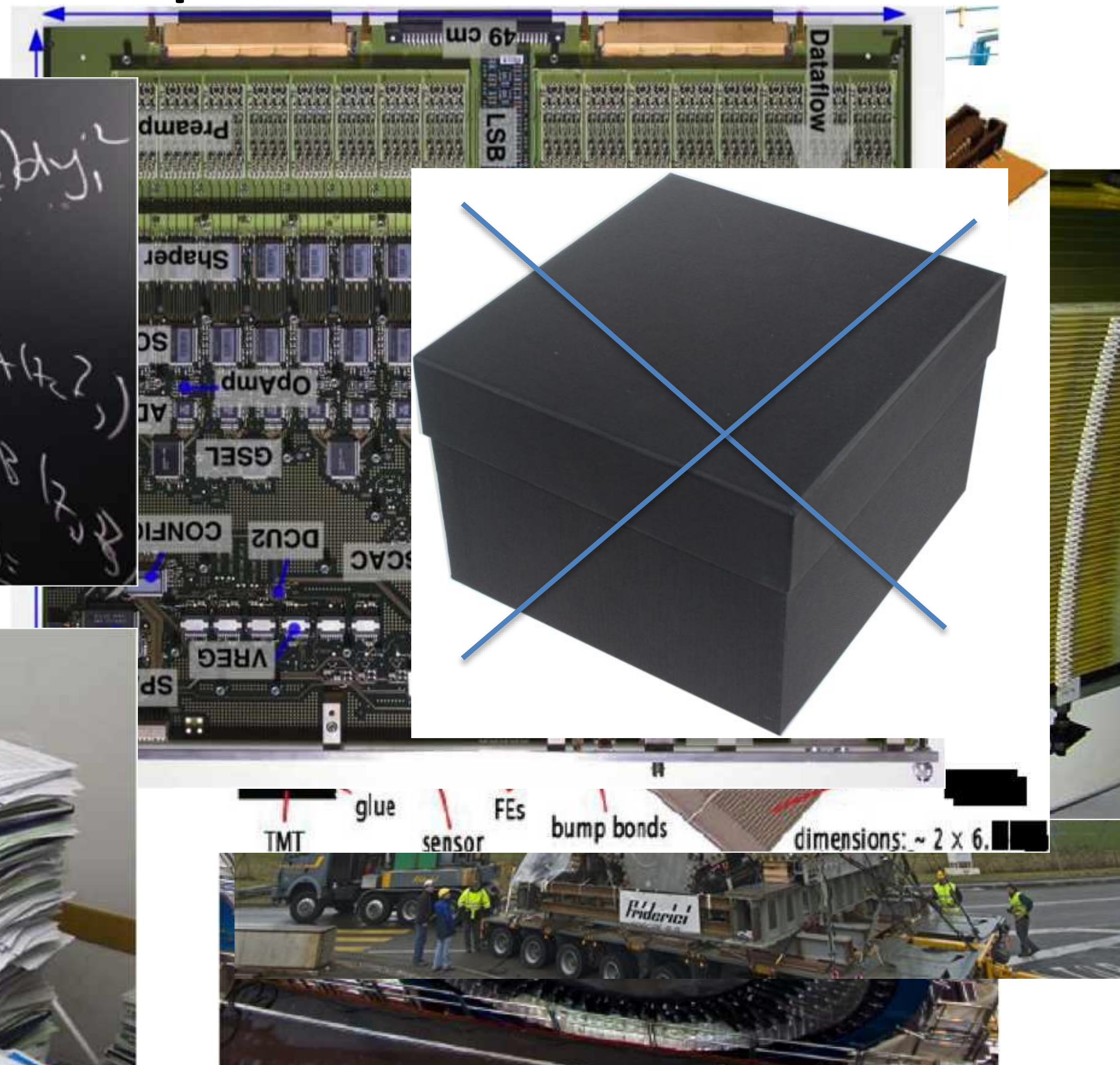
Près de 200 pages et plusieurs semaines de questions/réponses avant l'approbation de la publication !

Trust but verify : de la théorie à l'expérience

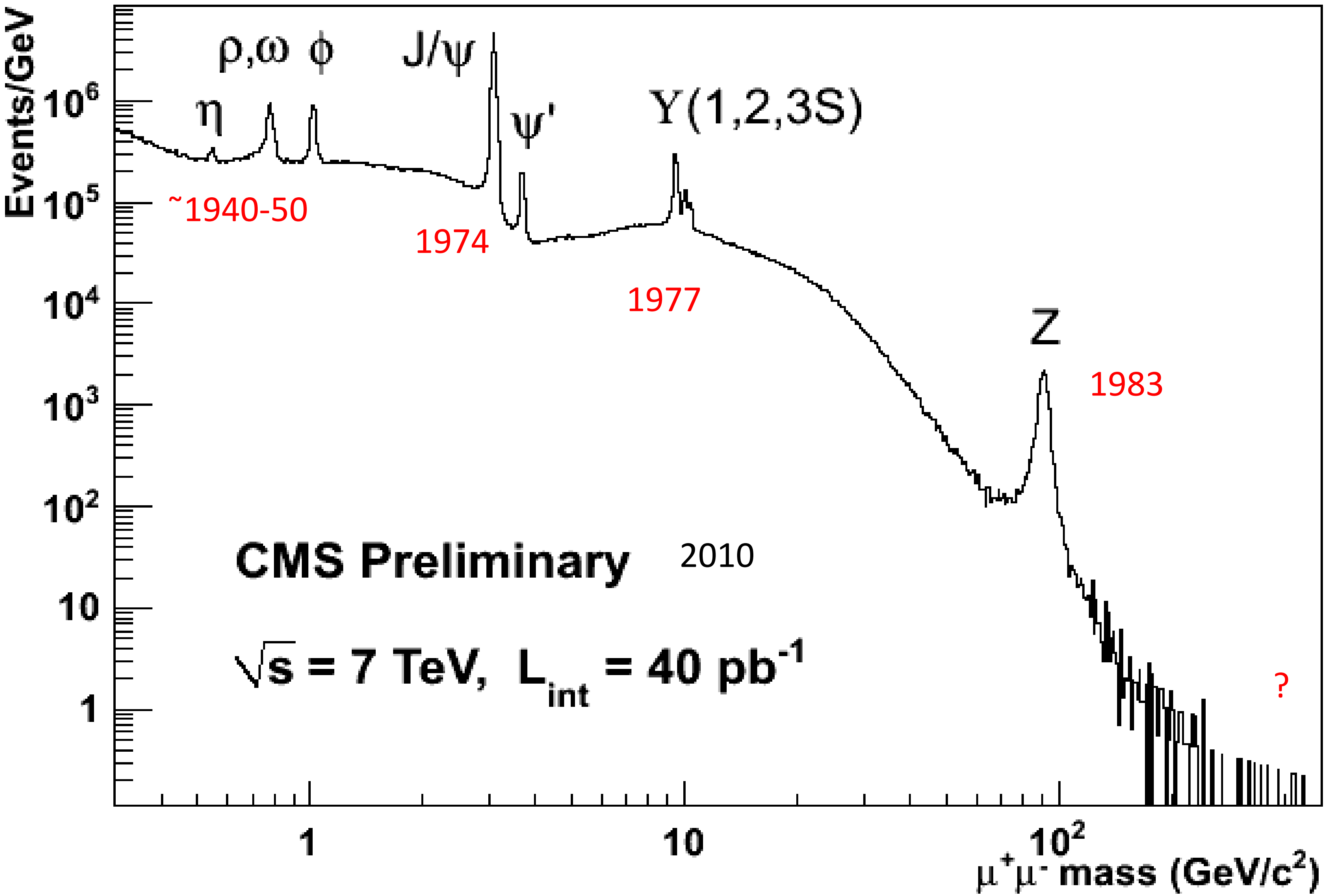


Lisa Randall, Harvard

John Ellis, CERN



Ré-observation de particules connues



ML et erreur systématique

- Nos papiers expérimentaux se concluent souvent par
  - measure = m ± σ(stat) ± σ(syst)
  - σ(syst) erreur systématique : inconnues connues, inconnues inconnues
- On cherche a minimiser la somme quadratique:

$$\sigma(\text{stat}) \pm \sigma(\text{syst})$$

- Technique ML standard minimise σ(stat)

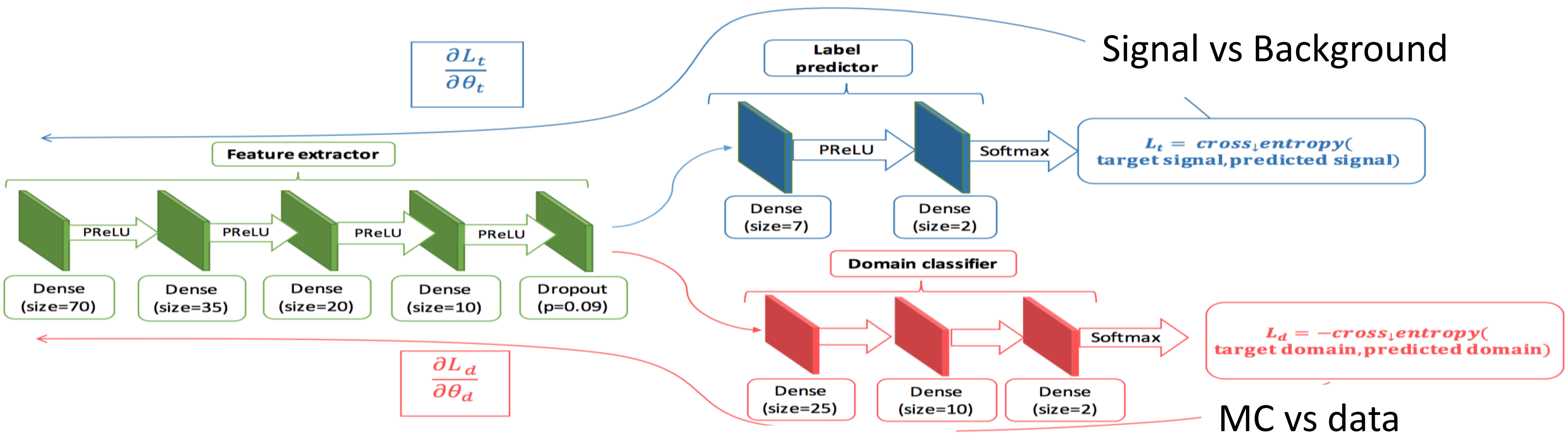
❑ Comment indiquer au ML de minimiser σ(stat) ± σ(syst) ?

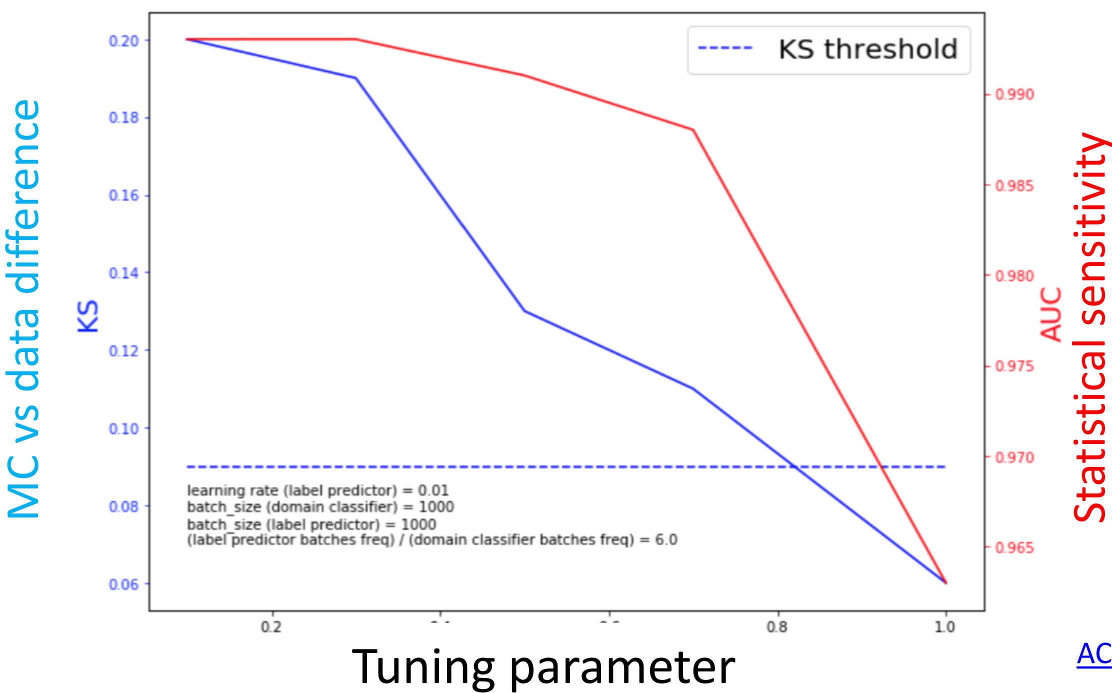
Effet systématique

Entrainement sur chiffres manuscrits. Effet de l’angle

Entrainement adversaire

Inspired from 1505.07818 Ganin et al :





[ACAT 2017 Ryzhikov and Ustyuzhanin]

Transfert de style

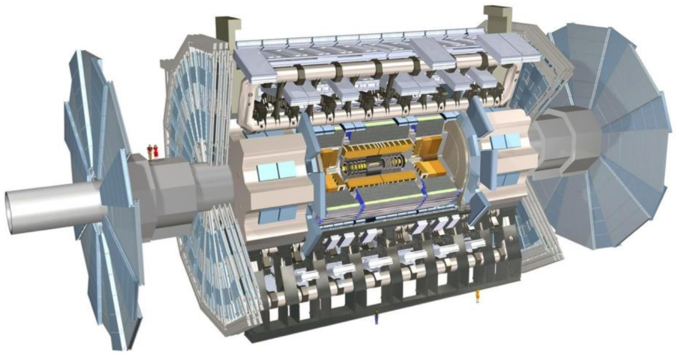
CycleGAN

photo → Monet

Monet → photo

CycleGAN

CycleGAN



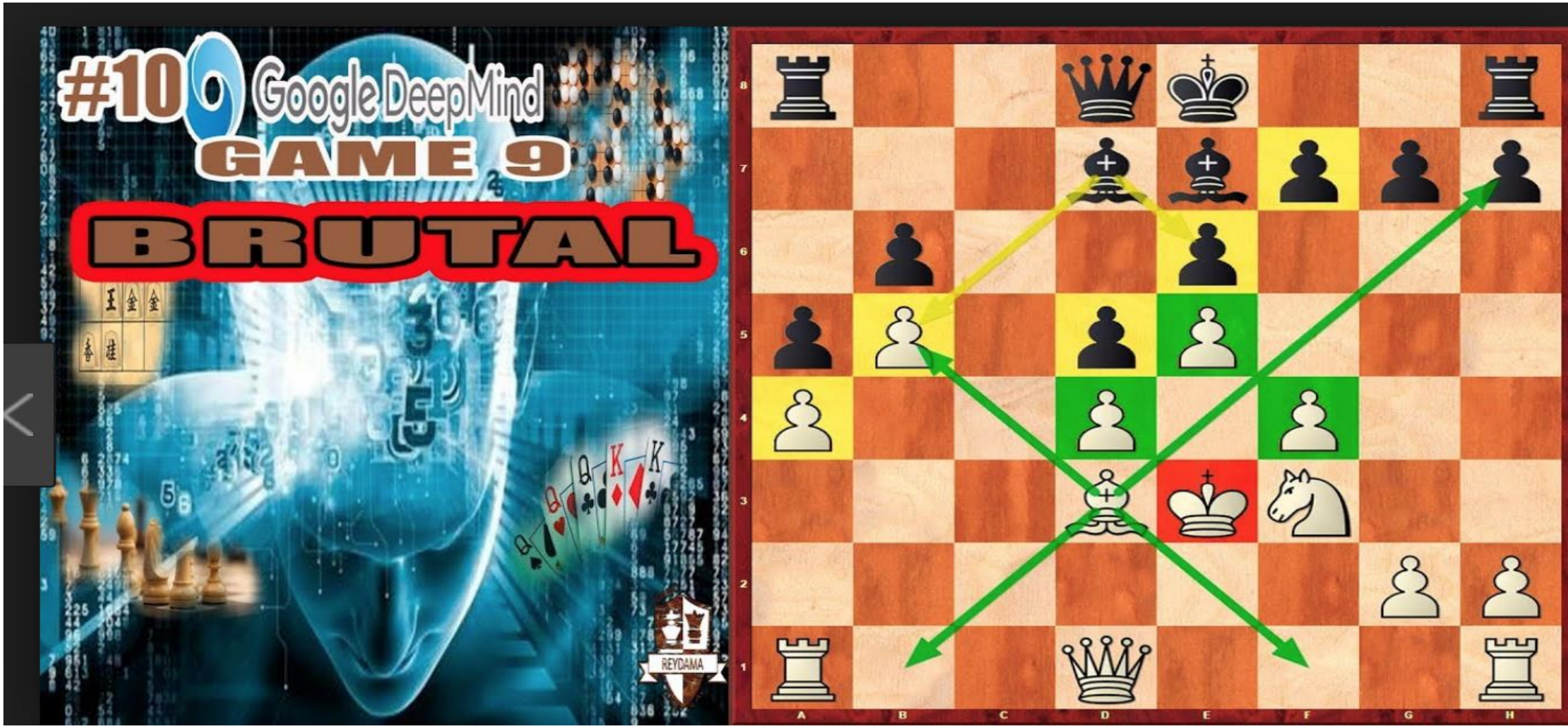
Données

Simulation

Pouvons nous faire confiance aux simulations ?

Apprentissage par renforcement

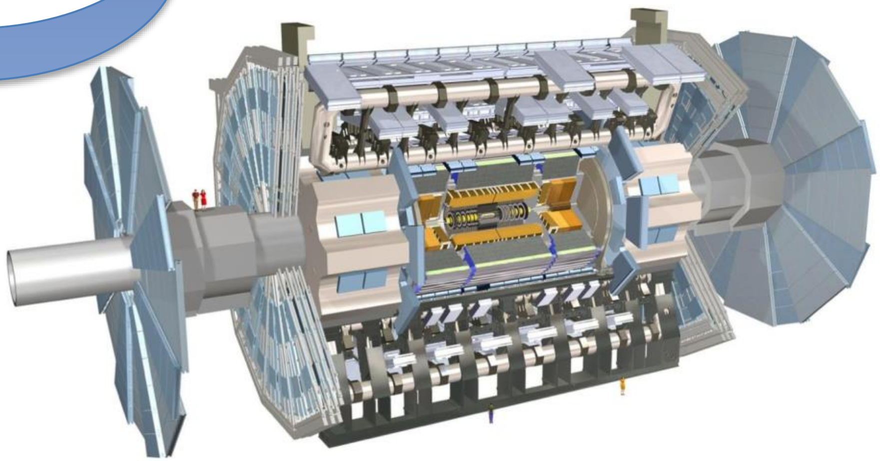
Apprentissage par renforcement



Alpha Zero : a appris tout seul contre lui-même en quelques jours à jouer (séparément) au Go, aux échecs ou aux échecs japonais, et bat tout le monde, homme ou machine

Application

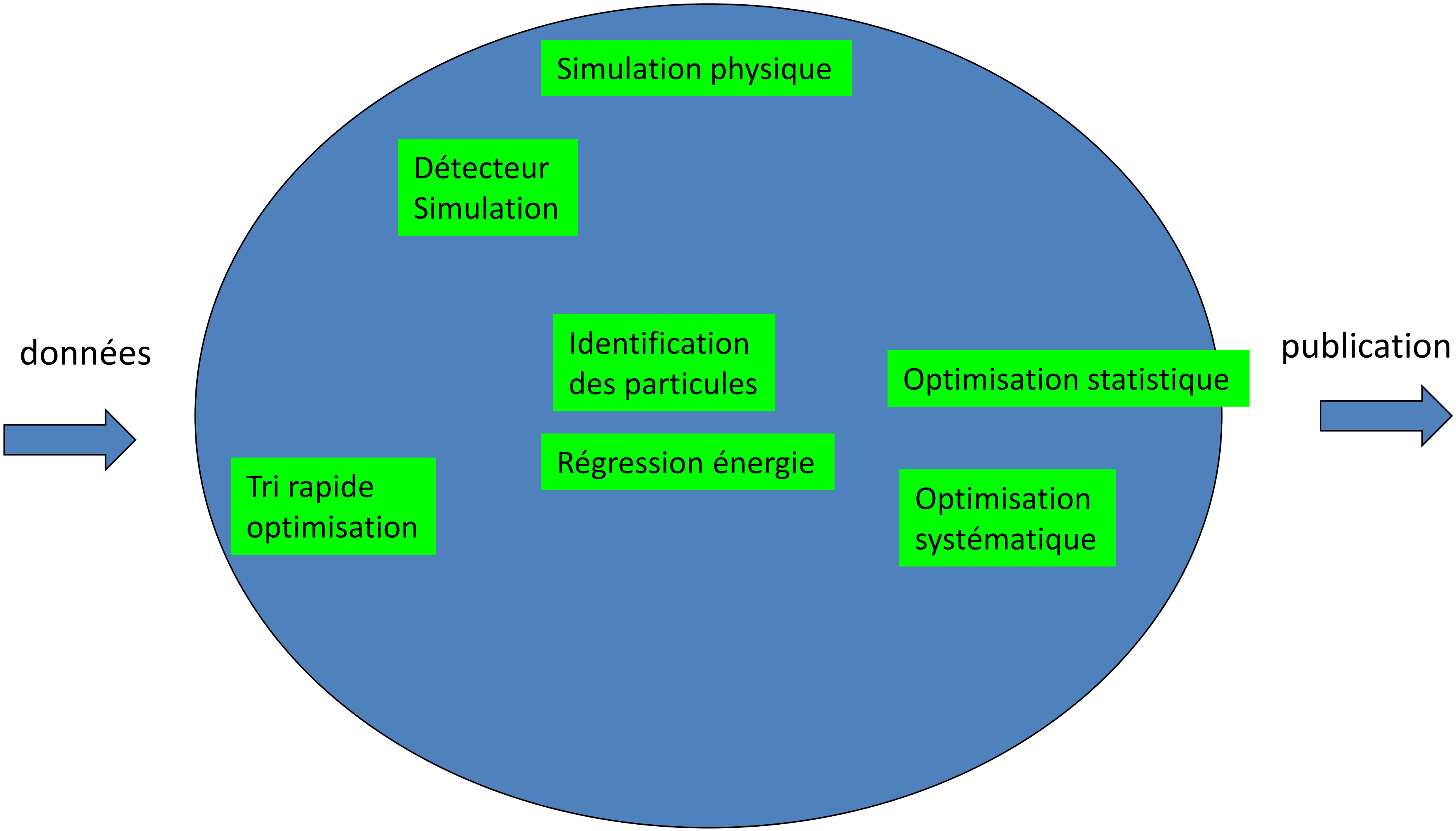
- Conception expérience
- Prise de données
- Echec/succès ?



➔En pratique, pourrait être utilisable pour le tri rapide d’événements, facilement paramètrable et se prêtant bien a la virtualisation de l’expérience

Finalement…

Terrain de jeu des algorithmes IA



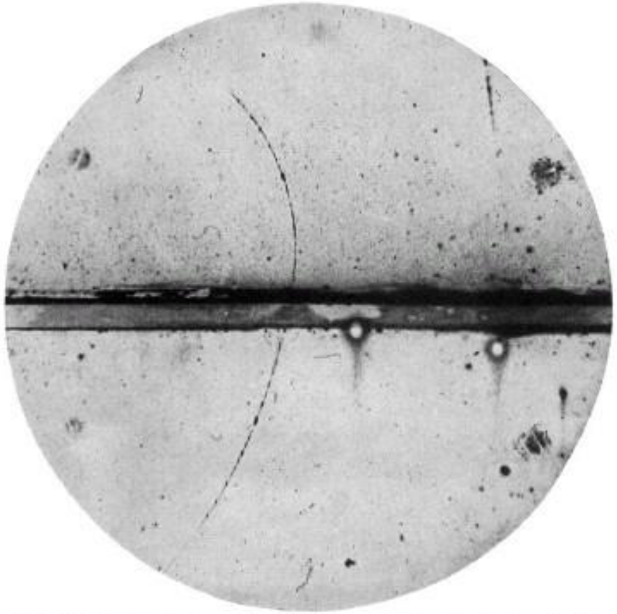
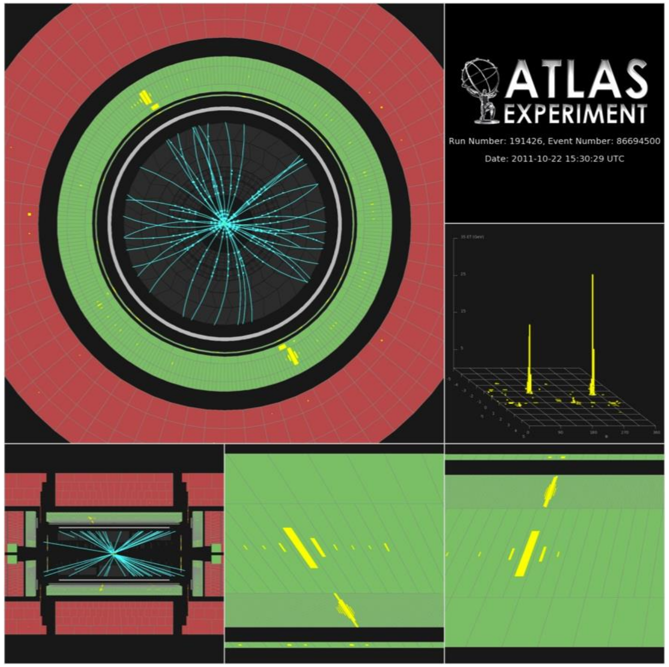
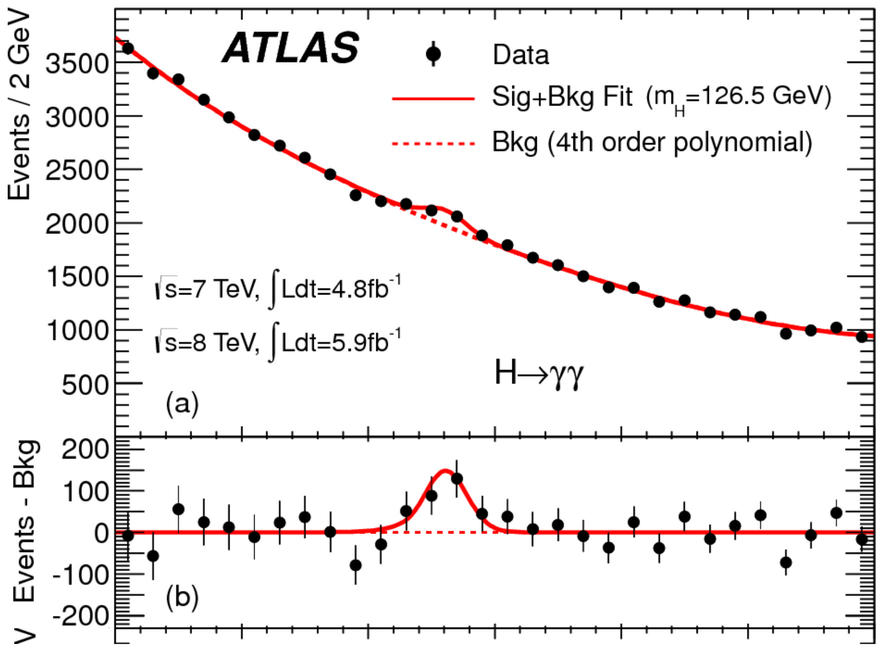
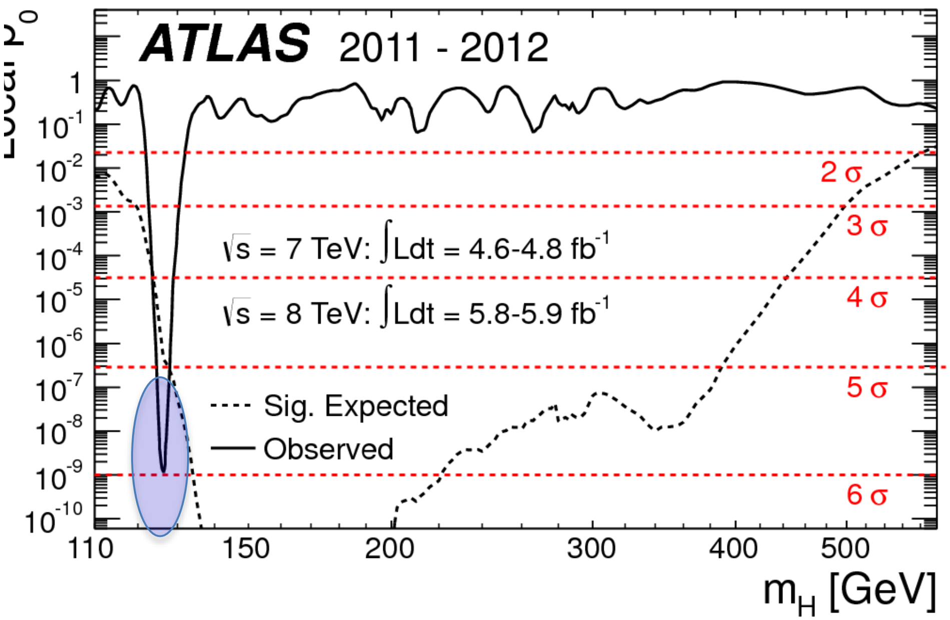


FIG. 1. A 63 million volt positron ($H\rho = 2.1 \times 10^5$ gauss-cm) passing through a 6 mm lead plate and emerging as a 23 million volt positron ($H\rho = 7.5 \times 10^4$ gauss-cm). The length of this latter path is at least ten times greater than the possible length of a proton path of this curvature.







Conclusion

- Intelligence Artificielle : beaucoup de promesses pour la physique des particules et la recherche scientifique
- Algorithmes très puissants
- Leur terrain de jeu s'aggrandit
- Jusqu'où…

- ... ça sera l'enjeu des prochaines années!