A novel technique to search for heavy stable charged particles with the CMS detector

Petar MAKSIMOVIC, Morris SWARTZ, Tamas Almos VAMI*

*Email: tvami@jhu.edu

GBO Exam

Outline

- Introduction to the Standard Model
- Beyond Standard Model (BSM)
- How to look for BSM physics: the LHC and the CMS detector
- Pixel detector and the template reconstruction algorithm
- A novel technique to search for BSM physics
- A proof of concept using simulations
- The struggle of big data and non-standard searches
- The further future: exploiting this technique for other searches

Standard Model (SM)

The SM is a Quantum Field Theory describing 3 out of the 4 known fundamental forces

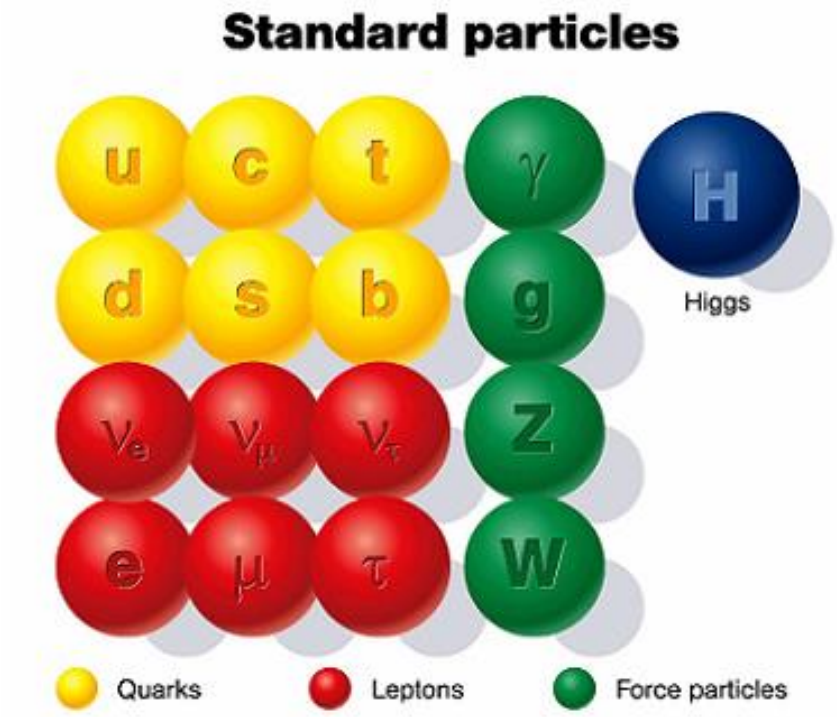
Standard Model (SM)

The SM is a Quantum Field Theory describing 3 out of the 4 known fundamental forces

- The electromagnetic force (photons)
- The strong interaction (gluons)
- The weak interaction (gauge bosons)
- Higgs boson is responsible for the mass of the elementary particles

Standard Model (SM)

The SM is a Quantum Field Theory describing 3 out of the 4 known fundamental forces

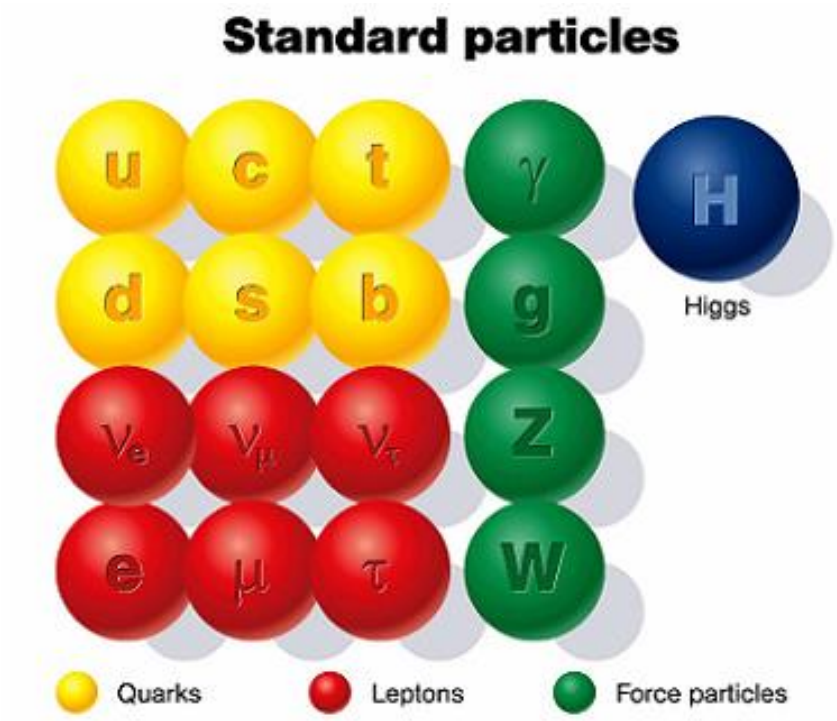


- The electromagnetic force (photons)
- The strong interaction (gluons)
- The weak interaction (gauge bosons)
- Higgs boson is responsible for the mass of the elementary particles

- Quarks and leptons make up matter
- Quarks interact with all the 3 forces
- (Charged) leptons interact with (EM and) weak interaction(s)

Standard Model (SM)

The SM is a Quantum Field Theory describing 3 out of the 4 known fundamental forces



- The electromagnetic force (photons)
- The strong interaction (gluons)
- The weak interaction (gauge bosons)
- Higgs boson is responsible for the mass of the elementary particles

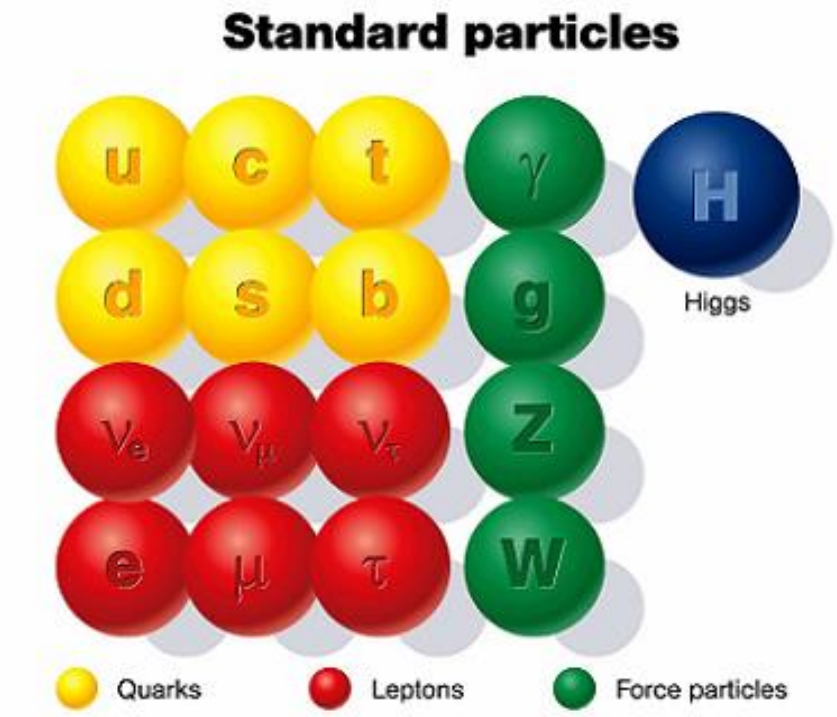
Bosons

- Quarks and leptons make up matter
- Quarks interact with all the 3 forces
- (Charged) leptons interact with (EM and) weak interaction(s)

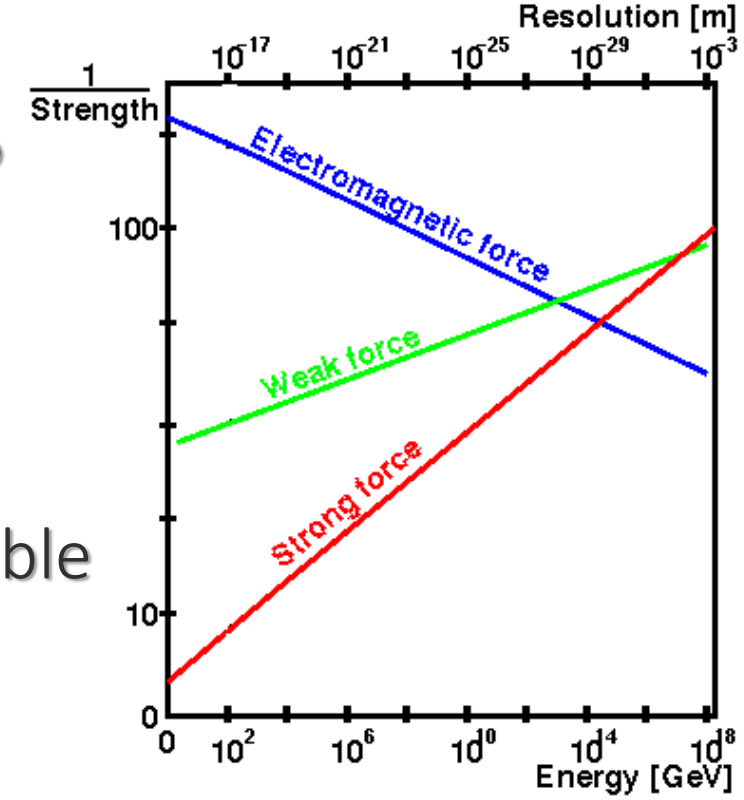
Fermions

Need for New Physics

Although all the success of the theory, we know that it cannot be the final theory



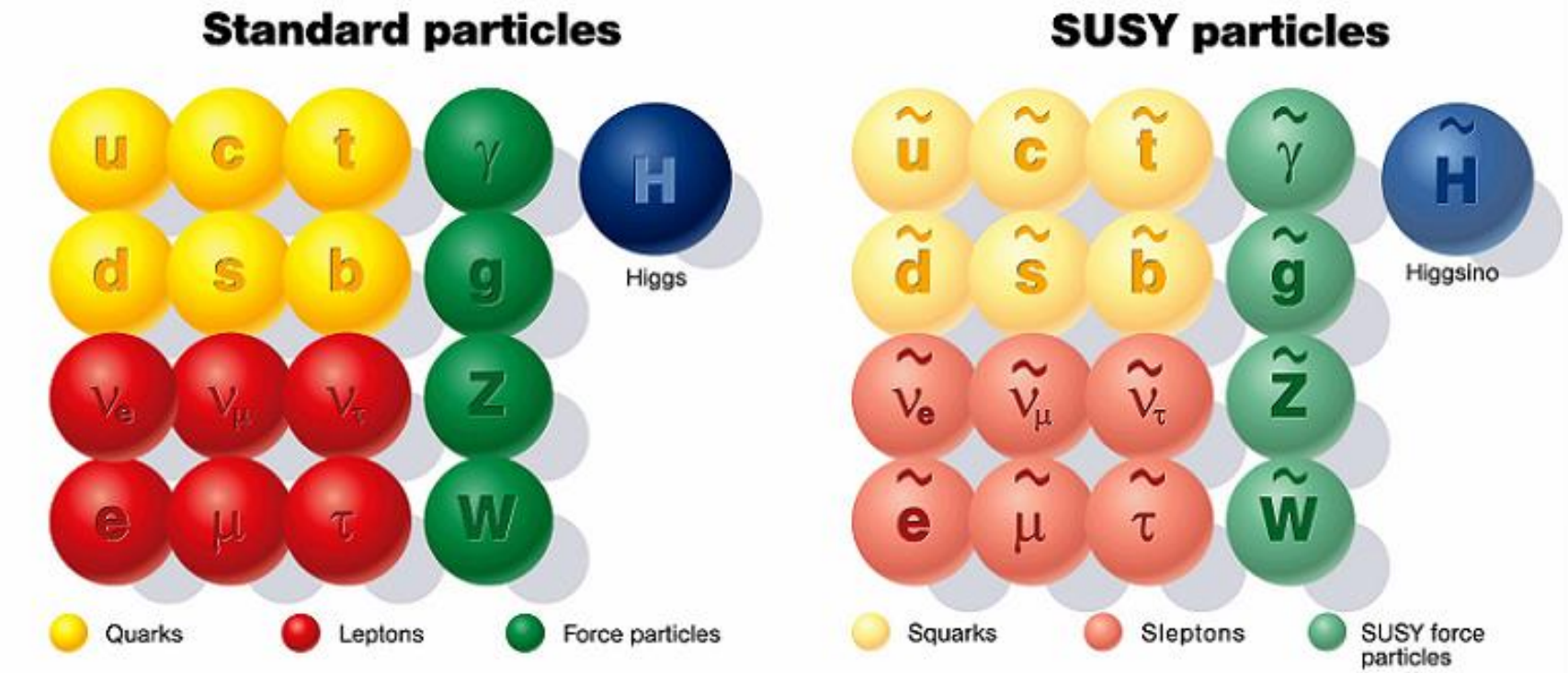
- Why is there a matter/antimatter asymmetry?
- What is Dark Matter?
- Why is gravity so weak?



- Unification of the interactions is not possible with the SM.

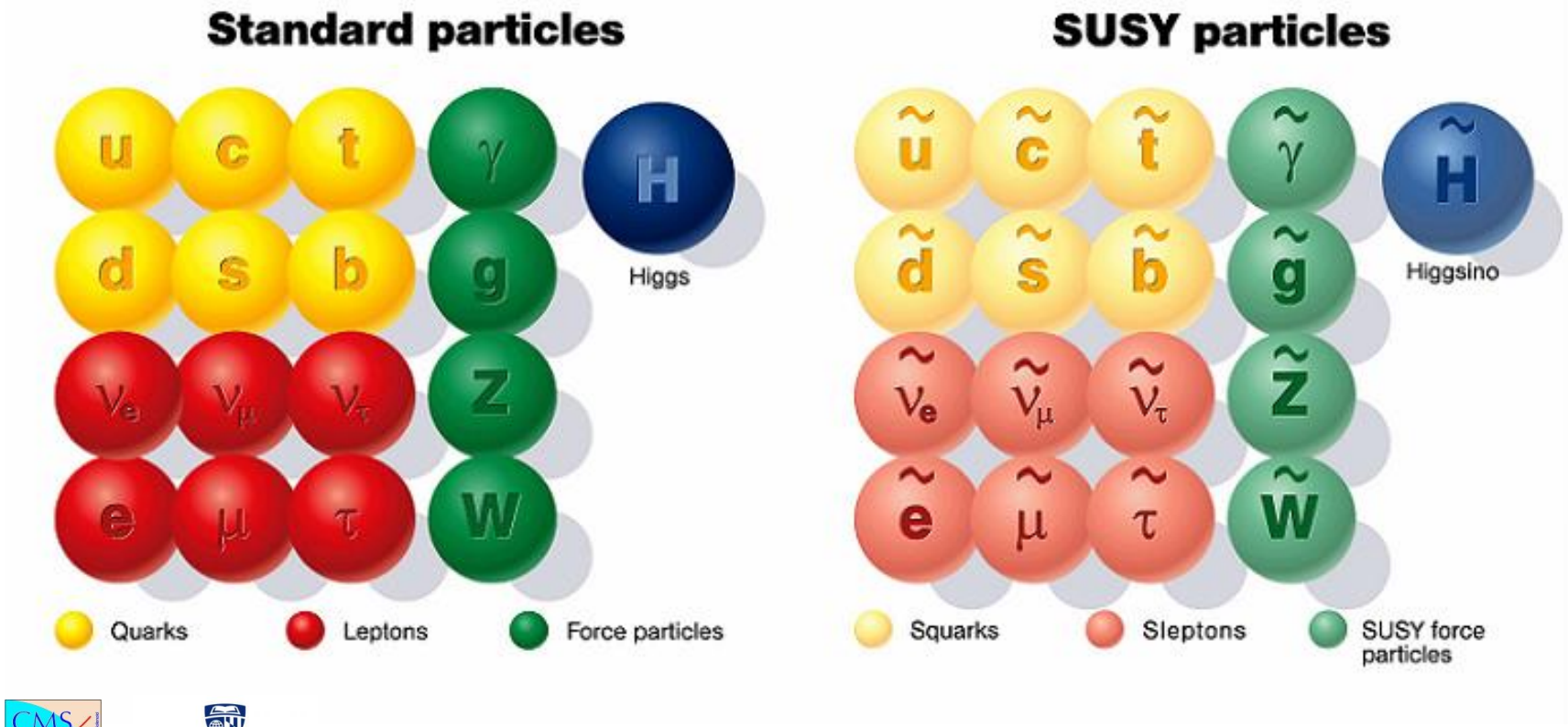
A possible solution: supersymmetry (SUSY)

Extending the SM with the superpartners of its particles (partners of the SM fermions are bosons and the SM bosons are fermions) seems to solve these problems theoretically



A possible solution: supersymmetry (SUSY)

Extending the SM with the superpartners of its particles (partners of the SM fermions are bosons and the SM bosons are fermions) seems to solve these problems theoretically



SM quark → squark ($\tilde{q}$)
SM lepton → slepton ($\tilde{l}$)

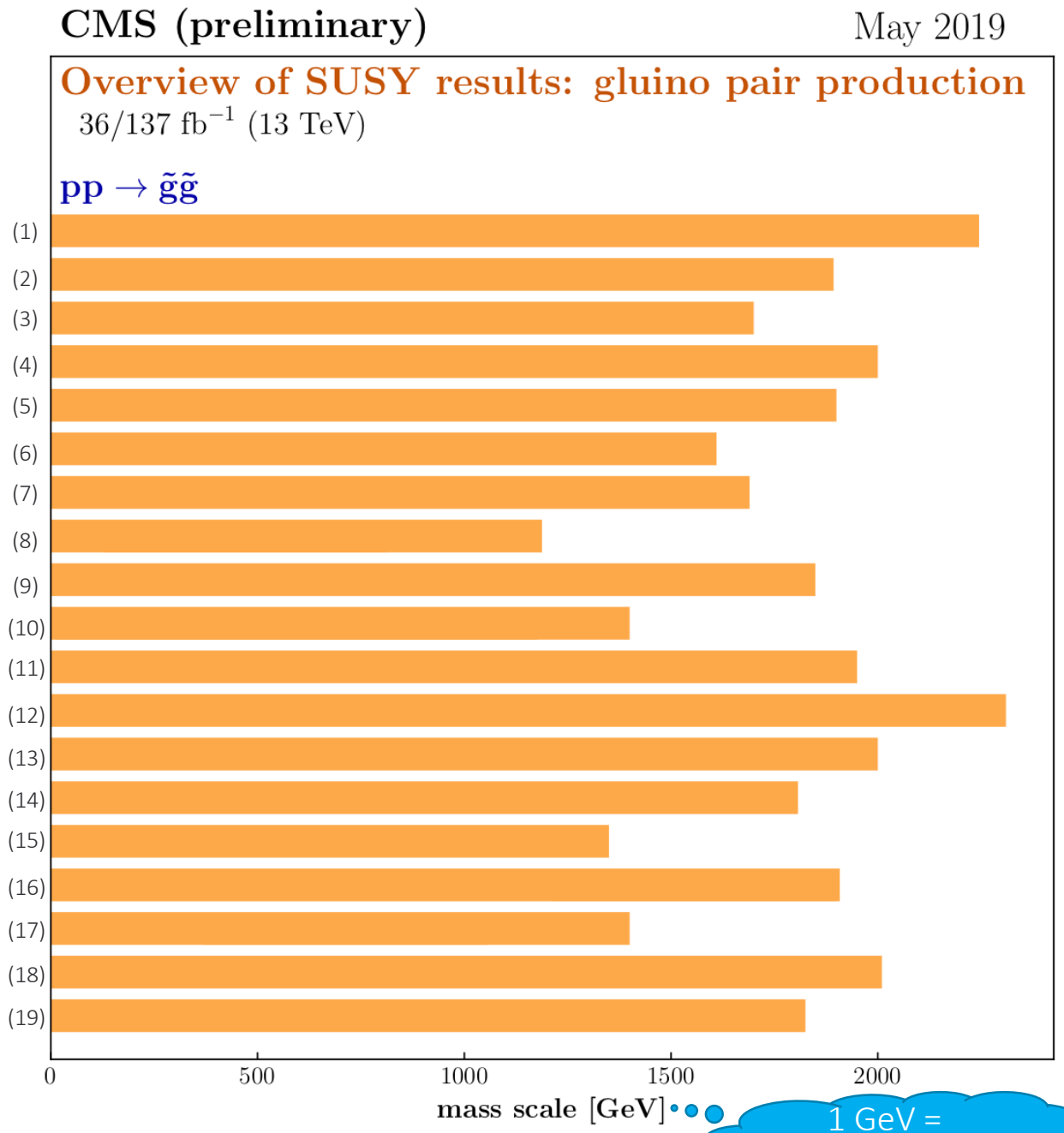
SM gluon → gluino ($\tilde{g}$)

Search for SUSY

There has been a tremendous effort put into finding SUSY particles.

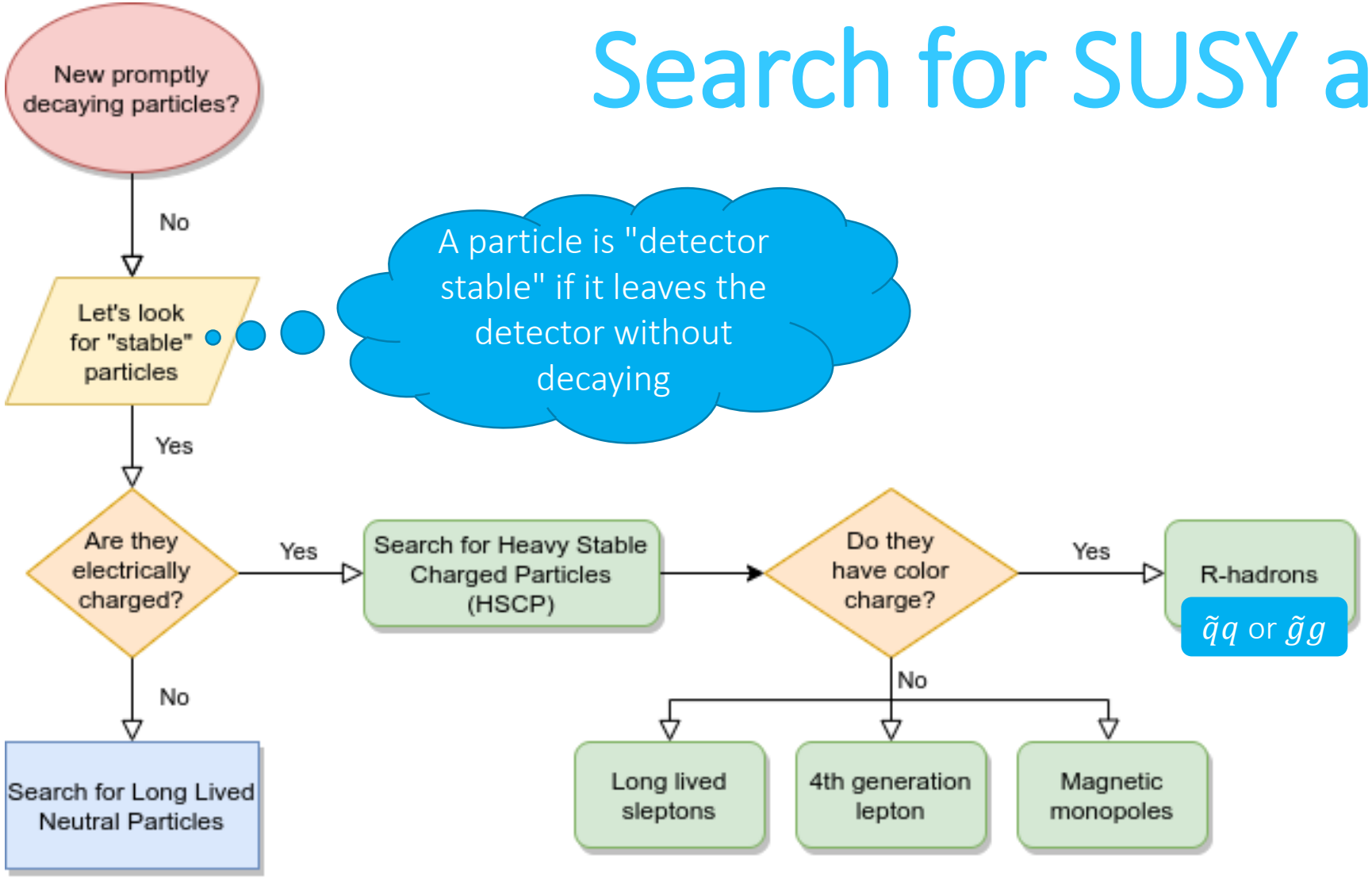
I was involved in these searches in the last several years [a,b,c,d]

In these cases the SUSY particles decay promptly

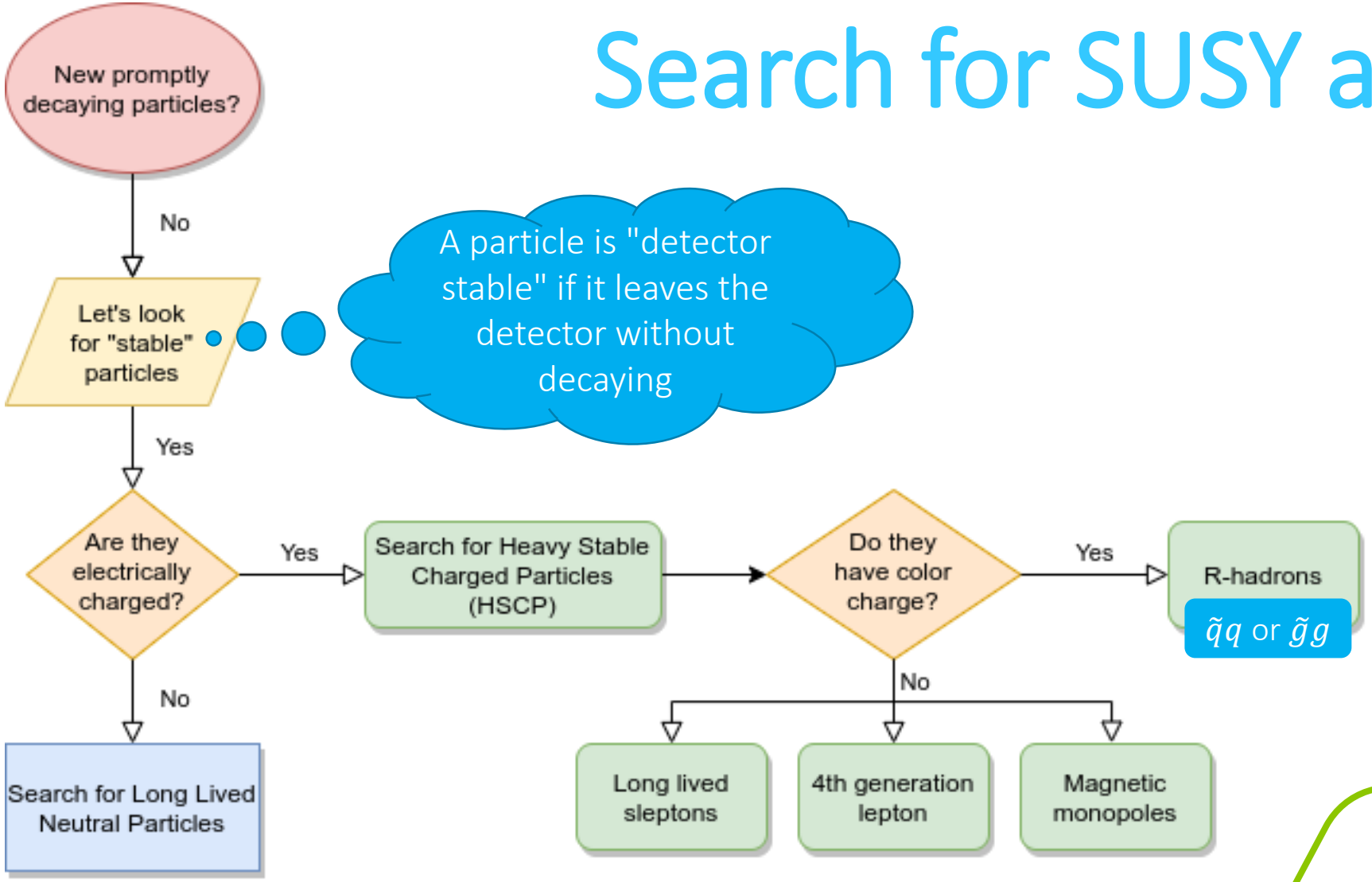


1 GeV = proton mass

Search for SUSY and beyond



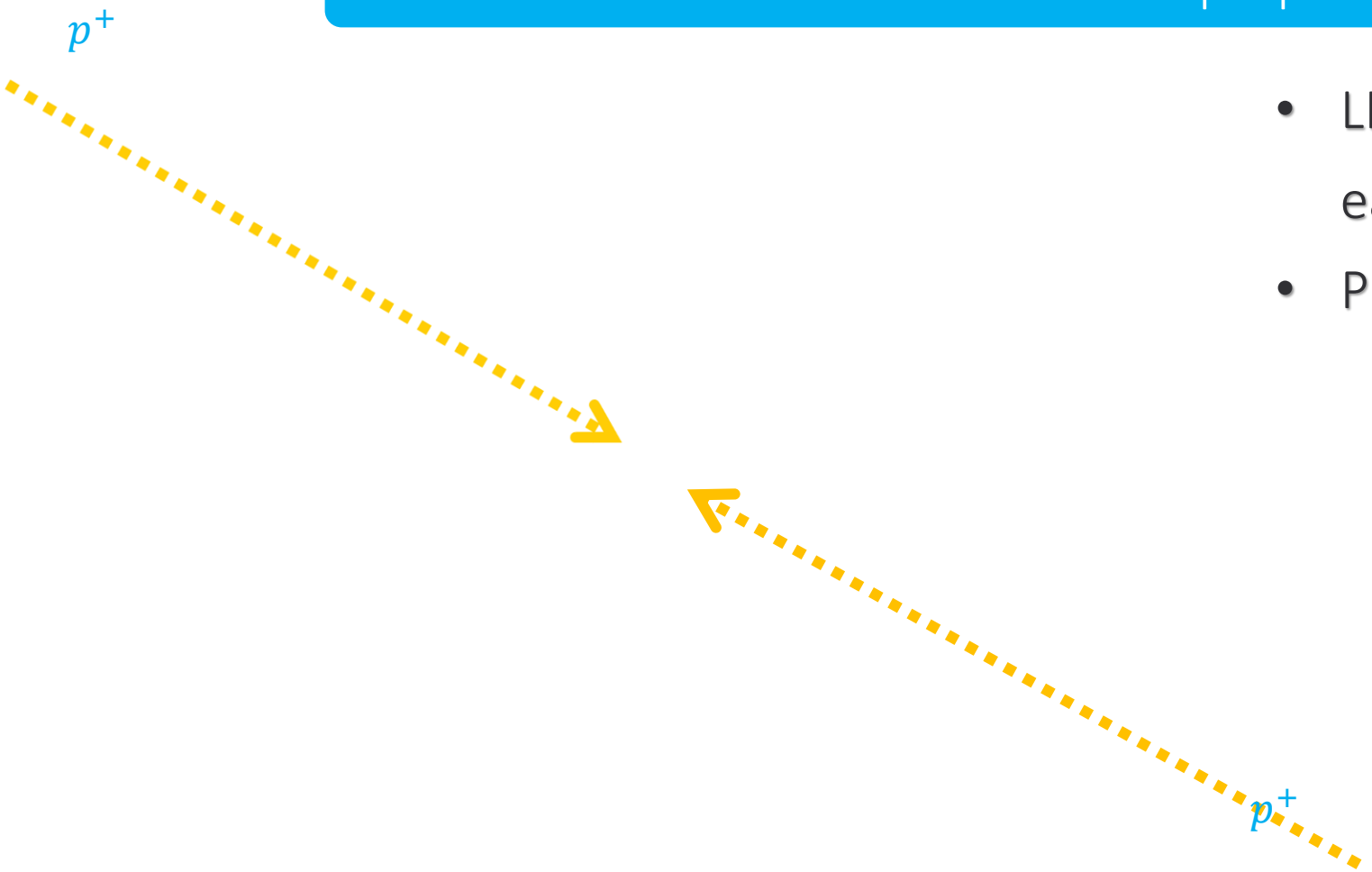
Search for SUSY and beyond



They have a distinctively different characteristic ionization loss (dE/dx) from the SM particles

CMS Detector at Large Hadron Collider

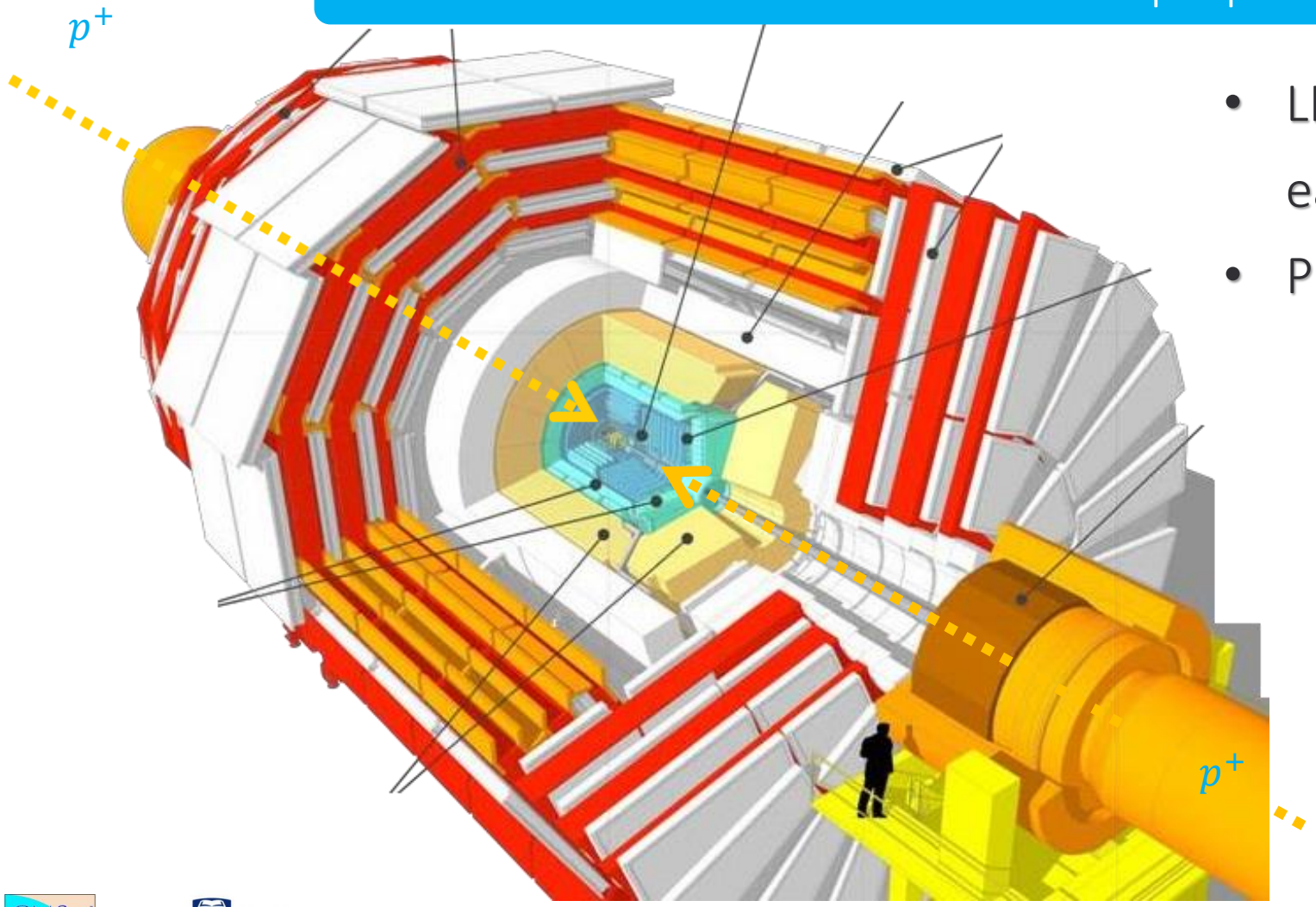
The measurements described in this proposal are pursued at CERN at the LHC



- LHC provides two proton beams, each with 6.5 TeV energy.
- Protons are collided in every 25 ns.

CMS Detector at Large Hadron Collider

The measurements described in this proposal are pursued at CERN at the LHC

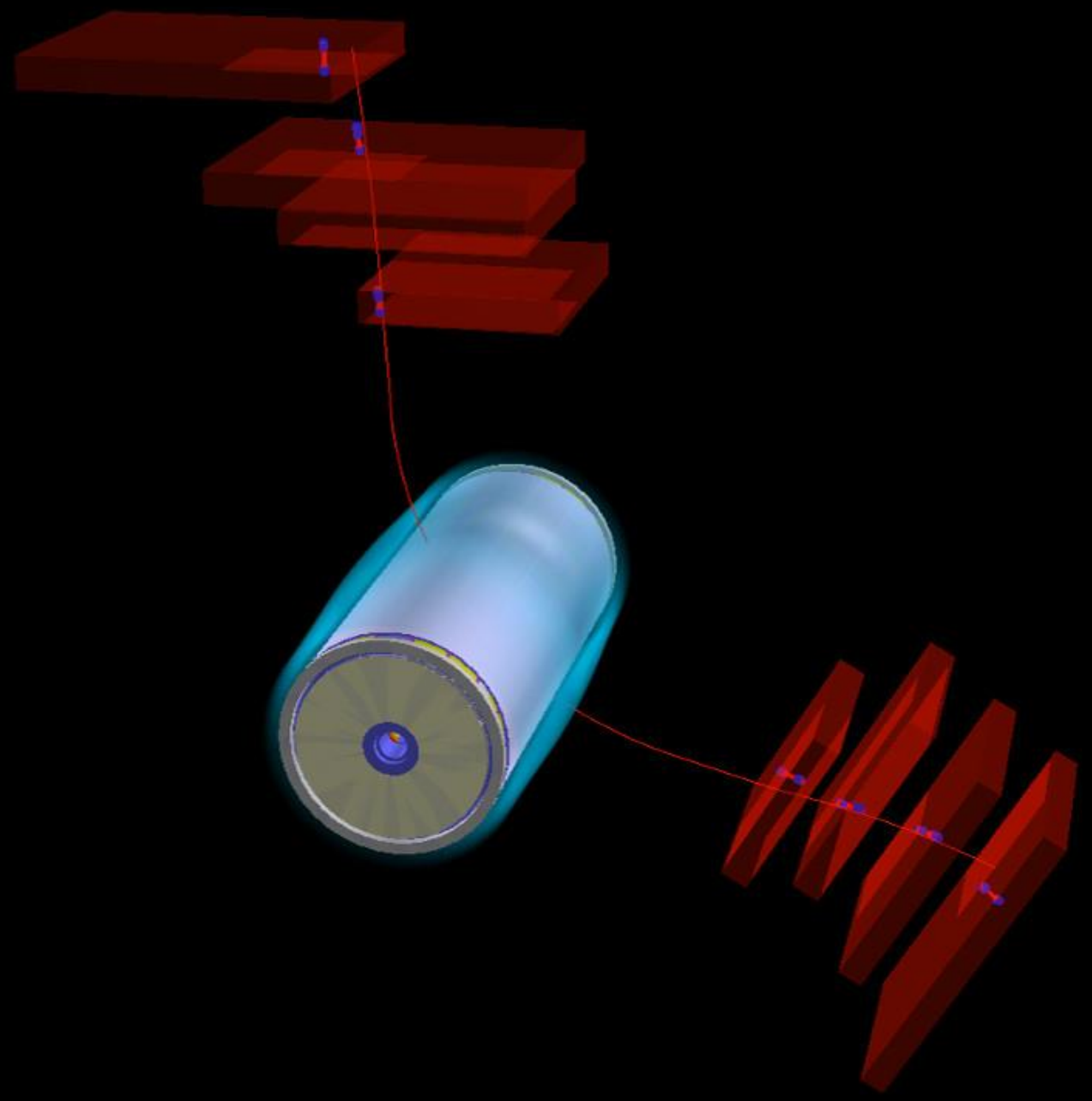


- LHC provides two proton beams, each with 6.5 TeV energy.
- Protons are collided in every 25 ns.

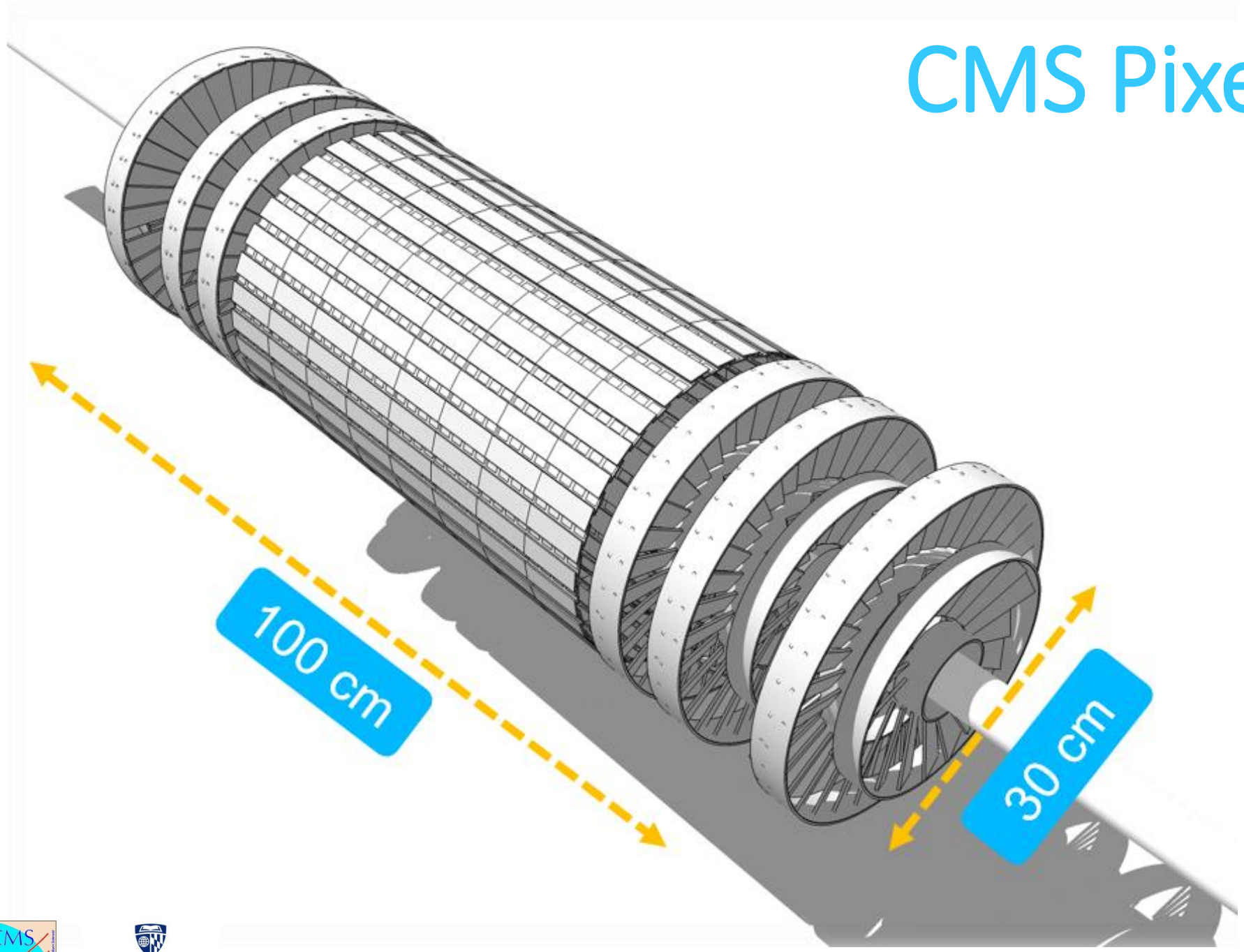
CMS detector reconstructs the products of these collisions using

- a 3.8 T solenoid,
- calorimeters (to measure the deposited energy),
- and tracking detectors

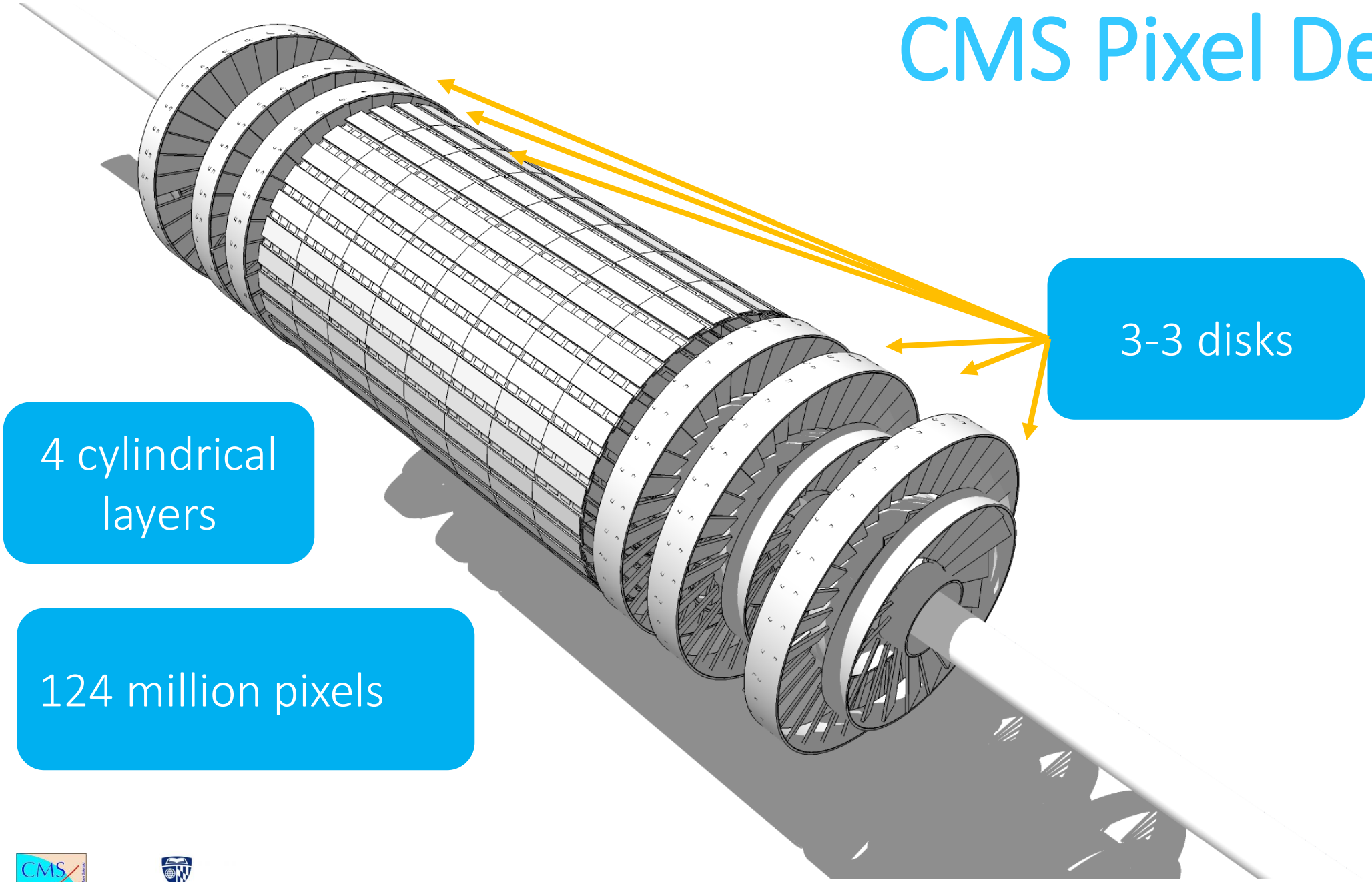
HSCP event visualization



CMS Pixel Detector



CMS Pixel Detector

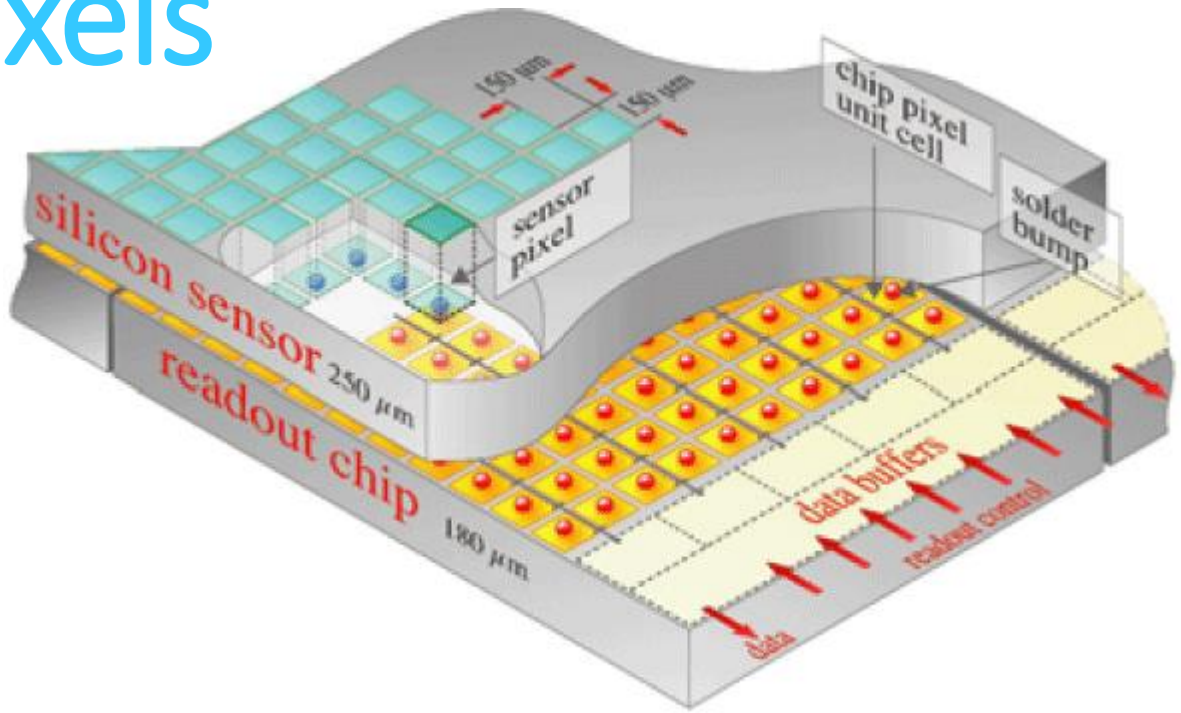


3-3 disks

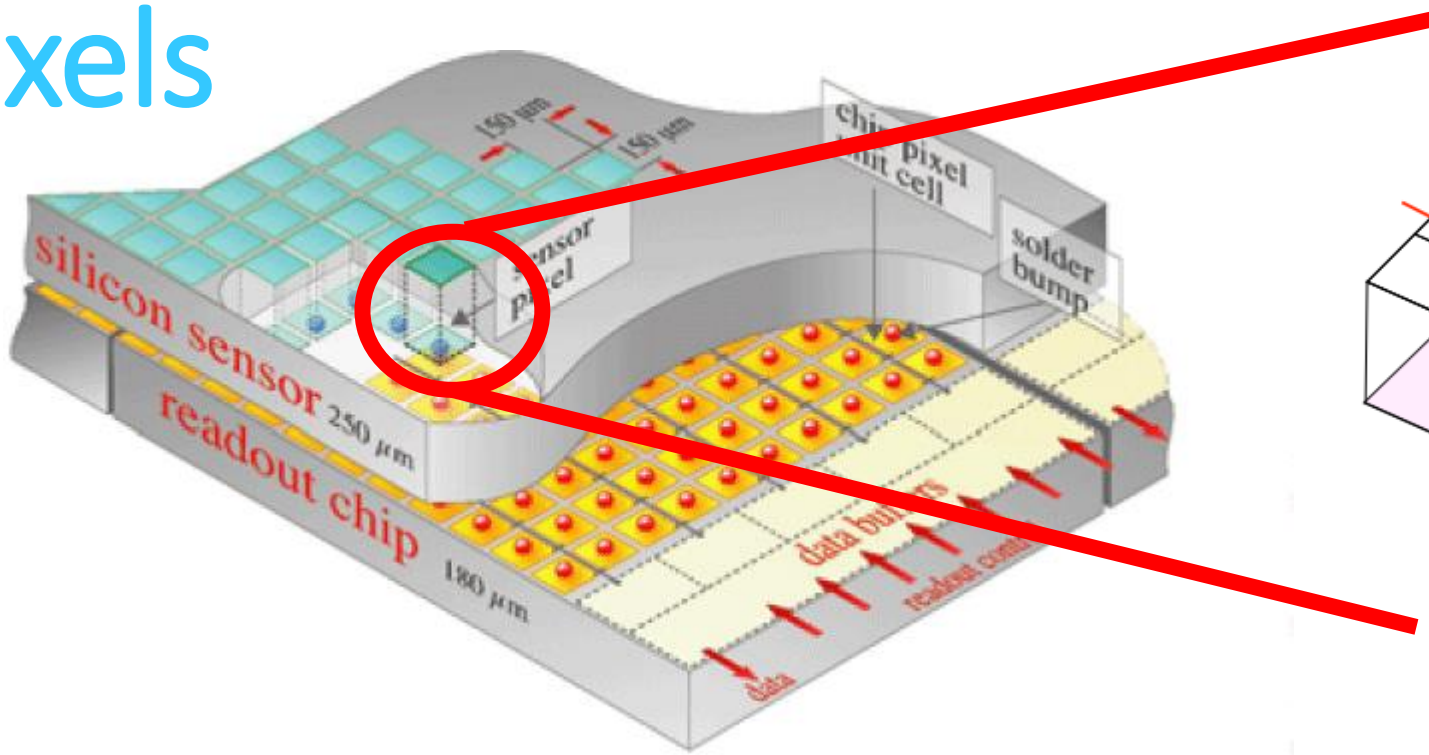
4 cylindrical layers

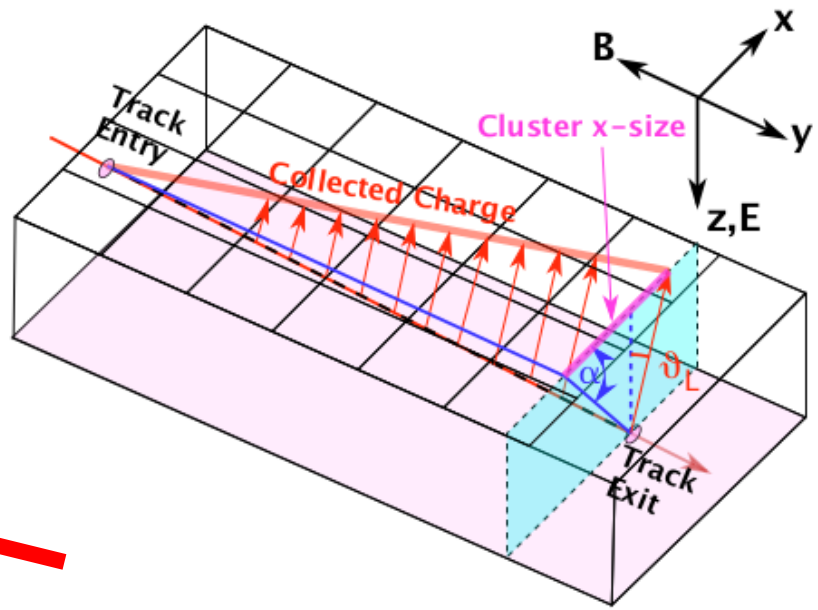
124 million pixels

Pixels

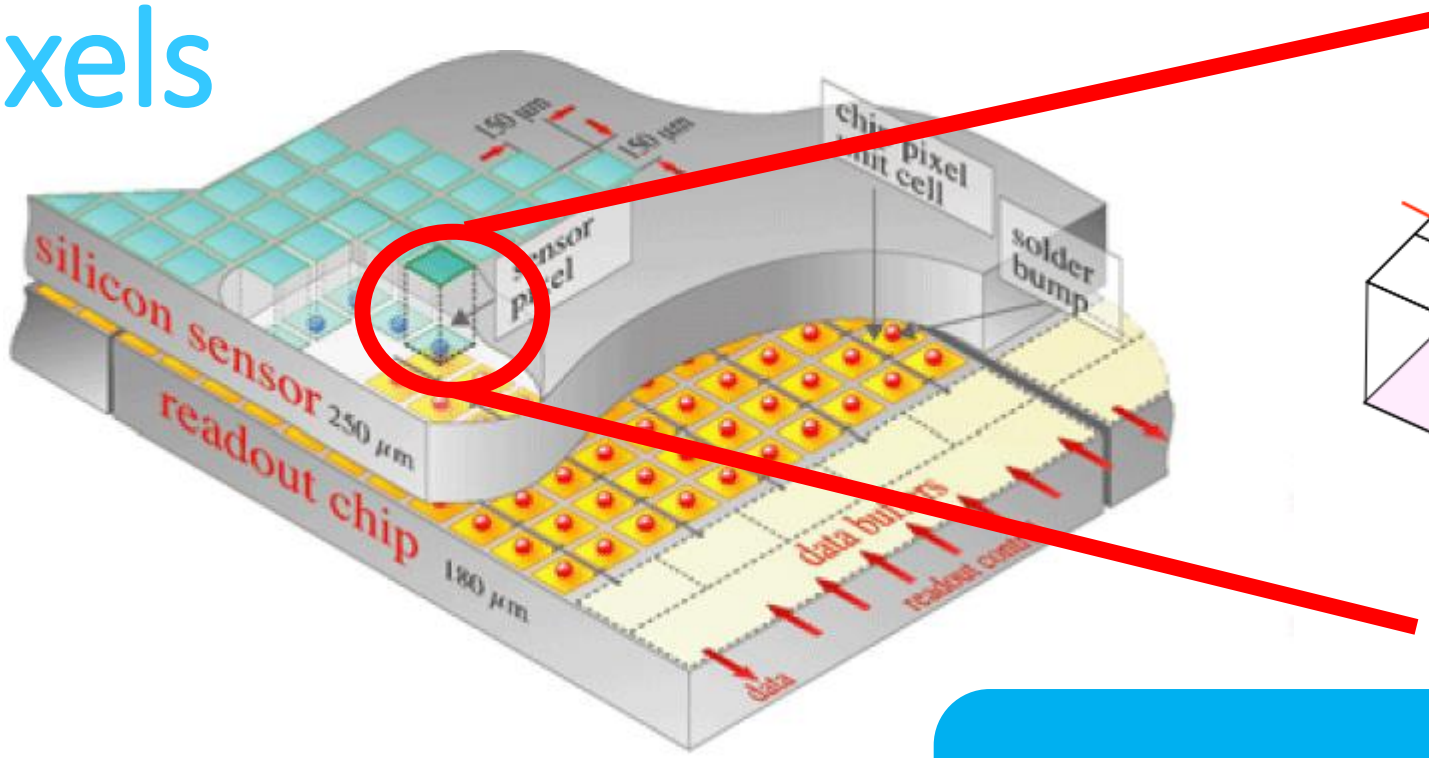

150 μm
150 μm
chip pixel
unit cell
sensor
pixel
solder
bump
silicon sensor 250 μm
readout chip 180 μm
data buffers
readout control
data

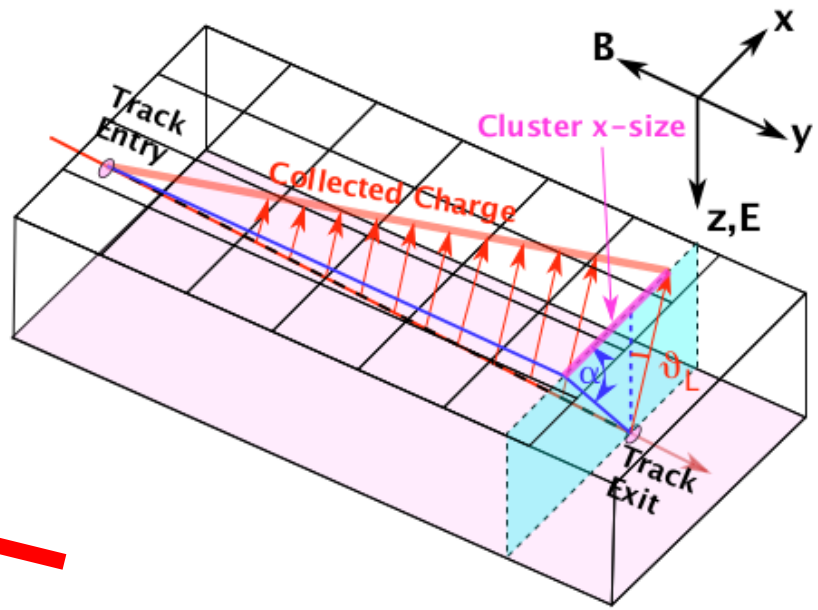
Pixels


150 μm
150 μm
pixel
solder
bump
silicon sensor 250 μm
readout chip 180 μm
data

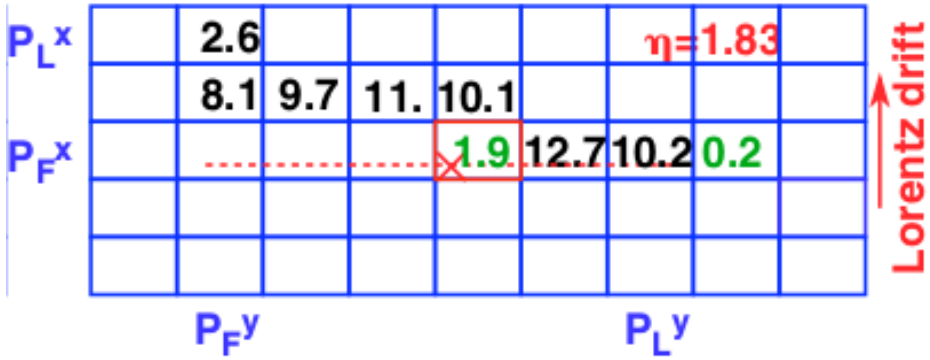

B
x
y
z,E
Track
Entry
Cluster x-size
Collected Charge
Track
Exit

Pixels





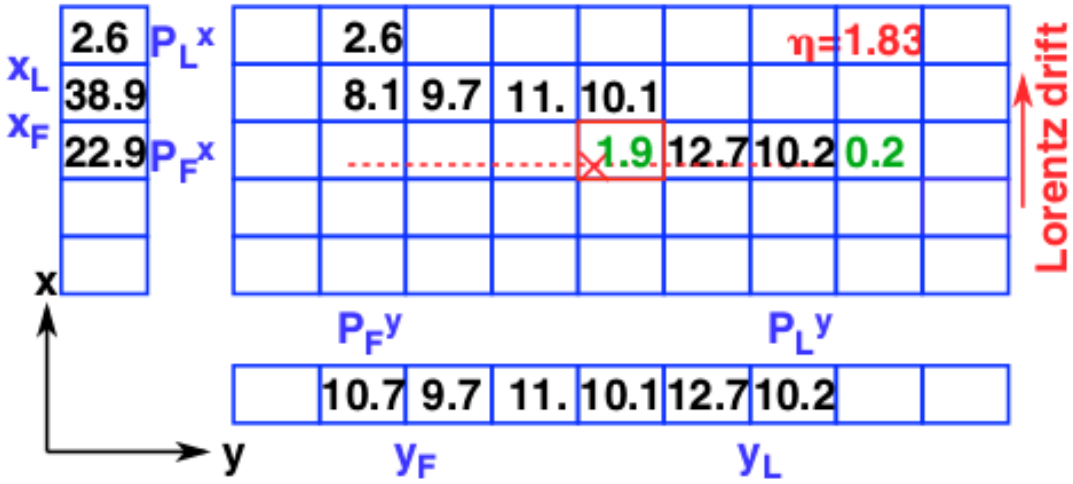
A traversing particle creates charges which are read out in the x-y plane



Example cluster shape. Signals in each pixel are given in ke$^-$.

Template reconstruction

The current track reconstruction uses a special reconstruction method that was developed to simulate irradiated sensors using an external software called PixelAV.



Cluster shape for the barrel hit ($\eta = 1.83$). The signals in each pixel are given in kiloelectrons.

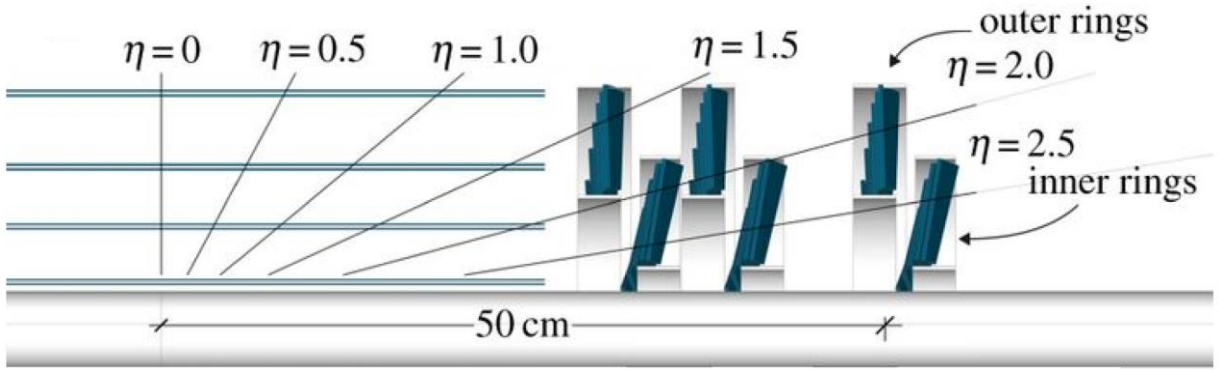
These simulated charged shapes are stored in a database as 1D (2D) projections, called (2D) templates.

The x- and y-projections are shown as one-dimensional arrays on the plot on the left.

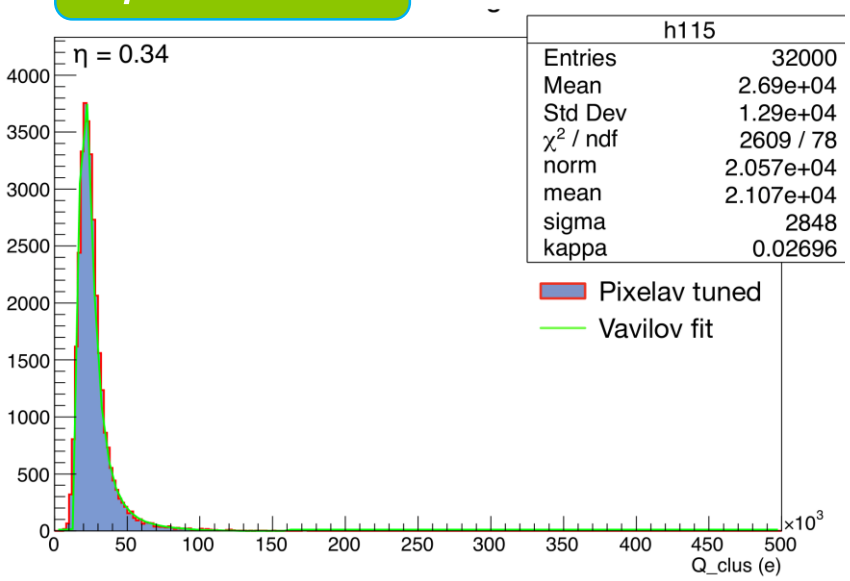
This technique leads to superior resolution, especially for irradiated sensors.

Template reconstruction

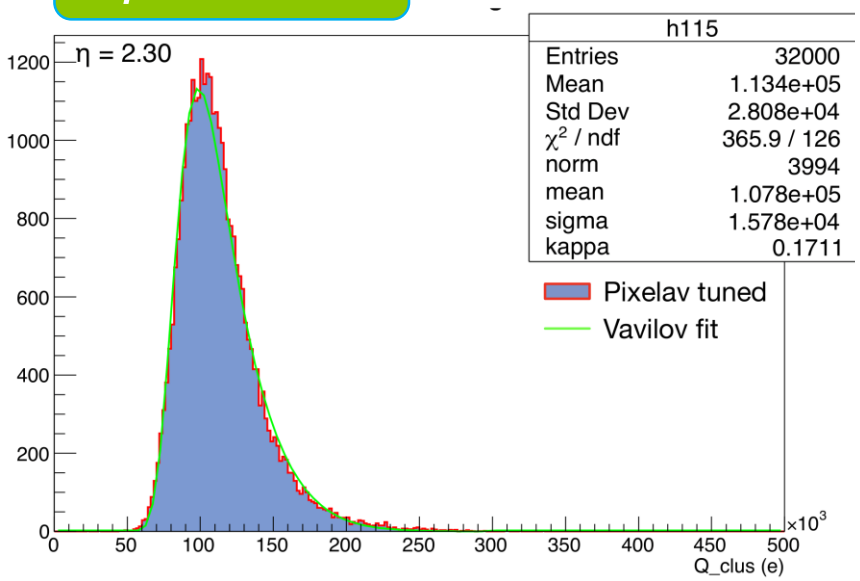
Cluster shapes are stored for several incident angles



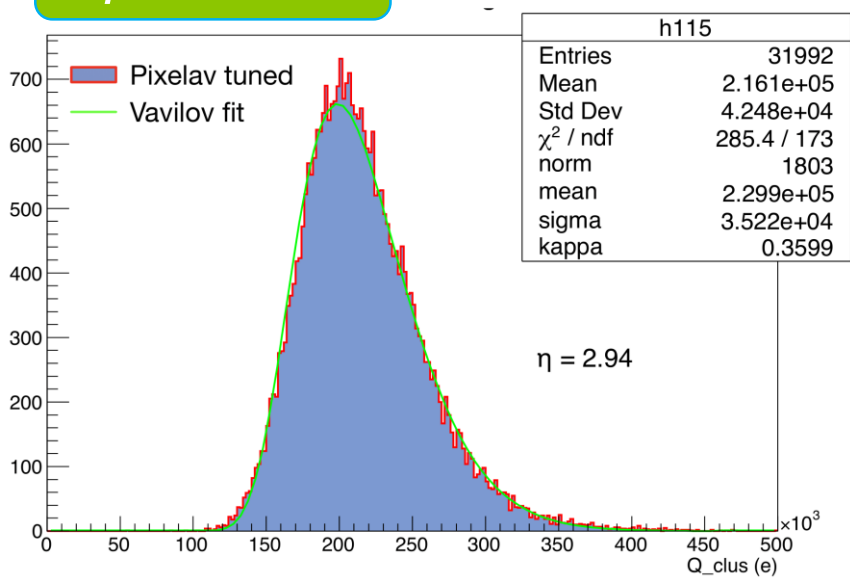
$\eta = 0.34$



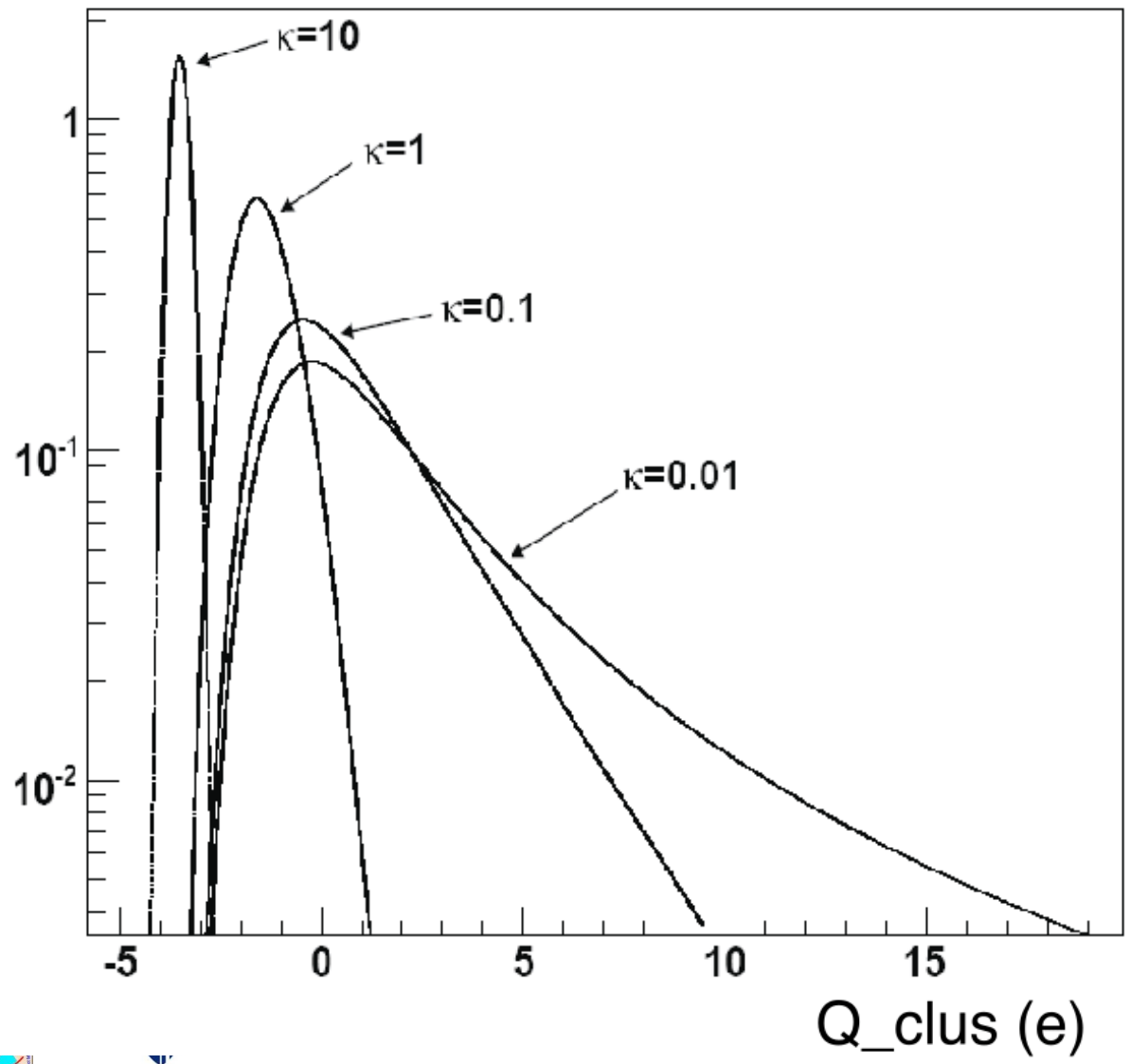
$\eta = 2.30$



$\eta = 2.94$



Template reconstruction



The distributions are fitted with a Vavilov function that has a parameter that describes the fluctuations in the ionization (κ):

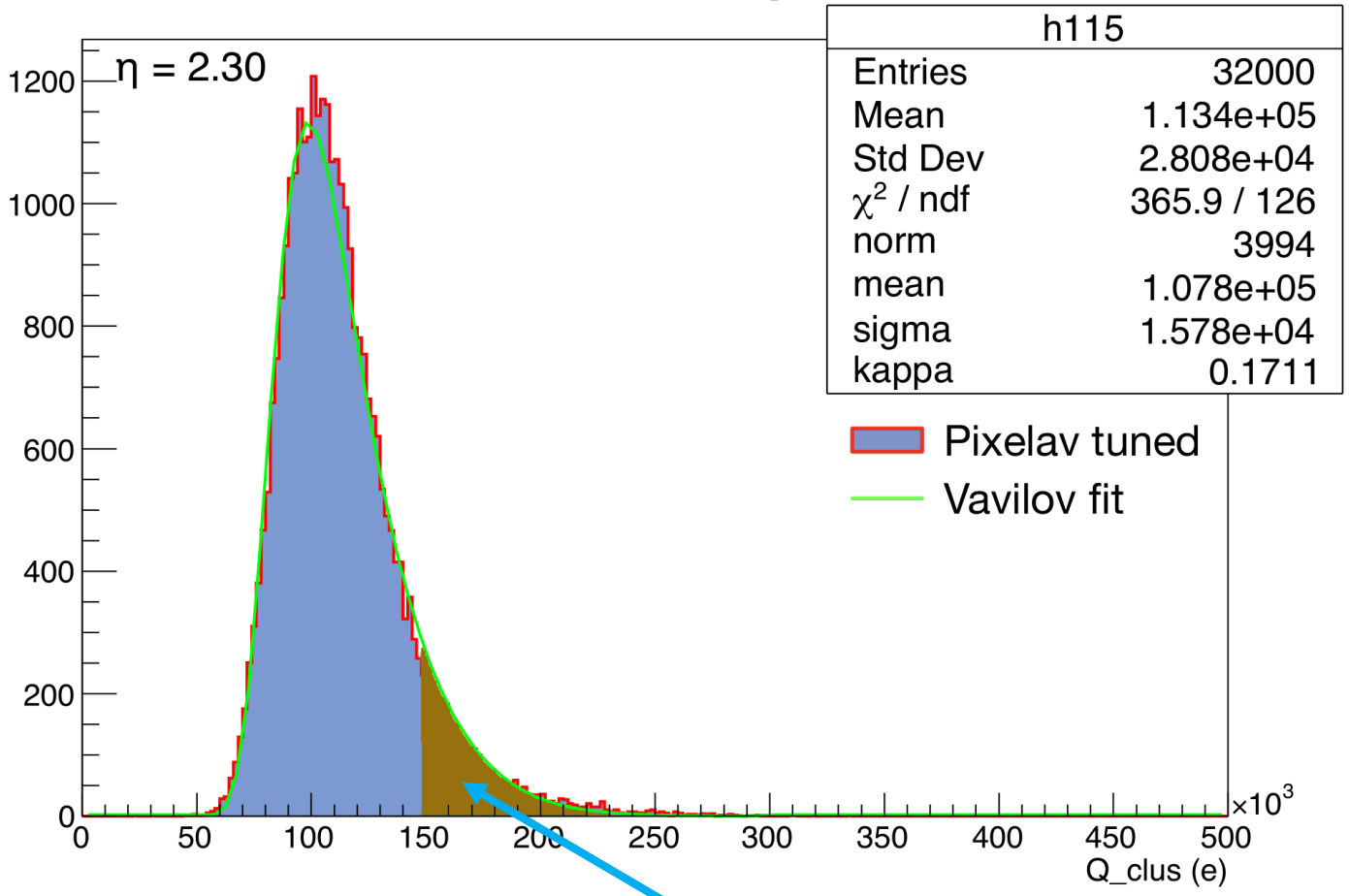
- $\kappa > 10$ means that almost all energy is deposited to the silicon: function is more Gaussian
- $0.01 < \kappa < 10$ is the case for thin absorbers, the Vavilov function is the only good description
- $\kappa < 0.01$ is the case for the limit with infinitely thin absorbers, the Landau-like situation

Template reco

Plan to exploit hitherto unused results of the template fit:

- ProbXY: describes cluster shape

- ProbQ: charge probability

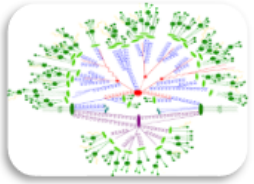


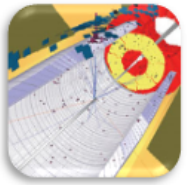
ProbQ is understood for each cluster as the integral from the given cluster charge in the tail

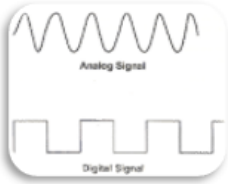
Low ProbQ thus means high ionization

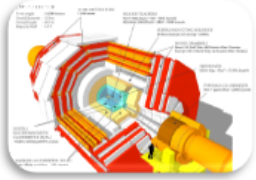
My simulation samples

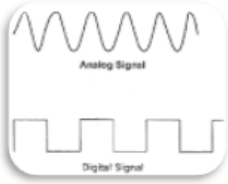
Simulations in the CMS software follow a multi step process.

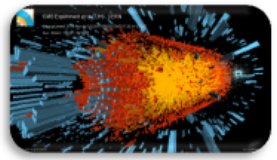












GEN	SIM	DIGI	DIGI2RAW	RAW2DIGI	RECO
Monte Carlo events with a physics generator	Particles' energy deposition in the detector (sim hits).	Sim hits converted into detector response	RAW data simulation equivalent to the detector output	Hits converted into detector response	Recon-structed hits, clusters and objects

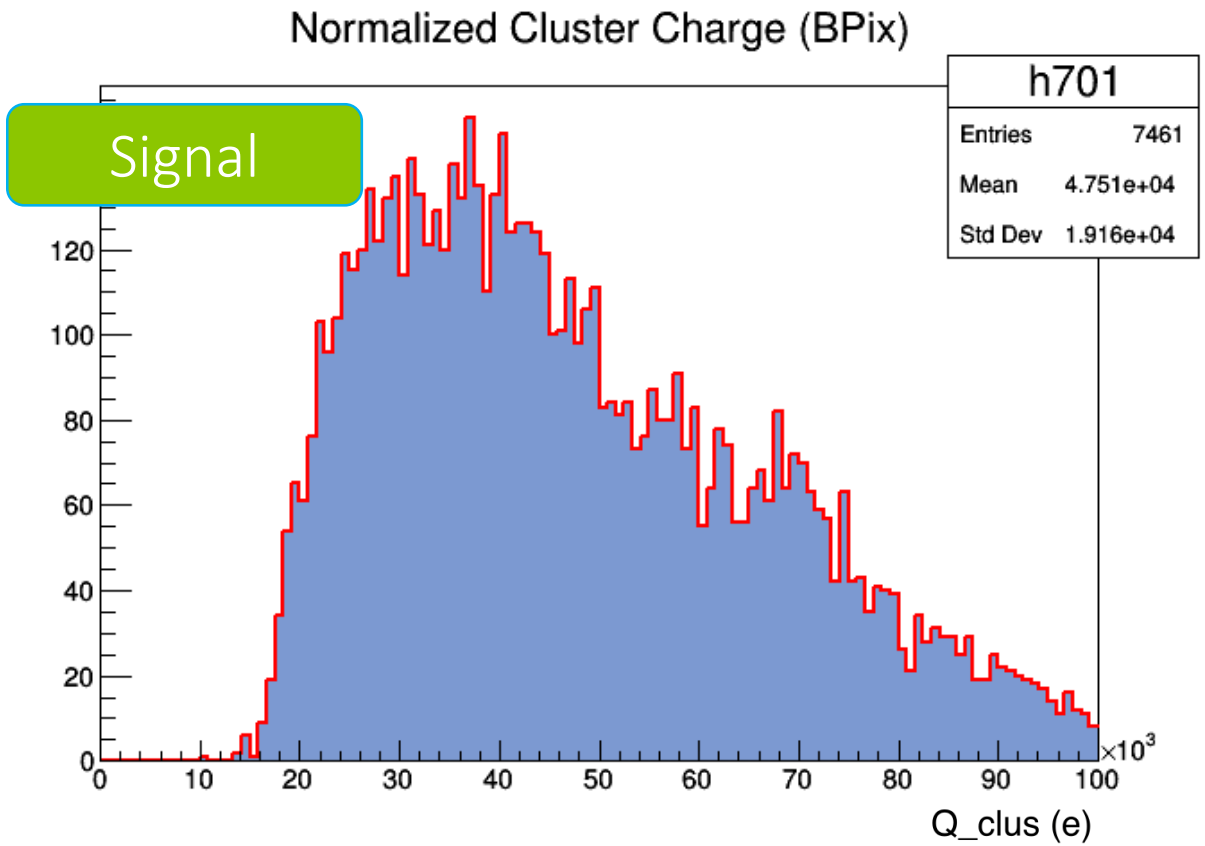
RECO	AODSIM	MINI AODSIM	NANO AODSIM

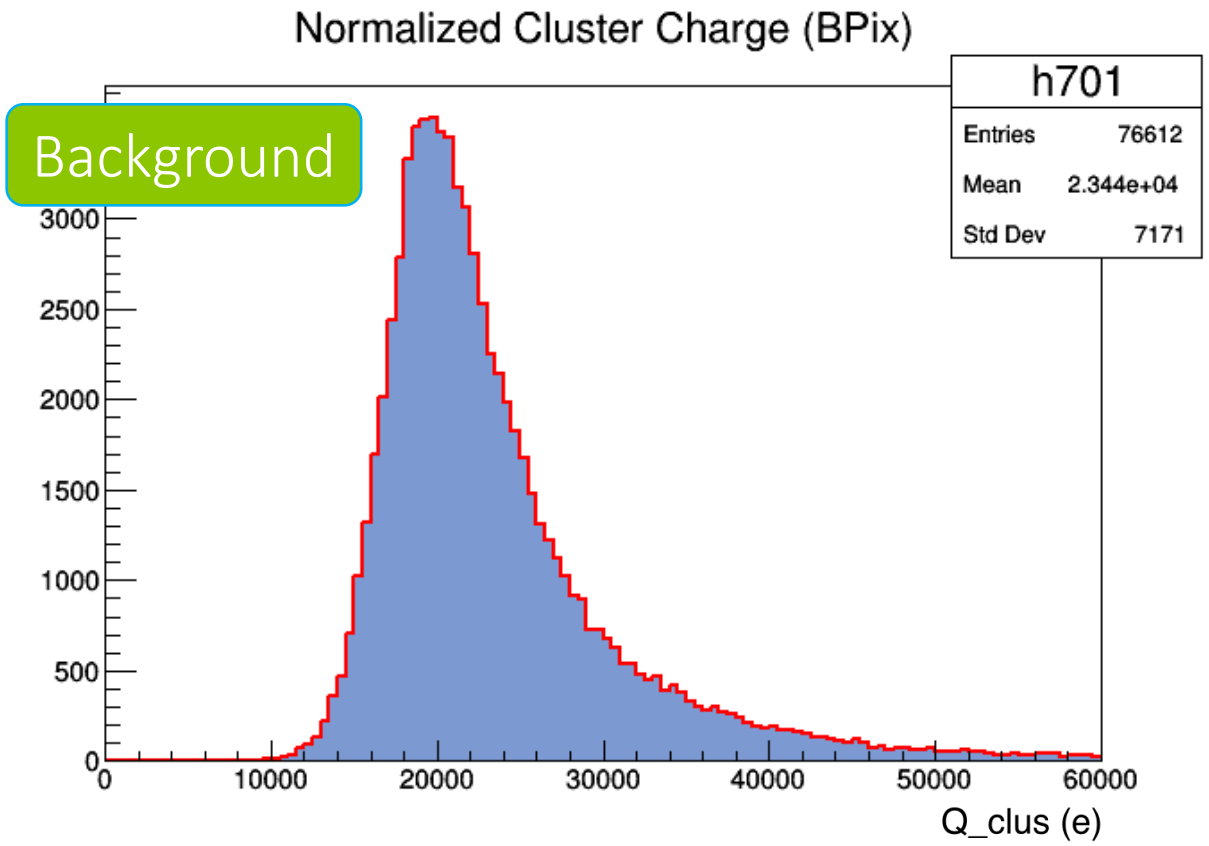
I simulated

- R-hadrons with 800 GeV and 1800 GeV masses,
- and also the forth-generation lepton with mass 800 GeV.
- SM background from a top quark pair sample.

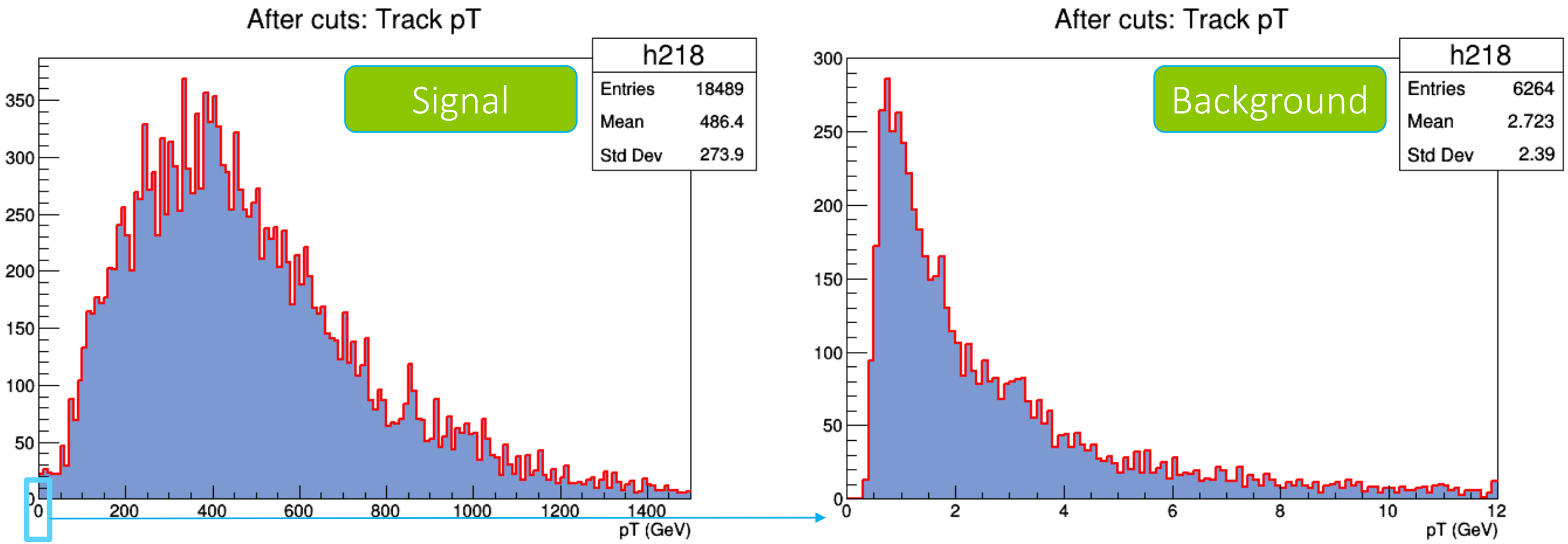
Cluster charges in my simulations

Let us plot the charge distributions in the barrel layers (normalized to track length).





Momentum distributions in MC



Please note the different scales for visibility, the left plot goes up to 12 GeV, while the right to 1500 GeV.

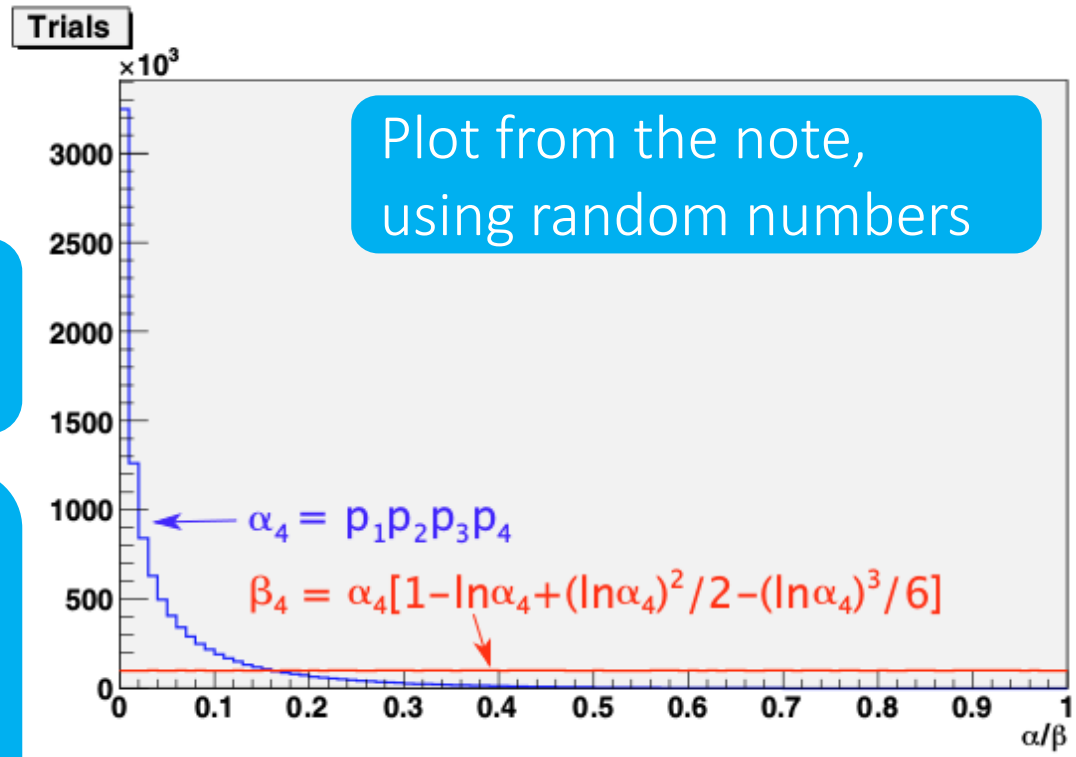
Combining probabilities

The probQ-s in the track are independent of each other and instead of multiplying them together we use a special combination technique introduced by Morris Swartz.

$$\beta_n = \alpha_n \sum \frac{[-\ln(\alpha_n)]^n}{n!}$$

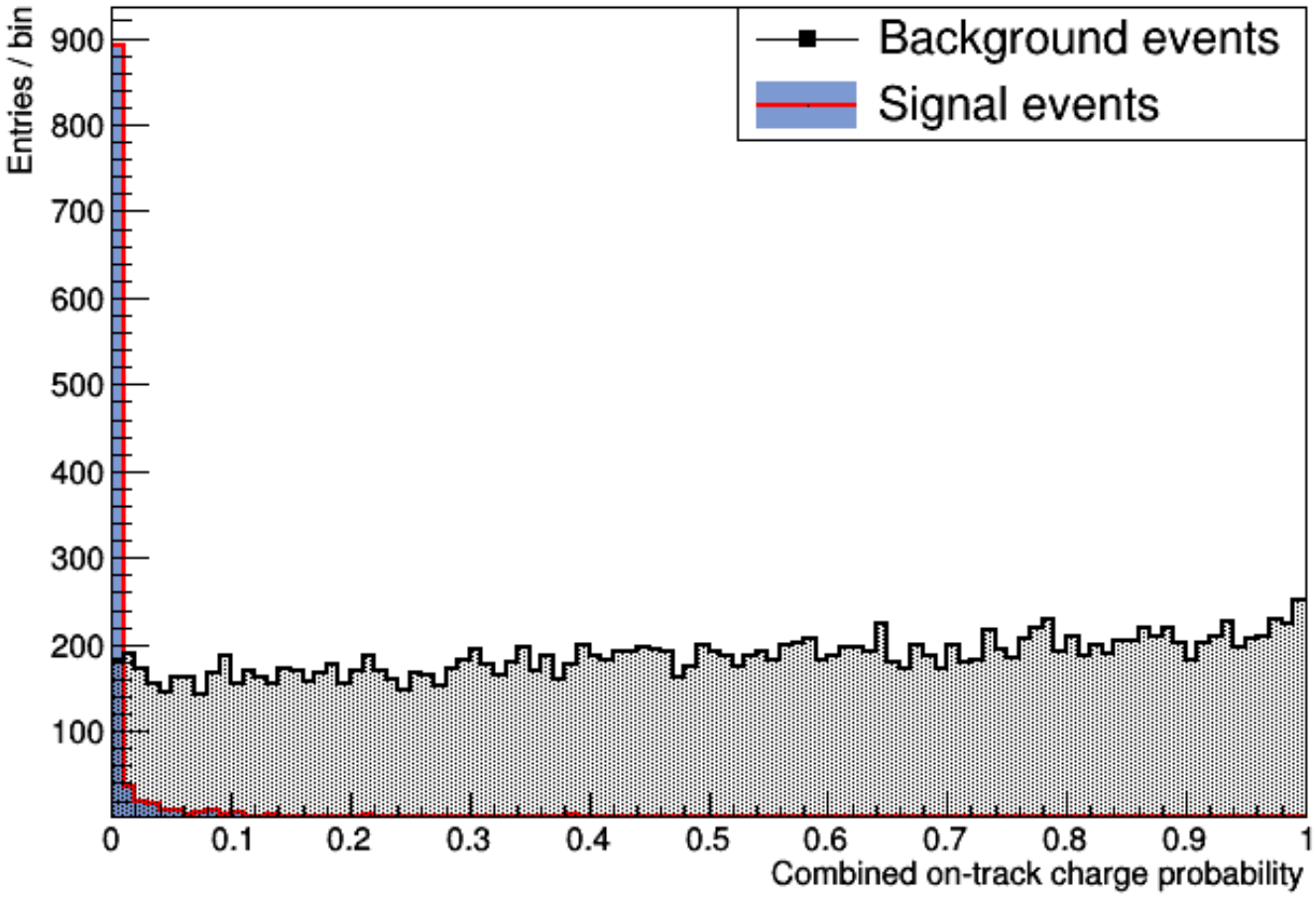
With this we have a variable for the clusters on track, whose distribution is flat for SM particles

<u>Important</u>: This is flat independently of the choice of n. n=4 is the situation with 4 pixel layers, but can be done with less too. Including the strips increases this number and this works like that too.



Plot from the note, using random numbers

Proof of concept



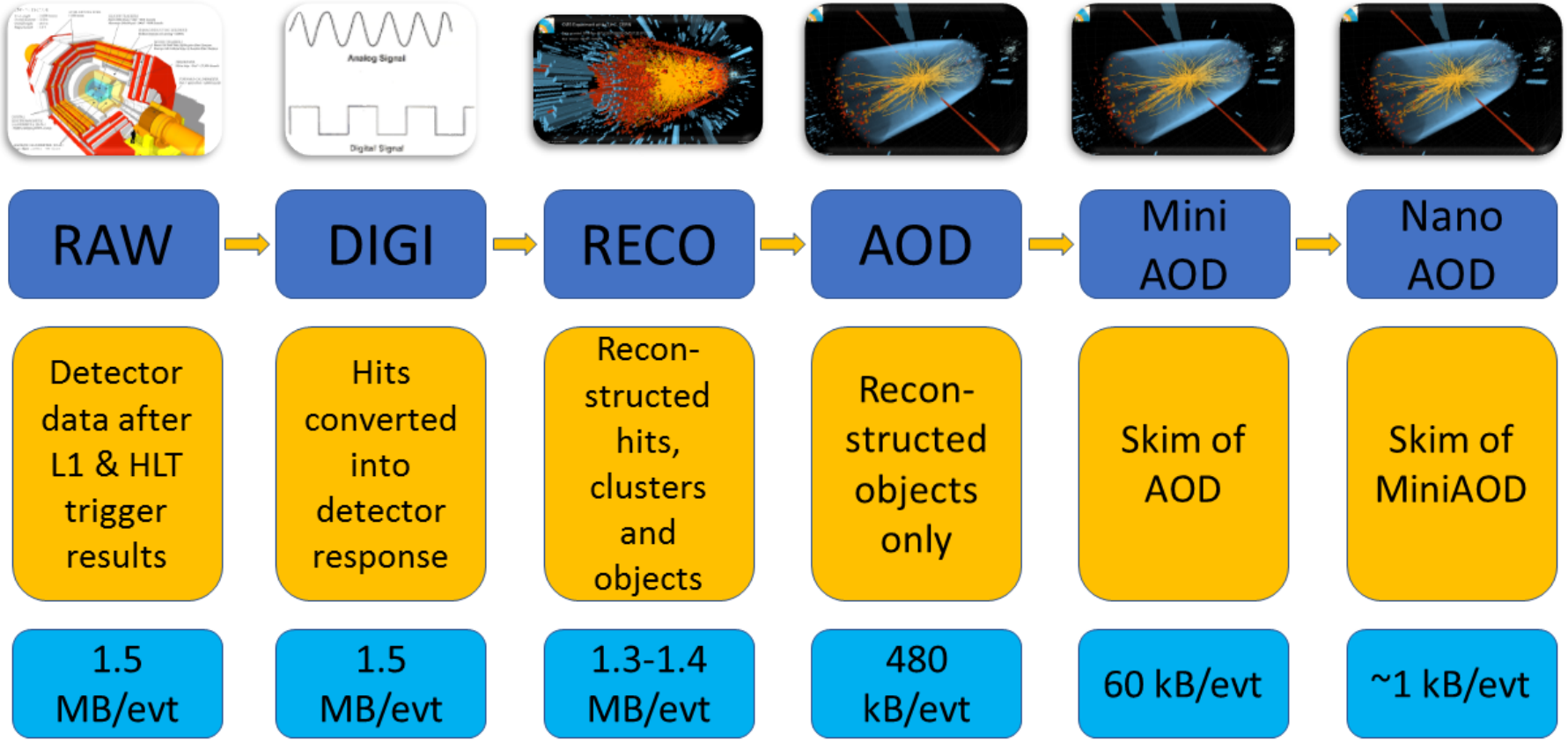
The on-track charge probabilities combined with the method described above

The black curve shows SM events, while the red is for the events containing R-hadrons.

This variable is proven to be a very powerful discriminator between signal and background.

How to analyze the data?

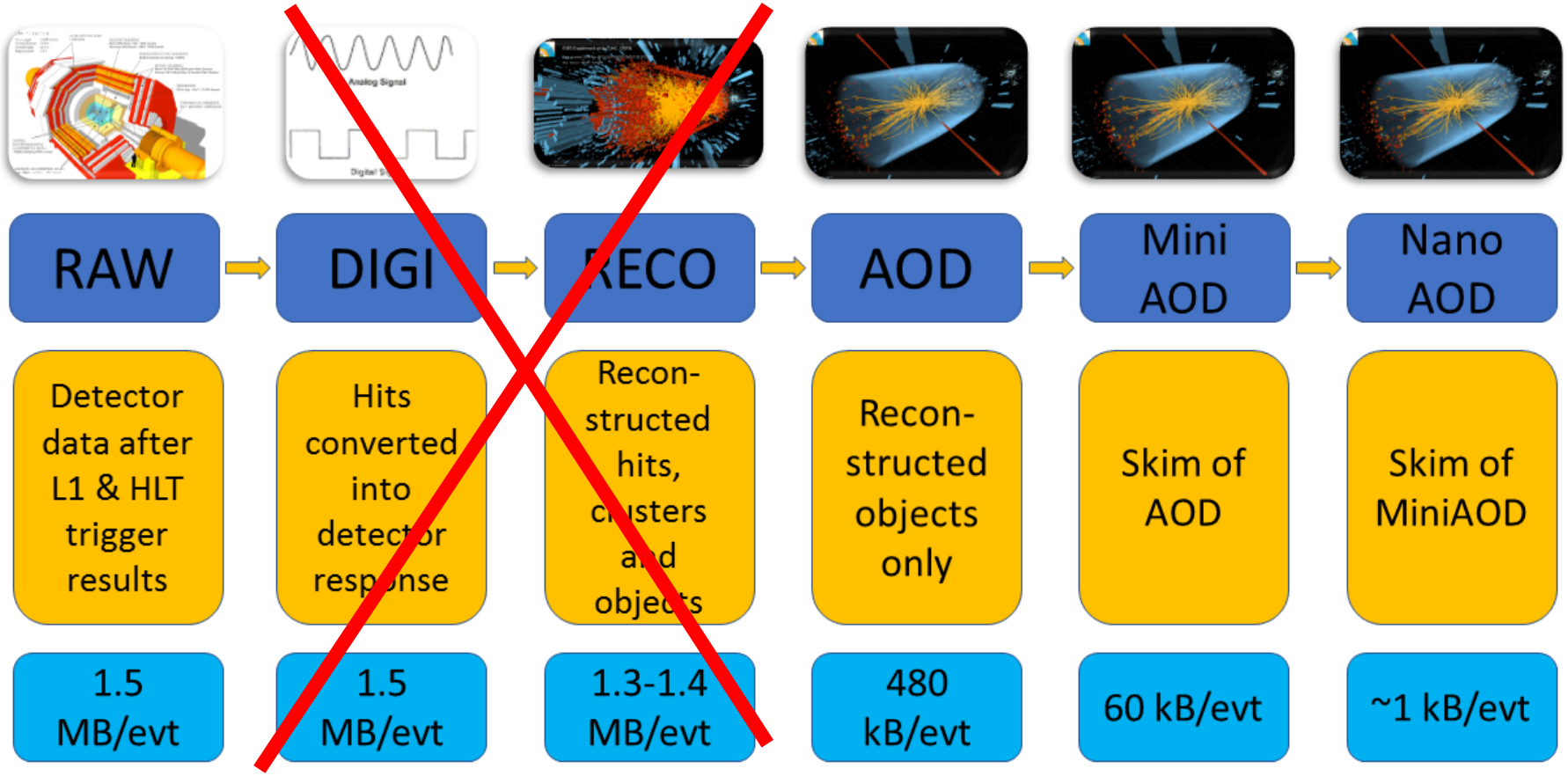
Next step is to analyze the data collected in 2017-2018.



- CMS has several intermediate data
- Most users use NanoAOD
- But this analysis needs reconstructed hits which are stored in RECO

How to analyze the data?

Unfortunately, CMS does not save the relevant intermediate data. → This could be a showstopper!



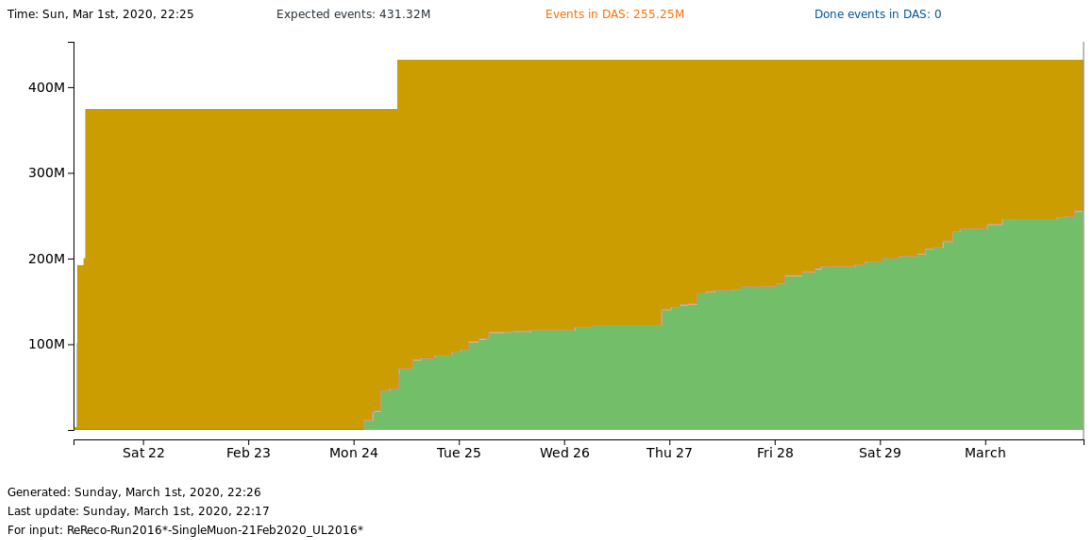
- 4212 TB of RAW data
- Just reading that in takes several months

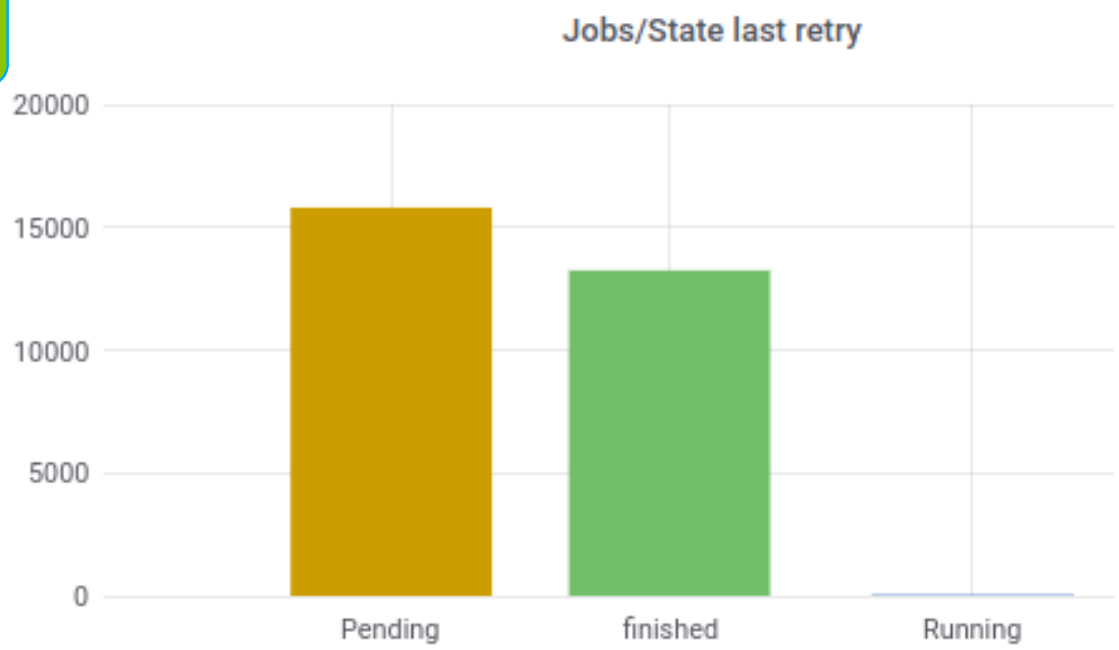
How to analyze the data?

I develop a RECO like format myself,
containing information about the tracker only.

It can be injected into the CMS central reprocessing system.

Re-reconstruction of the data is ongoing.

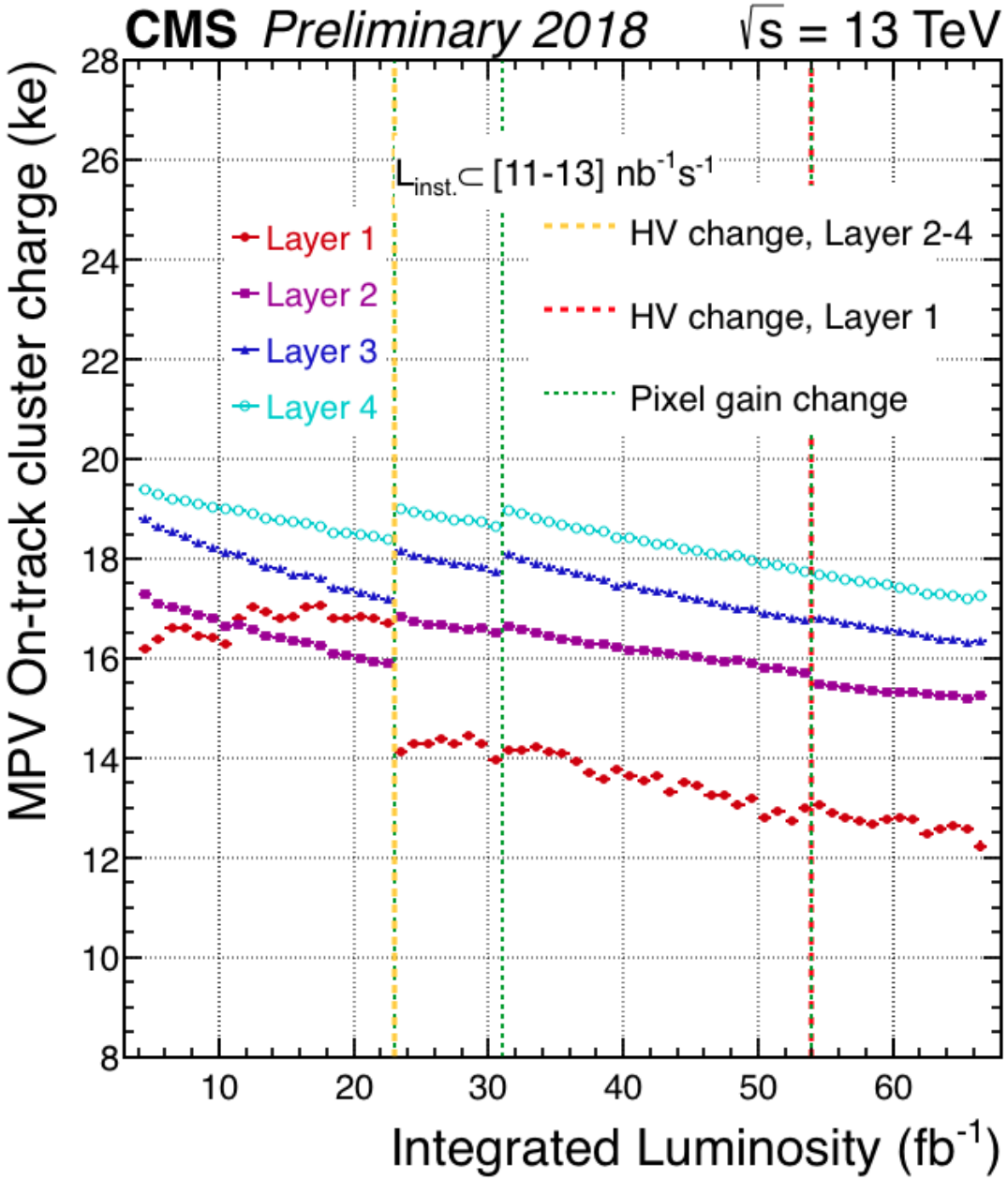




Next steps

In case no significant excess is found, limits for the mass of the HSCP will be set.

Charge has to be constant over the years



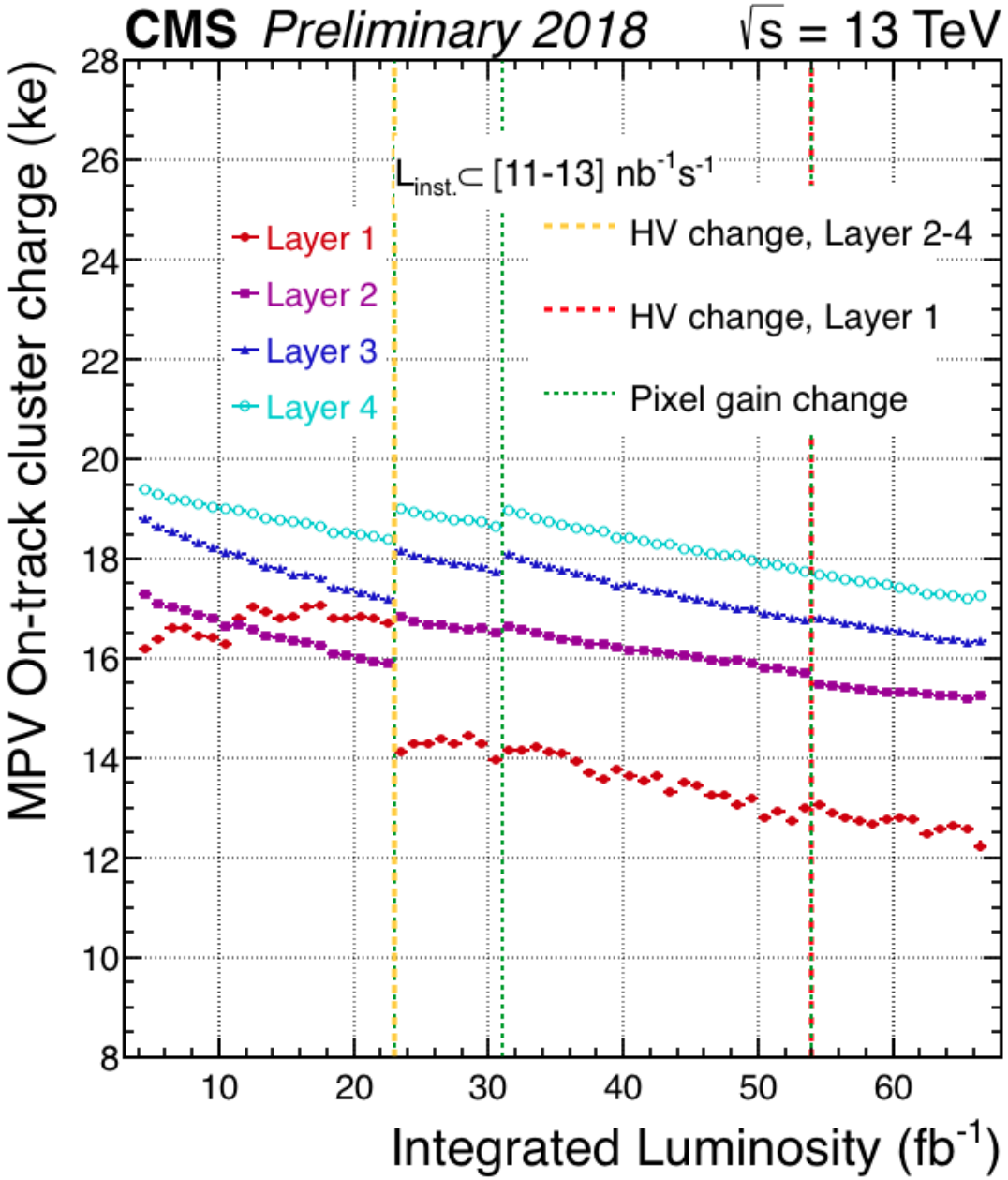
Next steps

In case no significant excess is found, limits for the mass of the HSCP will be set.

Charge has to be constant over the years

One could take the corrections from the Template DB Object:

$$\text{Charge} = \frac{\text{measured charge} * \text{charge scale}}{\text{correction factor}}$$



Next steps

In case no significant excess is found, limits for the mass of the HSCP will be set.

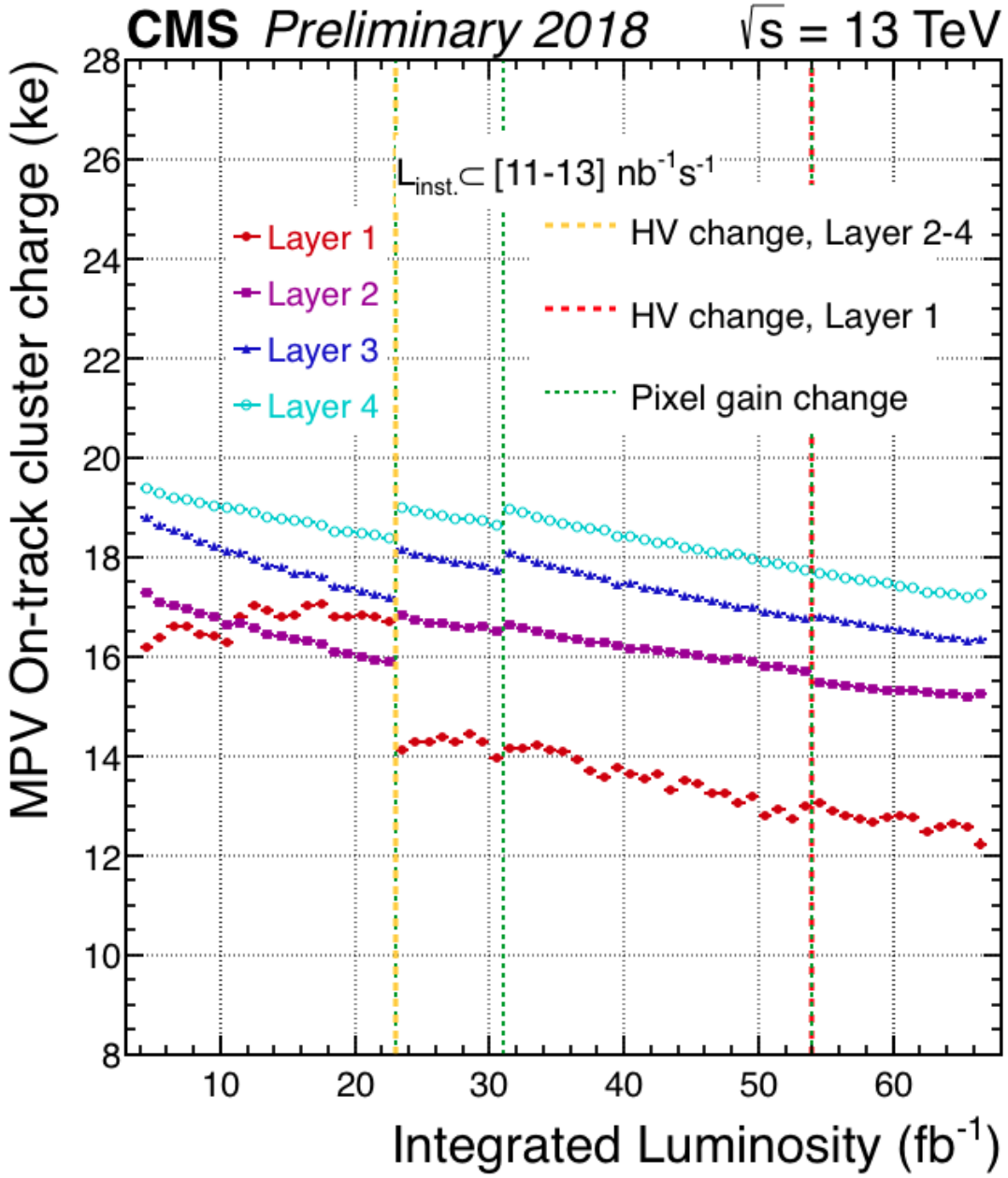
Charge has to be constant over the years

One could take the corrections from the Template DB Object:

$$\text{Charge} = \frac{\text{measured charge} * \text{charge scale}}{\text{correction factor}}$$

Then we define the dE/dx estimator as the generalized harmonic mean of the normalized charges (c_i)

$$\frac{\mathrm{dE}}{\mathrm{dx}} \propto \left(\frac{1}{N} \sum_{i}^{N} c_i^{-2} \right)^{-1/2}$$



Next steps: setting mass limits

$$-\left\langle \frac{\mathrm{dE}}{\mathrm{dx}} \right\rangle = \frac{4\pi n}{m_e c^2} \cdot \frac{z^2}{\beta^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \left[\ln\left(\frac{2 m_e c^2 \beta^2}{I(1-\beta^2)} \right) - \beta^2 \right]$$

Bethe-Bloch formula

Next steps: setting mass limits

$$-\left\langle \frac{\mathrm{dE}}{\mathrm{dx}} \right\rangle = \frac{4\pi n}{m_e c^2} \cdot \frac{z^2}{\beta^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \left[\ln\left(\frac{2 m_e c^2 \beta^2}{I(1-\beta^2)} \right) - \beta^2 \right]$$

Bethe-Bloch formula

$$= \frac{4\pi n}{m_e c^2} \cdot \frac{z^2}{\beta^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \left[\ln\left(\frac{2 m_e c^2 \beta^2}{I(1-\beta^2)} \right) \right] - \frac{4\pi n}{m_e c^2} \cdot \frac{z^2}{\beta^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \beta^2$$

$$= a \cdot \frac{1}{\beta^2} + B$$

$\beta = v/c$

$$= A \cdot \frac{1}{v^2} + B = A \cdot \frac{m^2}{p^2} + B$$

Next steps: setting mass limits

$$-\left\langle \frac{\mathrm{dE}}{\mathrm{dx}} \right\rangle = \frac{4\pi n}{m_e c^2} \cdot \frac{z^2}{\beta^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \left[\ln\left(\frac{2 m_e c^2 \beta^2}{I(1-\beta^2)} \right) - \beta^2 \right]$$

Bethe-Bloch formula

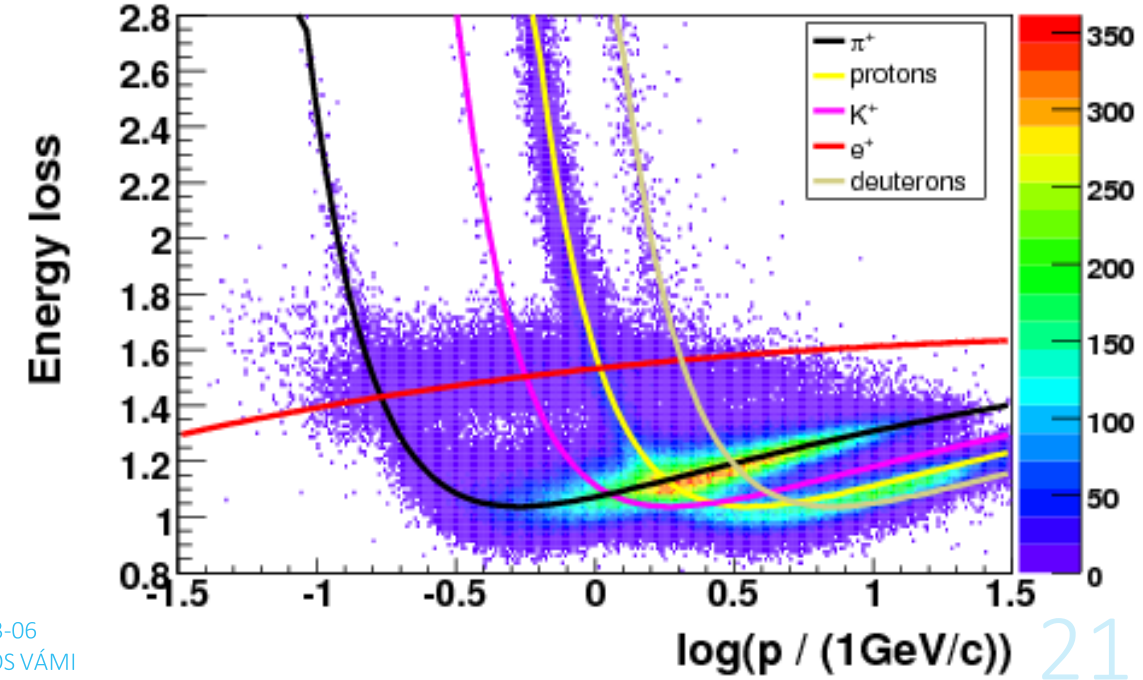
$$= \frac{4\pi n}{m_e c^2} \cdot \frac{z^2}{\beta^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \left[\ln\left(\frac{2 m_e c^2 \beta^2}{I(1-\beta^2)} \right) \right] - \frac{4\pi n}{m_e c^2} \cdot \frac{z^2}{\beta^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \beta^2$$

$$= a \cdot \frac{1}{\beta^2} + B$$

$\beta = v/c$

$$= A \cdot \frac{1}{v^2} + B = A \cdot \frac{m^2}{p^2} + B$$

A, B are determined using measurements with kaons and protons

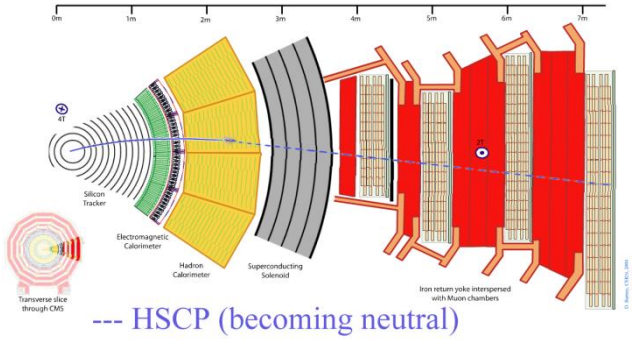


Further future – Other physics usage

Developing this method opens up several possibilities for BSM searches

In itself it can be exploited by further searches for

- Fractionally charged particles
- Multiply charged particles (double charged Higgs?)
- Long lived but not stable particles



Or it could be combined with the conventional discriminator based searches (and with this significantly reduce the phase space) like

- Low probQ + Razor variable based searches
- Low probQ with disappearing tracks
- …

Using it as the condition to record future data (called trigger in HEP)?

Conclusions

Many models predict the existence of Heavy Stable Charged Particles (HSCPs).

A proof of concept has been performed using Monte Carlo simulations to use low level tracker reconstruction quantities to discriminate between signal and background.

The re-reconstruction of the 2017-2018 data and the creation of the new data tier designed by me is in progress.

When this is done, I will use this technique to select and save the interesting events only. Afterwards, if no significant excess is found, observed limit for the masses of the HSCPs will be set.

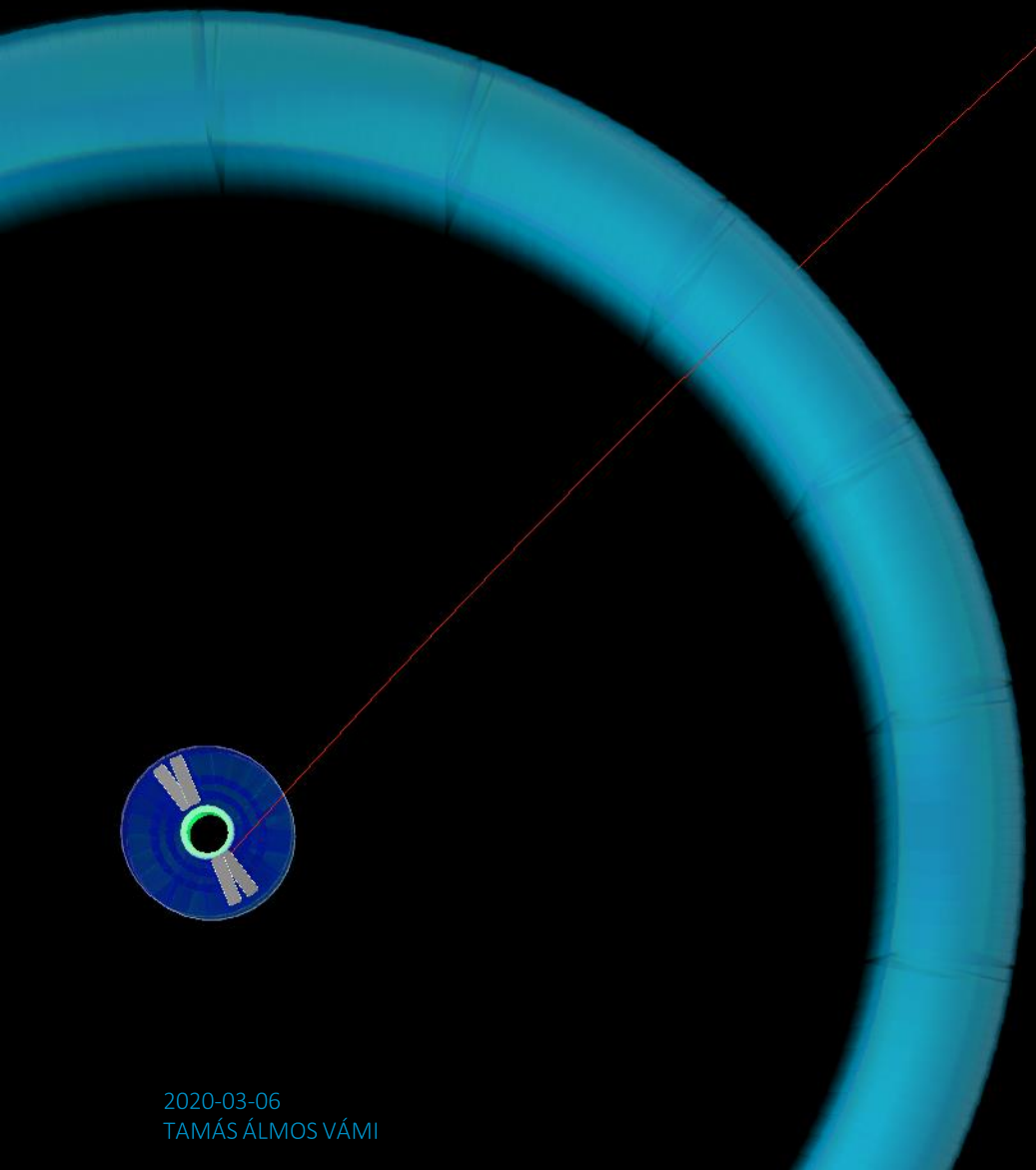
The technique is general and opens up many possibilities for BSM searches

Published papers / proceedings

[a] CMS Collaboration incl T.A.Vami. Inclusive search for supersymmetry using razor variables in pp collisions at √s = 13 TeV. CMS-PAS-SUS-16-017, CERN, Geneva, 2018.
[b] T. A. Vami for the CMS Collaboration. Searches for strong production of supersymmetry in CMS. In 18th Hellenic School and Workshops on Elementary Particle Physics and Gravity (CORFU2018) Corfu, Corfu, Greece, August 31-September 28, 2018, 2019, 1904.12148.
[c] T. A. Vami for the ATLAS and CMS Collaborations. Searches for gluinos and squarks. PoS, LHCP2019:168, 2019, 1909.11753.
[d] T.A. Vami. Master's thesis, Sensitivity study of a search for a simplified supersymmetric model using pp collisions at √s = 13 TeV with the CMS detector, 2019

* T. A. Vami for the CMS Collaboration. Calibration of the CMS Pixel Detector at the Large Hadron Collider. Jun 14, 2015. 3 pp.,Nucl.Part.Phys.Proc. 273-275 (2016) 2527-2529, 1506.05046
* T. A. Vami for the CMS Collaboration. Study of the CMS Phase-1 Pixel Pilot Blade Reconstruction. Apr 2, 2018. 4 pp., Springer Proc.Phys. 213 (2018) 366-369, 1804.00483
* T. A. Vami for the CMS Collaboration. Calibration and performance of the CMS pixel detector in LHC Run 2. Sep 9, 2019. 7 pp. PoS LHCP2019 (2019), 1909.12920

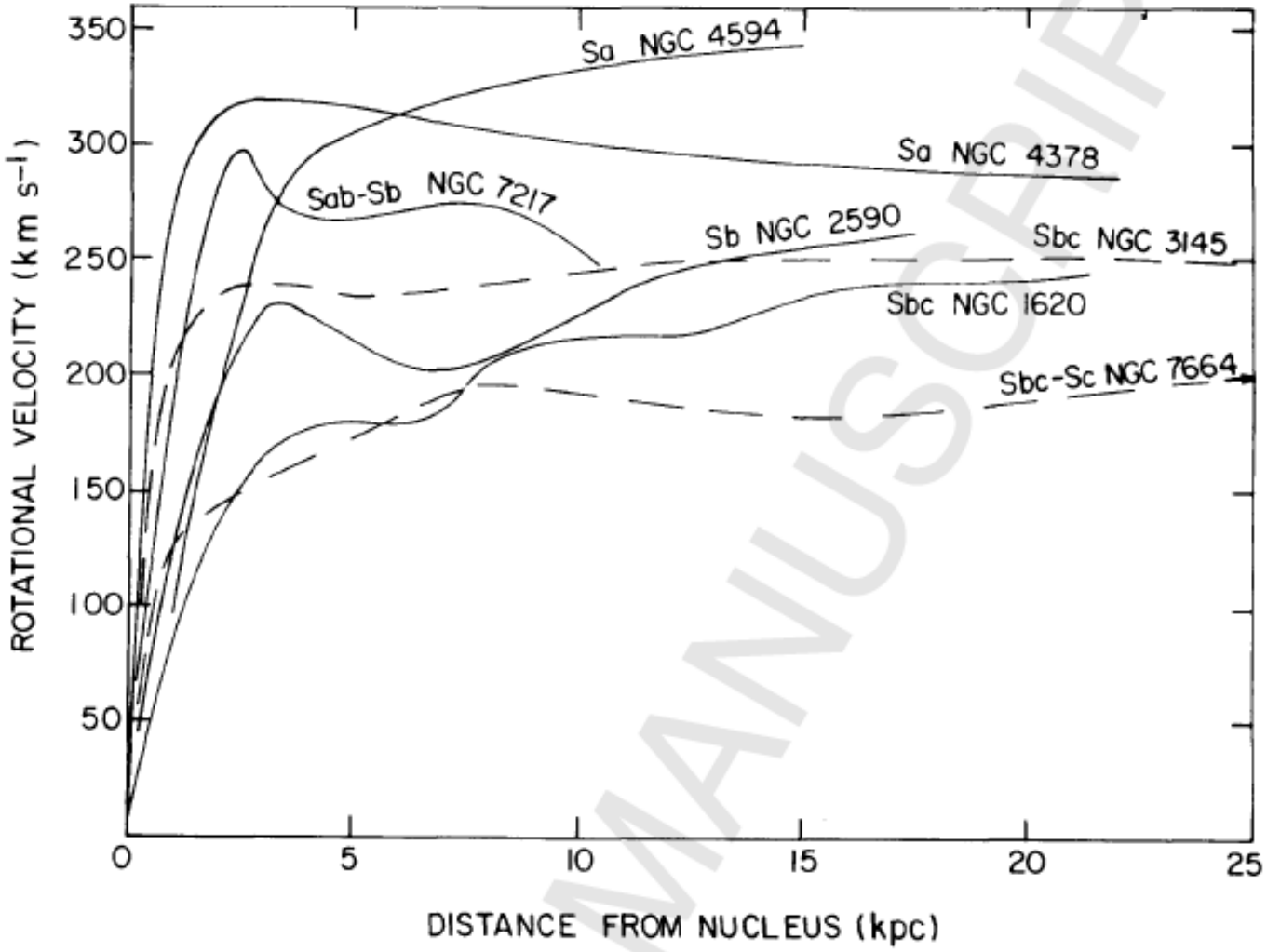
BACKUP



„Evidence” for dark matter

1932, Jan Oort -> dim stars

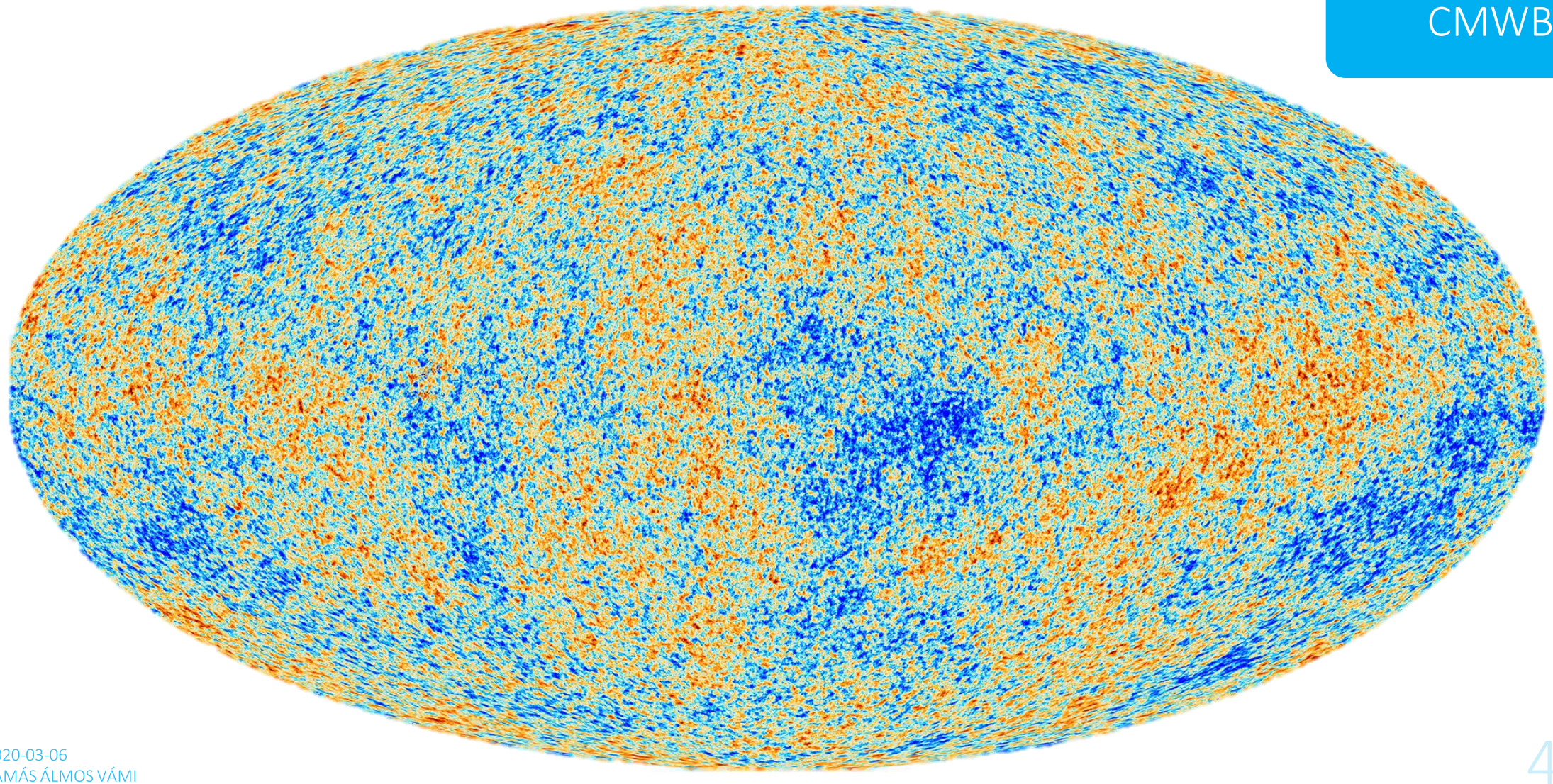
1933, Fritz Zwicky -> dark matter (DM)



MANUSCRIPT

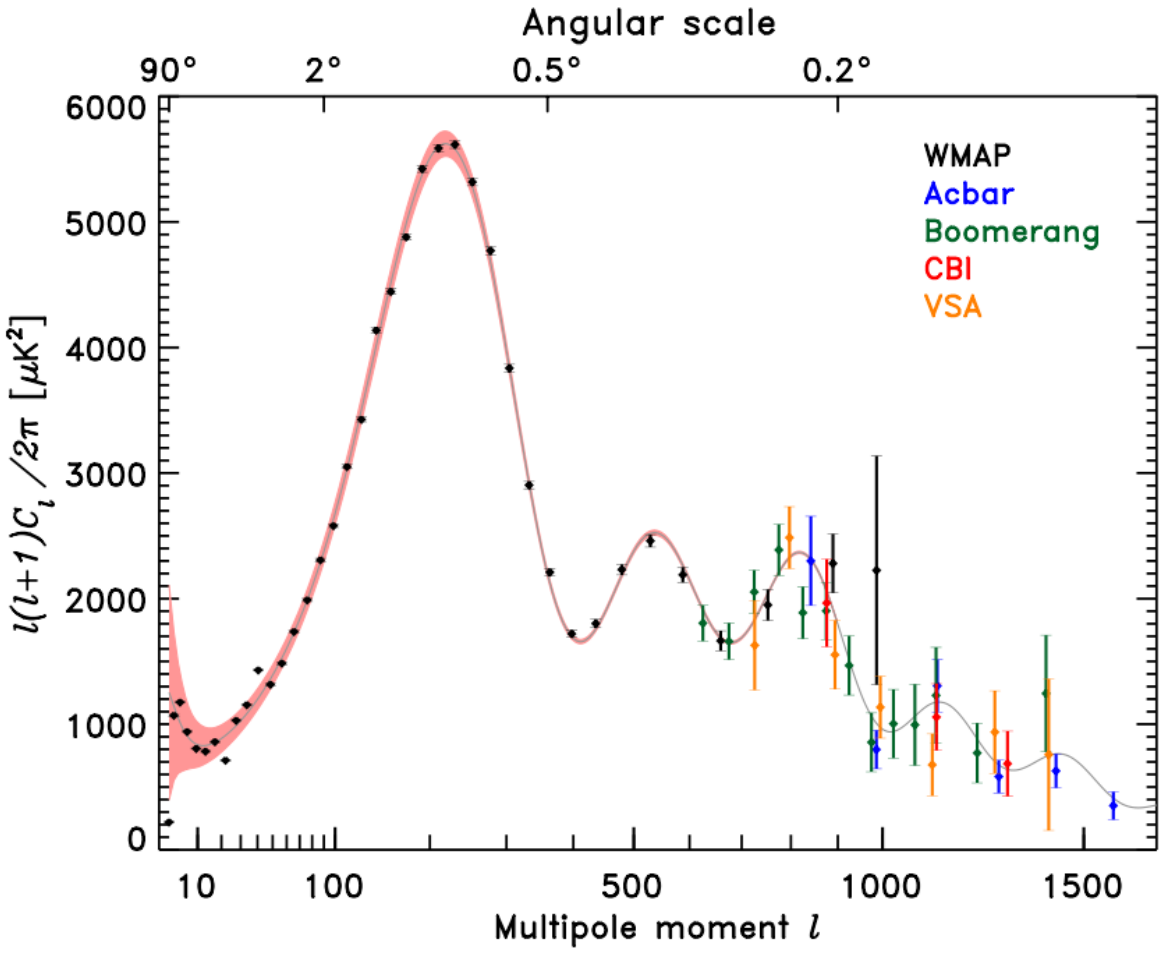
„Evidence” for dark matter/energy

CMWB



„Evidence" for dark matter/energy

Power spectrum:
-> each peak corresponds to the energy-mass density



SUSY solution to dark matter

In 1978 a new quantity was introduced, called
R-parity: $(-1)^{3(B-L)+2s}$
where s is the spin, B is the baryon number, and L is the lepton number

This allows the SM particles to be distinguished from the SUSY ones,

- as the former have R-parity of +1
- while the latter have R-parity of −1.

If R-parity is conserved then the lightest SUSY particle is stable and neutral, thus it is an excellent candidate for Dark Matter

Hierarchy problem

The electric force is 10^{42} times stronger than gravity.

The relative strengths of the forces are characterized by their dim-less coupling constants:

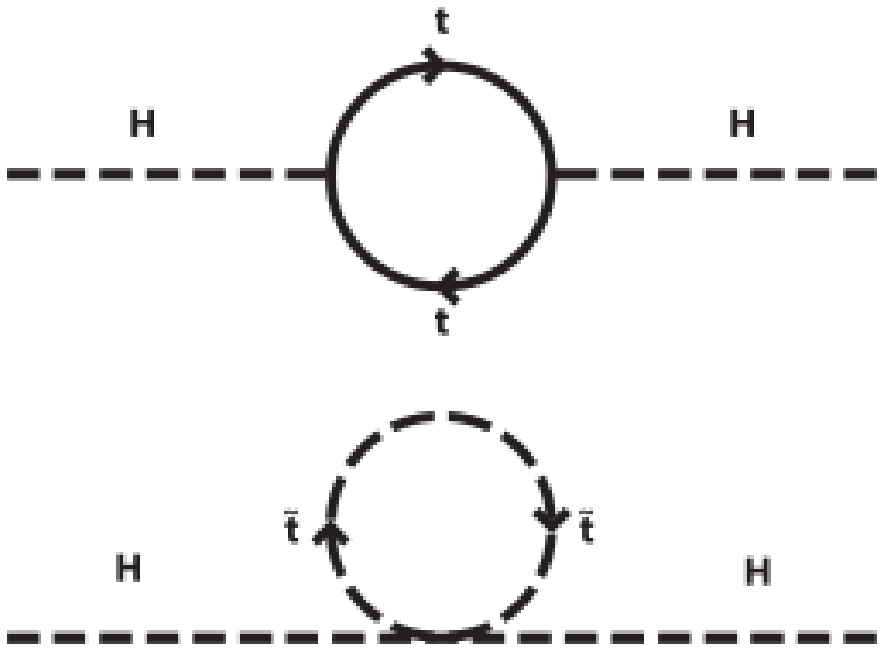
- $\alpha_{strong} \approx 1$
- $\alpha_{EM} \approx \frac{1}{137} = 7.29 * 10^{-3}$
- $\alpha_{weak} \approx 10^{-6}$
- $\alpha_{grav} \approx 10^{-45}$

- The SM explains the EM and weak forces in the framework of the Salam-Weinberg electroweak theory using the Higgs field.
- Thus the problem can be restated connected to the mass of the Higgs boson.
- In the SM only, the quantum contributions to the Higgs mass should make its mass huge

SUSY solution to the hierarchy problem

The mass of the Higgs is not on the Planck scale

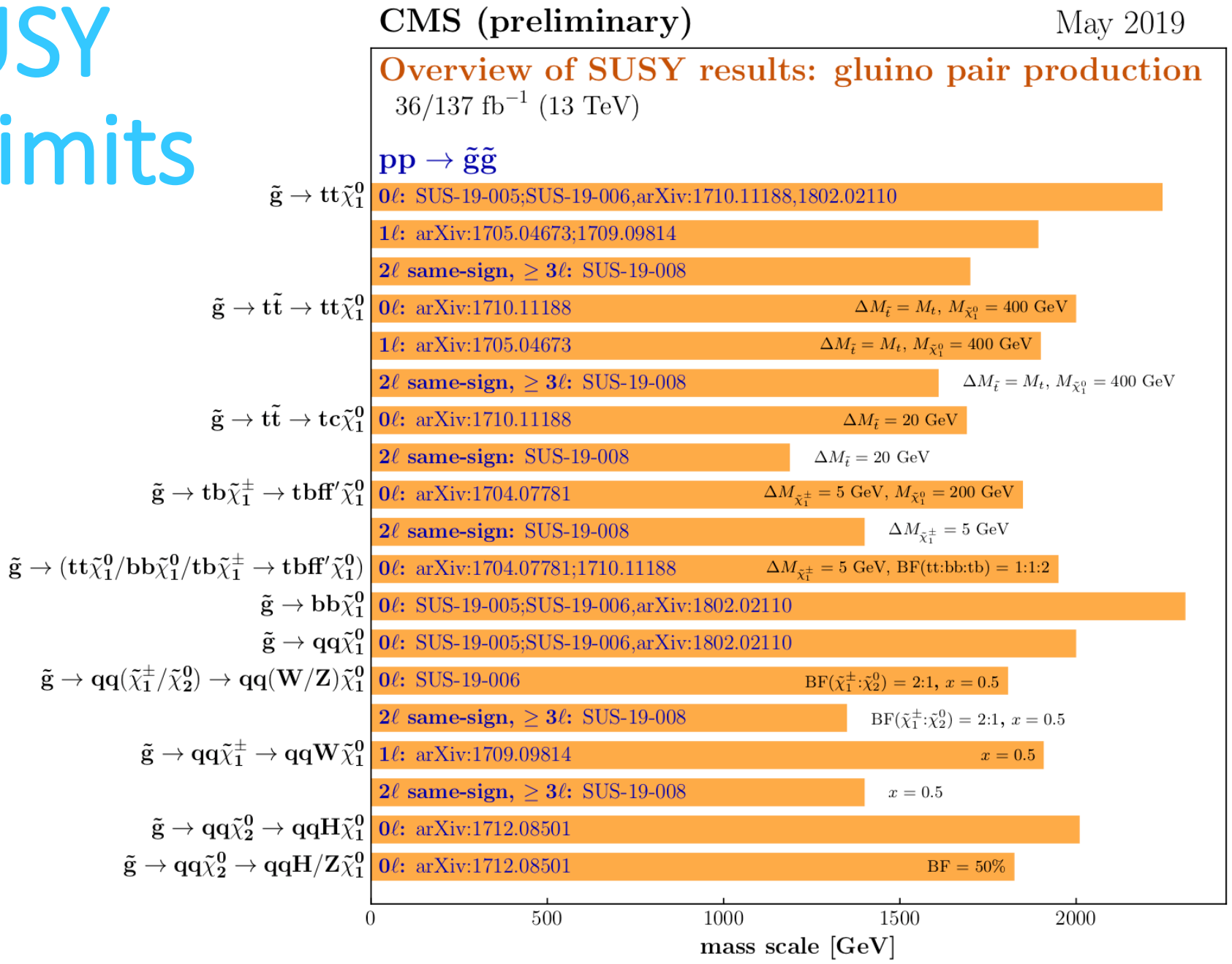
This means that the quantum contributions should be cancelled



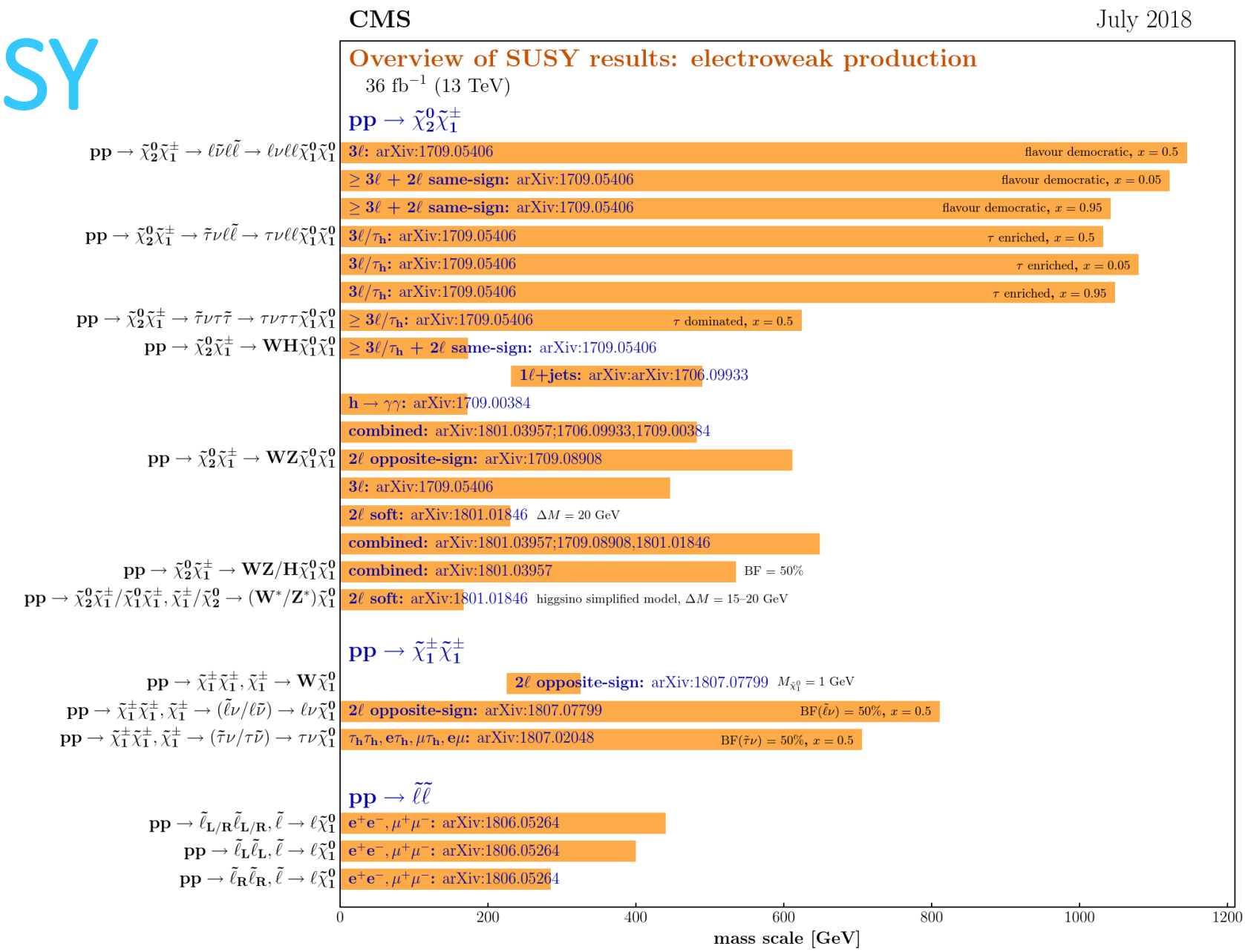
The most significant corrections to the Higgs mass will originate from the heaviest particle, the top quark

The fermionic top quark loop and scalar stop squark tadpole contributions cancel eachother

Search for SUSY gluino mass limits



Search for SUSY Electroweak production mass limits



Selection of observed limits at 95% C.L. (theory uncertainties are not included). Probe **up to** the quoted mass limit for light LSPs unless stated otherwise. The quantities ΔM and x represent the absolute mass difference between the primary sparticle and the LSP, and the difference between the intermediate sparticle and the LSP relative to ΔM, respectively, unless indicated otherwise.

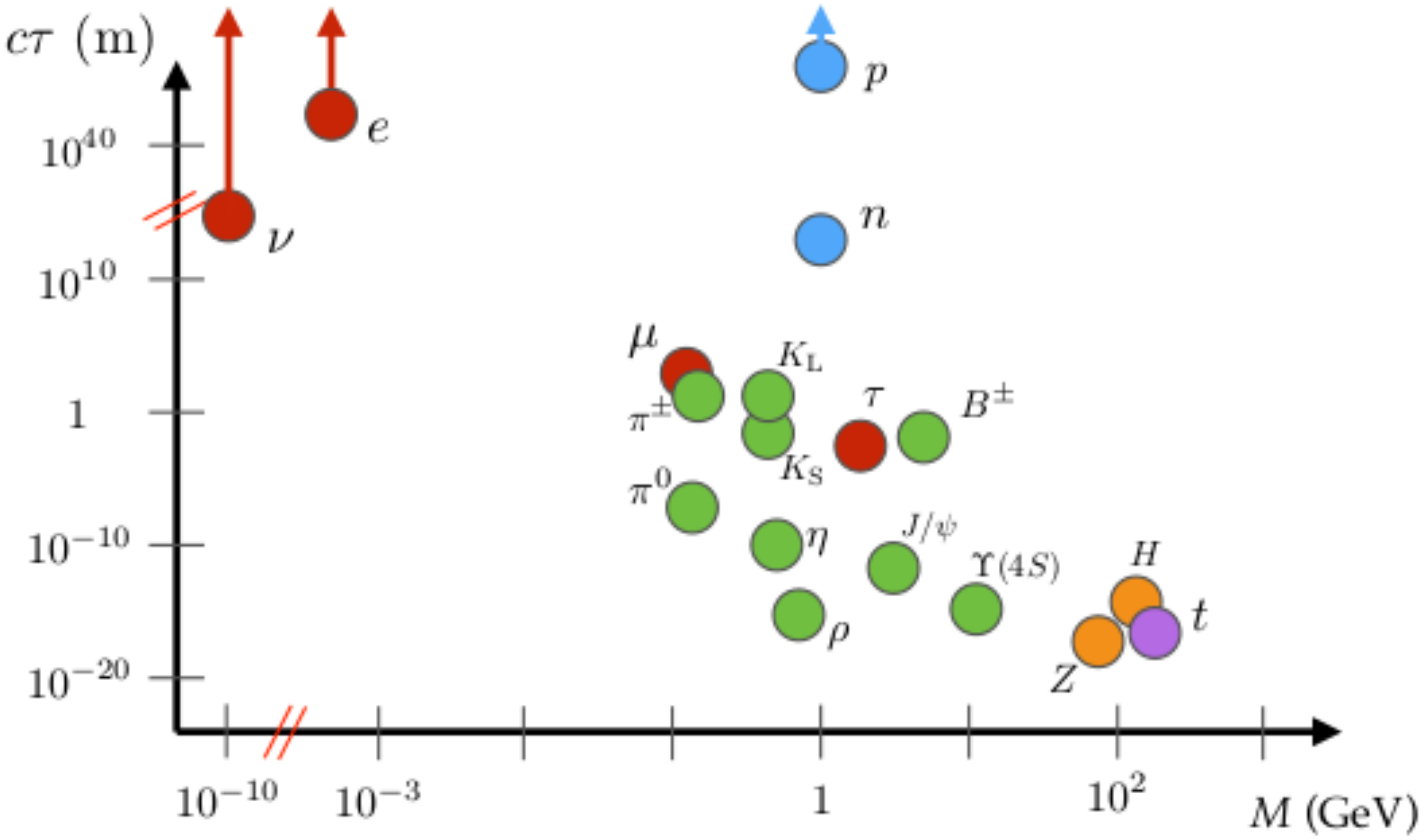
Particle lifetimes

In the SM lifetimes spanning an enormous range of magnitudes

$$d\Gamma \sim \frac{1}{M}|\mathcal{M}|^2 d\Pi.$$

Long lived if

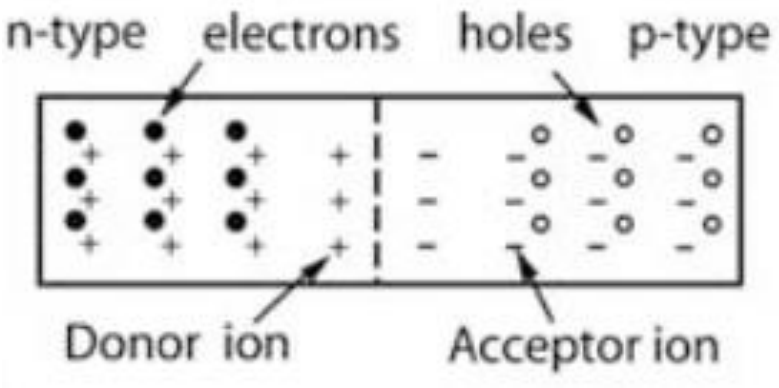
- there was an approximate symmetry which would forbid the decay if it was precise, or simply a small effective coupling constant
- small breaking of an approximate symmetry that splits otherwise degenerate states, or it can arise due to accidental degeneracies in the spectrum



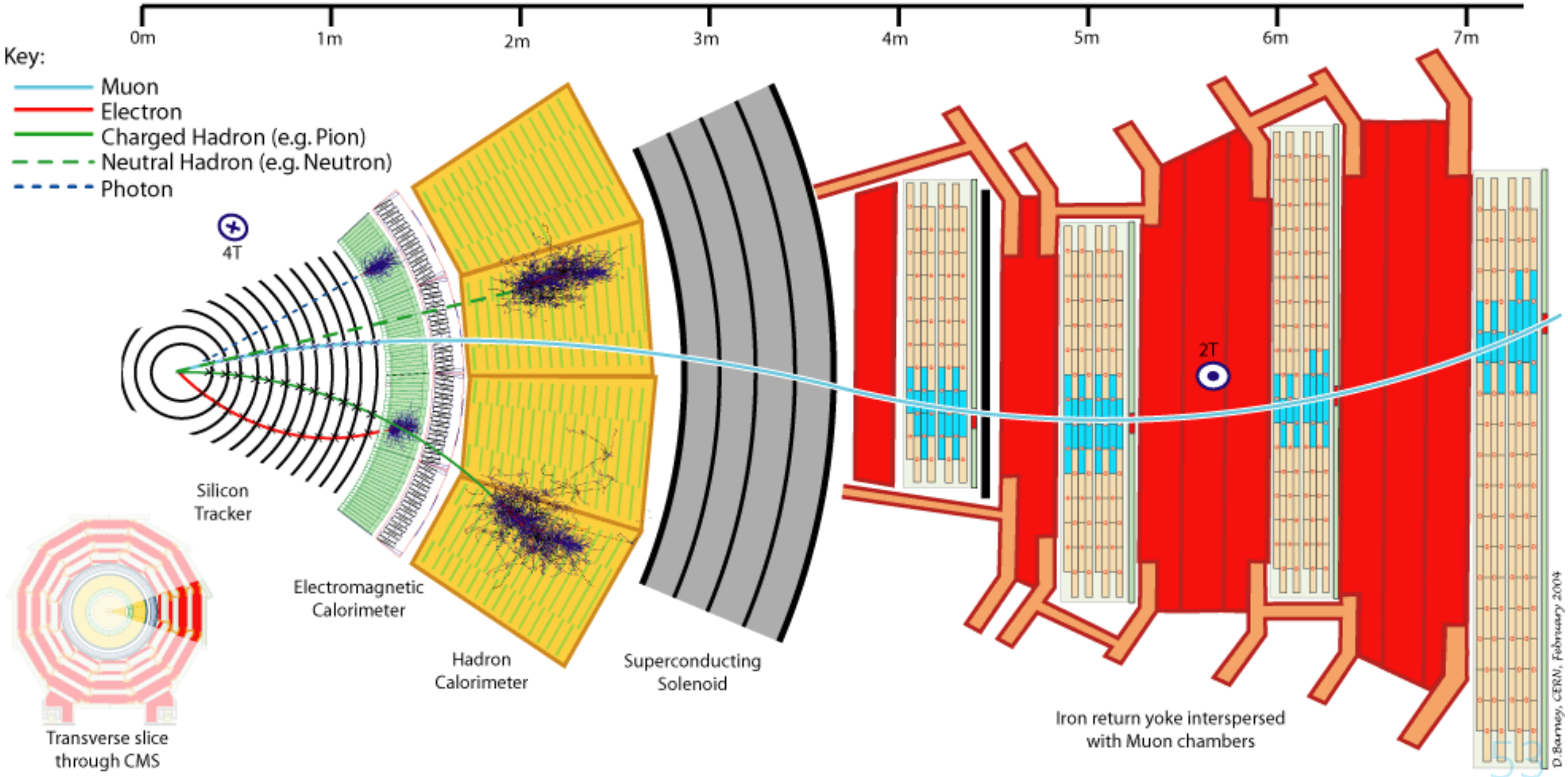
Semiconductor detectors

n-type silicons: Si doped with an electron donor atoms
p-type silicons: Si doped with an electron acceptor atoms
Semiconductor detector: a pn junction

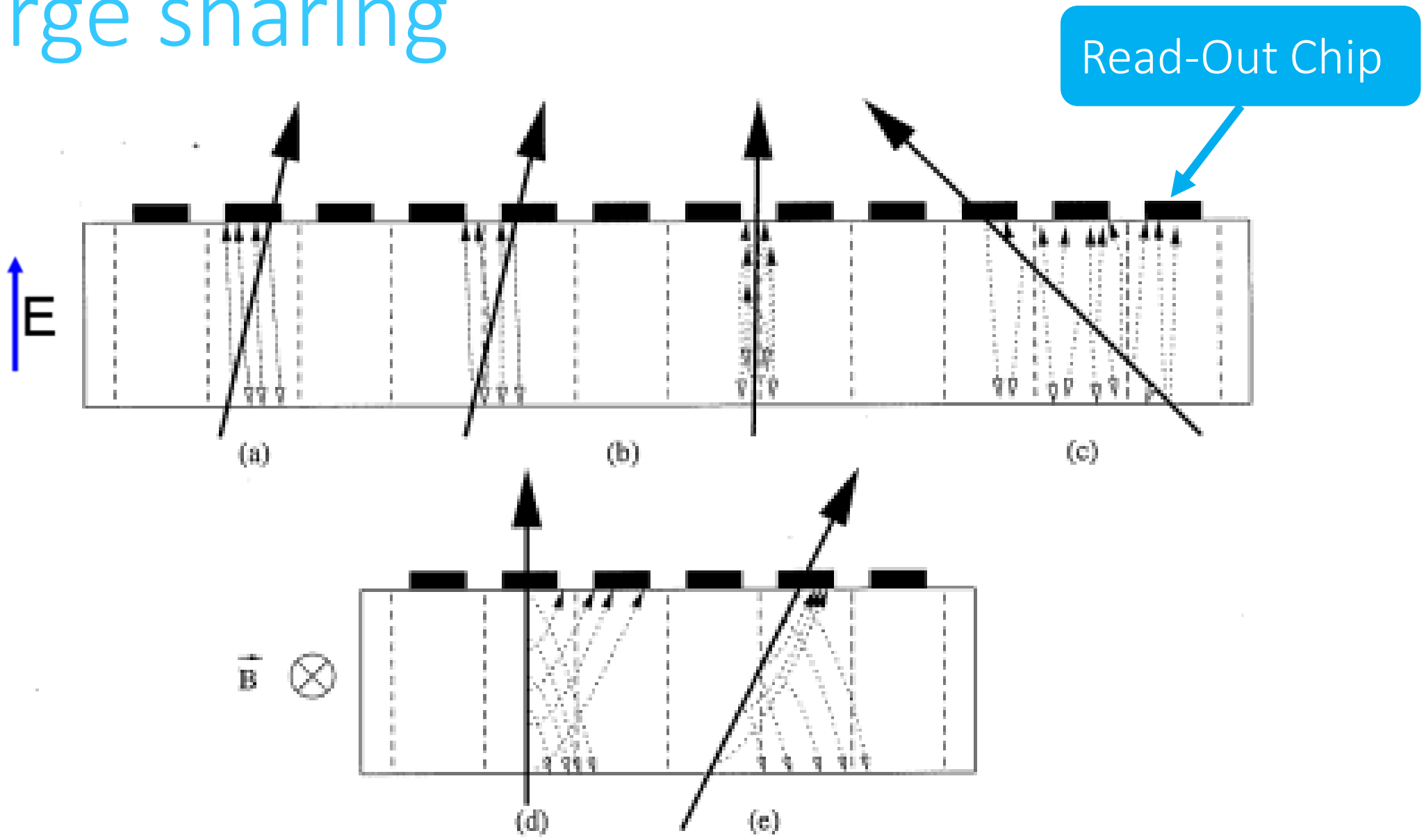
- The excess of negative (positive) charge in the p(n)-region builds up a potential difference and therefore an electric field over the junction region
- This process will stop when the electric field generated is sufficiently large to prevent any further build-up of a charge difference
- In this contact area (the depletion region), the number of free charge carriers will be strongly reduced because free charges are removed by the electric field.



The detector: Compact Muon Solenoid



Charge sharing



PixelAV

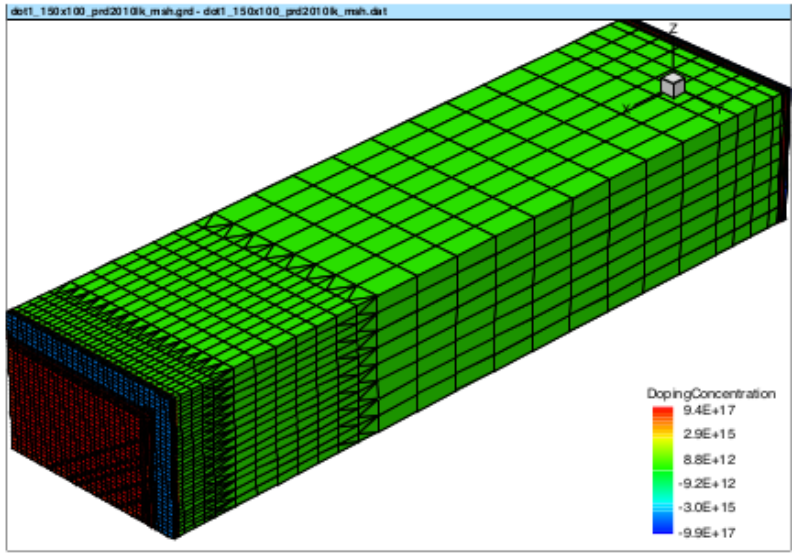
Created to interpret beam tests of irradiated sensors, now used to perform Lorentz calibrations and generate template profile shapes:

- charge deposition model based on Bichsel π-Si x-sections
- delta ray range: Continuous Slowing Down Approx + Nist Estar dedx
- plural scattering and magnetic curvature of delta ray tracks
- carrier transport from Runge-Kutta integration of saturated drift

$$\frac{d\vec{x}}{dt} = \vec{v} = \frac{\mu\left[q\vec{E} + \mu r_H \vec{E} \times \vec{B} + q\mu^2 r_H^2 (\vec{E}\cdot\vec{B})\vec{B}\right]}{1 + \mu^2 r_H^2 |\vec{B}|^2}$$

  - electric field map from ISE TCAD simulation of pixel cell
  - includes diffusion, trapping, and charge induction on implants

- Electronic Simulation: noise, linearity, thresholds, mis-calibration



3

Vavilov function

$$f(\lambda) = \frac{1}{2\pi i}\int_{c+i\infty}^{c-i\infty} \phi(s)\, e^{\lambda s} ds \qquad c \geq 0$$

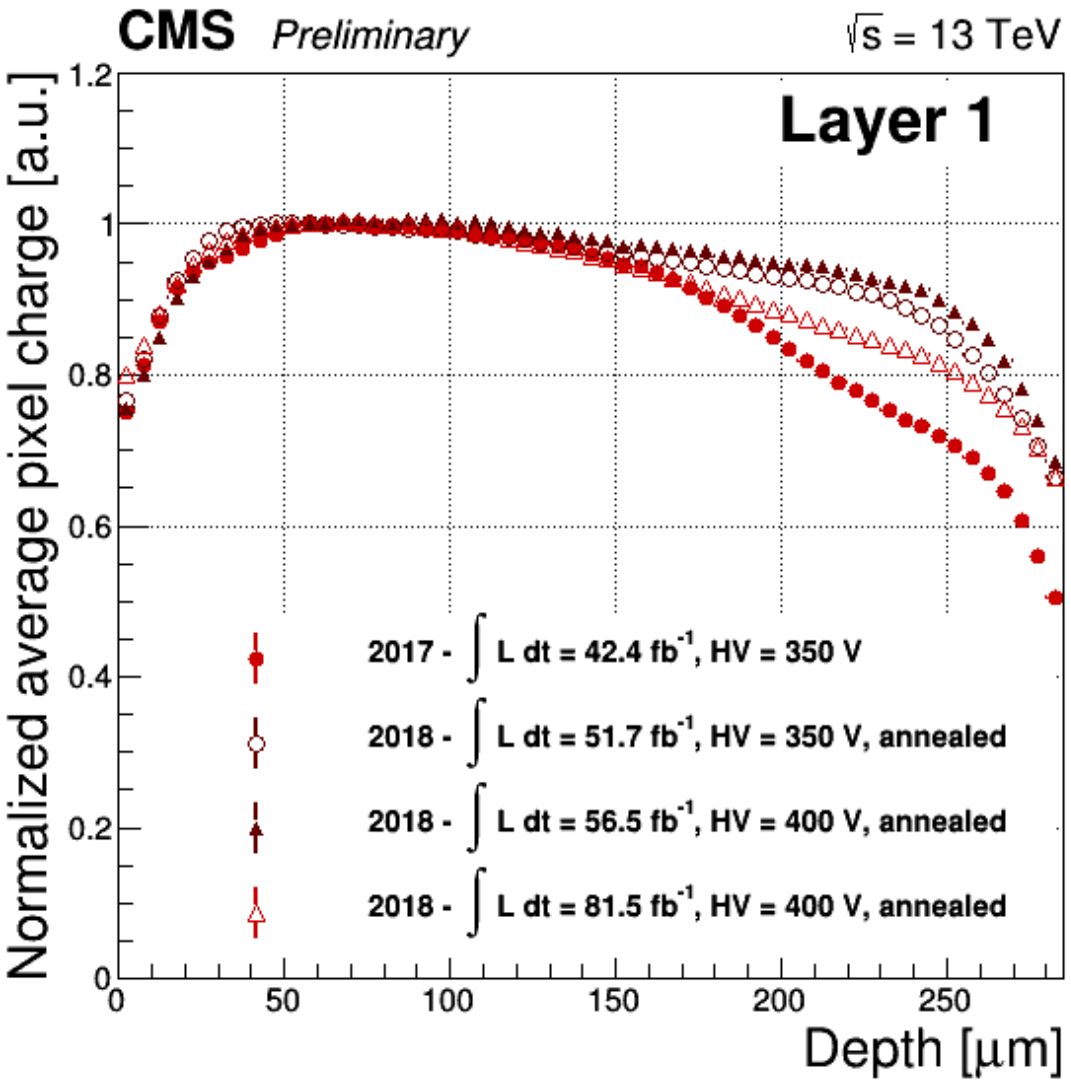
$$\phi(s) = \exp\left[\kappa(1+\beta^2\gamma)\right] \exp\left[\psi(s)\right],$$

$$\psi(s) = s\ln\kappa + (s+\beta^2\kappa)\left[\ln(s/\kappa) + E_1(s/\kappa)\right] - \kappa e^{-s/\kappa},$$

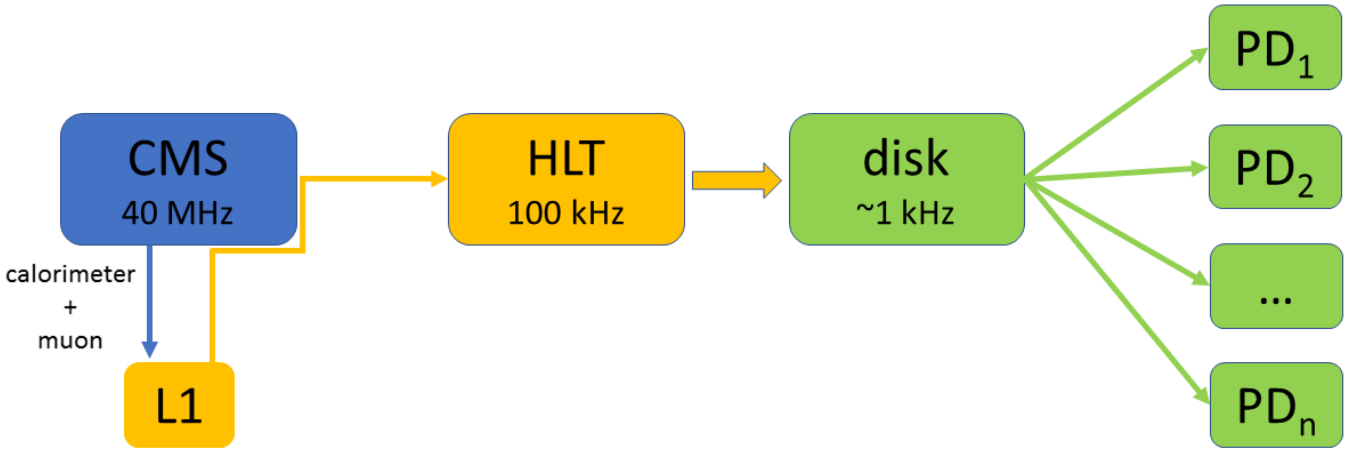
$$E_1(z) = \int_{\infty}^{z} t^{-1} e^{-t} dt$$

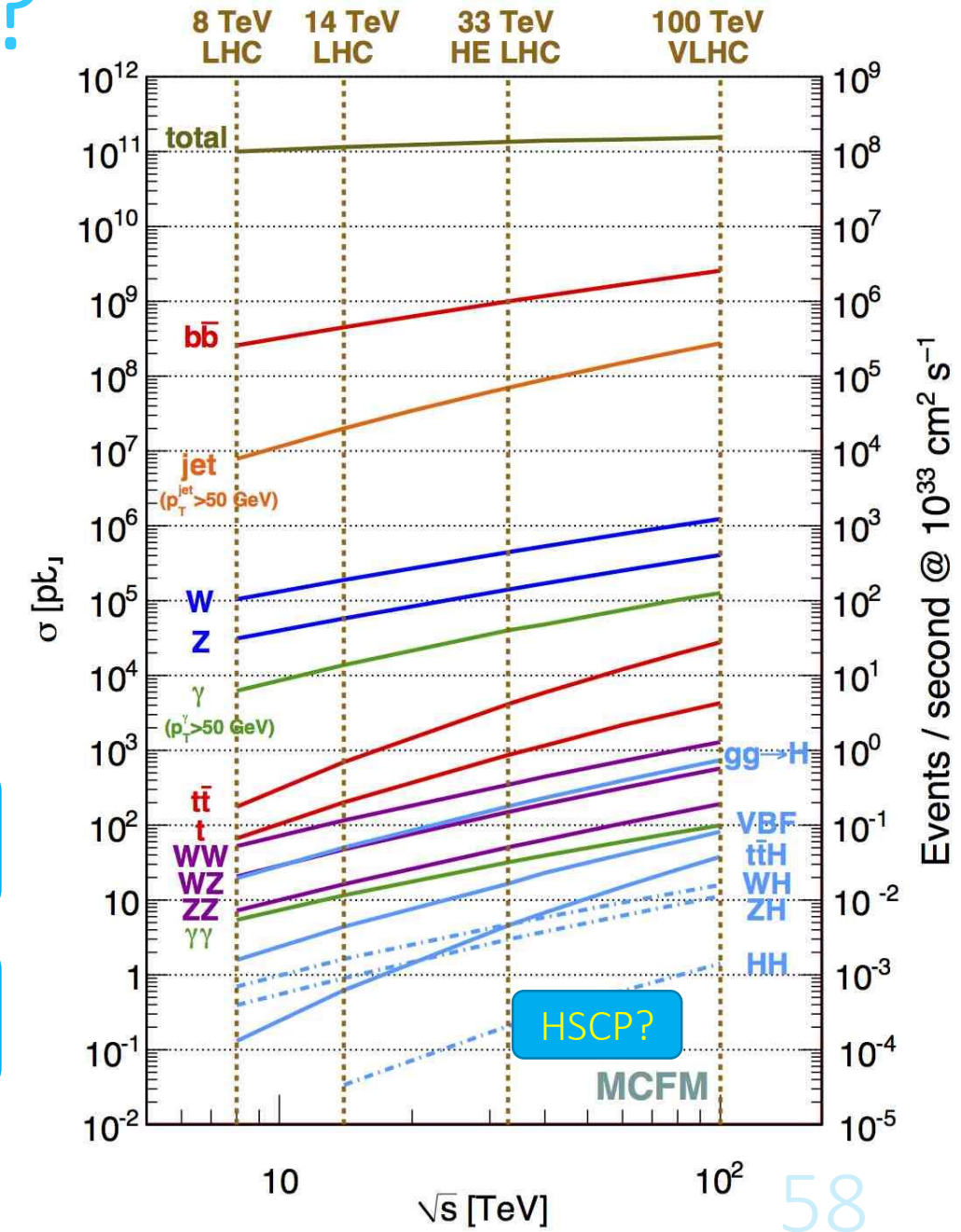
Why does the charge decrease?

- For non irradiated, fully depleted detector, the pixel charge profile [normalised average pixel charge as a function of the production depth] is expected to be flat as detector is fully efficient and all charge is collected,
- while for irradiated detector the losses are expected due to the trapping of carriers. The losses are larger for the charges released further from the readout plane



Further future – Trigger?





From 2021, HLT will use CPUs+GPUs, which may allow to run the template algorithm on HLT

Running template algo on HLT is a GAME CHANGER!

Among many things, we could use it to trigger on low ProbQ events

Combining probabilities (1)

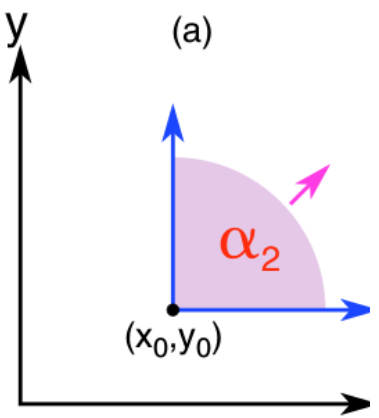
Let x and y to be two independent random variables, which are distributed according to the differential probability distributions f (x) and g(y), respectively. Let x_0 and y_0 be the outcome of a single observation.

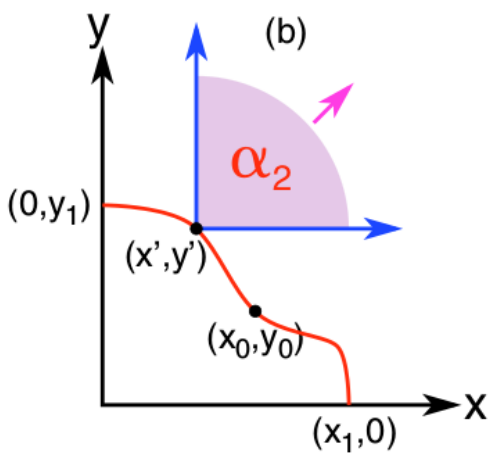
The simplest definition of the joint probability :

$$\alpha_2 = \int_{x_0}^{\infty} dx\ f(x) \int_{y_0}^{\infty} dy\ g(y) = \int_{(x_0,y_0)}^{(\infty,\infty)} dxdy\ f(x)g(y)$$

An alternate definition is to define a contour in x-y space of all points (x' , y ') that have a constant joint probability:

$$\int_{(x',y')}^{(\infty,\infty)} dxdy\ f(x)g(y) = \int_{(x_0,y_0)}^{(\infty,\infty)} dxdy\ f(x)g(y) = \alpha_2$$





Combining probabilities (2)

Note that the endpoints of the contour are thus defined as:

$$\int_{x_1}^{\infty} dx\, f(x) = \int_{y_1}^{\infty} dy\, g(y) = \alpha_2.$$

We therefore define the new probability β which is the integral of the joint distribution function:

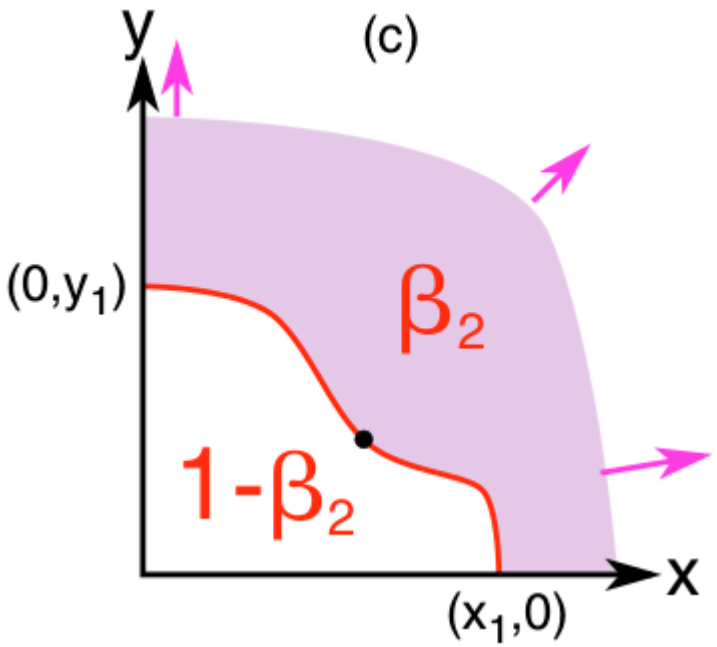
$$\beta_2 \equiv \int_{(x',y')}^{(\infty,\infty)} dxdy\, f(x)g(y).$$

The evaluation of the complementary integral is easier:

$$1 - \beta_2 = \int_0^{x_1} dx\, f(x) \int_0^{y'} dy\, g(y)$$

And yields to

$$\beta_2 = \alpha_2 \left[1 - \ln(\alpha_2)\right].$$

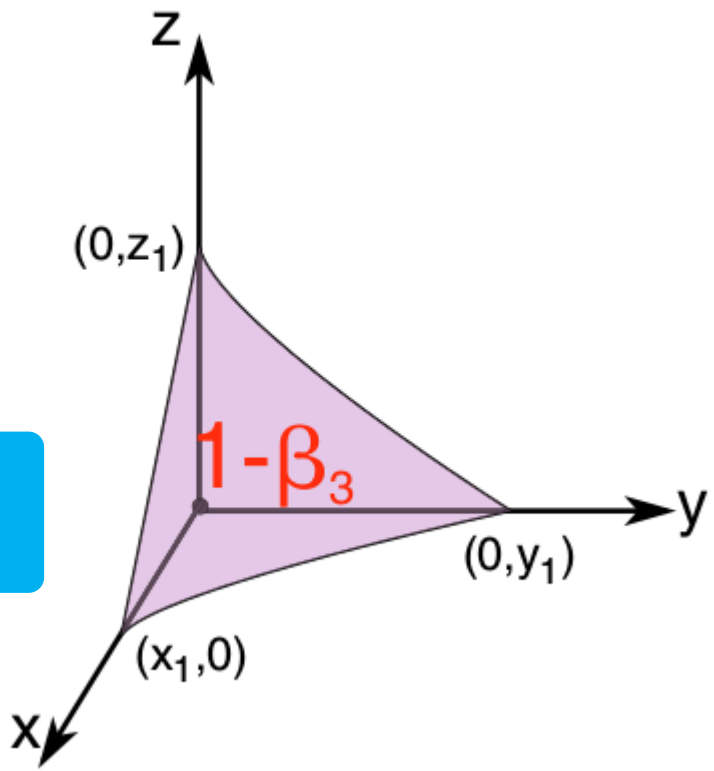


Combining probabilities (3)

Generalization to N dimensions leads to the formula

$$\beta_N = \alpha_N \sum_{n=0}^{N-1} \frac{[-\ln(\alpha_N)]^n}{n!}$$

Hypersurface for 3D:



Systematic uncertainties

The table on the right shows the 2016 uncertainties

Source of Systematic Uncertainties	Relative Uncertainty (%)	
Signal acceptance	Trk-only	Trk+TOF
- Trigger efficiency	13	13
- Track momentum scale	< 5	< 14
- Track reconstruction	< 2	< 2
- Ionization energy loss	< 13	< 7
- HIP background effect	< 4	< 4
- Time-of-flight	-	< 6
- Muon reconstruction	-	2
- Pile-up	< 2	< 2
Total uncertainty on signal acceptance	< 20	< 20
Collision background uncertainty	20	20
Luminosity uncertainty	2.5	

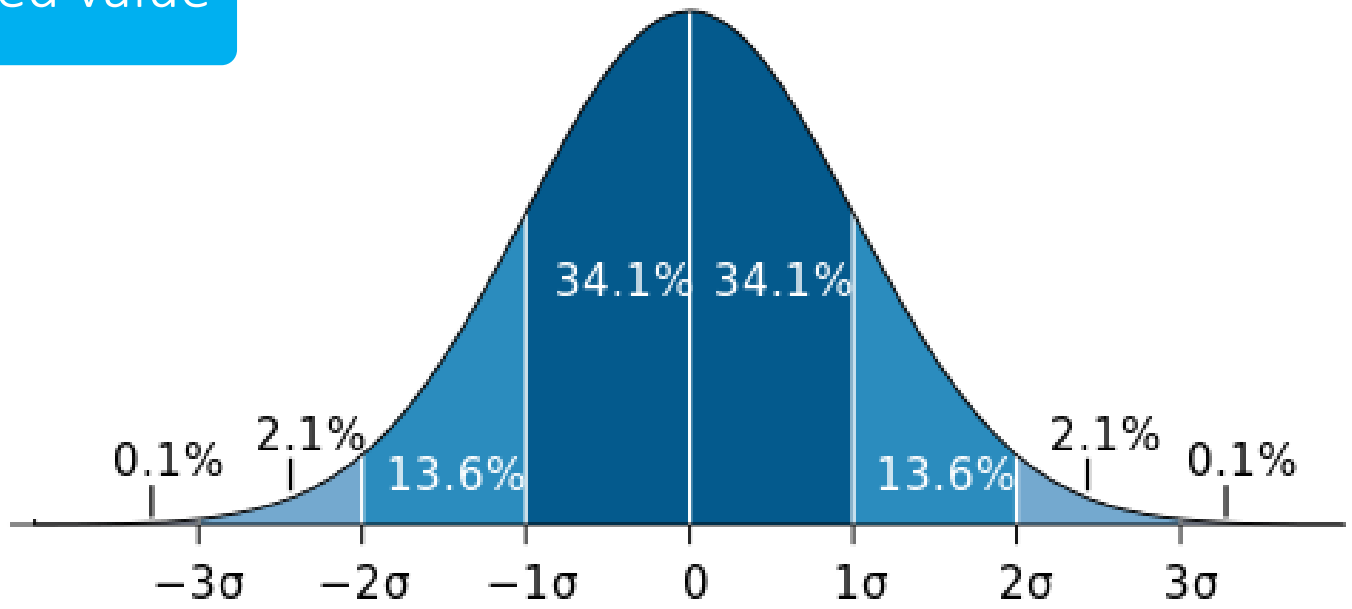
With ProbQ this should be better

Also with probQ we cut out 90 % of the background. The 2016 method cut away 35 %

What counts as significant excess?

σ: standard deviation from the expected value

1-sigma: 68.2 %
2-sigma: 95.5 %
3-sigma: 99.73 %
4-sigma: 99.993 %
5-sigma: 99.99994 %



Purely statistical fluctuations will give a result way out in the 5-sigma range 0.00006 % of the time. 5-sigma result thus means that there's a 1 in 3.5 million chance that it was the result of a statistical fluctuation (due to the statistical nature of QFT)

Mass limits from 2016 AN

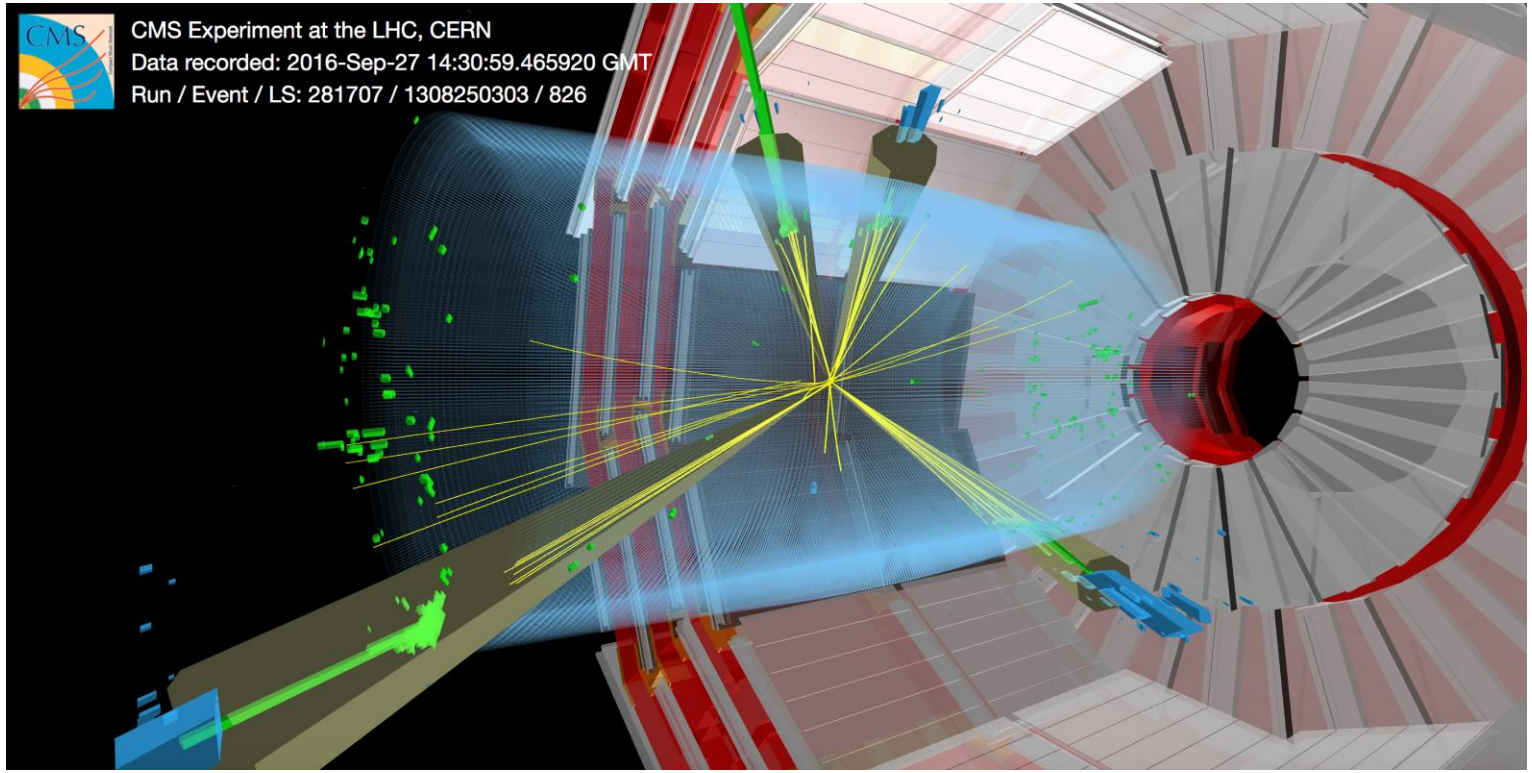
The table on the right shows the 2016 data based mass limits

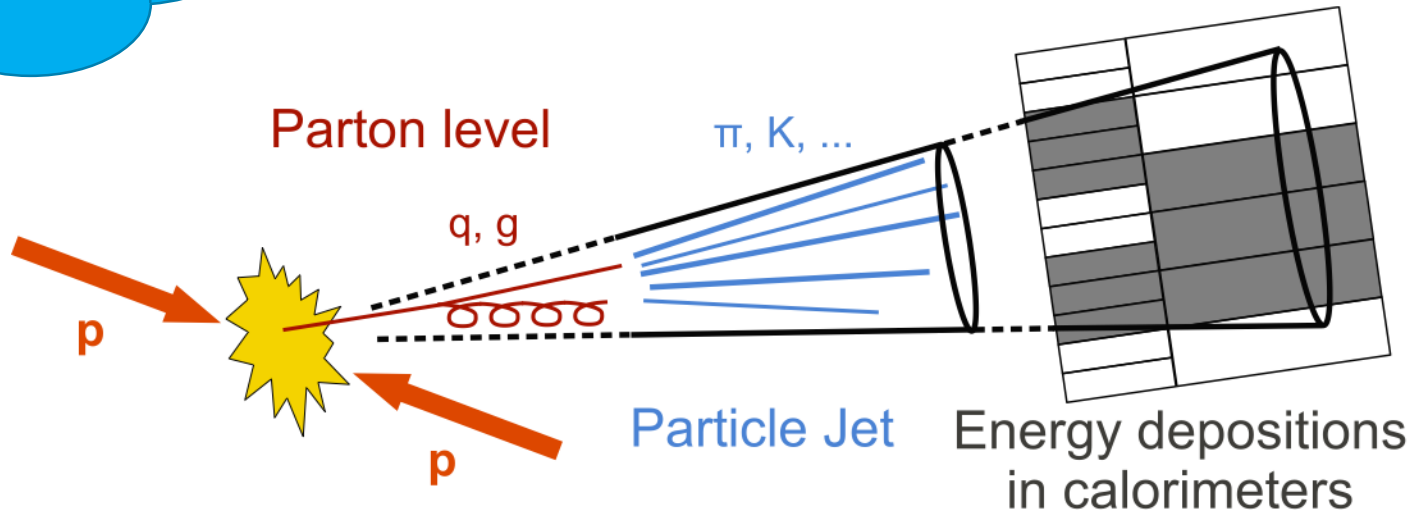
Model	2015 analysis	2016 analysis
Gluino $f = 0.1$	$M > 1610$ GeV	$M > 2081(2167)$ GeV
Gluino $f = 0.1$ CS	$M > 1580$ GeV	$M > 2071(2158)$ GeV
Gluino $f = 0.5$	$M > 1520$ GeV	$M > 1981(2058)$ GeV
Stop	$M > 1040$ GeV	$M > 1389(1441)$ GeV
Stop CS	$M > 1000$ GeV	$M > 1349(1412)$ GeV
GMSB Stau	$M > 480$ GeV	$M > 718(742)$ GeV
Pair Prod. Stau		$M > 310(312)$ GeV
DY $Q = 1e$	$M > 510$ GeV	$M > 795(837)$ GeV
DY $Q = 2e$	$M > 680$ GeV	$M > 976(985)$ GeV

Main observables

- $\vec{p}_T$
- Number of (b-)jets

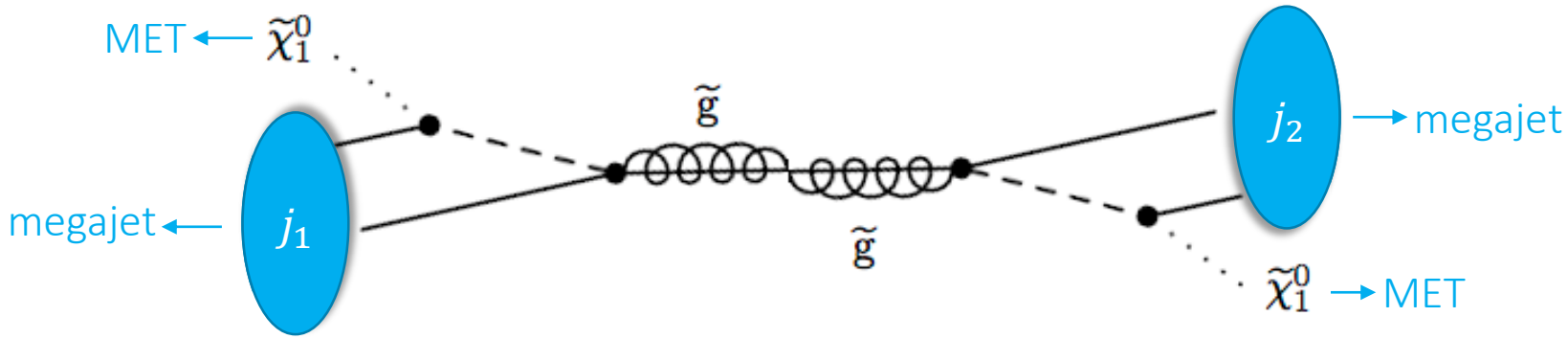
A jet is a number of particles originating from a single point, propagating in almost the same direction, and forming a cone.





Main observables

- $\vec{p}_T$
- Number of (b-)jets
- MET = $|-\sum_{all}\vec{p}_T|$
- $H_T = \sum_{jets}|\vec{p}_T|$



The search is performed in bins of the razor variables M_R and R^2 .

$$M_R = \sqrt{(|p^{j_1}| + |p^{j_2}|)^2 - \left(p_z^{j_1} + p_z^{j_2}\right)^2}$$

$$M_T^R = \sqrt{\frac{E_T^{miss}\left(p_T^{j_1} + p_T^{j_2}\right) - \vec{p}_T^{\,miss}(\vec{p}_T^{\,j_1} + \vec{p}_T^{\,j_2})}{2}}$$

$$R = \frac{M_T^R}{M_R}$$

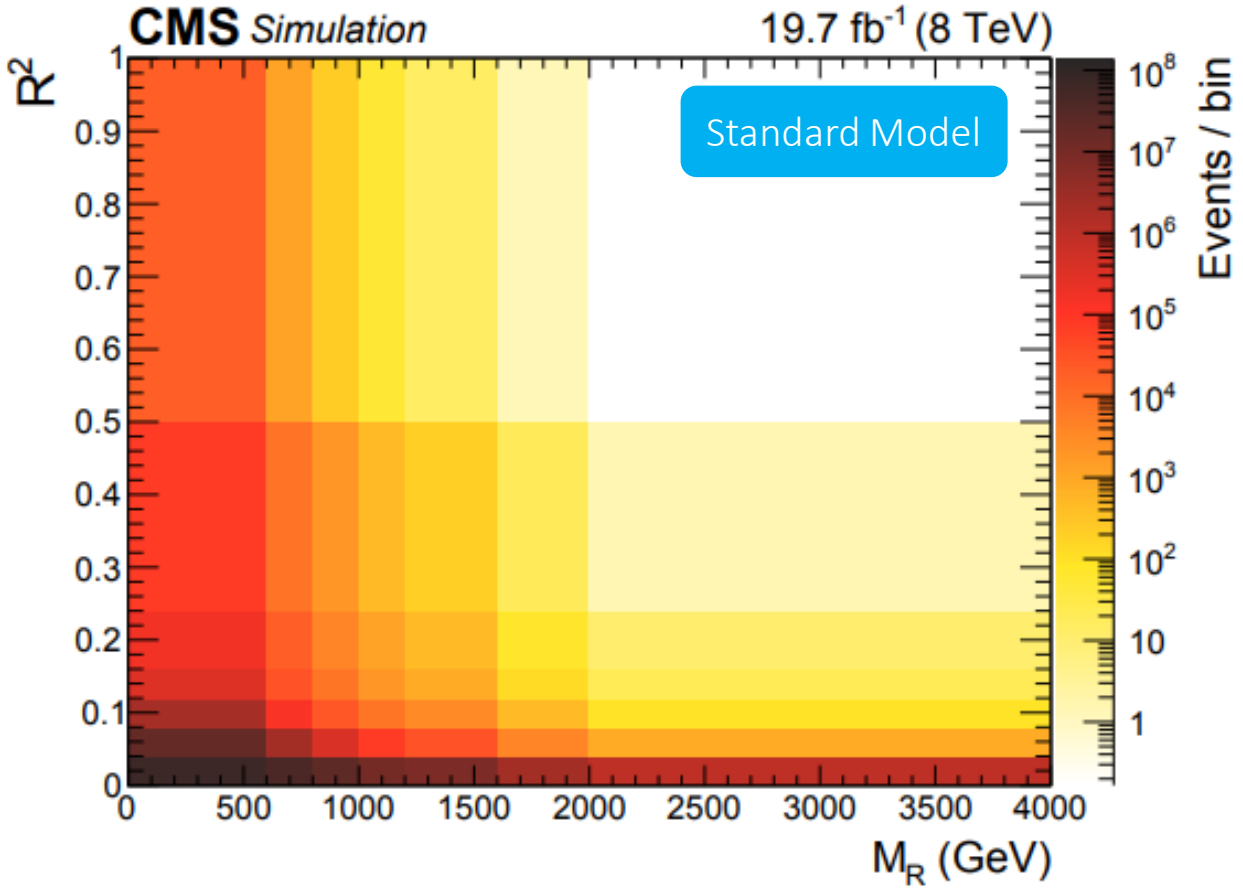
M_R is like H_T and

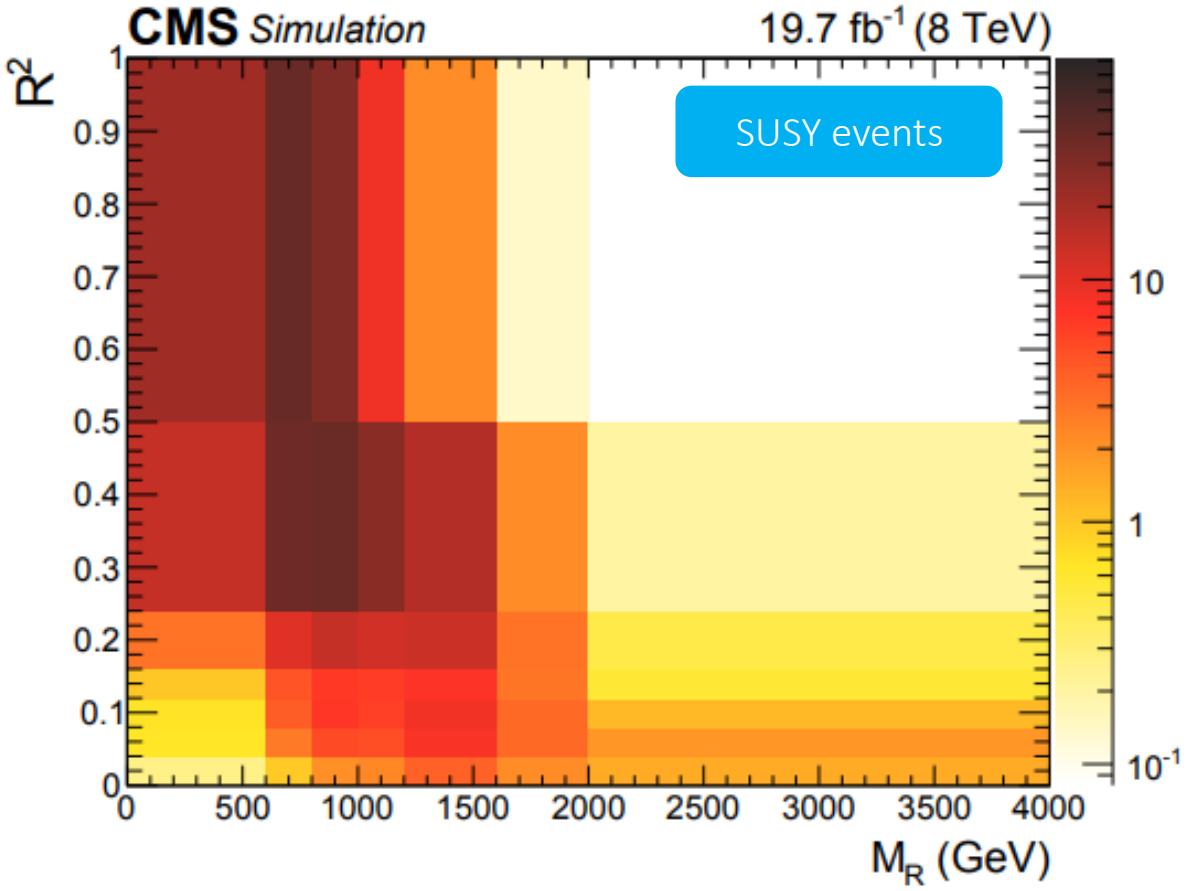
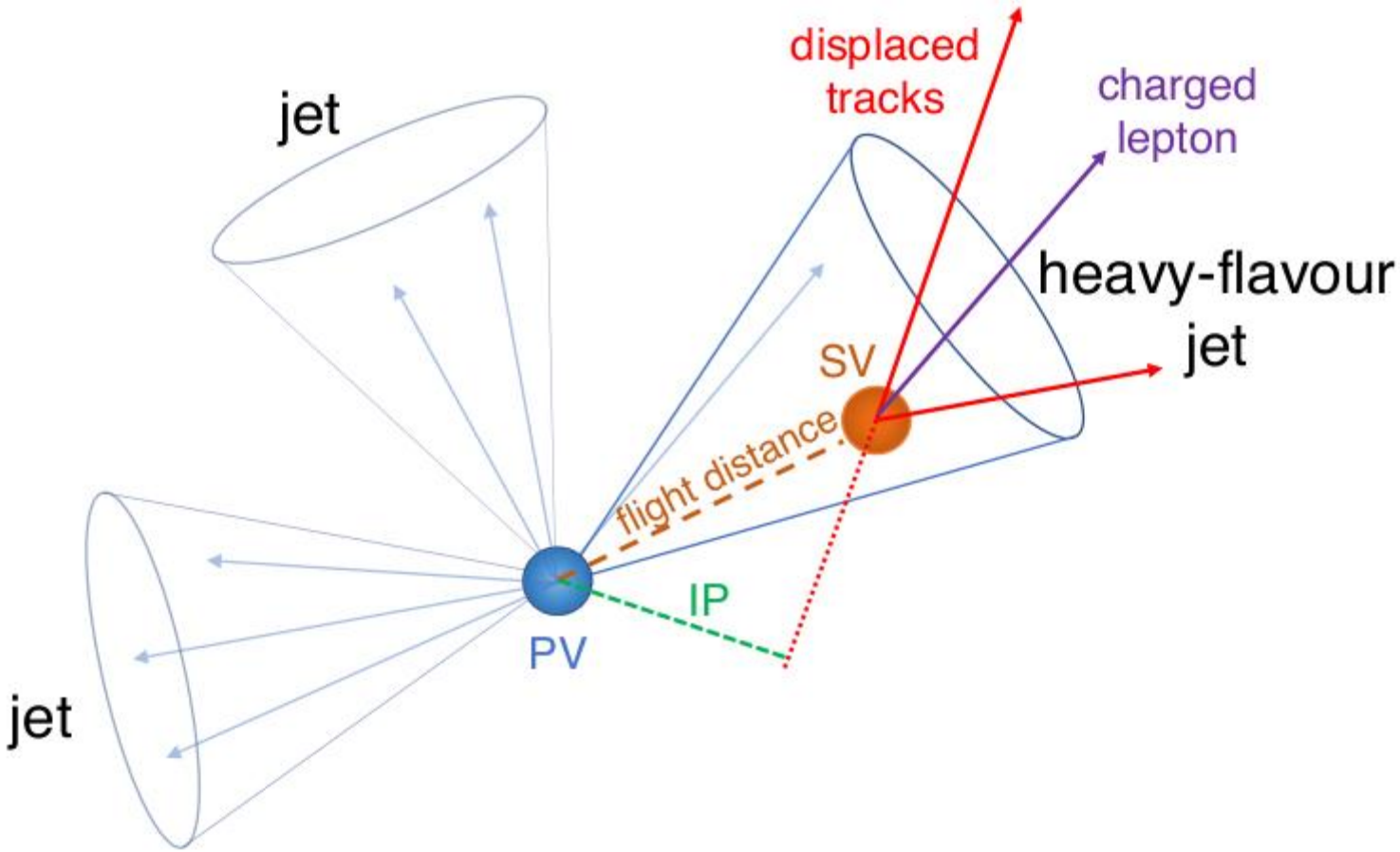
M_T^R is like MET

R then is the ratio of the invisible and visible energy in the event.

Main observables

How good discriminators are the Razor variables?





Backup – B tagging