

# 素粒子物理における深層学習の応用

高エネルギー加速器研究機構  
野尻美保子

# 素粒子物理学とは

“物質の根源や宇宙の成り立ちについて考える物理学”

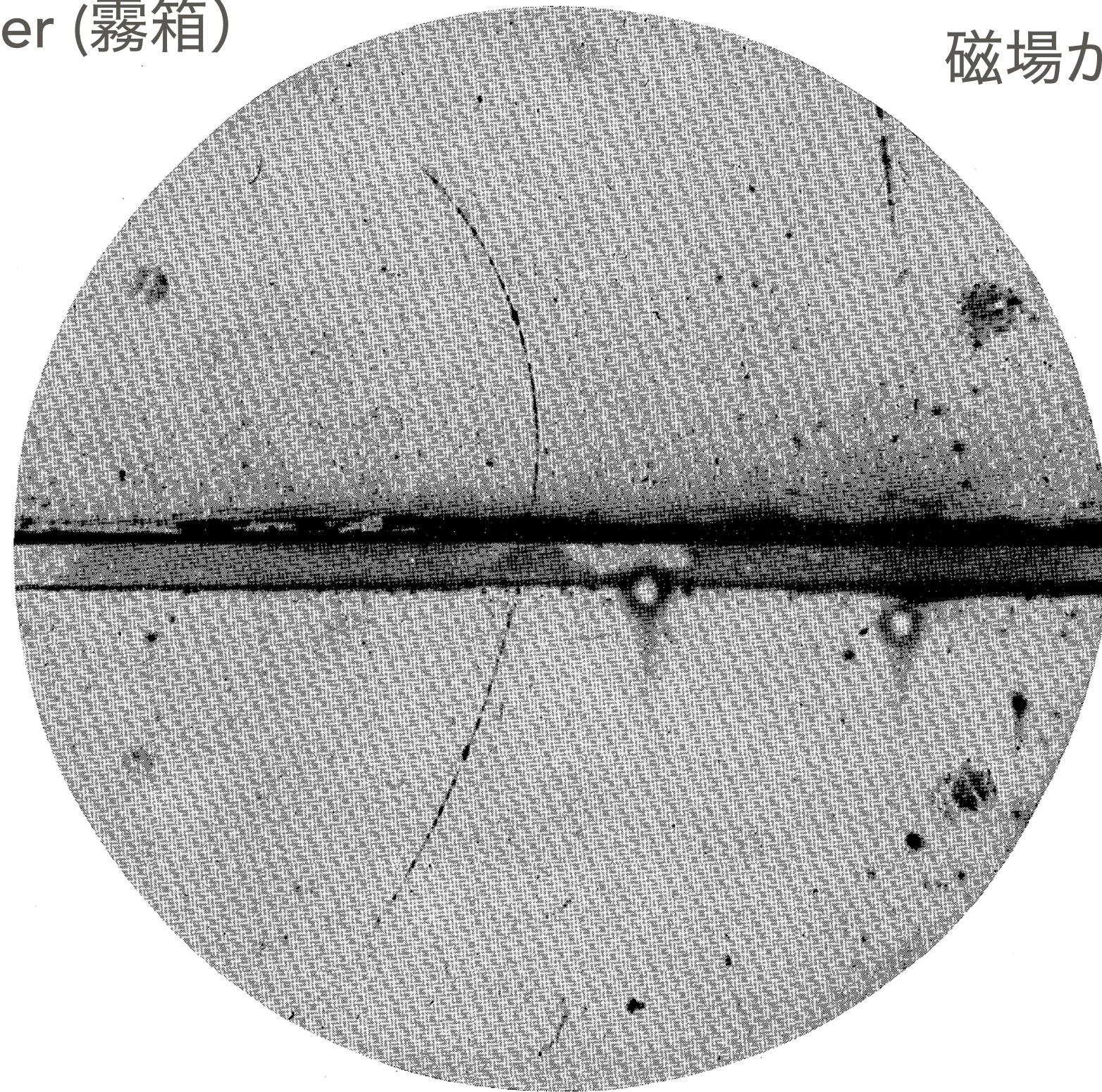
- それ科学なの？
- 物質の性質は基本方程式でかける
- 基本方程式を使って、短距離でおこることや、過去、未来におこることを推測すると、理屈に合わない →今の基本方程式と矛盾のない、新しい基本法則の提案→実験的に確認

昔の例 反粒子の発見→粒子が生成、消滅する (相対論的場の理論)

## The Positive Electron (陽電子)

C.D Anderson , PRD 43,491, (March 15, 1933) 宇宙線の観測から発見

Cloud Chamber (霧箱)



磁場がかかっている。

金属プレート

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.

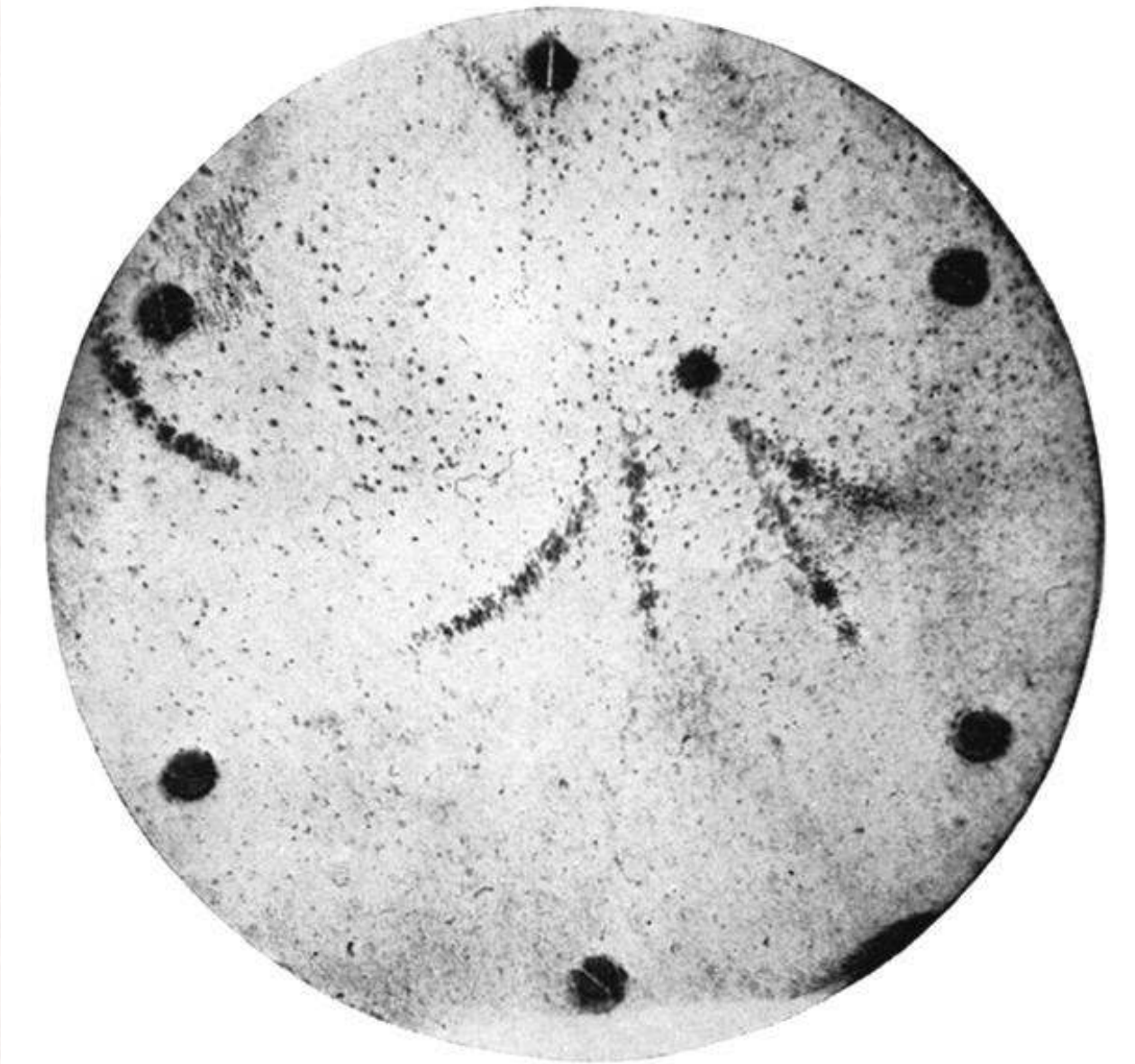
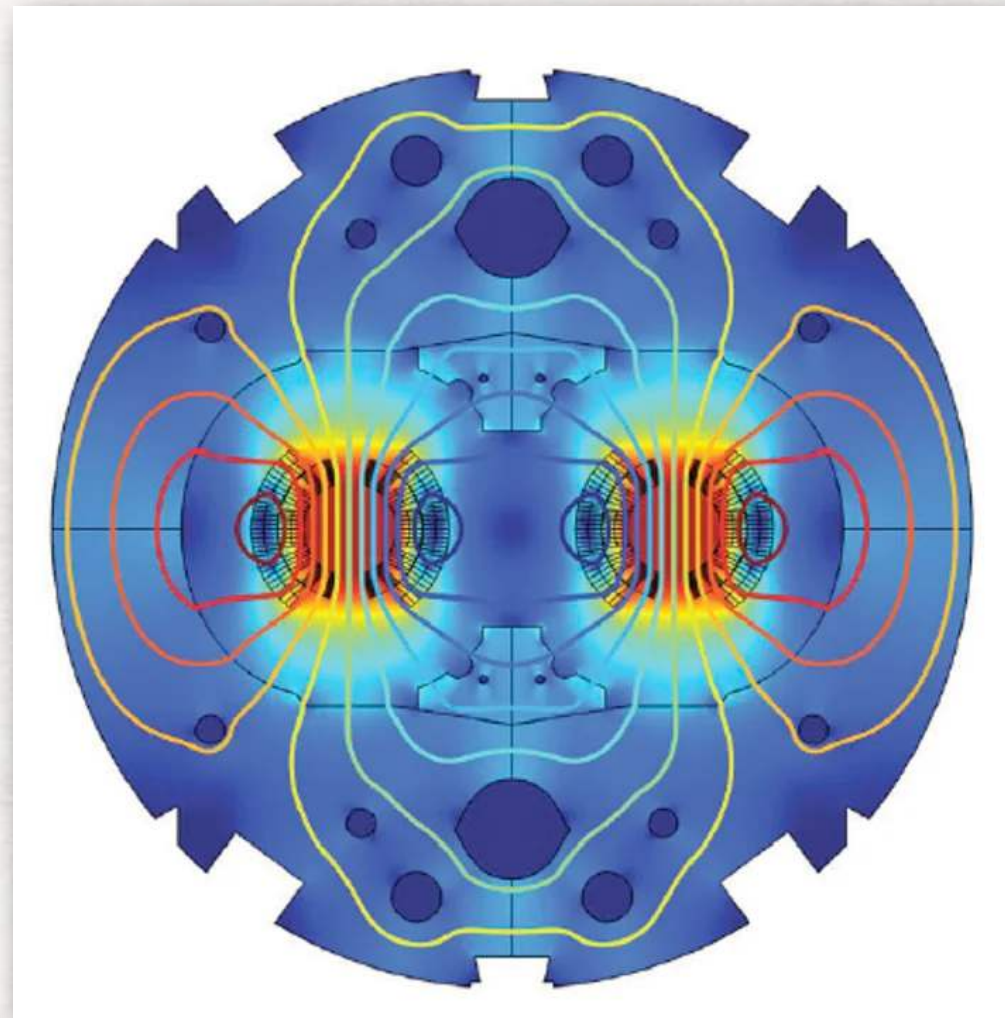


FIG. 3. A group of six particles projected from a region in the wall of the chamber. The track at the left of the central group of four tracks is a negatron of about 18 million volts energy ( $H\rho=6.2\times 10^4$  gauss-cm) and that at the right a positron of about 20 million volts energy ( $H\rho=7.0\times 10^4$  gauss-cm). Identification of the two tracks in the center is not possible. A negatron of about 15 million volts is shown at the left. This group represents early tracks which were broadened by the diffusion of the ions. The uniformity of this broadening for all the tracks shows that the particles entered the chamber at the same time.

# コライダー実験

曲げる(2重極)

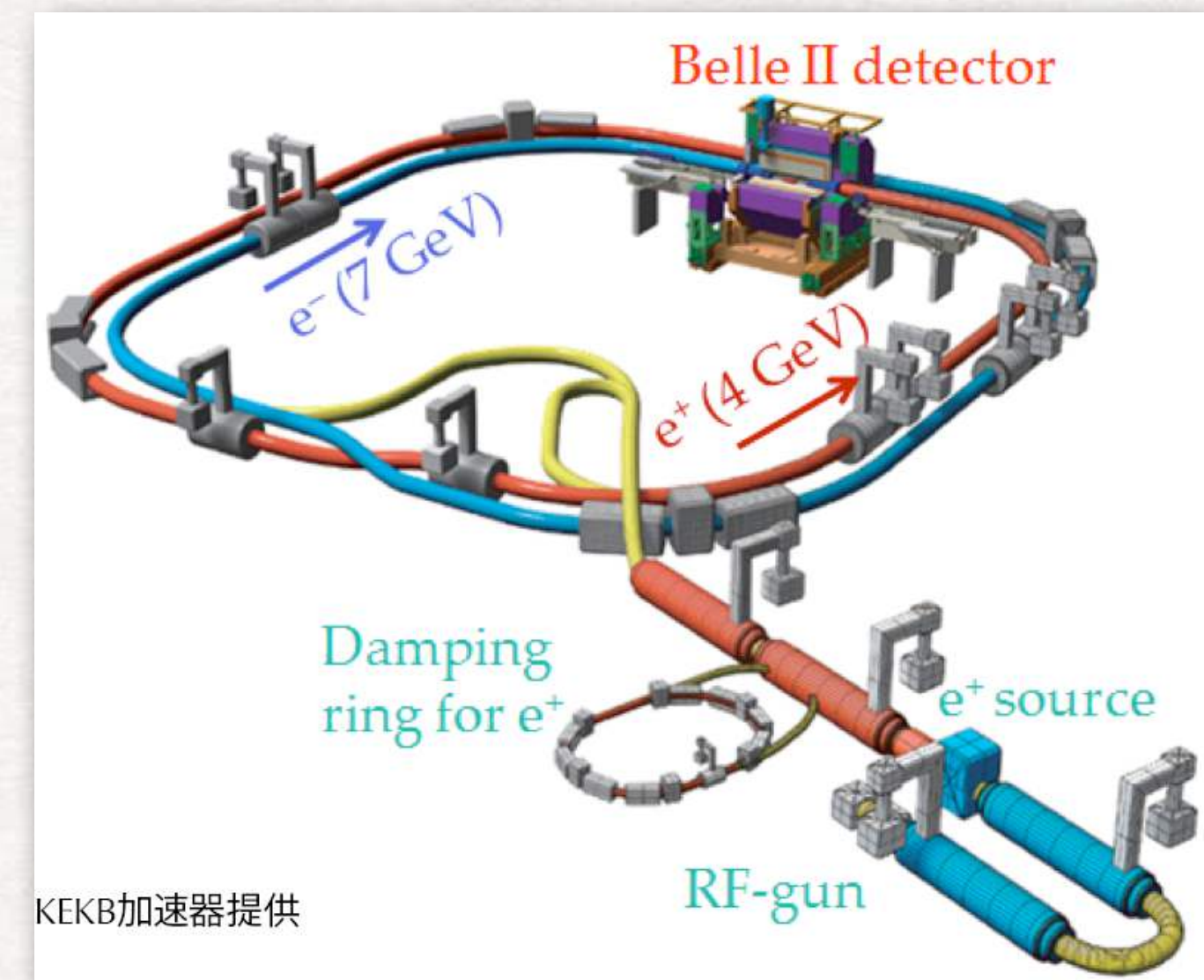
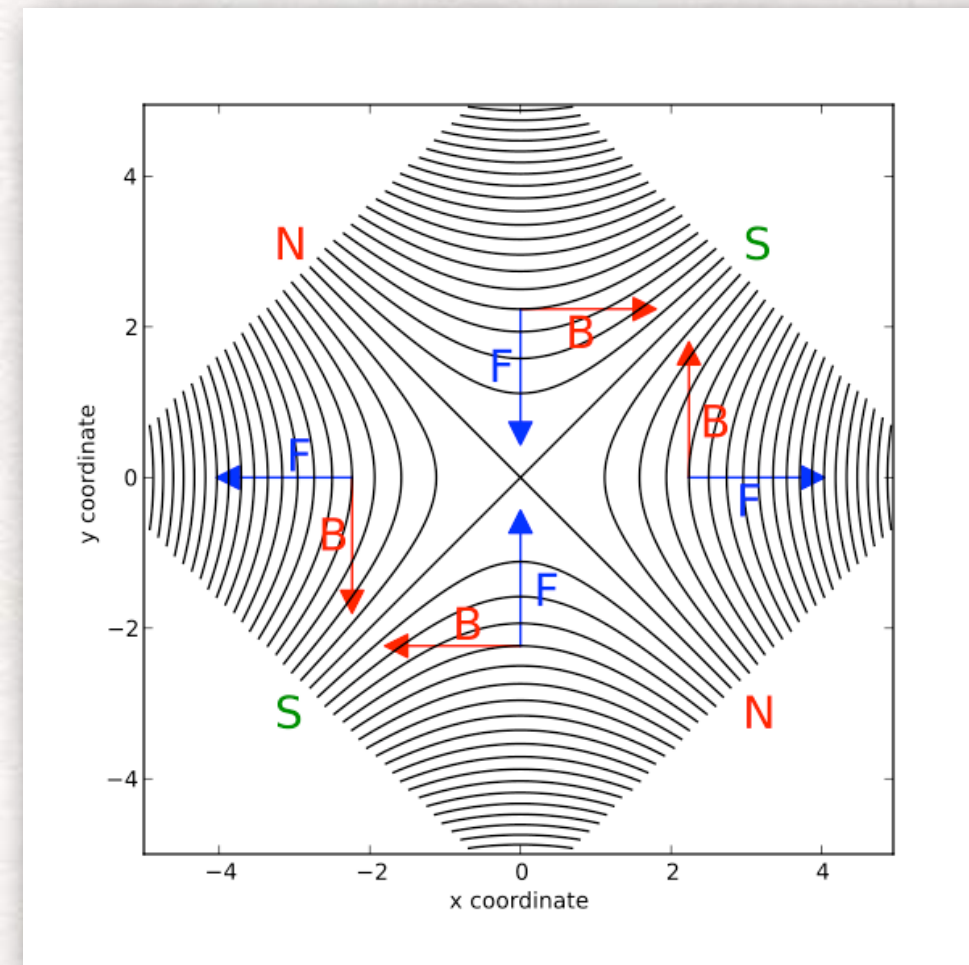
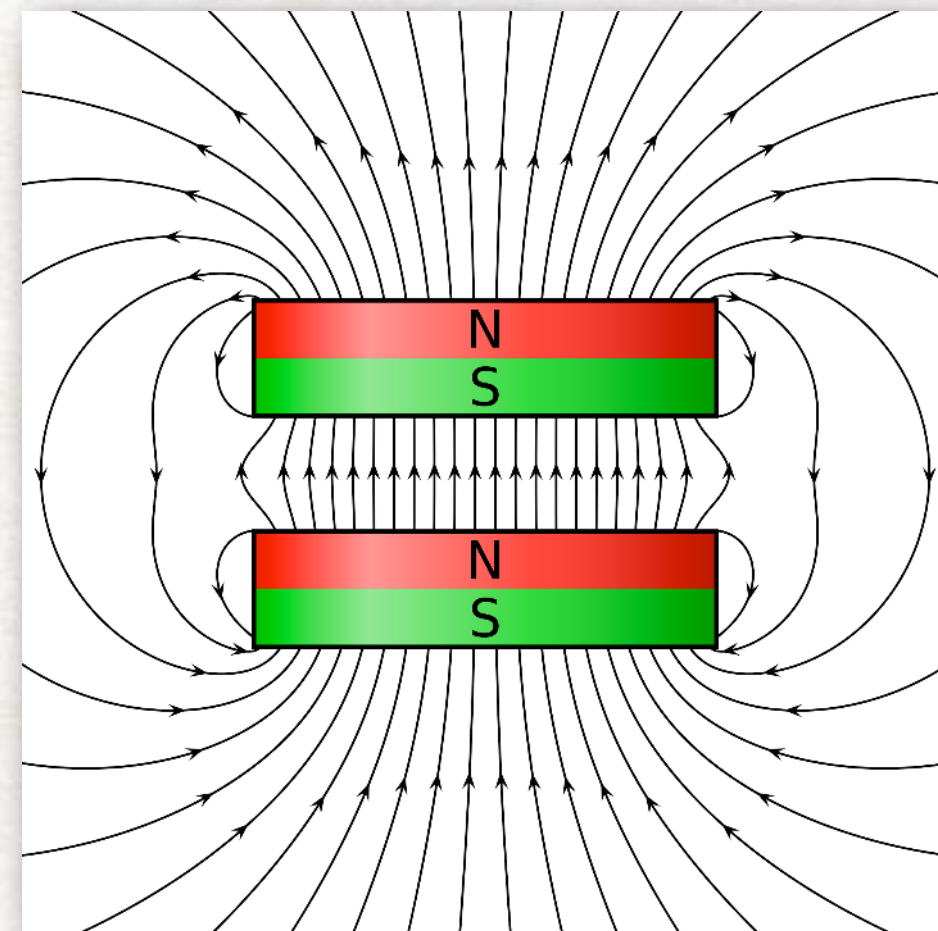


収束させる(四重極)



$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}),$$

KEK B



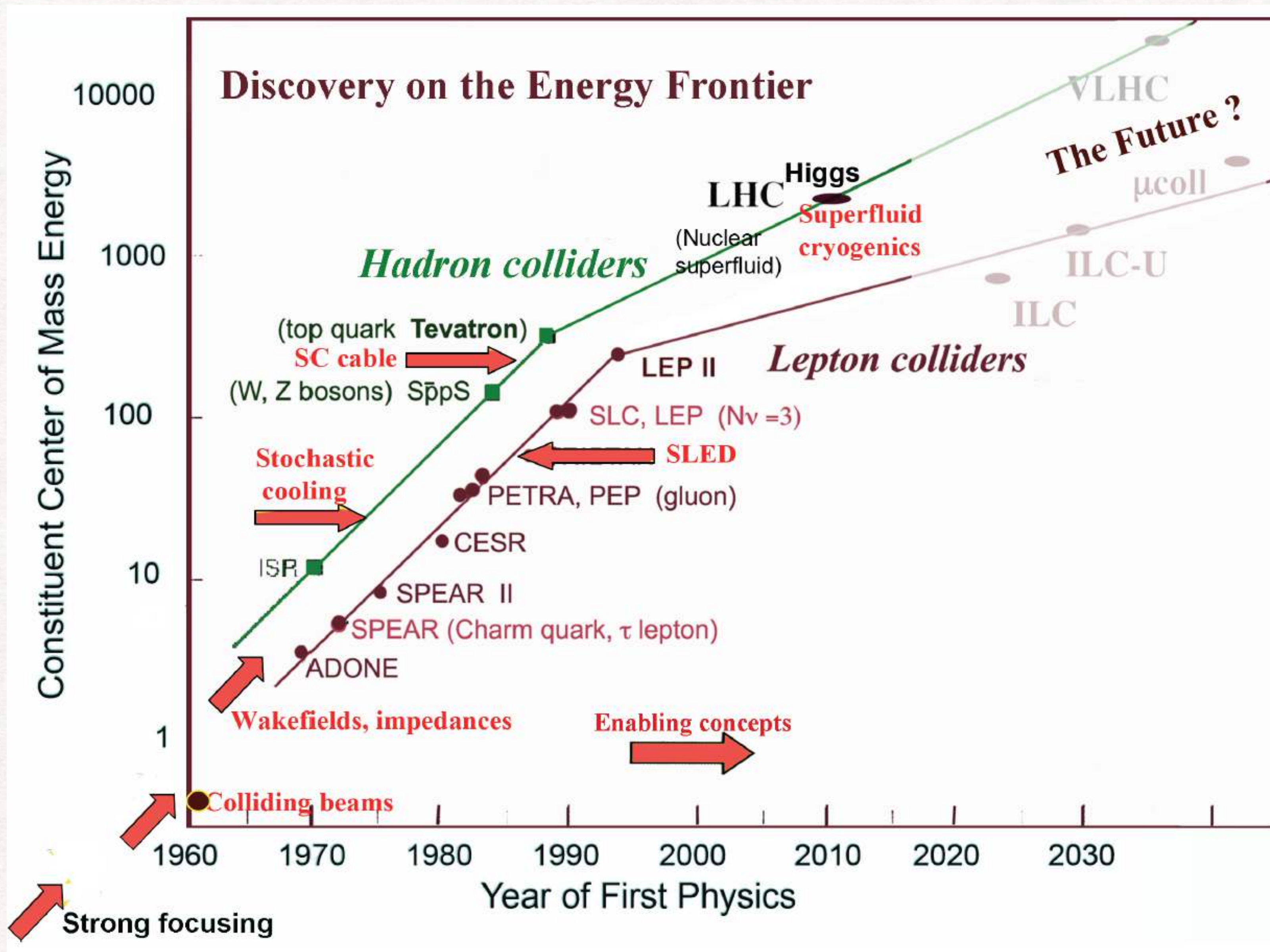


Figure 6-1. The so-called Livingston plot illustrates how history of discovery on the energy frontier has been enabled by the history of invention (red arrows) in accelerator science and technology.

# 素粒子の標準模型

	物質の三世代 (フェルミ粒子)			相互作用 / 力の伝搬 (ボース粒子)	
	I	II	III		
質量	$\approx 2.2 \text{ MeV}/c^2$	$\approx 1.28 \text{ GeV}/c^2$	$\approx 173.1 \text{ GeV}/c^2$	0	$\approx 124.97 \text{ GeV}/c^2$
電荷	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	0	0
スピン	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	0
	<b>u</b> アップ	<b>c</b> チャーム	<b>t</b> トップ	<b>g</b> グルーオン	<b>H</b> ヒッグス粒子
	$\approx 4.7 \text{ MeV}/c^2$	$\approx 96 \text{ MeV}/c^2$	$\approx 4.18 \text{ GeV}/c^2$	0	
	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	0	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
	<b>d</b> ダウン	<b>s</b> ストレンジ	<b>b</b> ボトム	<b><math>\gamma</math></b> 光子	
	$\approx 0.511 \text{ MeV}/c^2$	$\approx 105.66 \text{ MeV}/c^2$	$\approx 1.7768 \text{ GeV}/c^2$	$\approx 91.19 \text{ GeV}/c^2$	
	-1	-1	-1	0	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
	<b>e</b> 電子	<b><math>\mu</math></b> ミュー粒子	<b><math>\tau</math></b> タウ粒子	<b>Z</b> Zボソン	
	$< 1.0 \text{ eV}/c^2$	$< 0.17 \text{ MeV}/c^2$	$< 18.2 \text{ MeV}/c^2$	$\approx 80.433 \text{ GeV}/c^2$	
	0	0	0	$\pm 1$	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
	<b><math>\nu_e</math></b> 電子 ニュートリノ	<b><math>\nu_\mu</math></b> ミュー ニュートリノ	<b><math>\nu_\tau</math></b> タウ ニュートリノ	<b>W</b> Wボソン	

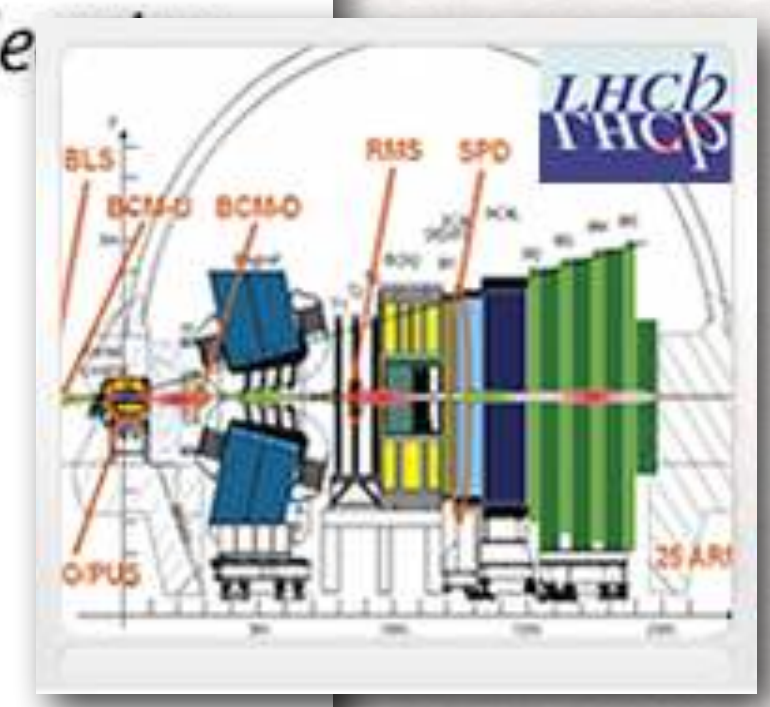
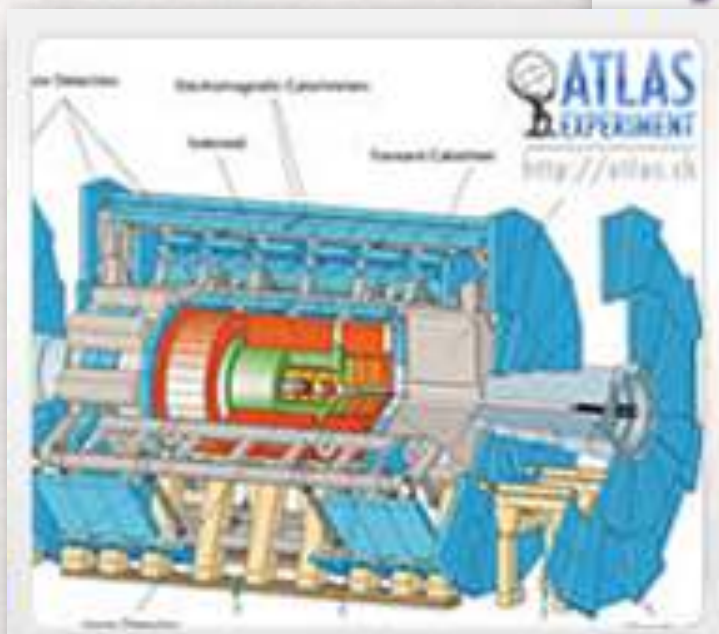
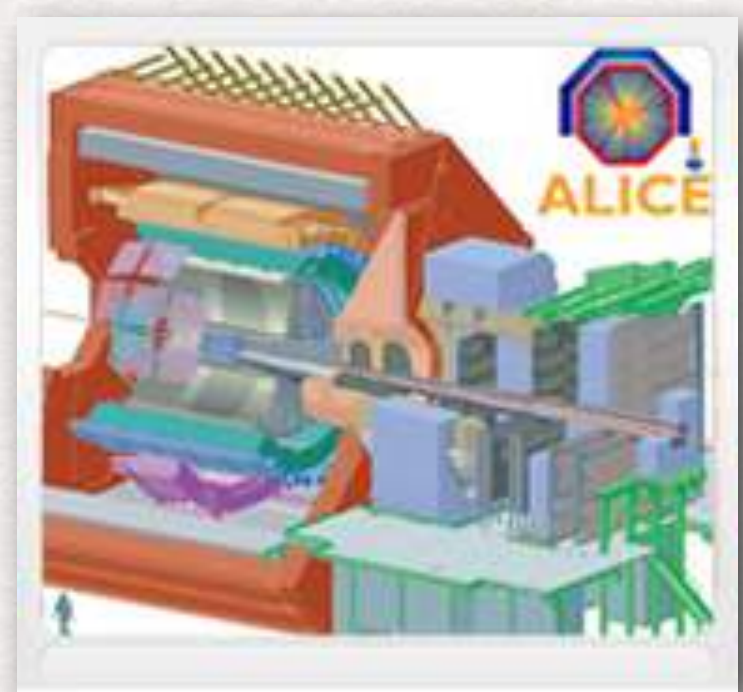
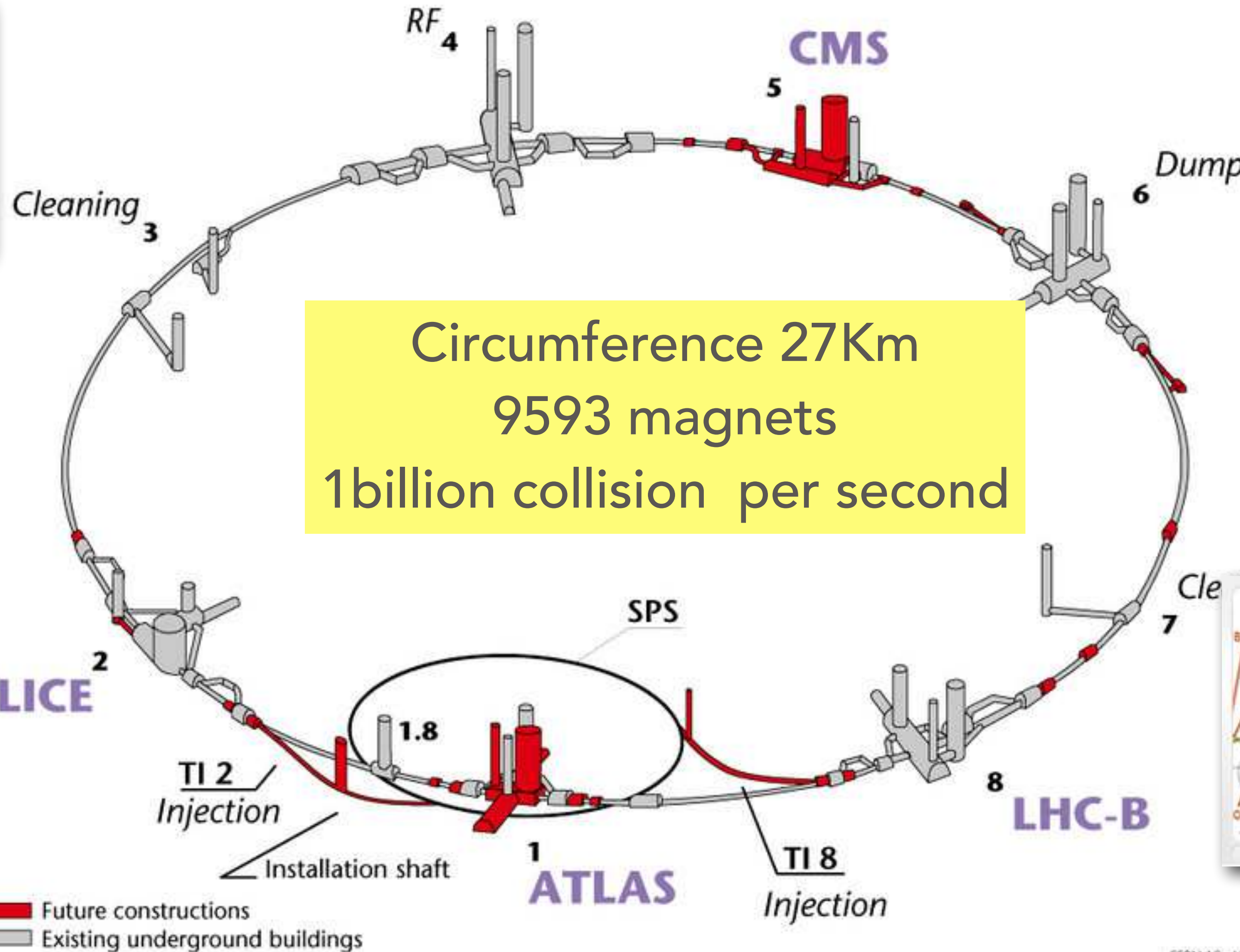
クォーク

レプトン

スカラール粒子

ゲージ粒子  
ベクトル粒子

# Layout of the LEP tunnel including future LHC infrastructures.



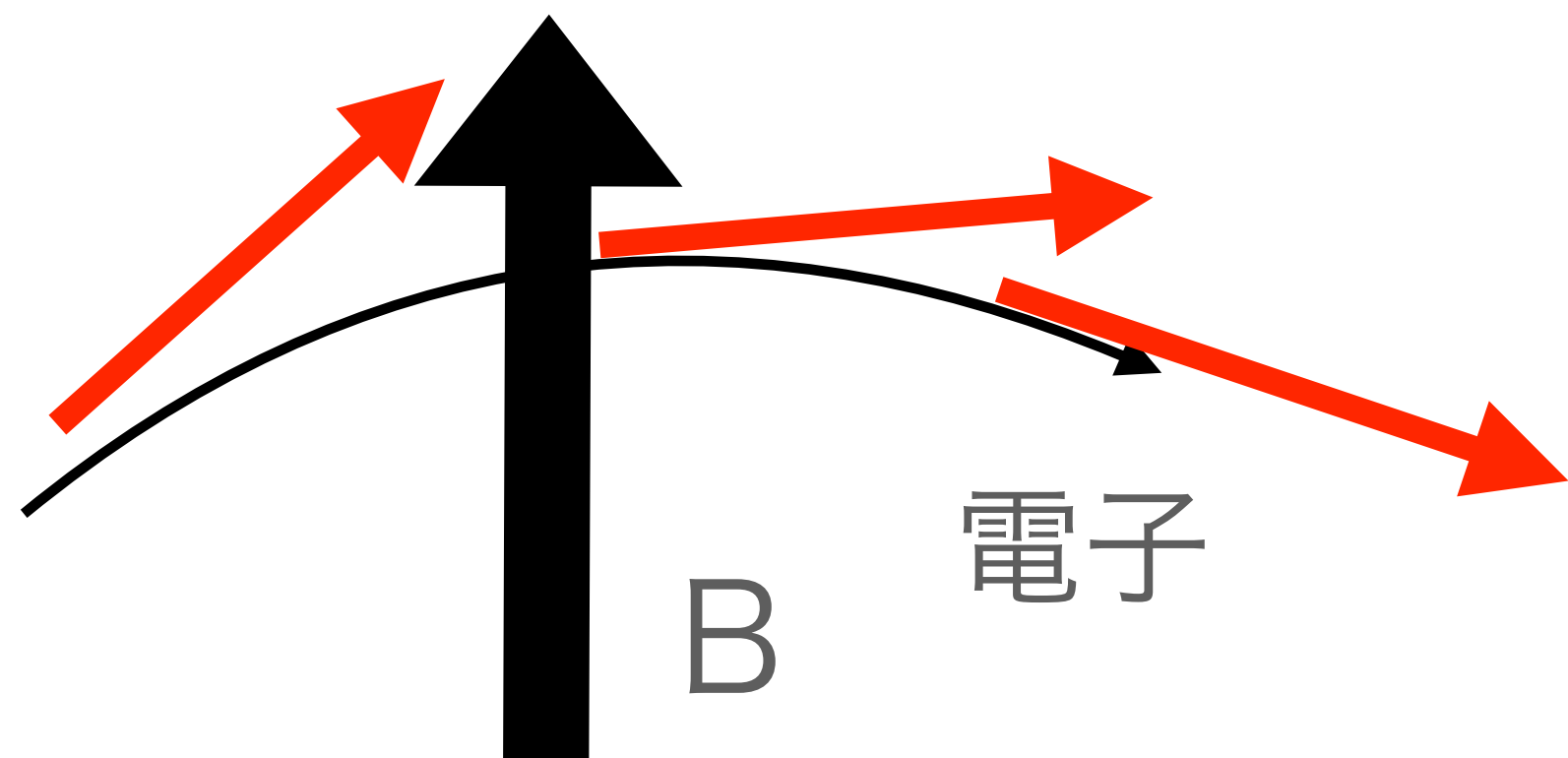
# LHC実験の成果: Higgs 粒子の発見

- 電子と光の相互作用

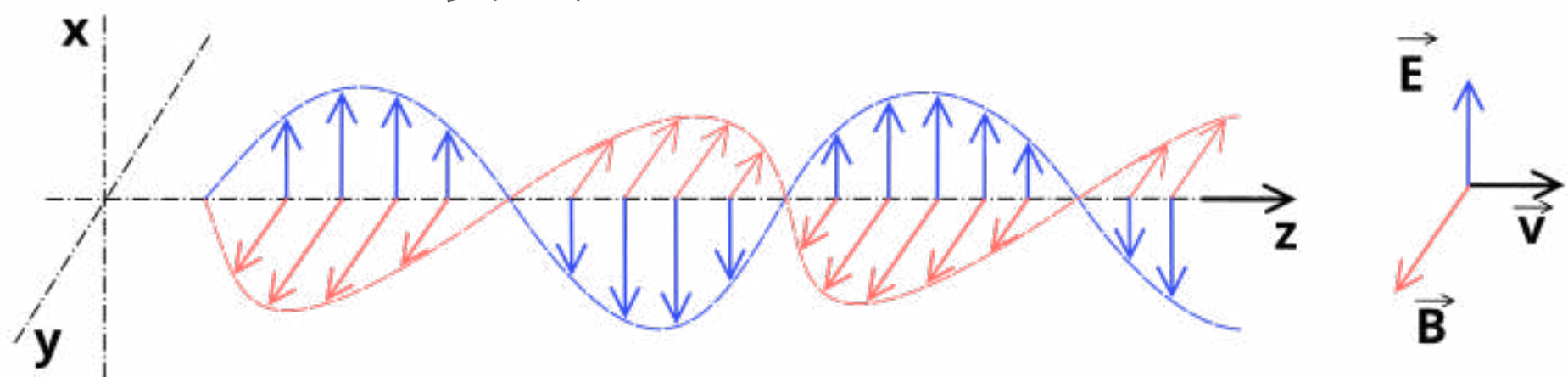
$$L = e\bar{\psi}\gamma^\mu\psi A_\mu$$

Spin 1 の粒子

電磁波 (シンクロトロン放射)



磁場



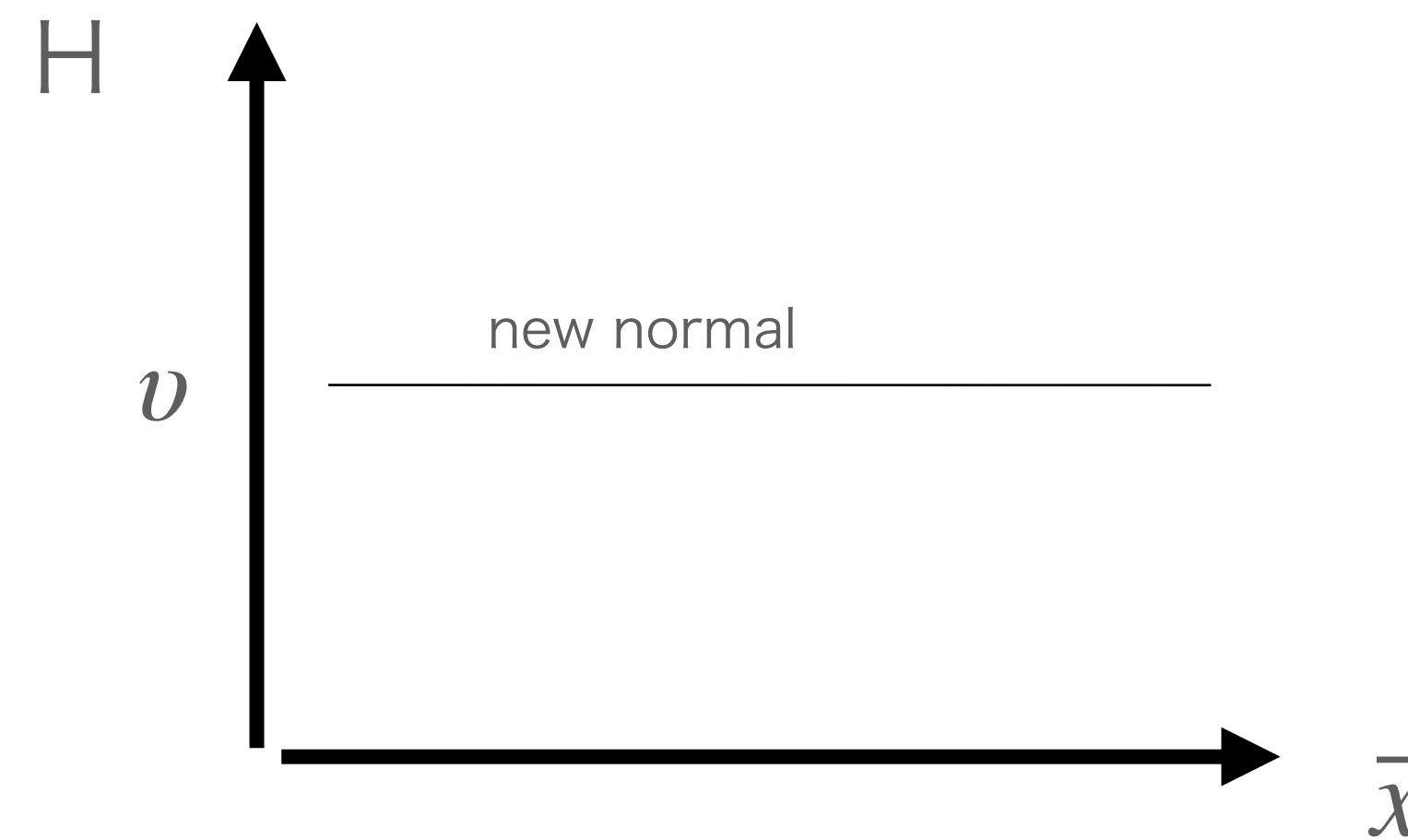
粒子とHiggs粒子の相互作用

$$L = \frac{m}{v}\bar{\psi}\psi H$$

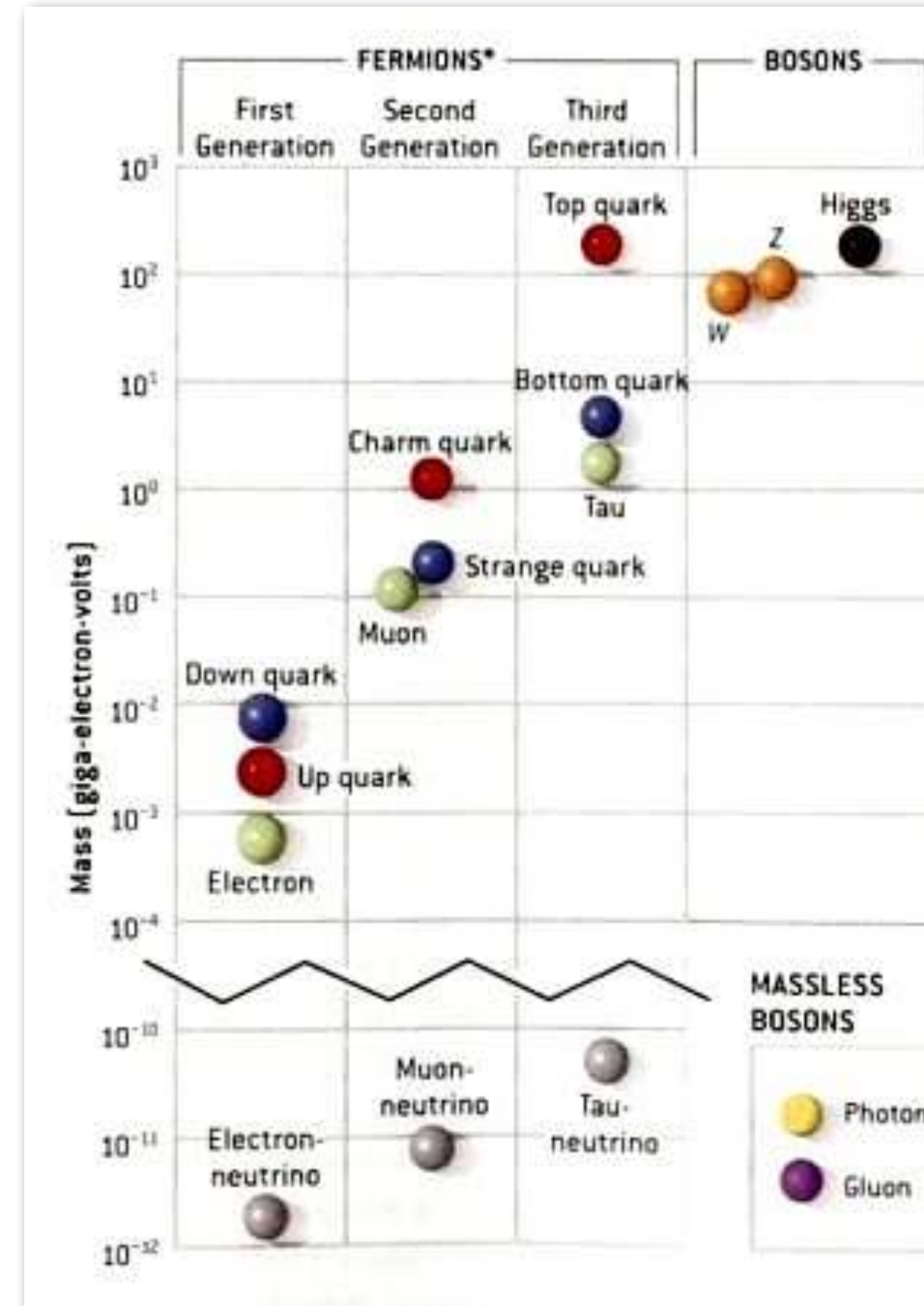
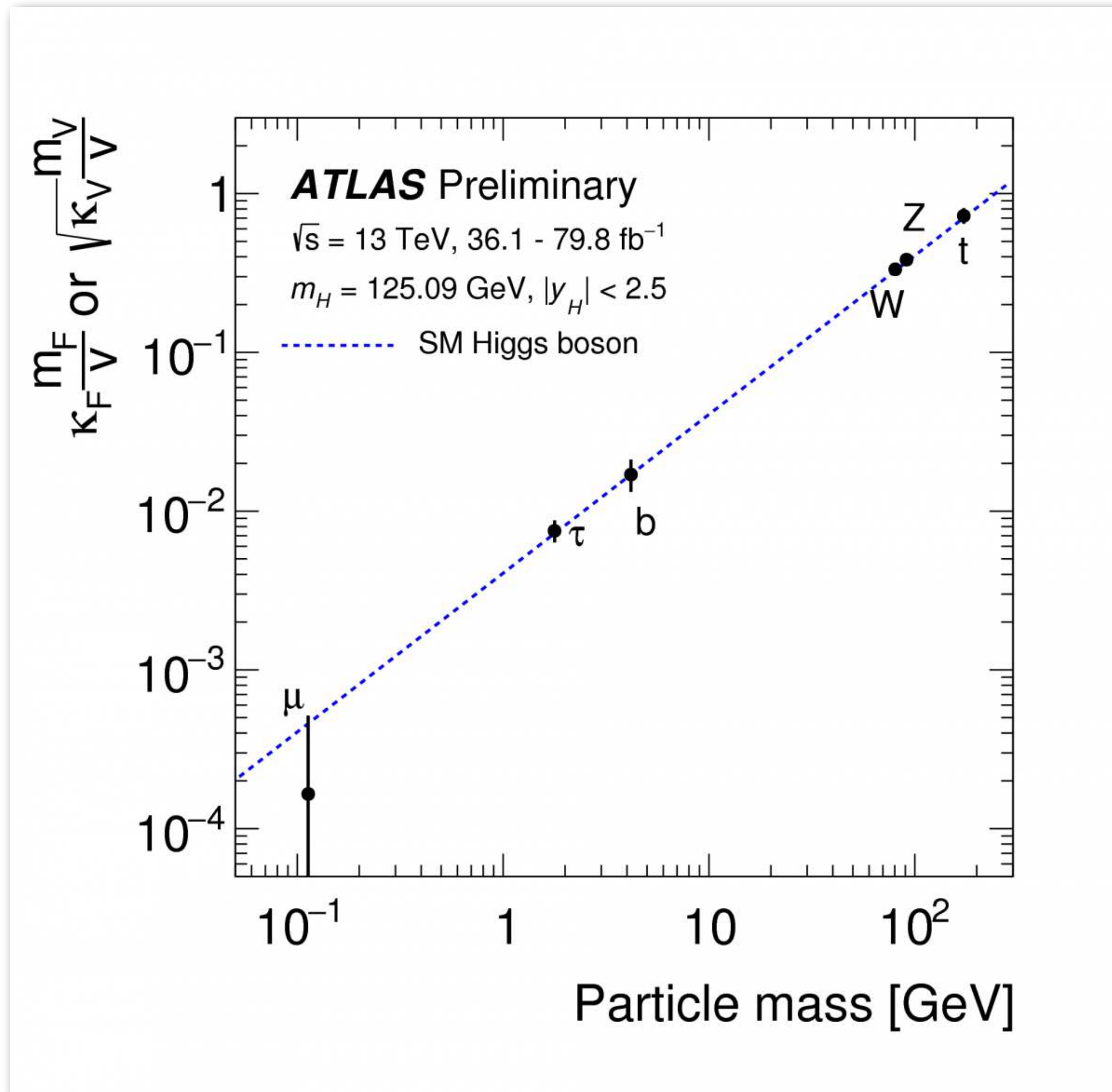
質量に比例した相互作用  
→質量を作り出した相互作用

スピン0の粒子

運動方程式の解  $H = v + h(x)$

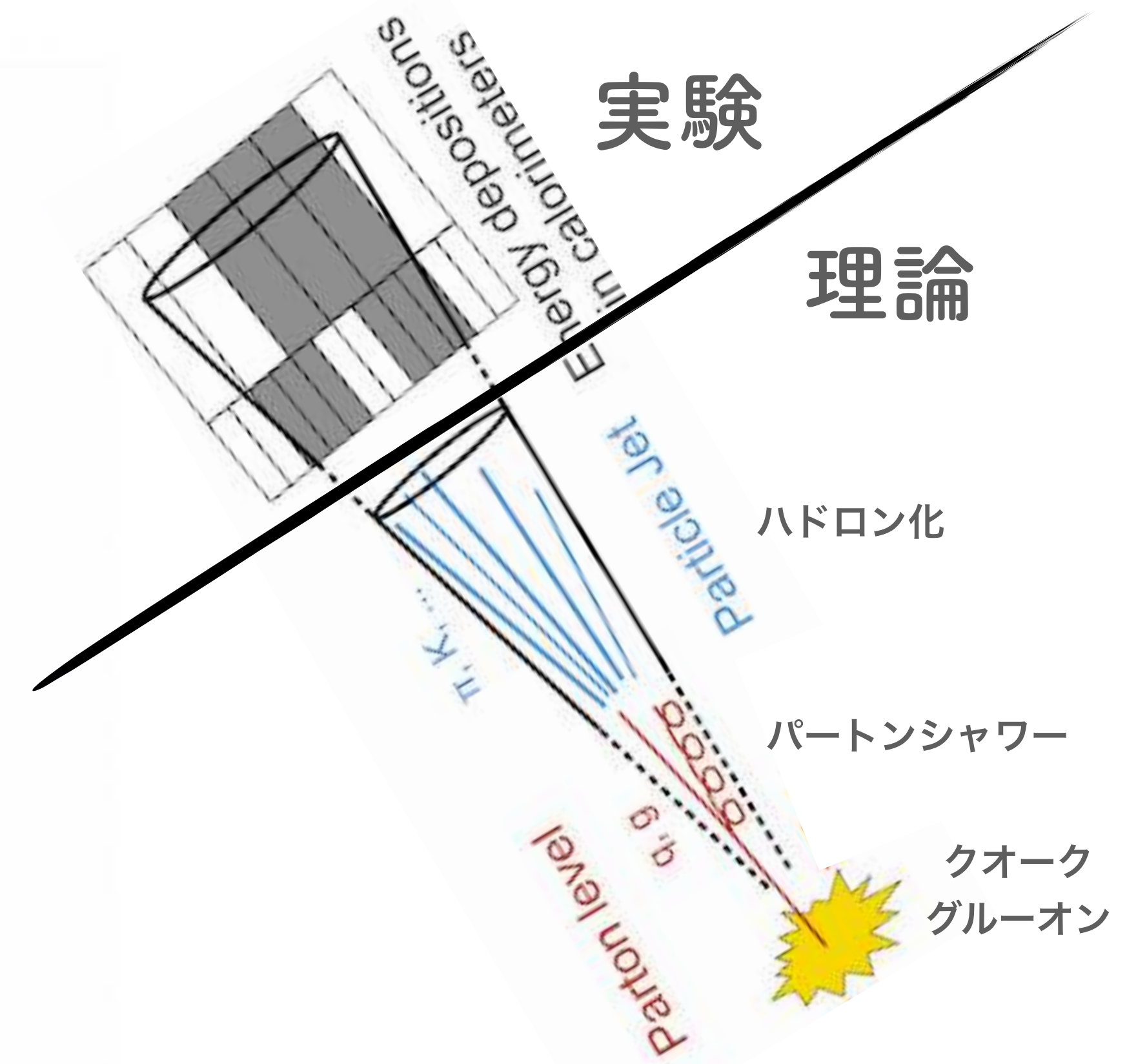
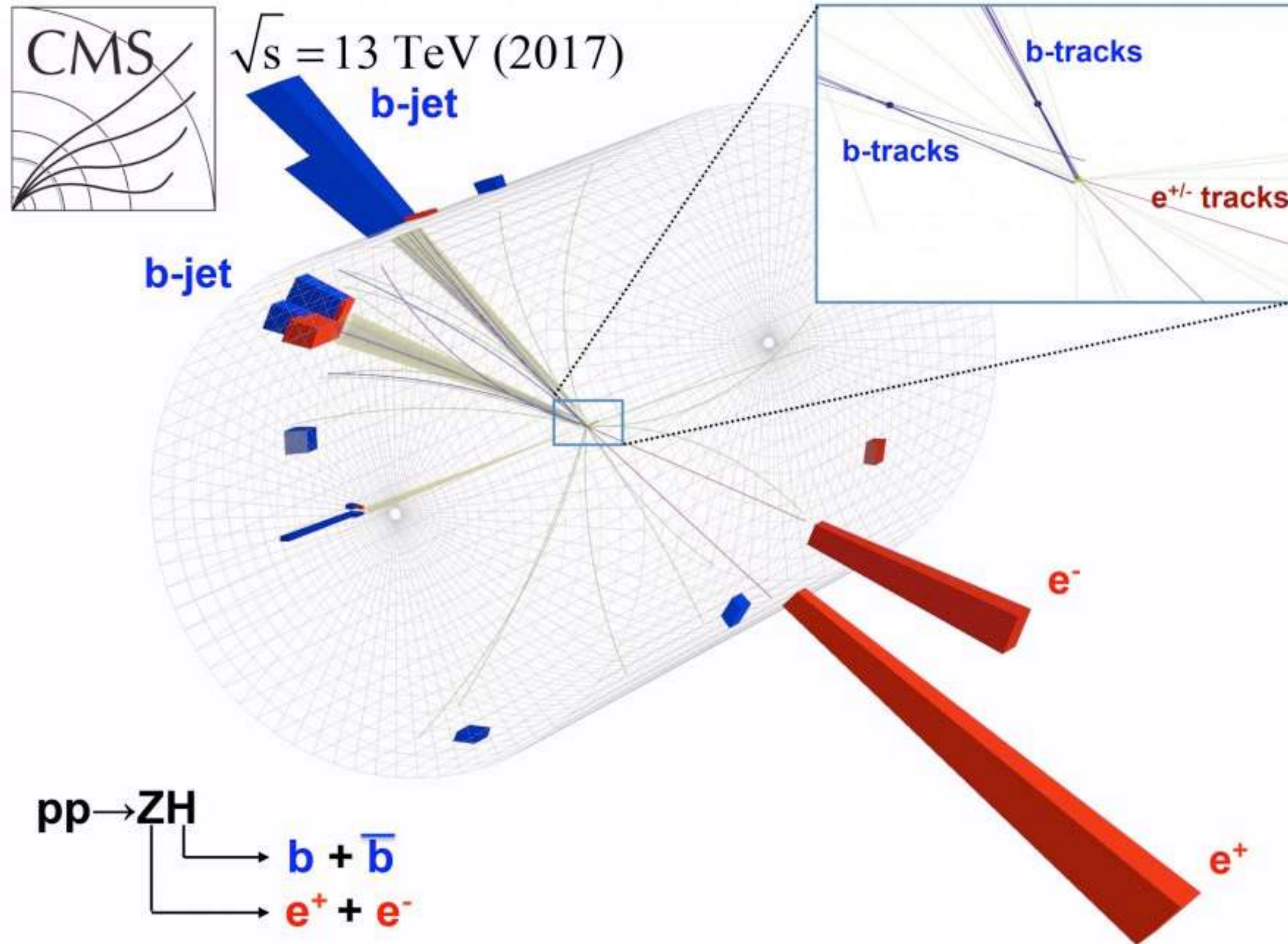


# HIGGS boson



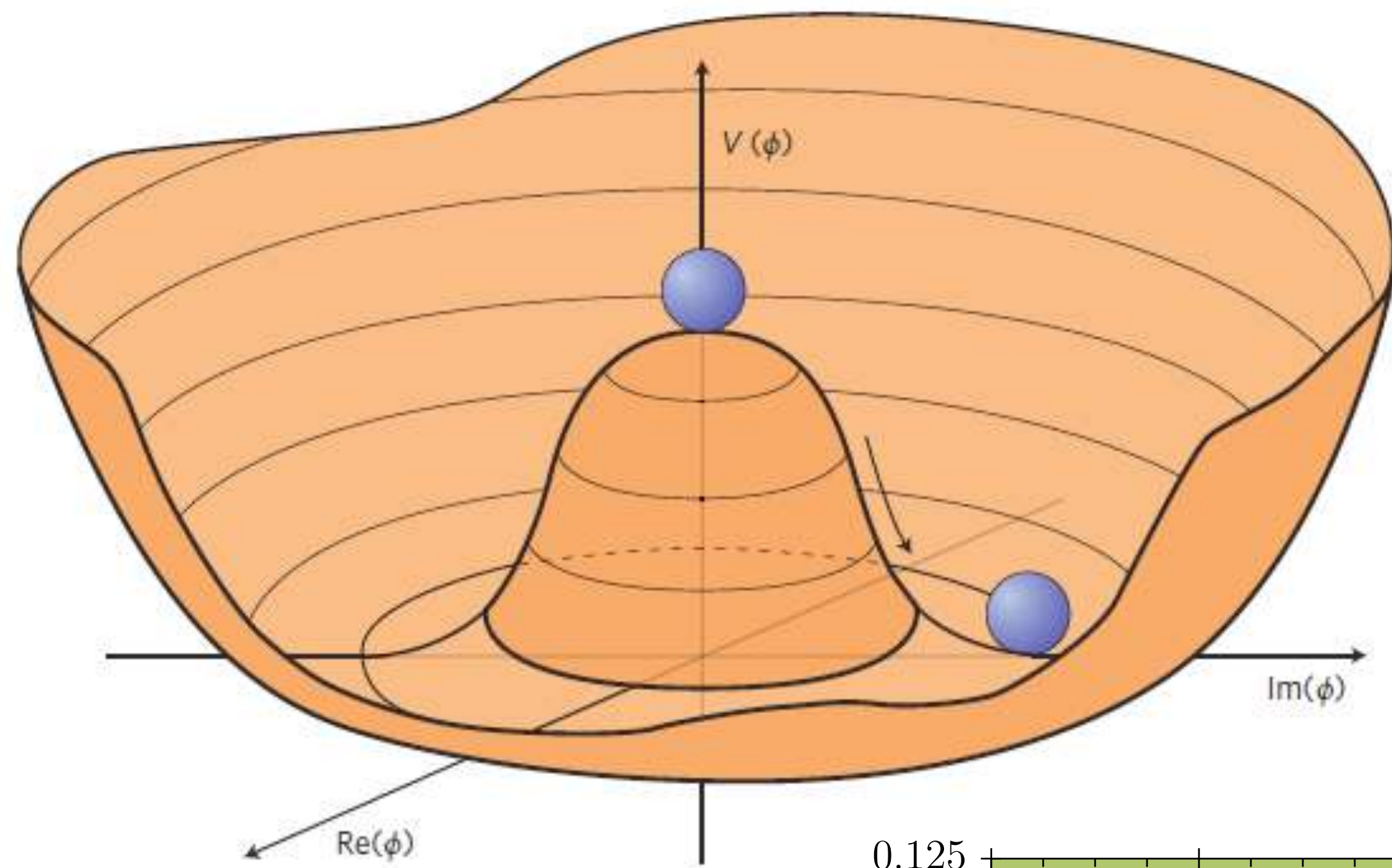
# Higgs 粒子はどんなふうに見えるか

測定器と粒子の相互作用

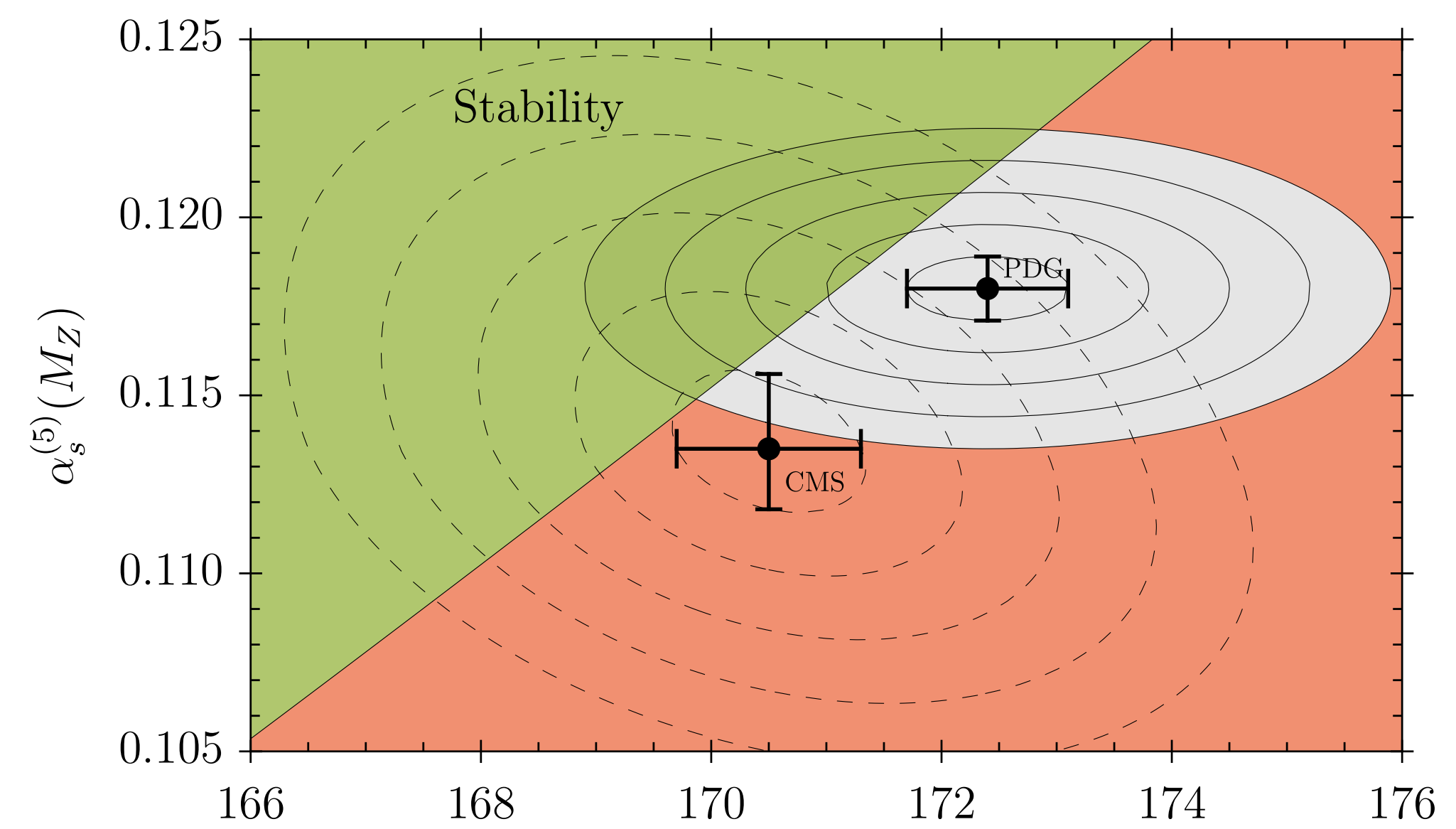
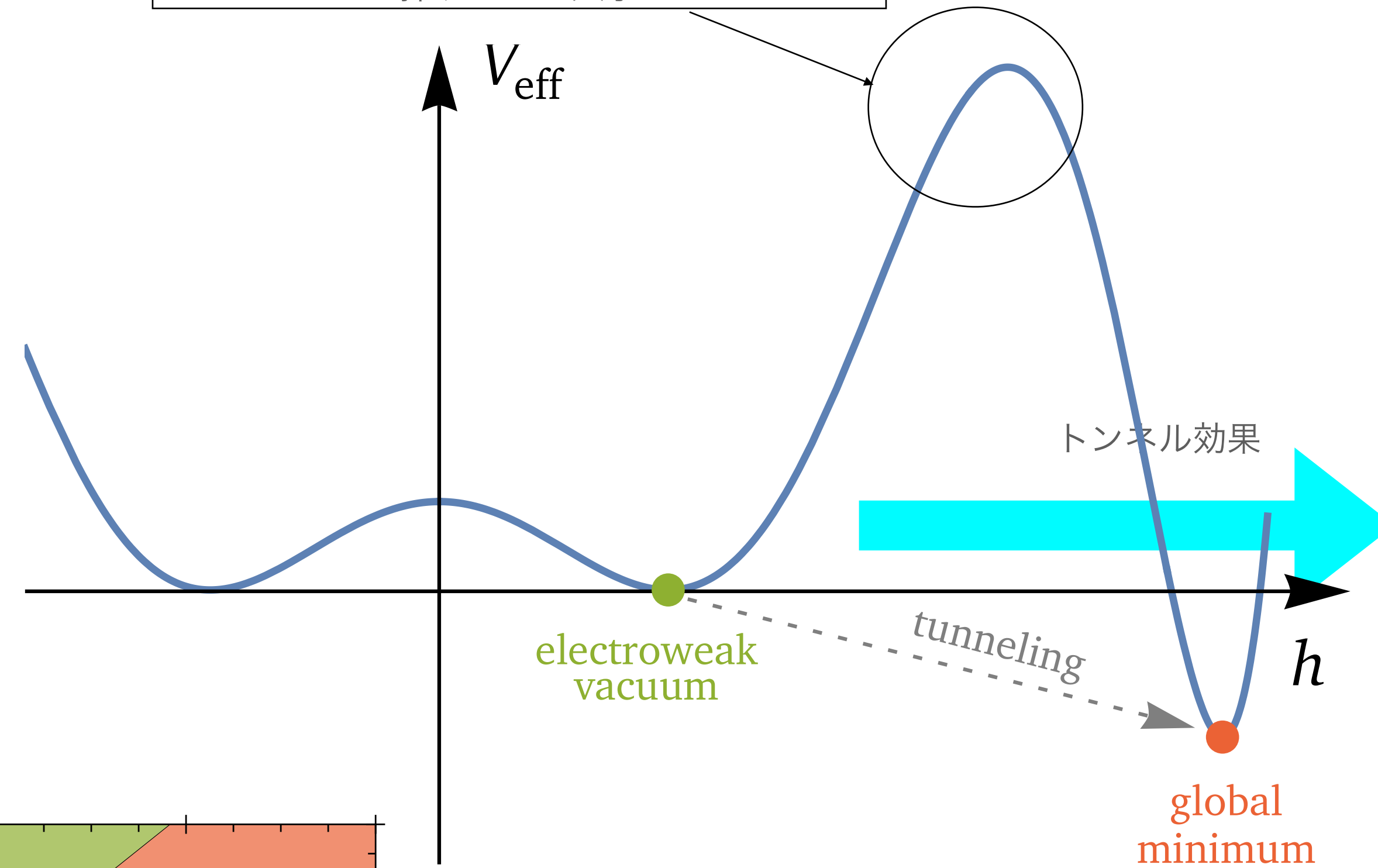


クォークは単体では見えず  
同じ方向にいく粒子の束  
jet として見える

# 真空期待値の起源と真空不安定性



このあたりのことを考えるためには、  
Higgsとtop の結合やHiggs が複数ある状態を  
探ることが大事



2412.09241

# ハドロンコライダーの理論と実験の関係

精密測定、宇宙論

素敵な  
究極理論

高精度  
原理的

理論的不定性  
現象論的な取り扱い

高いエネルギースケール

中間的なエネルギースケール

低いエネルギー

プロセス計算  
(振幅計算)

パートン  
シャワー

ハドロン化

Higgs 粒子を発見したい!

Higgs 粒子の重さ  
相互作用

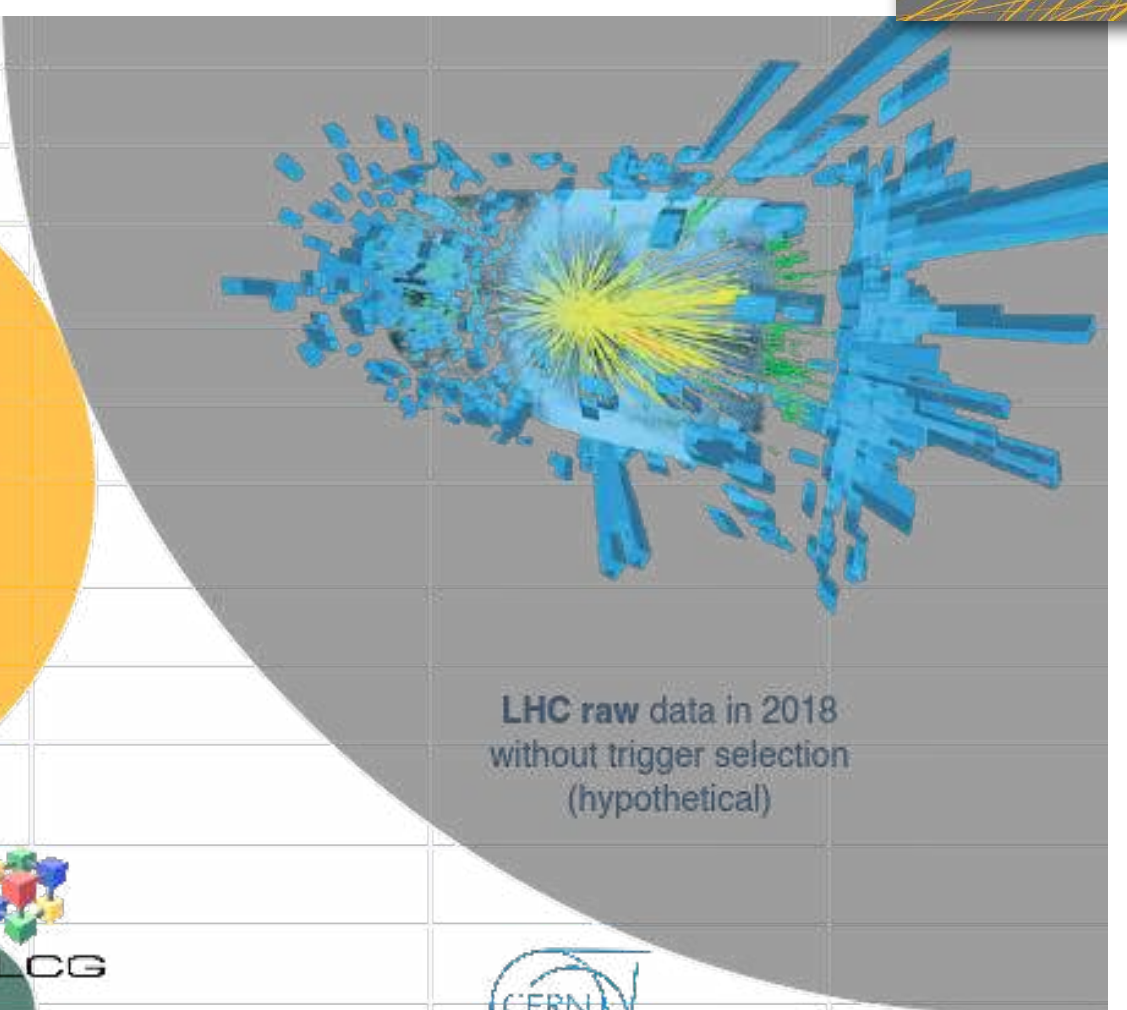
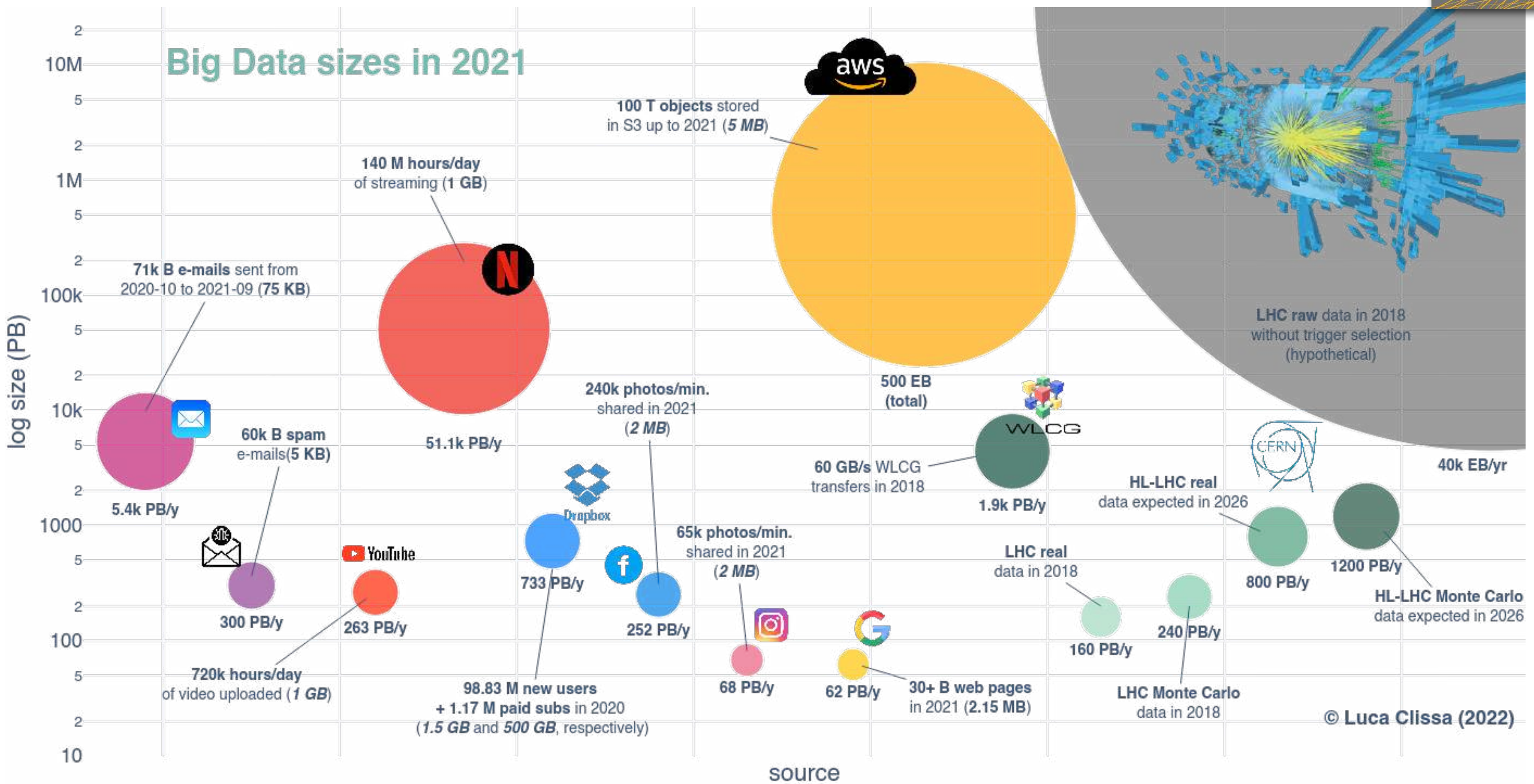
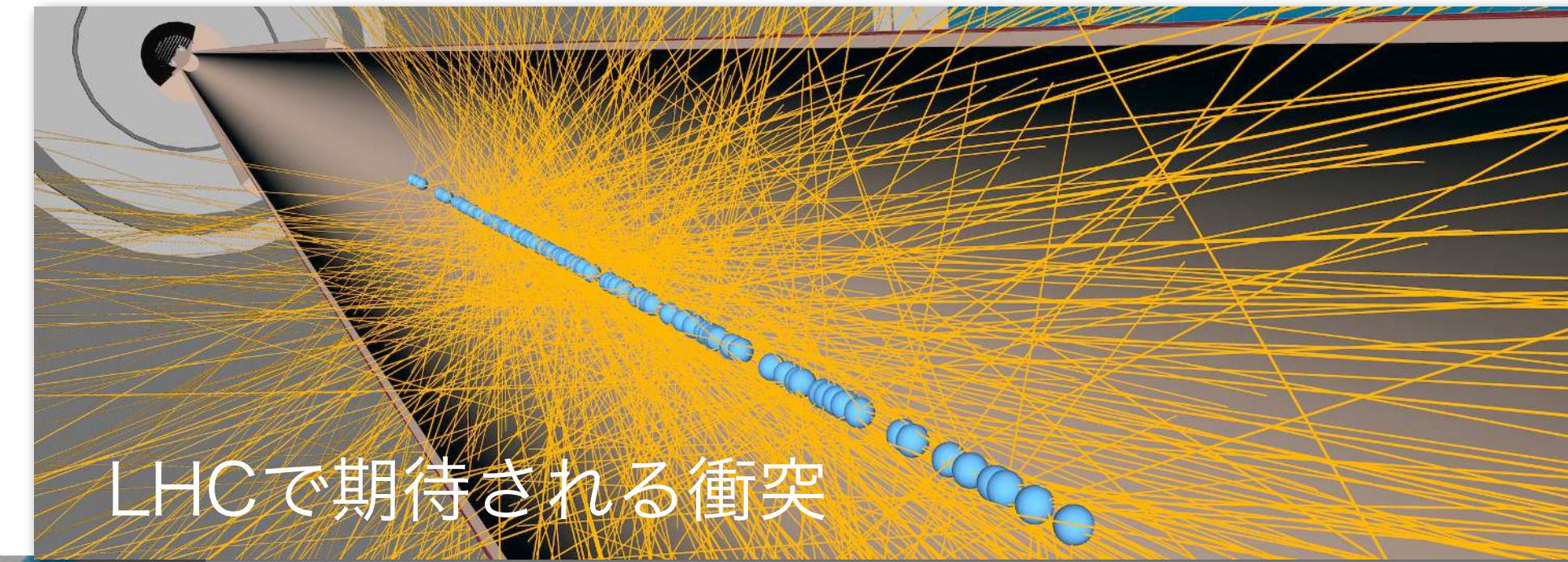
データ解析

測定器  
シミュレーション

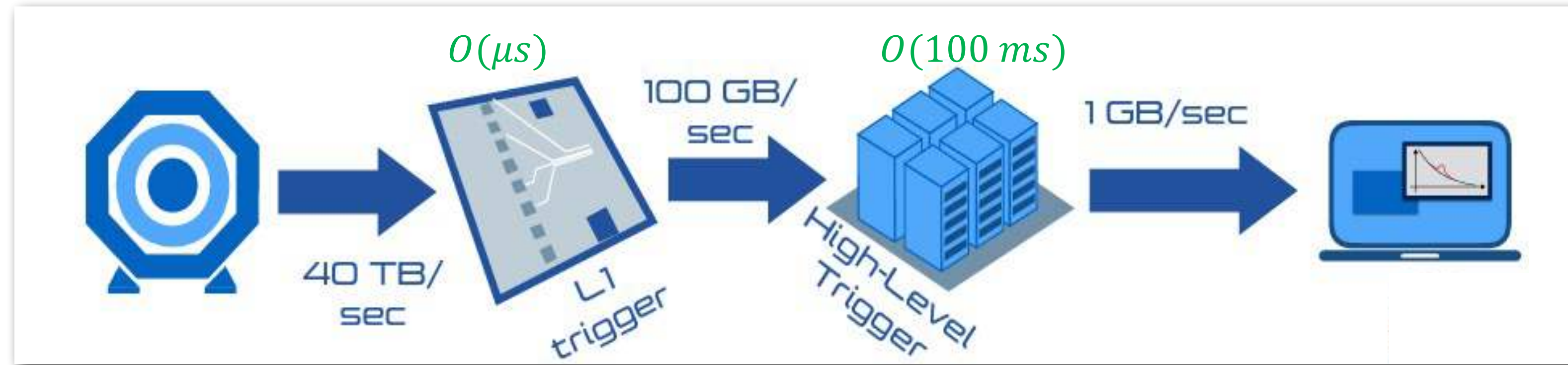
# LHCからHL(High Luminosity) -LHCへ

LHC(2025 年 run 3 終了 300fb-1 )

HL-LHC(2029年から 高輝度化 3000fb-1 )



# HL-LHC実験の計算コスト



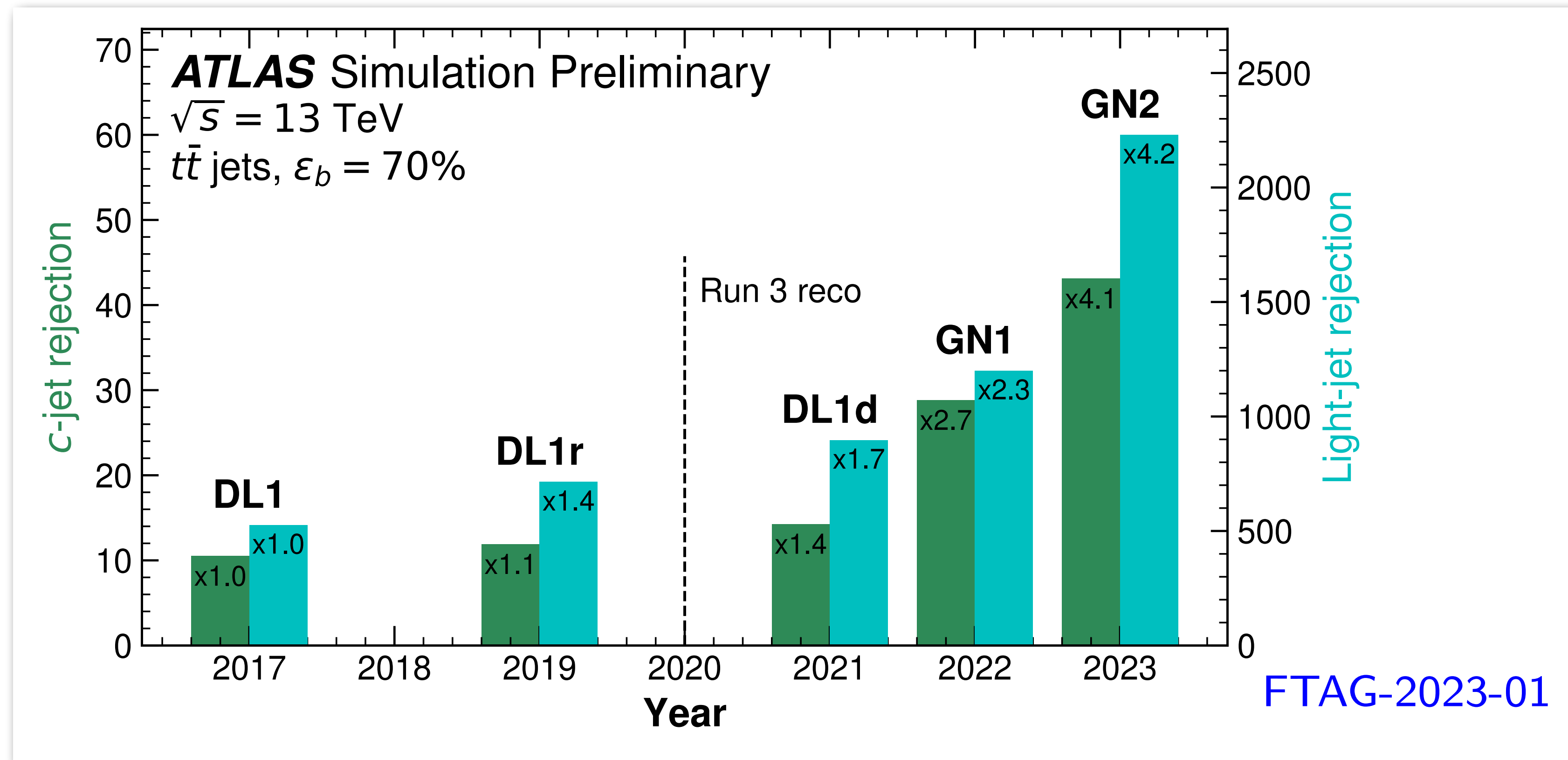
- 40TB/sec (factor 100 increase)
- 10倍くらい計算を高速化することを想定している。(測定器のシグナルを元の粒子と同定する、重要な事象を認識して保存するなど。)
- 計算量を確保するための予算が多額にならないようにしたいが、単純に考えると足りていない。今の工程表は、研究開発によって、計算予算が削減できることを仮定している。

# HL-LHC

- 衝突エネルギーは上がらないがより高いエネルギーの粒子衝突の数は増える
- 物理目標
  - Higgs 粒子や、top 粒子の性質を調べたい
  - 高いエネルギーの衝突からきた、top 粒子、Higgs 粒子の性質に変化が現れると期待されるので、これらの粒子に対する感度を向上させたい。
  - 深層学習が目立って活躍するようになってきた

# 深層学習学習の活躍：粒子再構成

深層学習によるボトムクオークを起源とするジェットと、より軽いジェットとの識別



$t \rightarrow b W$  崩壊や  $H \rightarrow bb$  崩壊を捉えるのとても重要

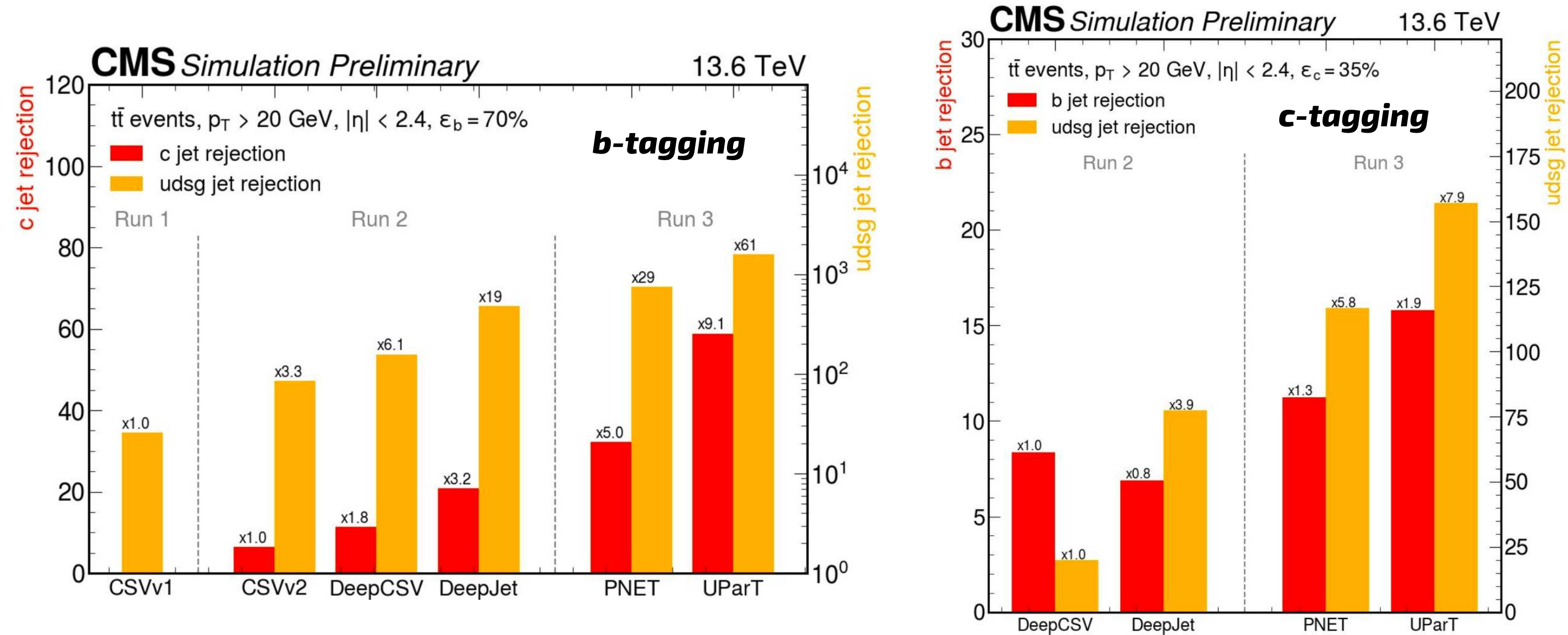
top 発見のころの技術の進歩と並ぶ勢い

# Flavor Tagging Performance



## b/c-tagging performance

- Promising performance compared to previous taggers
  - x3 better light jet rejection (at b-jet eff 70%) than DeepJet
  - x2 better light rejection + x2 better b-jet rejection (at c-jet eff 35 %) than DeepJet

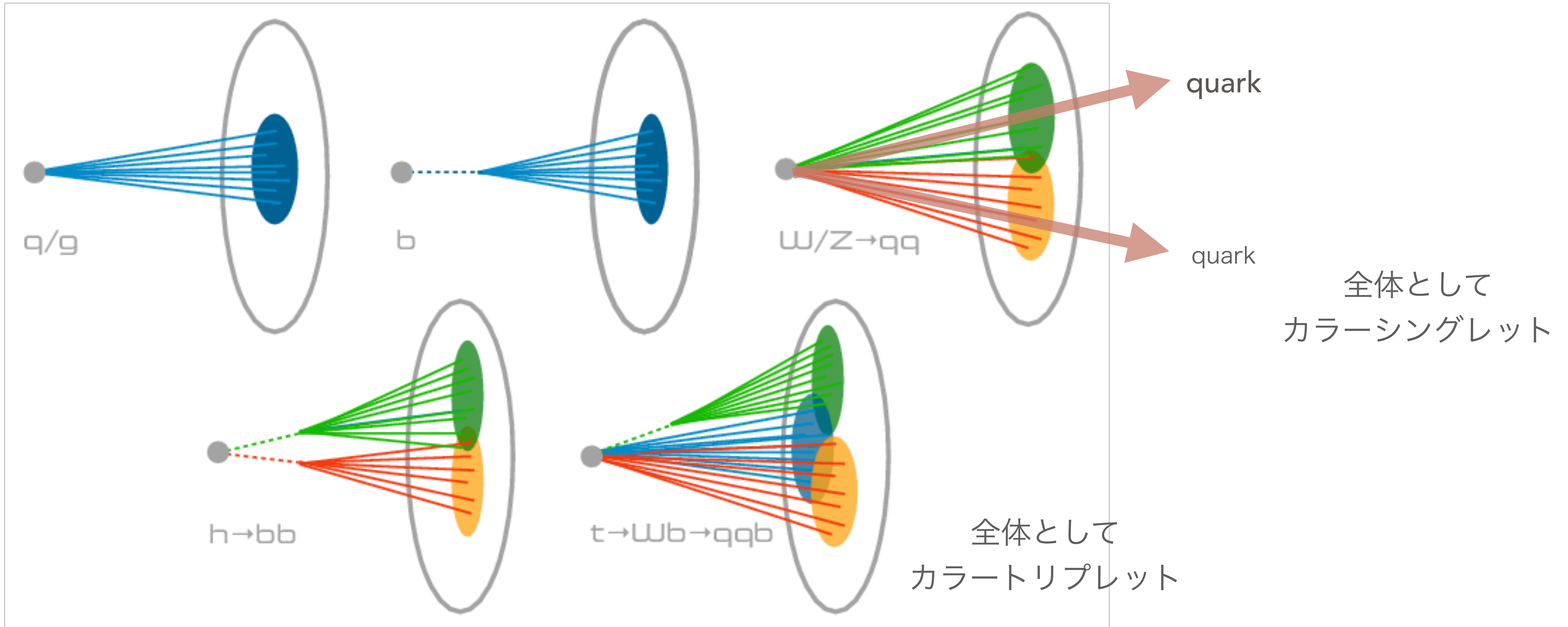


talk Jieun Choi, Top 2024 on behalf of CMS

vertex detector → top discovery(1995)

b tag and c tag improvement →??

# ジェット同定と深層学習



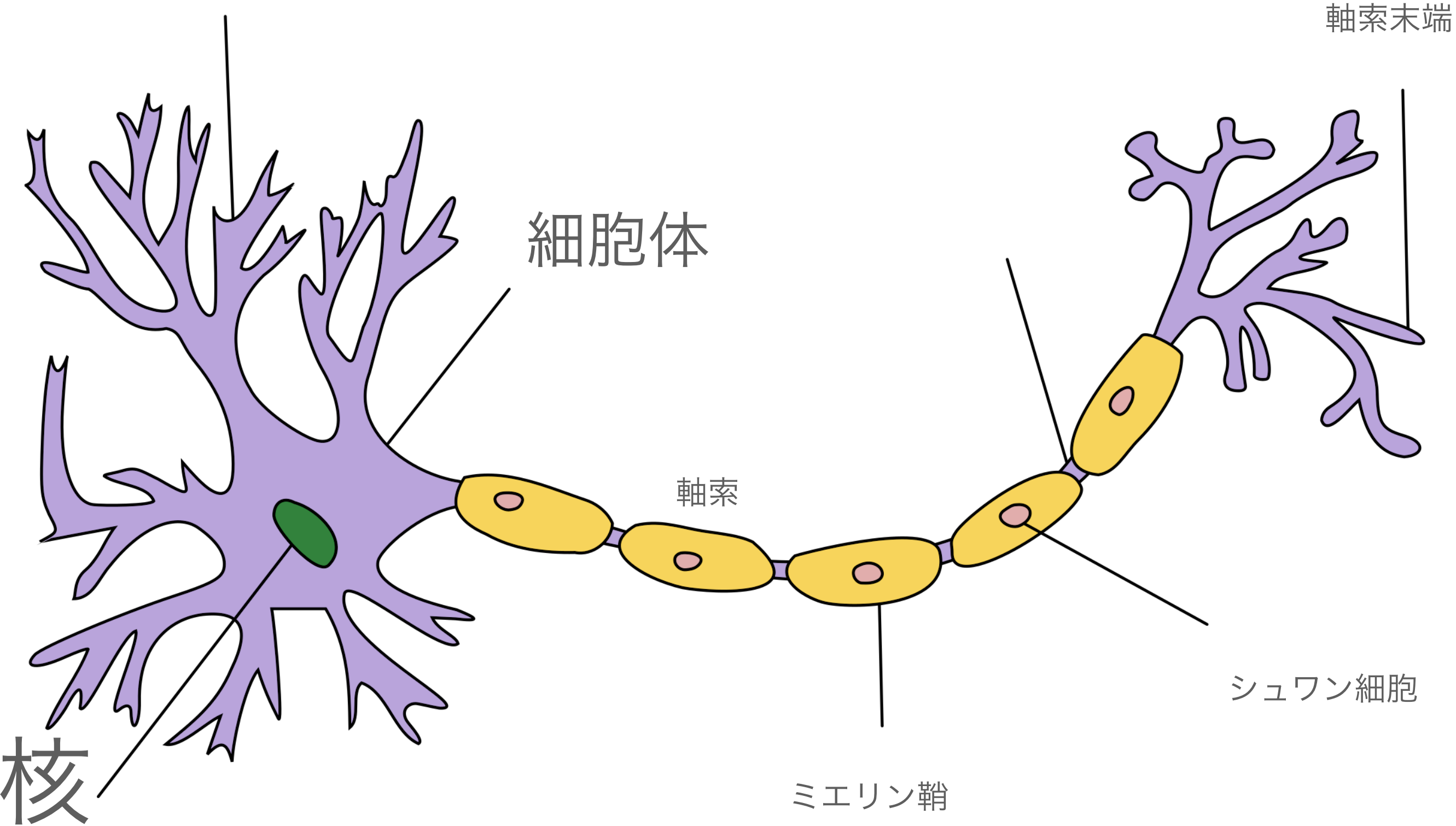
# 親粒子とジェットの性質

- カラーの違い: クォーク 3 グルーオン 8 ソフトな粒子の数はカラーが大きいほど多い
- W, Z, h クォークと反クォークに崩壊。内部構造が かならずある。
- W, Z, h: 長距離ではカラーがない 大きい角度方向の放出は少ない。クォークと反クォークの間に粒子を放出しがち
- t: 内部構造が3つ、 $t \rightarrow W b \rightarrow qq b$  W の質量をもつカラー0のジェットペア。

**深層学習とは何か ← 学習とは何か**

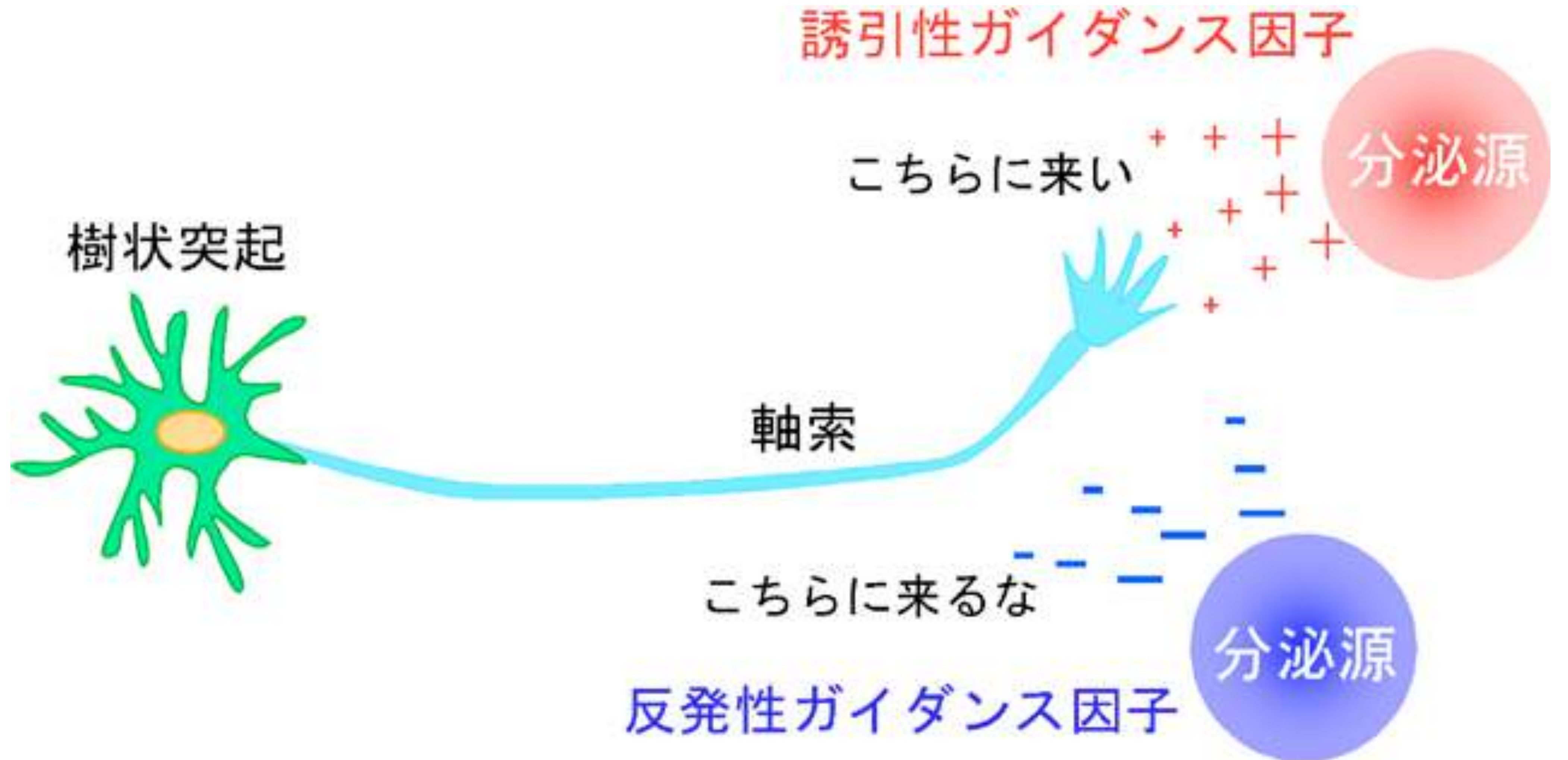
# 基本要素

## 樹状突起

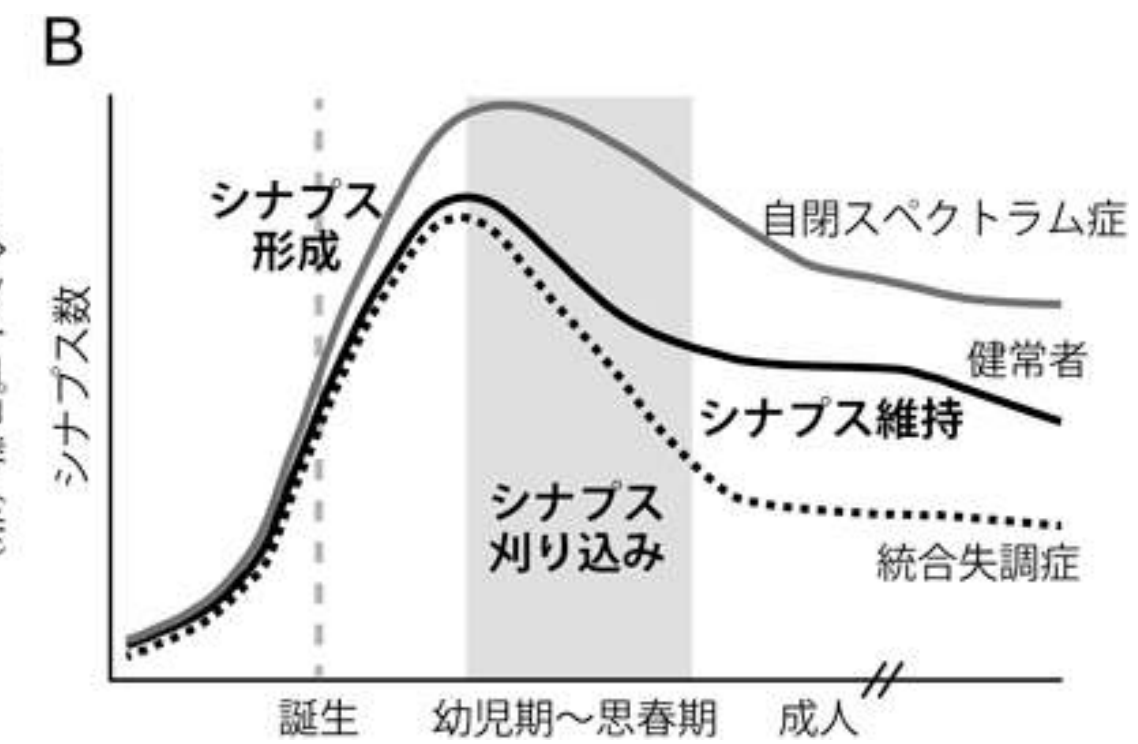
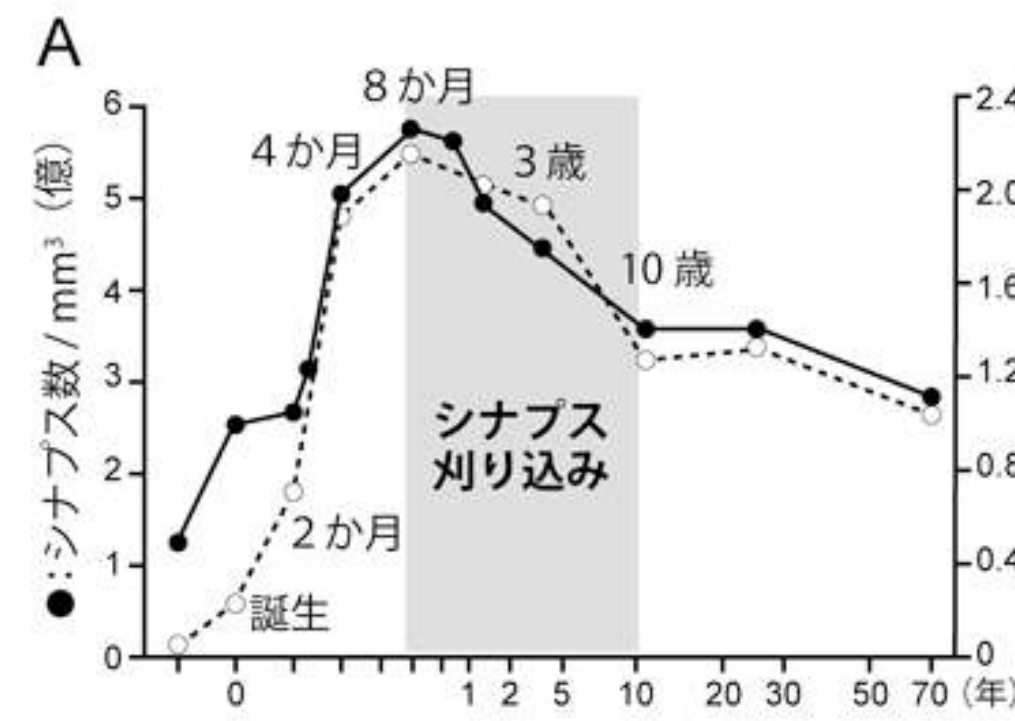
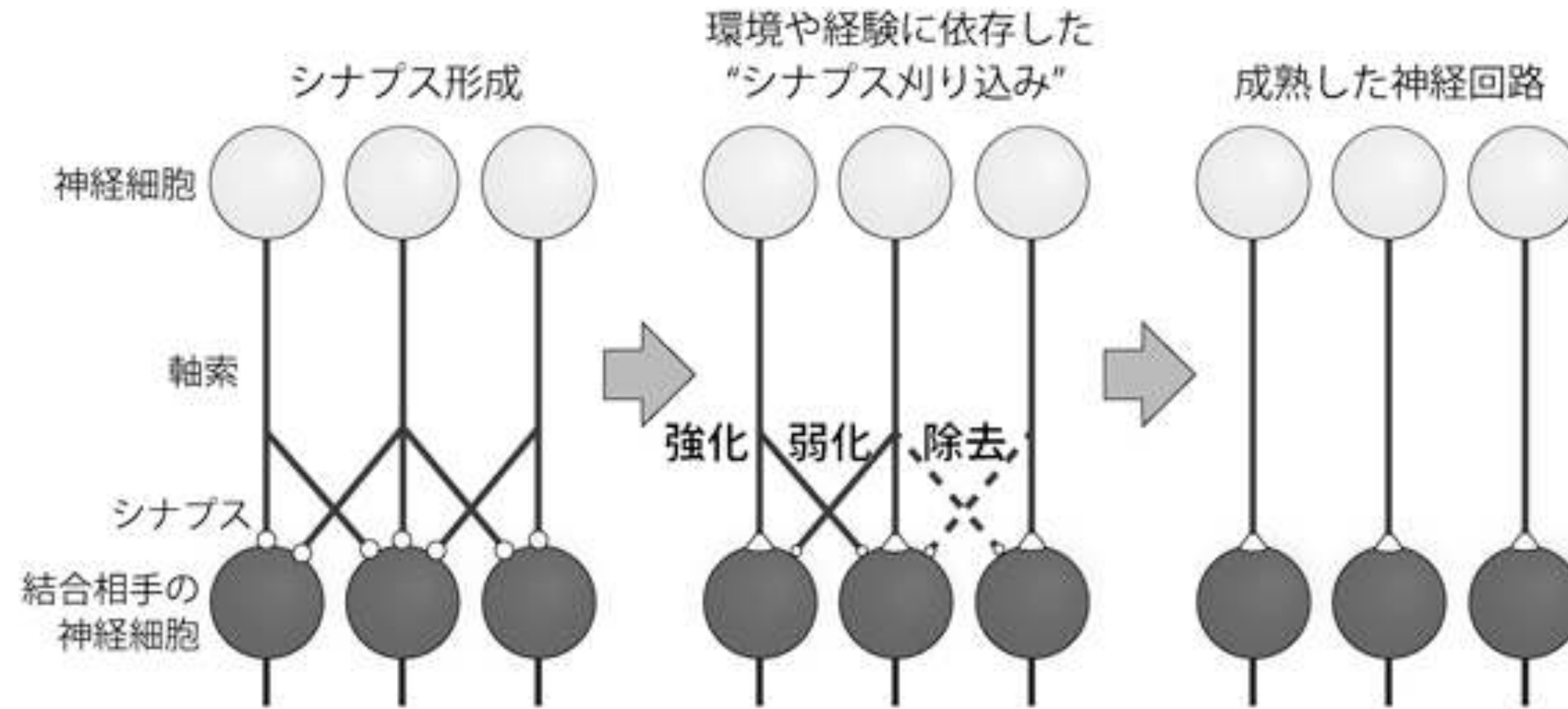


# ネットワークの形成

(報酬系)

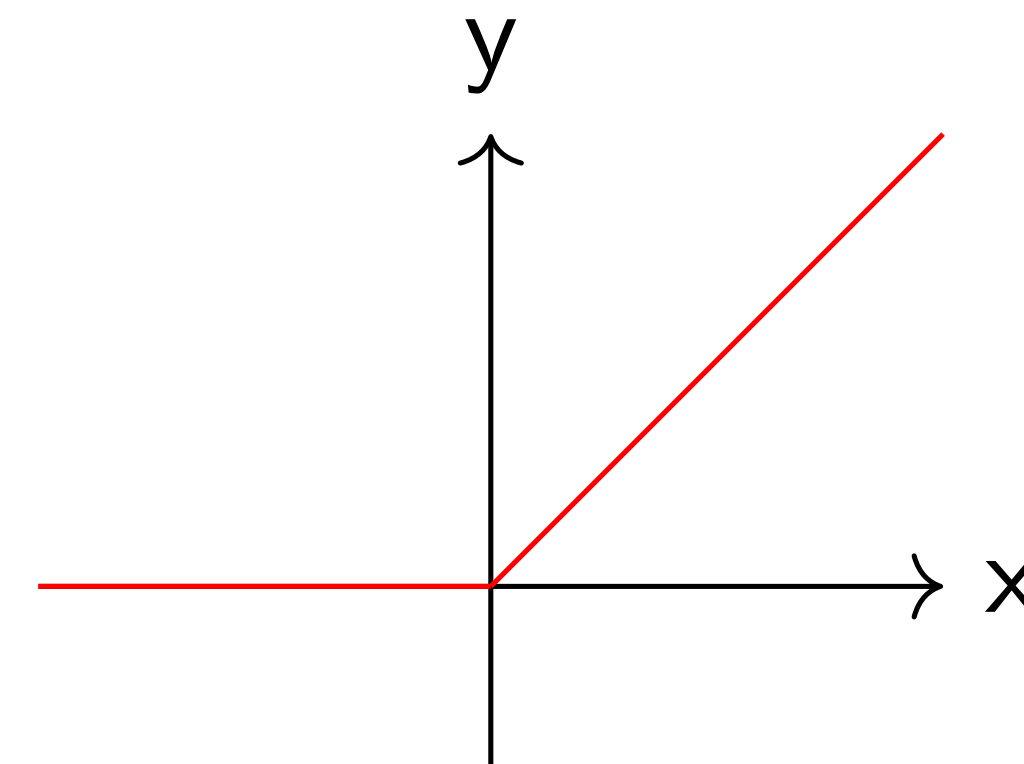


# ネットワークの形成



# 深層学習ネットワークの基本要素

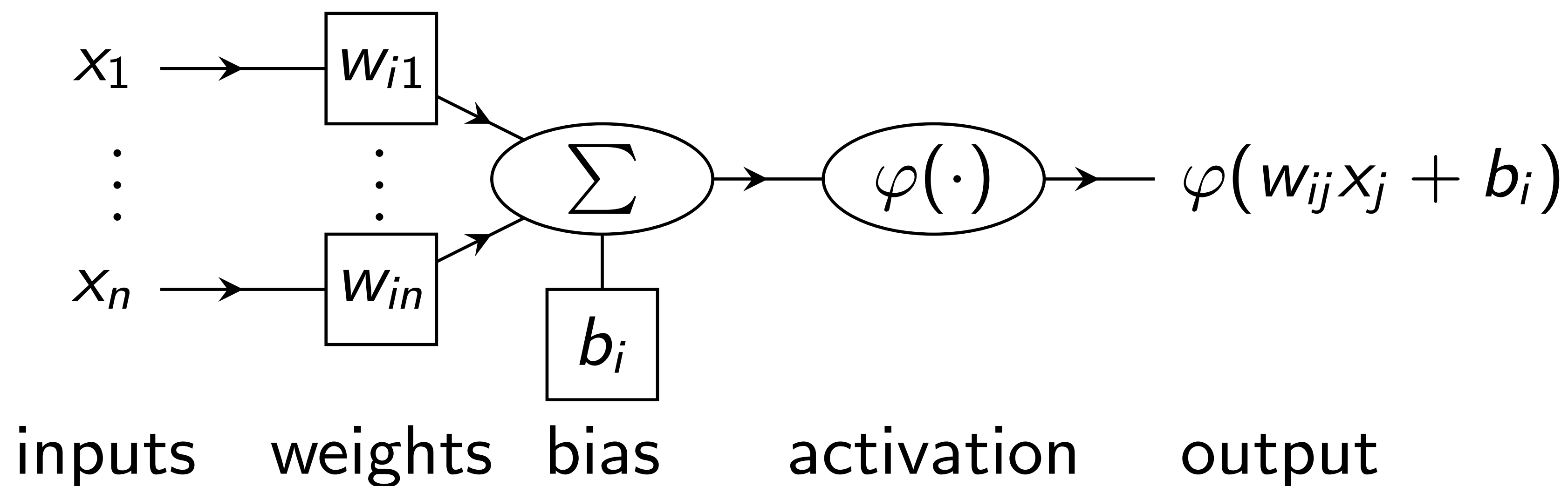
まず神経細胞を用意



Rectified Linear Unit(ReLU)

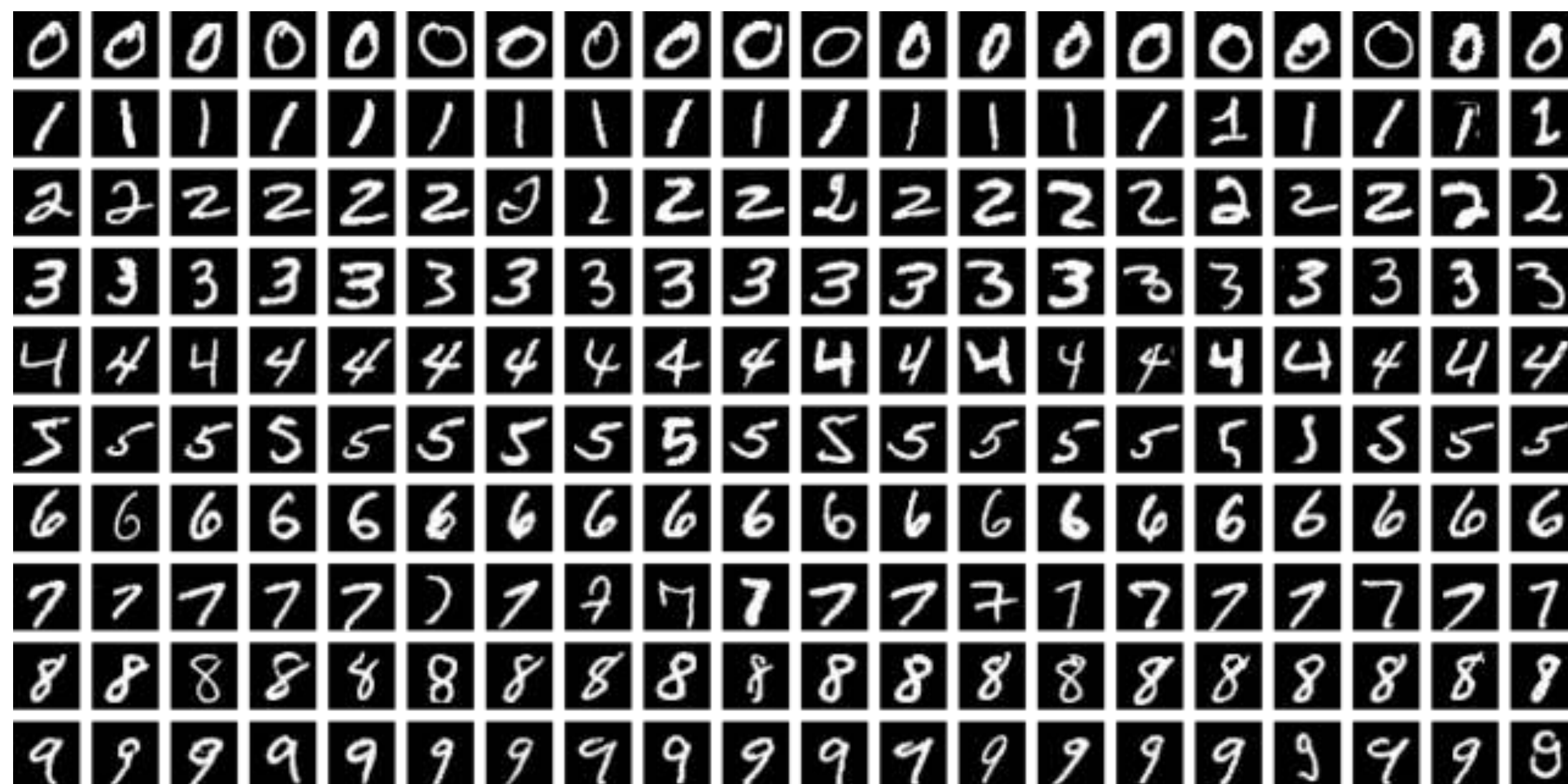
$$\varphi_{\text{ReLU}}(x) = x\theta(x)$$

$\varphi$ : 活性化関数 (非線形性)

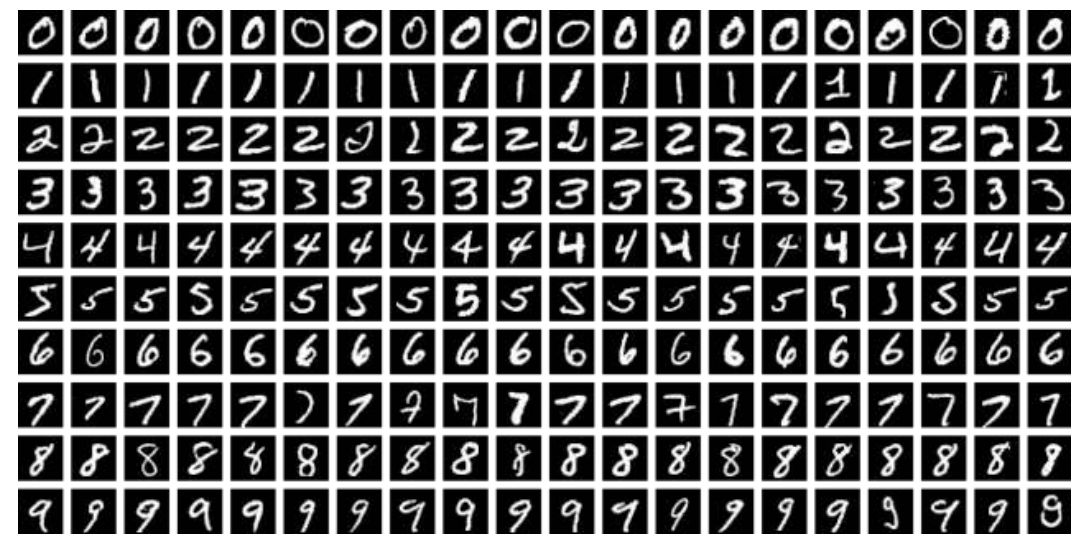


# 学習データ

- 手書き文字と数字
- 猫の絵と犬の絵と犬(0) か猫 (1) かのデータ
- 色々なものが書かれた絵と、書かれているものの名前



# 分類問題を扱うMLP (multiple layer perceptron )



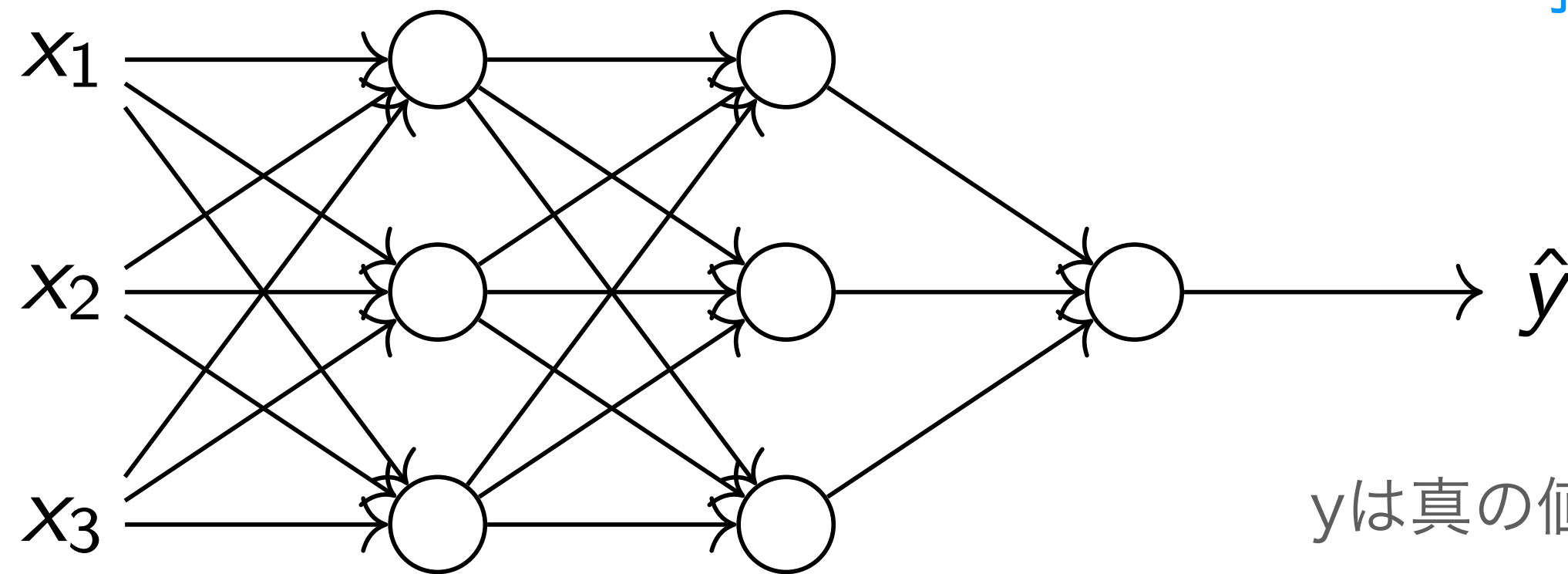
画像データプール

(28x28) の画像データをn 個

訓練による潜在変数の決定:  $w_{ij}, b_i$



損失関数 (loss function の最適化)



yは真の値  
 1 → (1, 0, 0, ..)  
 2 → (0, 1, 0, ...)  
 3 → (0, 0, 1, ...)

$$L(y, \hat{y})$$

ネットワークアウトプット  
 (n 個の10次元ベクトル)

"推定された確率"

$$\hat{y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_{10}), \sum_i \hat{y}_i = 1$$

$$\hat{y}_i = \exp(x_i) / \sum_j \exp(x_j)$$

- ✓表現力 expressive power
- ✓データを学習 learn from data
- ✓微分可能 Simple linear algebra + activation

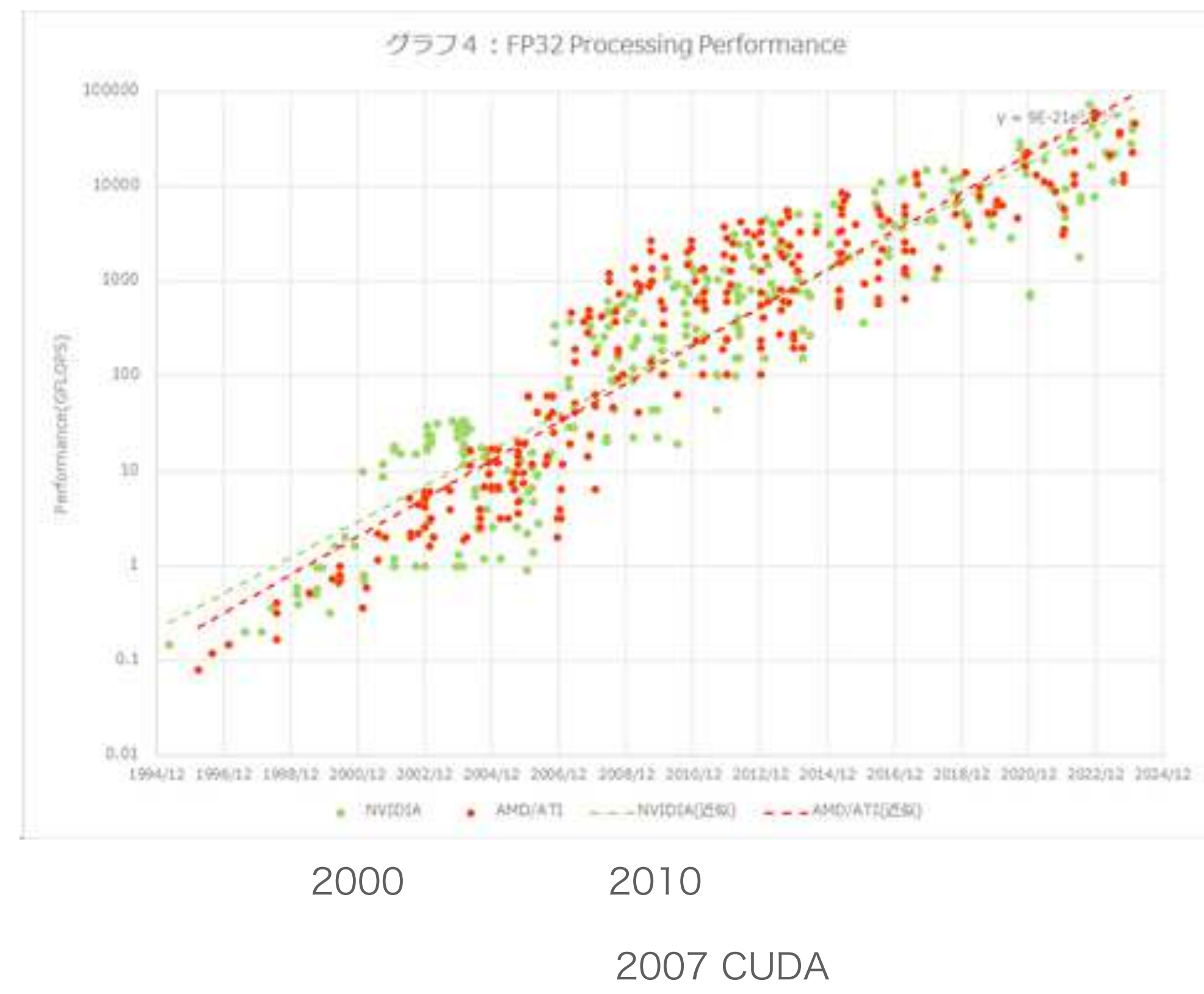
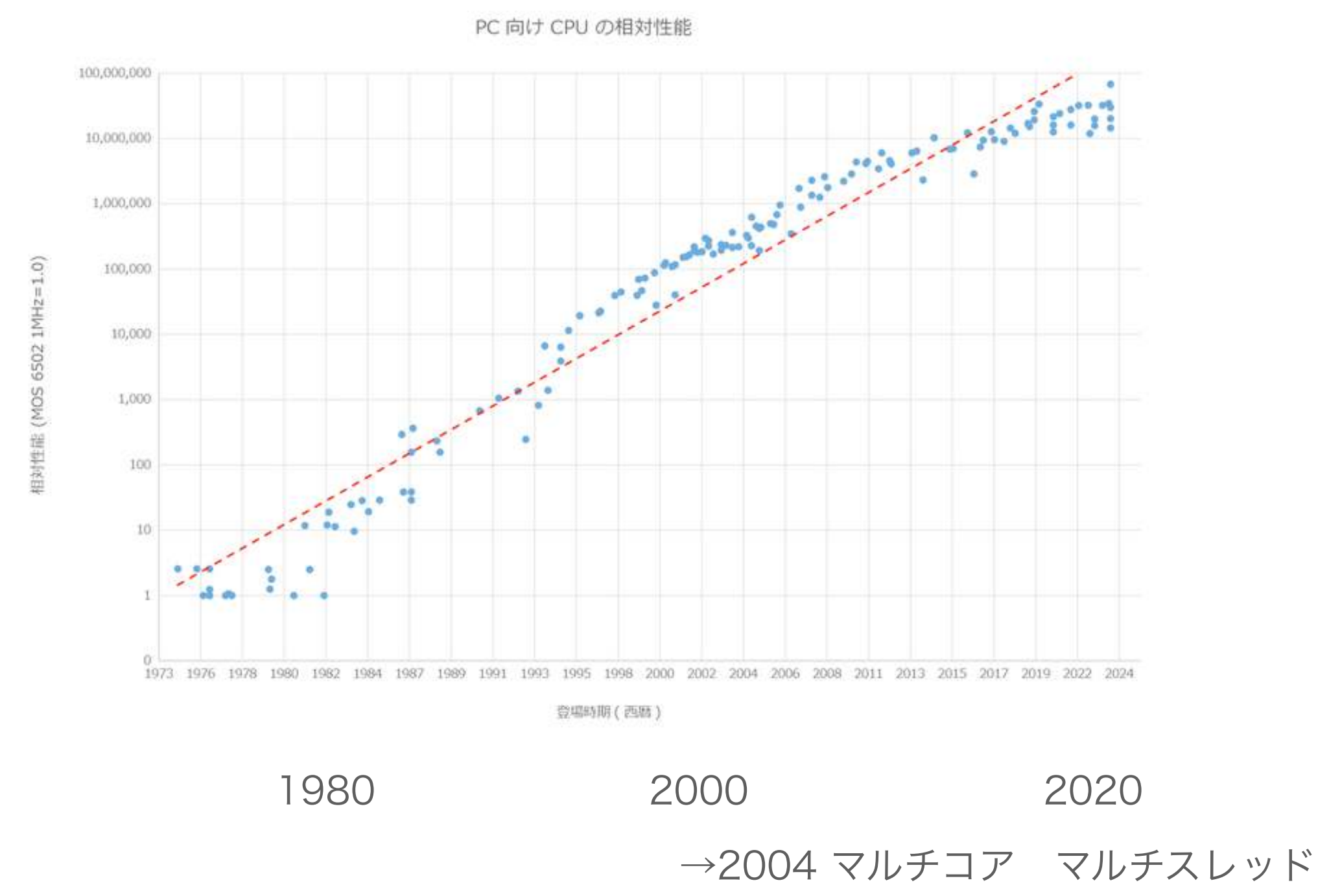
訓練後  $\hat{y}$  を使って "らしさ"を推定

# コンピューターの性能も重要

<https://singula2045.hatenablog.com/entry/2024/08/18/113413>

## PC 向けCPU の計算力の進化

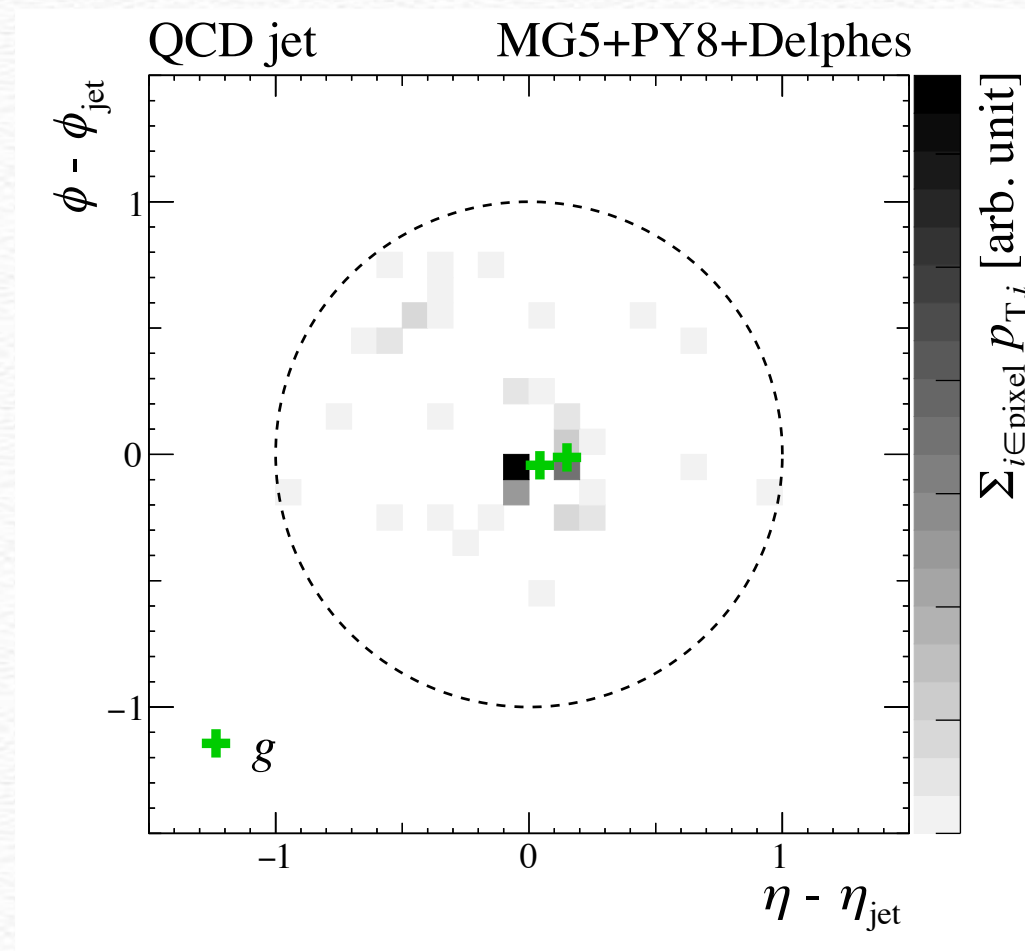
## G(raphics) PUの計算力の進化



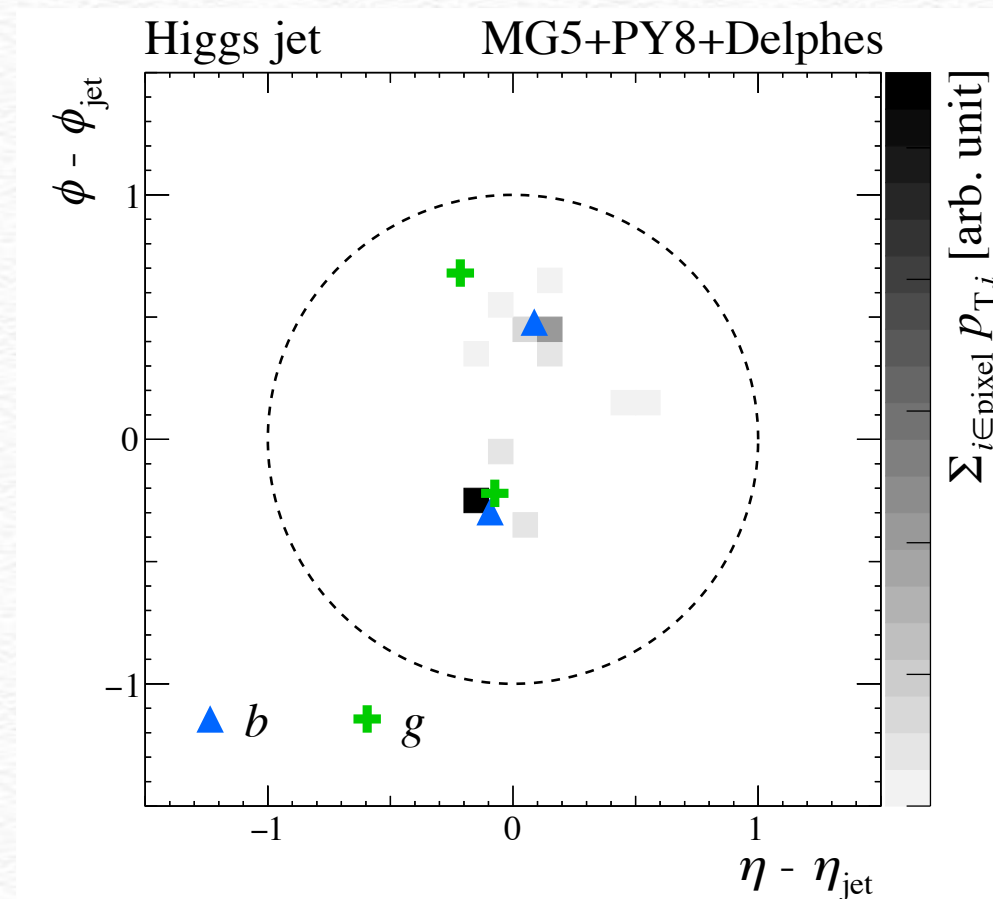
# ジェットの同定問題

Input: Jet 内部の粒子の位置情報

QCD (j+Z( $\rightarrow$ vv))

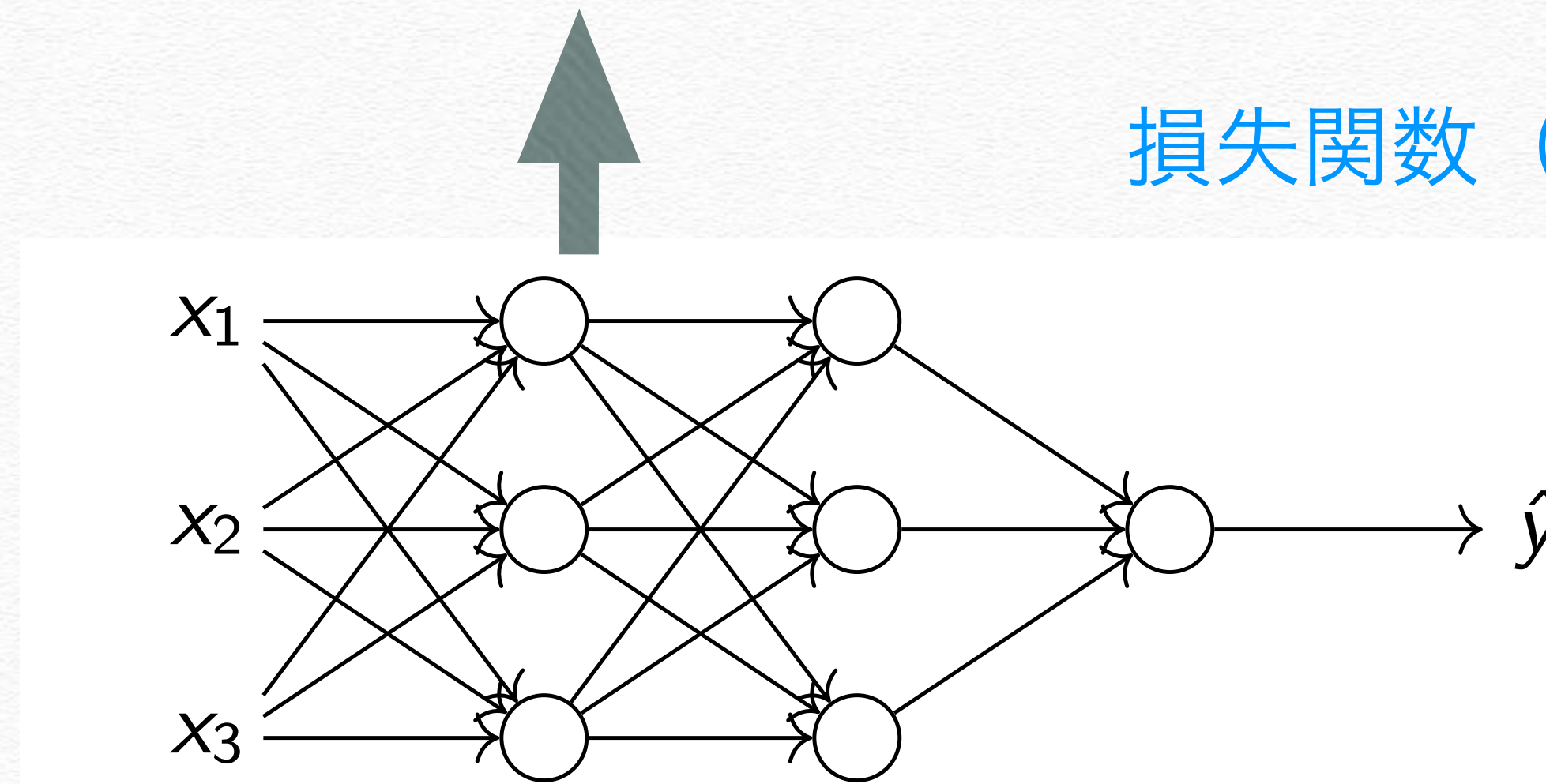


Higgs h+Z( $\rightarrow$ vv)



訓練による潜在変数の決定:  $w_{ij}, b_i$

損失関数 (loss function の最適化)

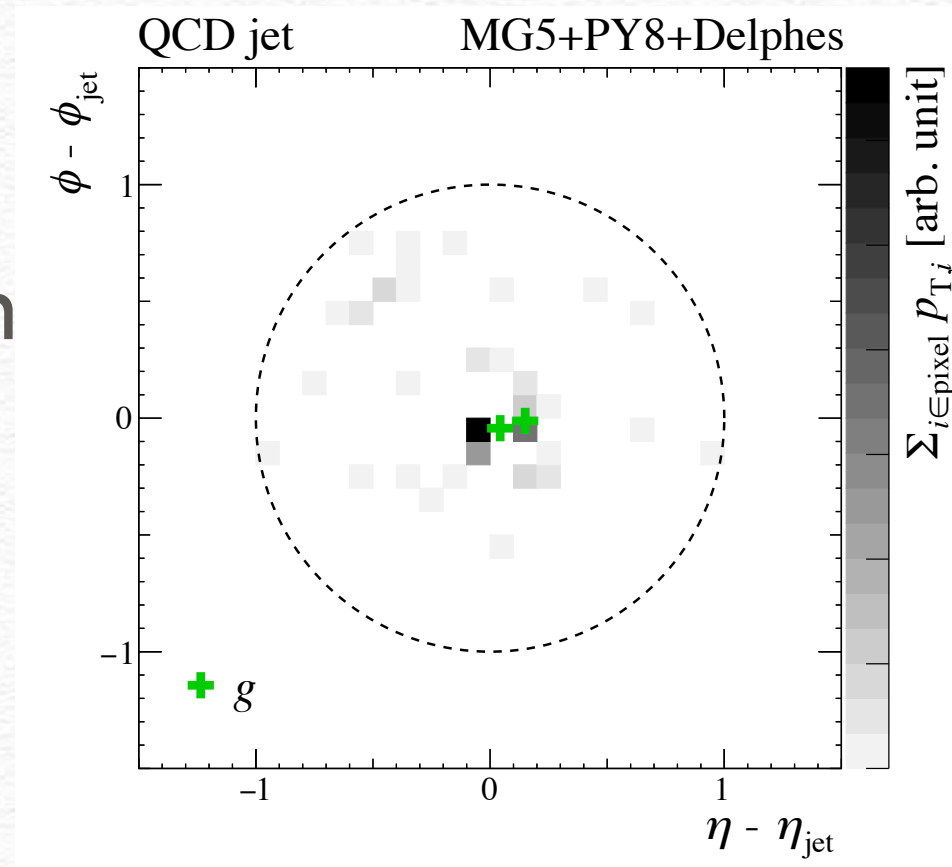


Higgs  
 $y = (1, 0)$   
 QCD  
 $y = (0, 1)$   
 $L(y, \hat{y})$

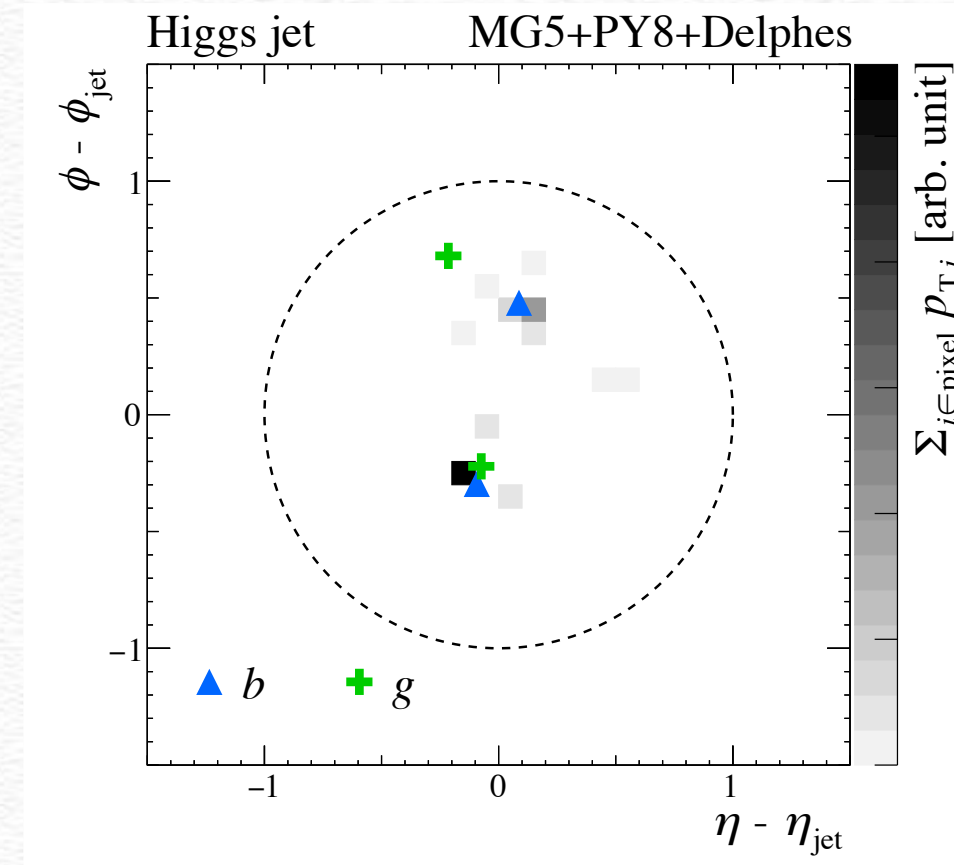
訓練後  $\hat{y}$  を使って、粒子らしさを推定

# 深層学習を使ったジェット同定

QCD jet  
(in W mass region)



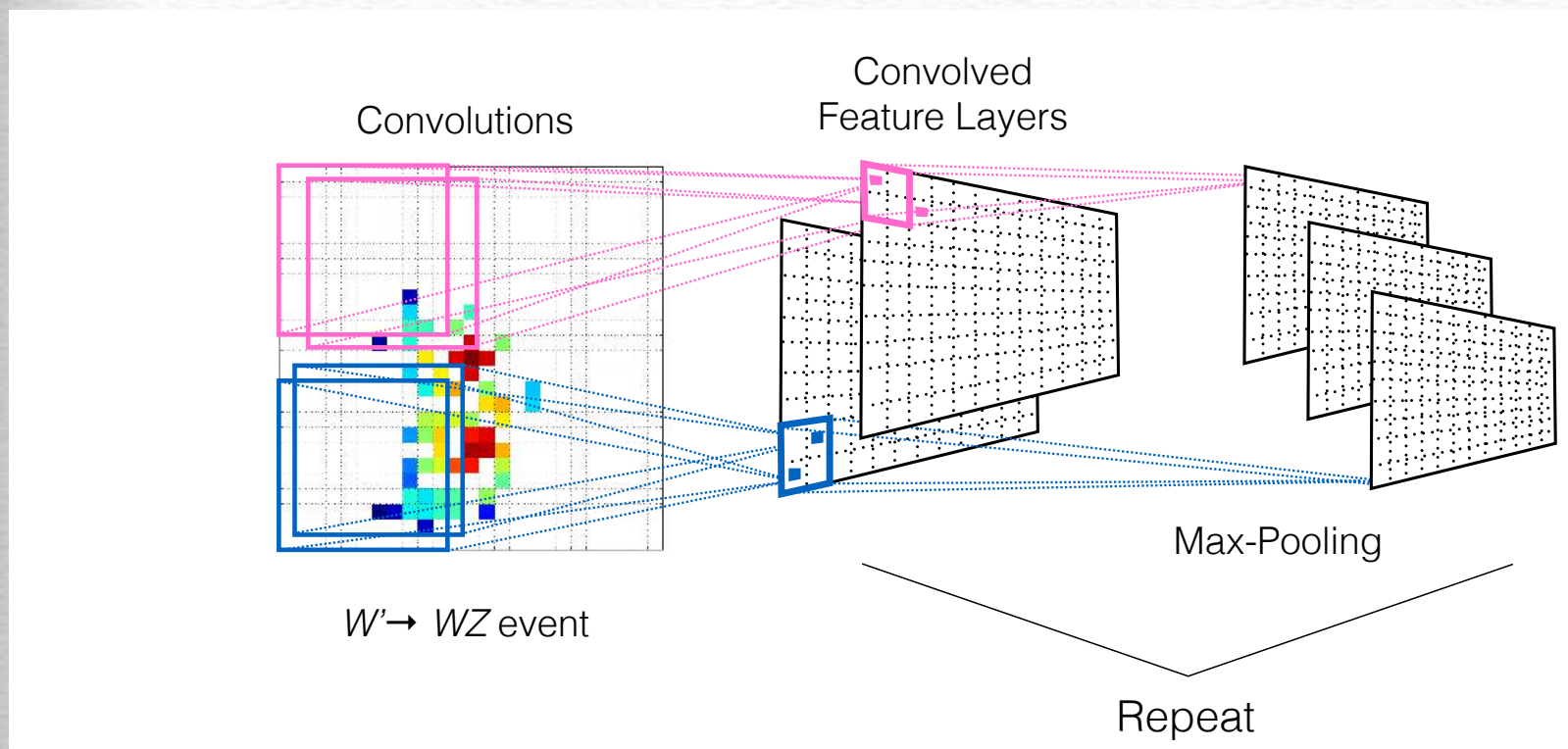
H jet



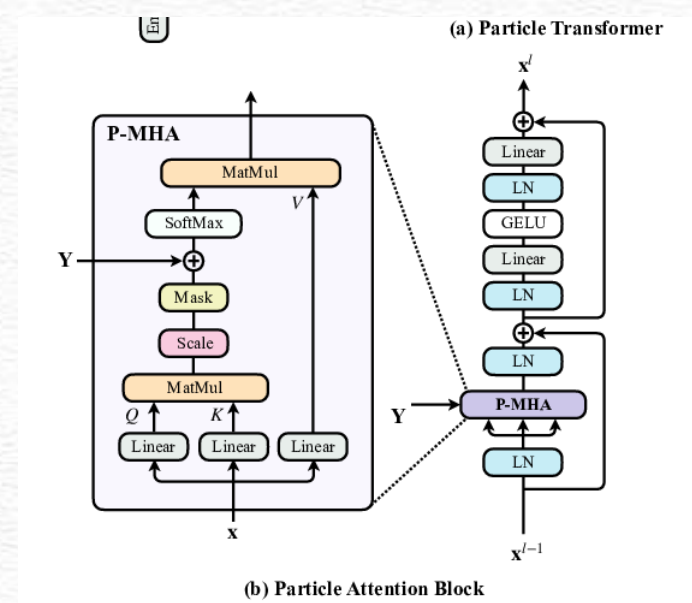
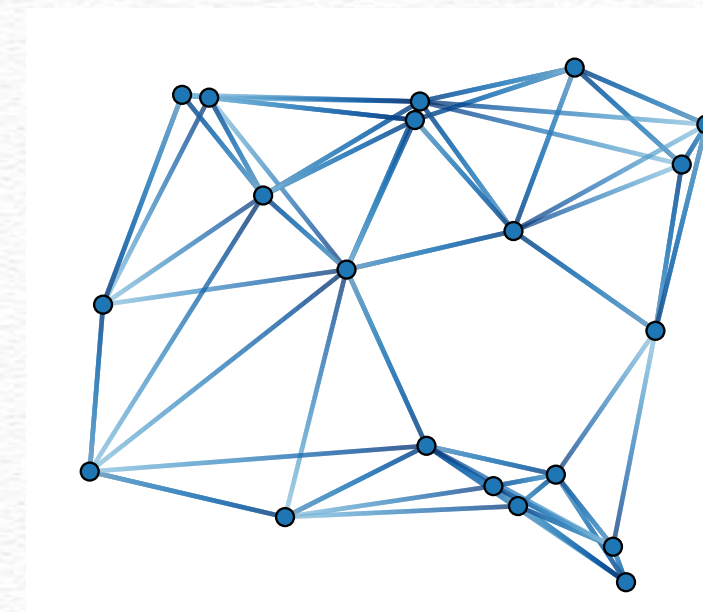
transformer  
ジェット全体から  
まんべんなく

画像データ (位置とエネルギー) ← CNN

Graph Neural network  
(入力データの節約、他の情報も入れられる)



→ sparse data



CNN Oliverira et al  
(1511.05190)

1902.08570 Particle Net

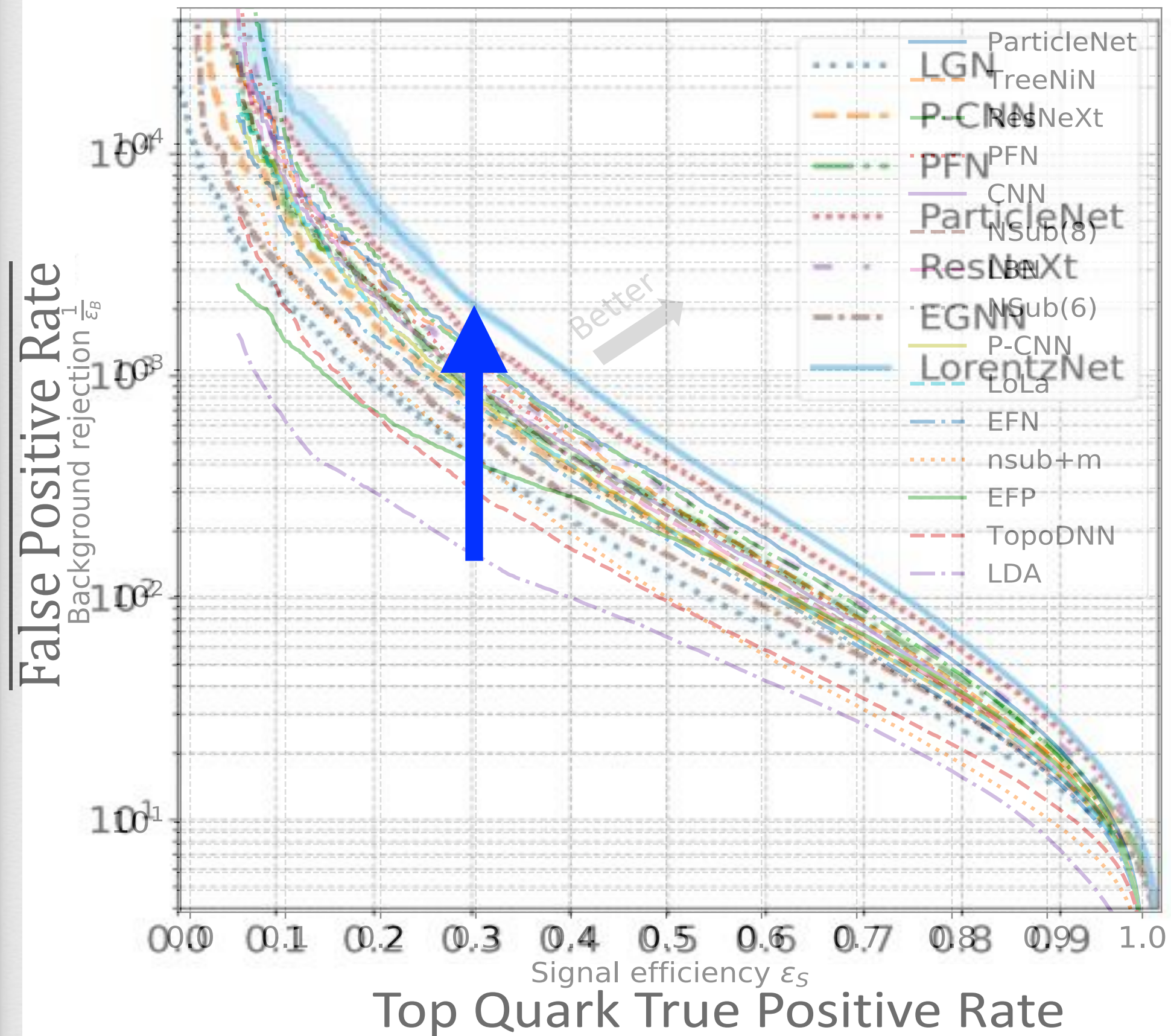
Particle transformer  
2202.03772

Dreyer et al LundNet (1807.04758)  
Gong et al LorentzNet(2201.08187)

Bogatskiy et al PELICAN (2211.00454)

2018 to 2021

## Improved Performance



[2201.08187](#)

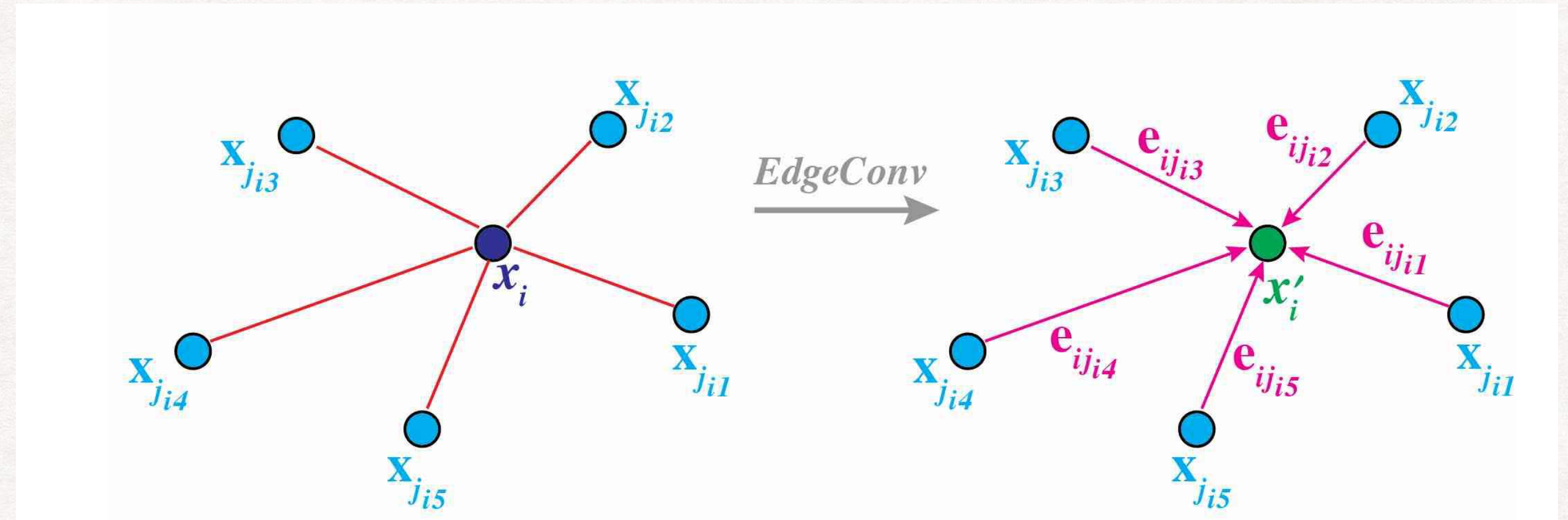
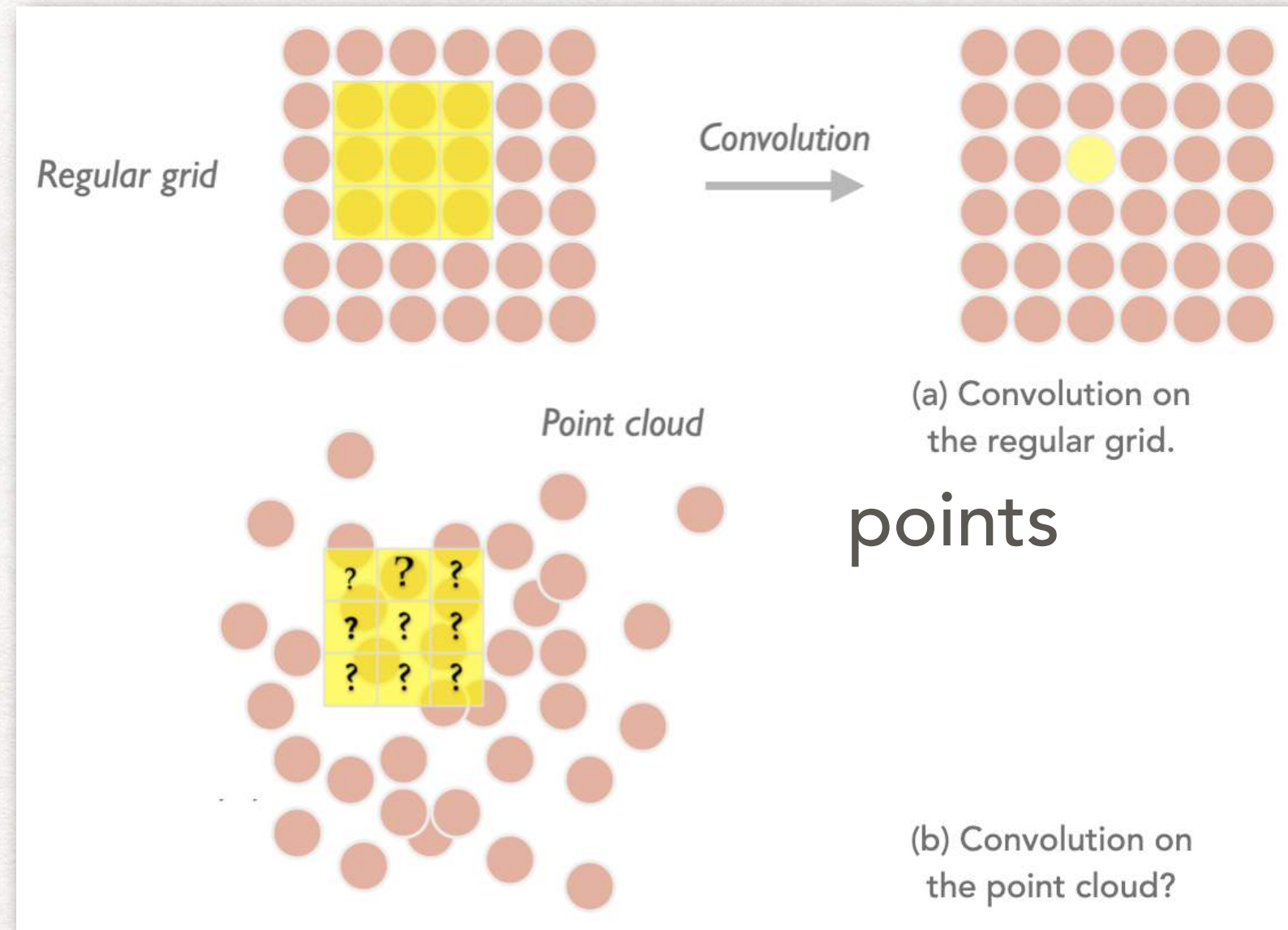
良いネットワークを使うことで  
パフォーマンスは良くなる

## ポイント

1. 粗いデータをどういうふうに扱うか
2. 対称性を重視→入力データが少なくても収束  
粒子の入れ替えに対する不変性  
並進に対する不変性  
ローレンツ変換に対する共変性
3. 物理的な構造を反映したネットワーク  
QCD の相互作用のパターン

# PARTICLE NET(1902.08570)

image(CNN)



$$x'_i = \square_{j=1}^k h_{\Theta}(x_i, x_{ij}) \quad x_i \in R^F$$

エッジ関数  $h_{\Theta}(x_i, x_{ij}) : R^F \times R^F \rightarrow R^{F'}$

$\square$  : channel wise symmetric operation (最大、平均、和など)

k ; 近傍の粒子数 7~16

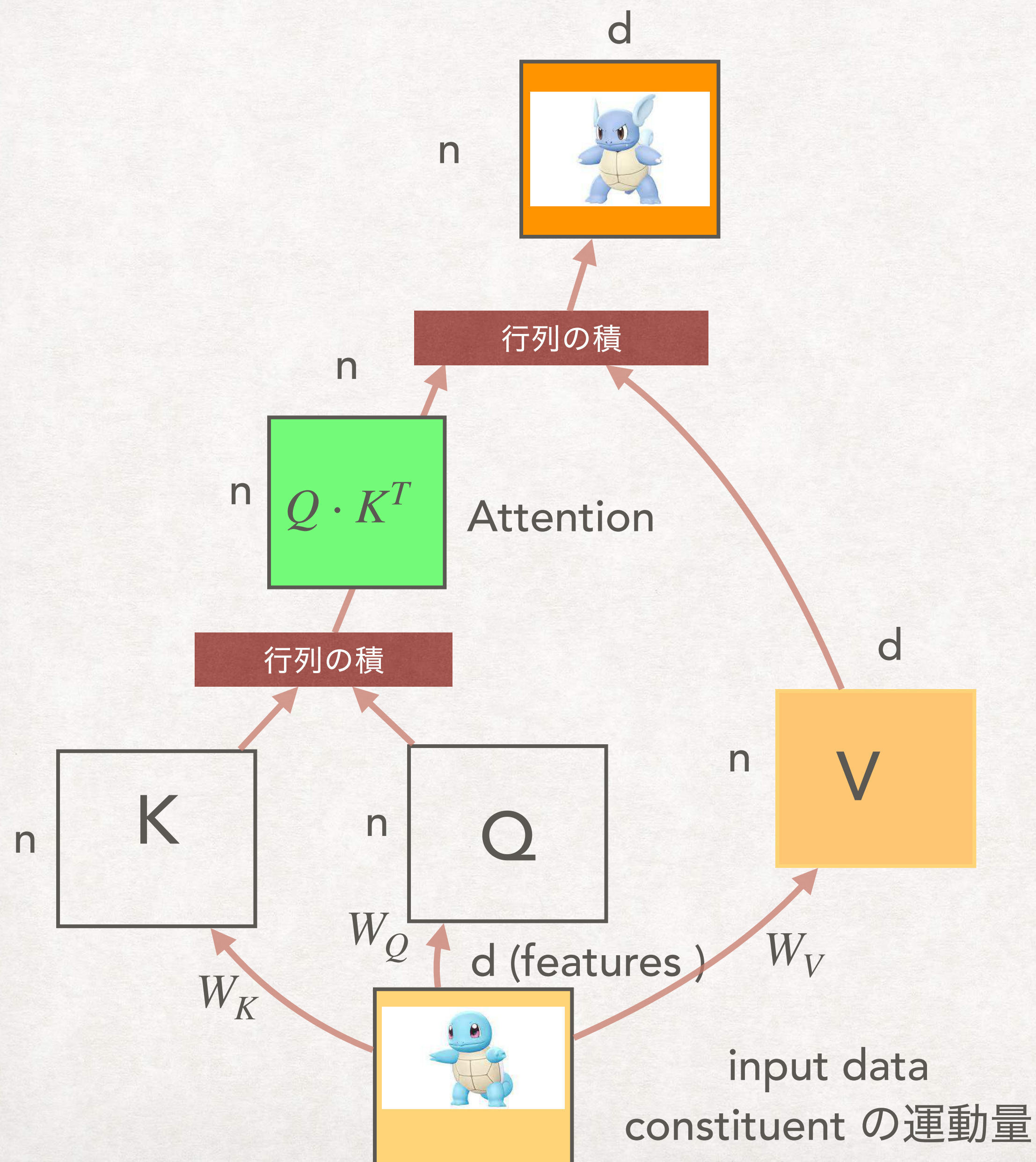
ジェットは疎なデータ(sparse data)なので画像と捉えるのはあまり賢くない。

GNN は近傍という概念に囚われて、グローバルな状況を掴むのに難がある

output size = input size

# “TRANSFORMER” : SELF ATTENTION

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$



- トランスフォーマーのコア部分は「Attention Matrix」
- 全ての粒子の間の訓練 (W)可能な相関を特徴すべてを使って構成する。
- アテンションブロックの入力と出力の次元が同じ。
- $X \rightarrow X' \rightarrow X''$ と入力が増減されずにどんどん進化していき、お互いの関係が Attention に整理されていく

強力だけど、計算、メモリーとも大変

# なんでもデータとしていられるのが楽しい

Table 2. Particle input features used for jet tagging on the JETCLASS, the top quark tagging (TOP) and the quark gluon tagging (QG) datasets. For QG, we consider two scenarios: QG<sub>exp</sub> is restricted to use only the 5-class experimentally realistic particle identification information, while QG<sub>full</sub> uses the full set of particle identification information in the dataset and further distinguish between different types of charged hadrons and neutral hadrons.

運動量

電荷、particle ID など

衝突点から飛んだ距離

Category	Variable	Definition	JETCLASS	TOP	QG <sub>exp</sub>	QG <sub>full</sub>
Kinematics	$\Delta\eta$	difference in pseudorapidity $\eta$ between the particle and the jet axis	✓	✓	✓	✓
	$\Delta\phi$	difference in azimuthal angle $\phi$ between the particle and the jet axis	✓	✓	✓	✓
	$\log p_T$	logarithm of the particle's transverse momentum $p_T$	✓	✓	✓	✓
	$\log E$	logarithm of the particle's energy	✓	✓	✓	✓
	$\log \frac{p_T}{p_{T(\text{jet})}}$	logarithm of the particle's $p_T$ relative to the jet $p_T$	✓	✓	✓	✓
	$\log \frac{E}{E(\text{jet})}$	logarithm of the particle's energy relative to the jet energy	✓	✓	✓	✓
	$\Delta R$	angular separation between the particle and the jet axis ( $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ )	✓	✓	✓	✓
Particle identification	charge	electric charge of the particle	✓	—	✓	✓
	Electron	if the particle is an electron ( $ \text{pid} ==11$ )	✓	—	✓	✓
	Muon	if the particle is an muon ( $ \text{pid} ==13$ )	✓	—	✓	✓
	Photon	if the particle is an photon ( $\text{pid}==22$ )	✓	—	✓	✓
	CH	if the particle is an charged hadron ( $ \text{pid} ==211$ or $321$ or $2212$ )	✓	—	✓	✓ <sup>a</sup>
	NH	if the particle is an neutral hadron ( $ \text{pid} ==130$ or $2112$ or $0$ )	✓	—	✓	✓ <sup>b</sup>
Trajectory displacement	$\tanh d_0$	hyperbolic tangent of the transverse impact parameter value	✓	—	—	—
	$\tanh d_z$	hyperbolic tangent of the longitudinal impact parameter value	✓	—	—	—
	$\sigma_{d_0}$	error of the measured transverse impact parameter	✓	—	—	—
	$\sigma_{d_z}$	error of the measured longitudinal impact parameter	✓	—	—	—

<sup>a</sup> ( $|\text{pid}|==211$ ) + ( $|\text{pid}|==321$ )\*0.5 + ( $|\text{pid}|==2212$ )\*0.2

<sup>b</sup> ( $|\text{pid}|==130$ ) + ( $|\text{pid}|==2112$ )\*0.2

# KEK グループの研究 パートンと粒子との間の関係を解析するCROSS ATTENTION

“Ahmed Hammad , MN “Streamlined jet tagging network assisted by jet prong structure” arXiv 2404 14677 JHEP 06 (2024) 176

- ・ 高いエネルギースケール クォークやグルーオン  $\{y\}$
- ・ フラグメンテーション:  $P(\text{ジェット中の粒子} | \text{パートン} \sim \text{jet}) = P(\{x_i\} | \{y\})$
- ・ 構造をもつジェット  $P(\{x_i\} | \{y_\alpha\})$

QCD ではsubject の運動量 $\sim$ パートンの運動量なので、 subject との相関を解析すれば良い。

# ネットワークのパフォーマンスの比較

グラフニューラルネットワーク  
計算量が膨大

ローレンツ対称性をリスペクトすると  
良い結果に。実験データではうまくいか  
ないという報告も

トランスフォーマー

	Accuracy	AUC	$1/\epsilon_B(\epsilon_s = 0.5)$	$1/\epsilon_B(\epsilon_s = 0.3)$	Parameters
ParticleNet[52]	0.940	0.9858	$397 \pm 7$	$1615 \pm 93$	370K
PFN[51]	—	0.9819	$247 \pm 3$	$888 \pm 17$	86.1K
rPCN[72]	—	0.984	$364 \pm 9$	$1642 \pm 93$	—
<b>Lorentz invariance based networks</b>					
PELICAN[35]	0.9426	0.987	—	$2250 \pm 75$	208K
LorentzNet[73]	0.942	0.9868	$498 \pm 18$	$2195 \pm 173$	224K
L-GATr[74]	0.942	0.9870	$540 \pm 20$	$2240 \pm 70$	—
<b>Attention based networks</b>					
ParT[49]	0.940	0.9858	$413 \pm 6$	$1602 \pm 81$	2.14M
MIParT[50]	0.942	0.9868	$505 \pm 8$	$2010 \pm 97$	720.9K
Mixer[21]	0.940	0.9859	$416 \pm 5$	—	86.03K
OmniLearn[75]	0.942	0.9872	$568 \pm 9$	$2647 \pm 192$	1.6M
Plain Transformer*	0.927	0.979	$362 \pm 7$	$780 \pm 73$	1.7M
<b>IAFormer*</b>	<b>0.942</b>	<b>0.987</b>	<b><math>510 \pm 6</math></b>	<b><math>2012 \pm 30</math></b>	<b>211K</b>