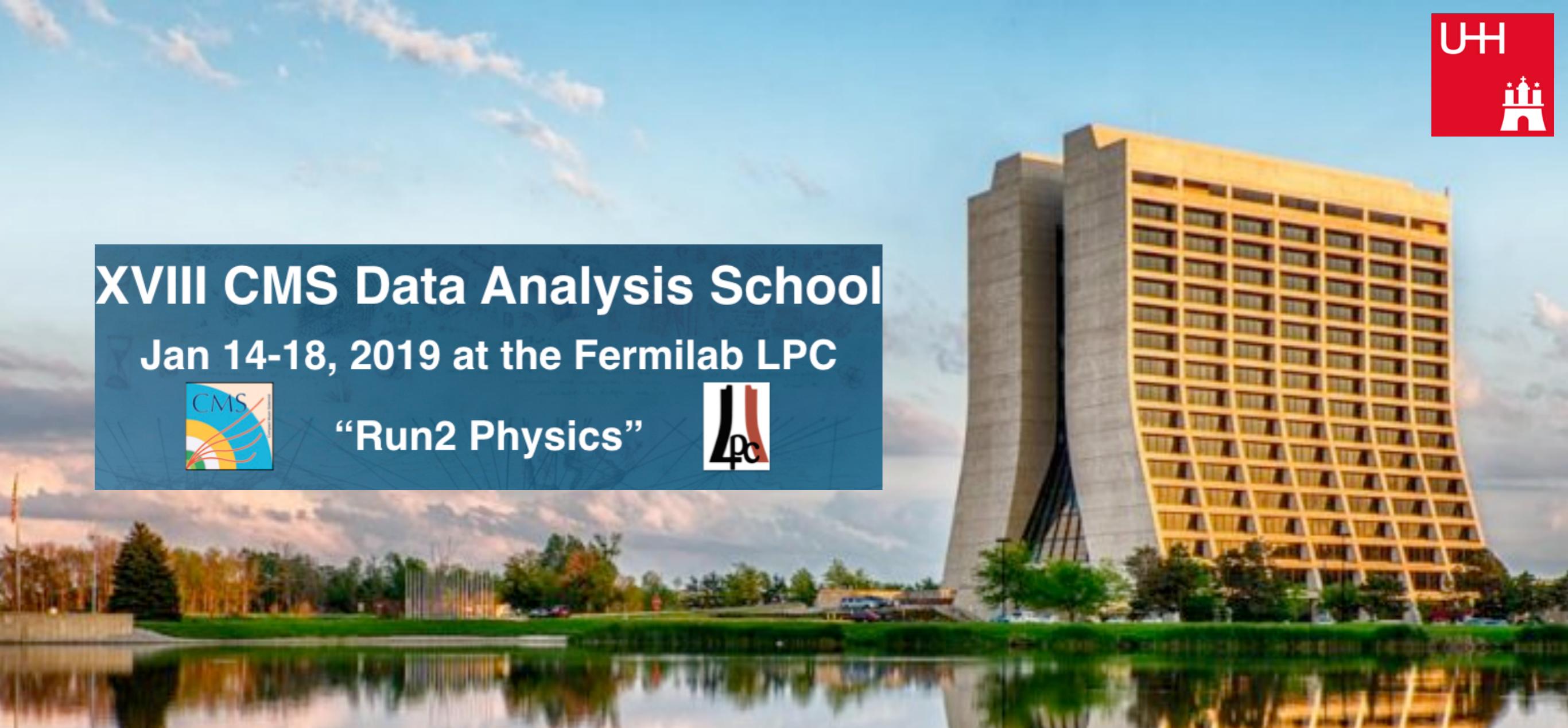


# XVIII CMS Data Analysis School

Jan 14-18, 2019 at the Fermilab LPC



“Run2 Physics”



## Long Exercise: Search for SUSY with disappearing tracks

Jan 16 2019

Sam Bein, Viktor Kutzner, Sezen Sekmen, Loukas Gouskos, Colin Fallon, Caleb Smith, Zhenbin Wu

## 1. Introduction to tracking

## 2. Search for disappearing tracks

1. track tag
2. signal region selection
3. prompt background estimation
4. non-prompt background estimation
5. limit estimation

### XVIII CMS Data Analysis School

Jan 14-18, 2019 at the Fermilab LPC



“Run2 Physics”



A school designed to teach CMS members how to perform data analyses with the CMS analysis software.

Hands-on exercises will cover all physics objects, trigger, visualization, statistics and participants will engage in full-fledged physics analyses with CMS data collected during LHC Run2, ranging from standard model measurements to searches for physics beyond the standard model.

<https://indico.cern.ch/e/cmsdas2019>

#### Organizing Committee:

Gabriele Benelli (Brown)  
Cecilia Gerber (UIC)  
Sergo Jindariani (Fermilab)  
Marguerite Tonje (UIC)



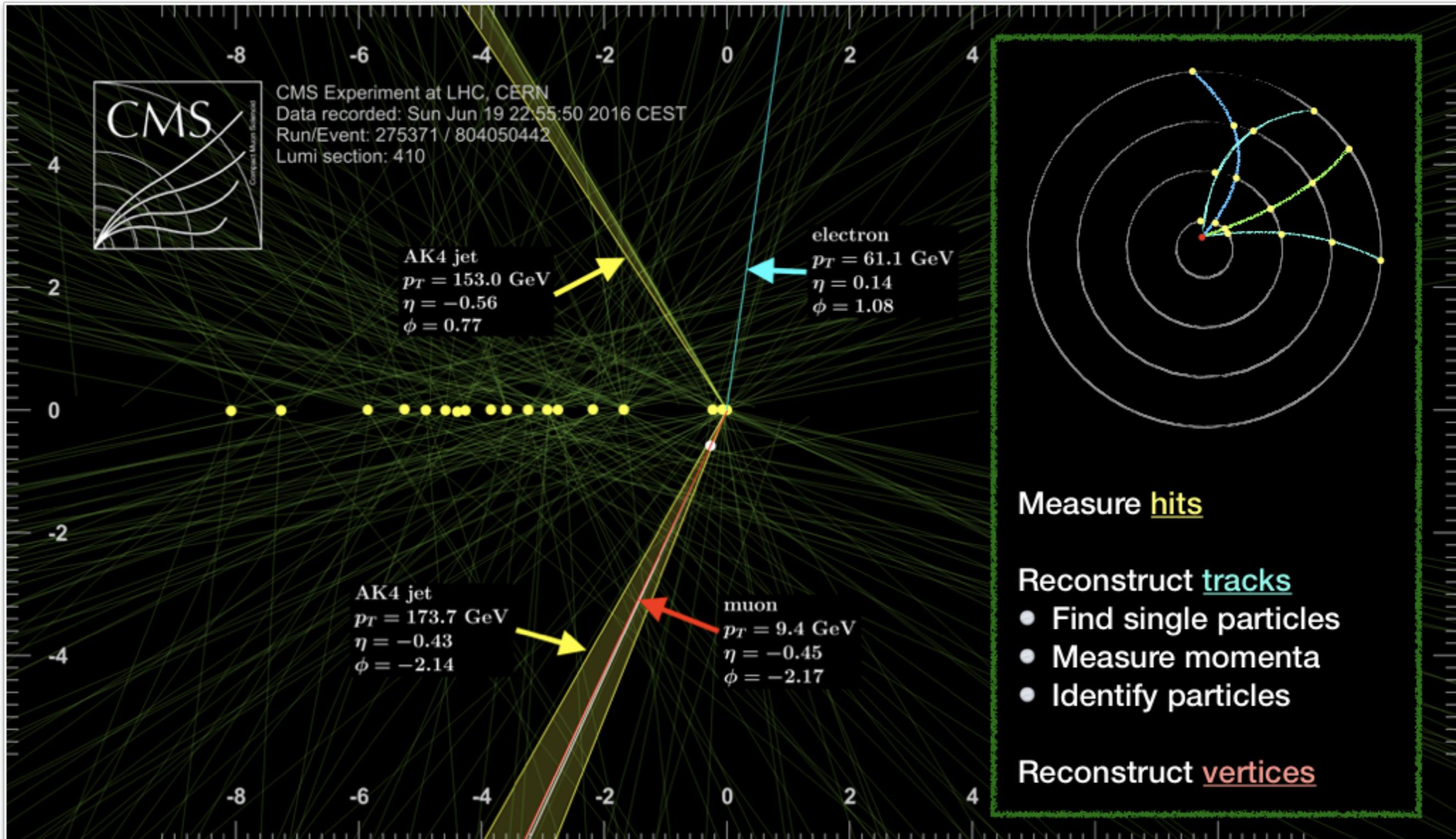
CMS Schools Committee:  
Lothar Bauerdick (Fermilab)  
Jack Chen (NTU)  
Nicola De Filippis (Bari, co-Chair)  
Elisabetta Gallo (DESY)  
Cecilia Gerber (UIC)  
Sudhir Malik (UPRM, co-Chair)  
Martijn Mulders (CERN)  
Fabrizio Palla (Pisa)  
Gigi Rolandi (Pisa)  
Phat Srimanobhas (Chulalongkorn)

#### Administrative Support:

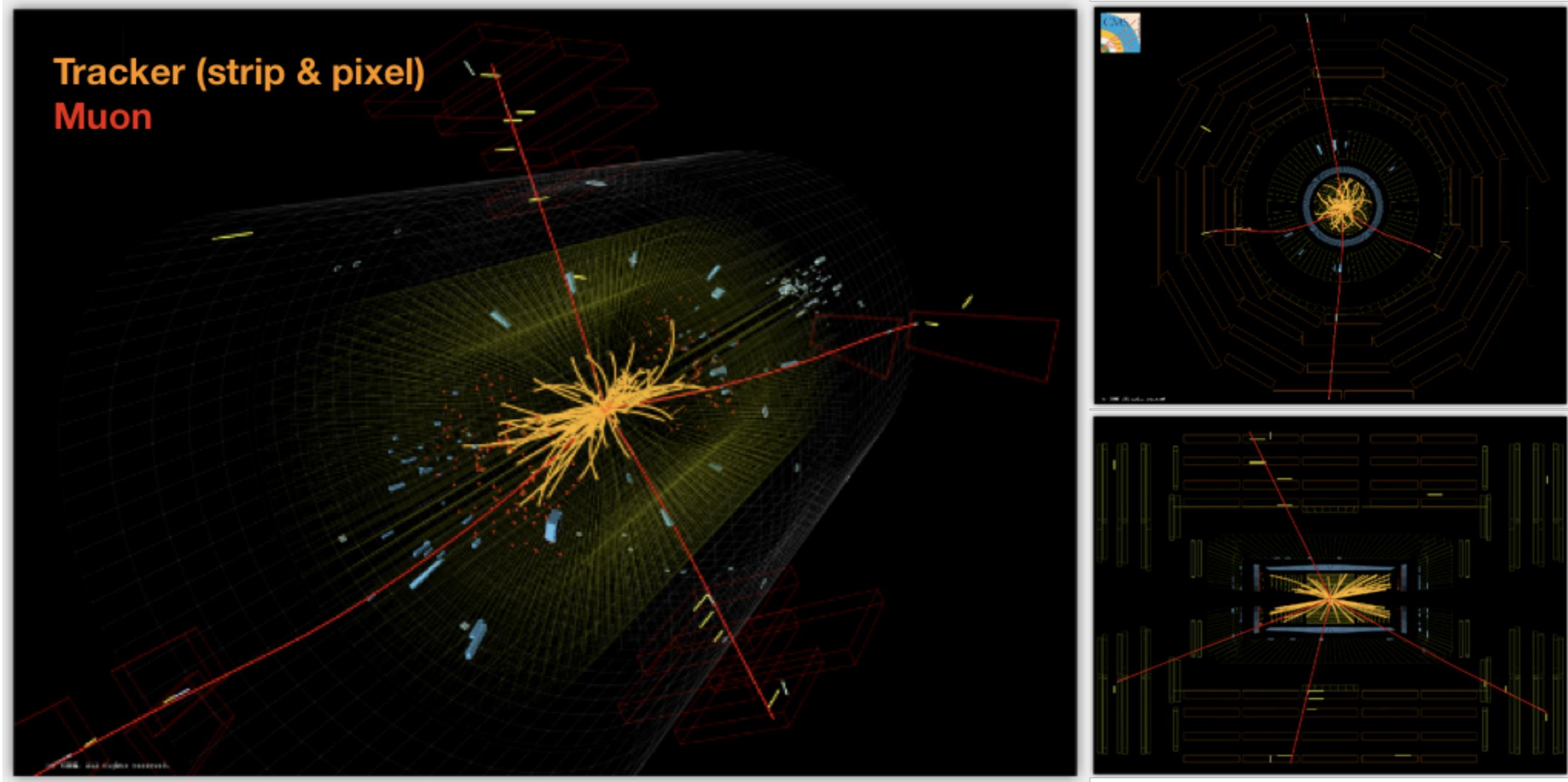
Carrie Farver, Terry Grozis, Terry Read, Sonya Wright



# Tracking at CMS



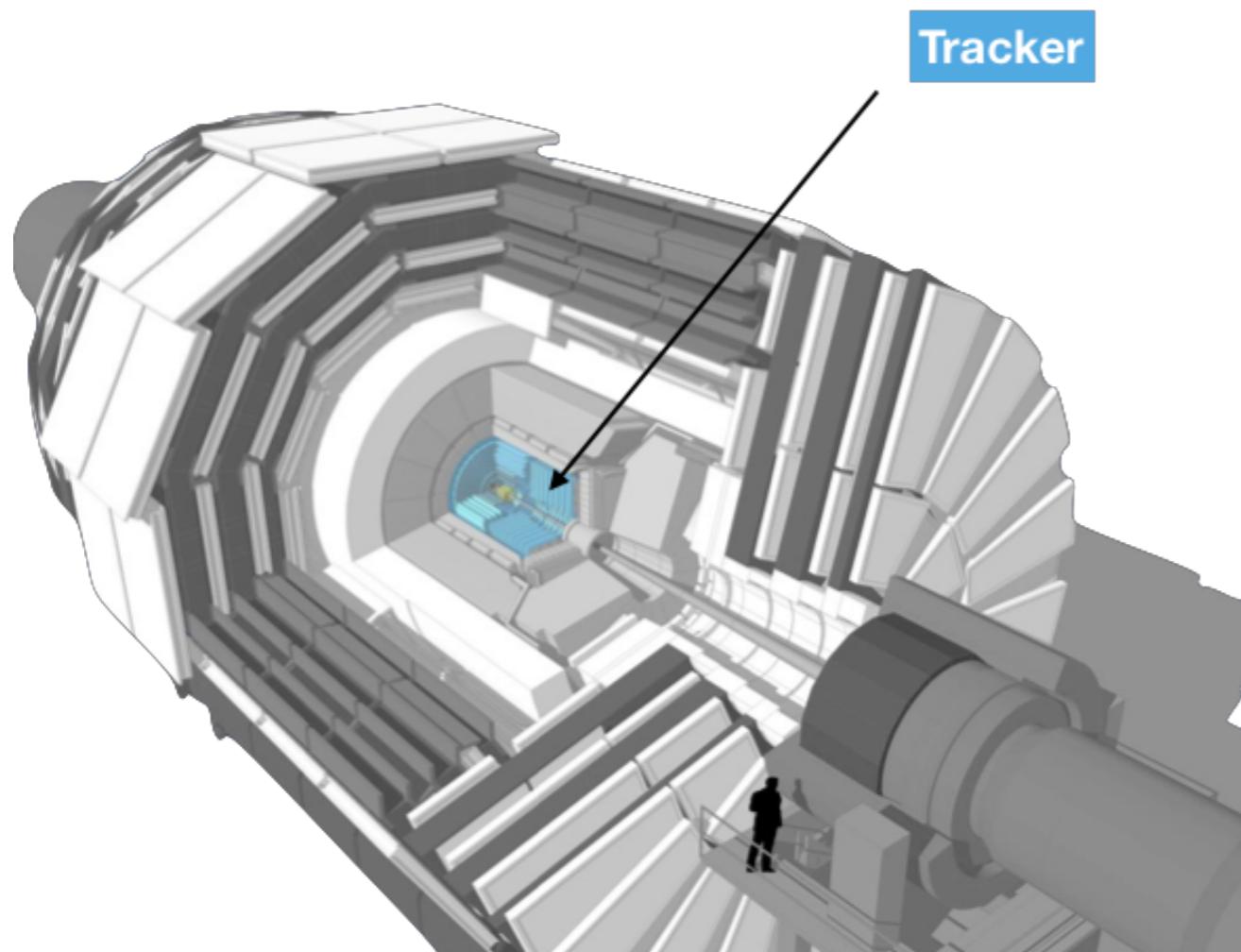
# Tracking components



# CMS tracker

Position information from finely segmented silicon sensors:

- record the path of charged particles
- measure momentum from bending radius in a magnetic field
- reconstruct primary and secondary vertices



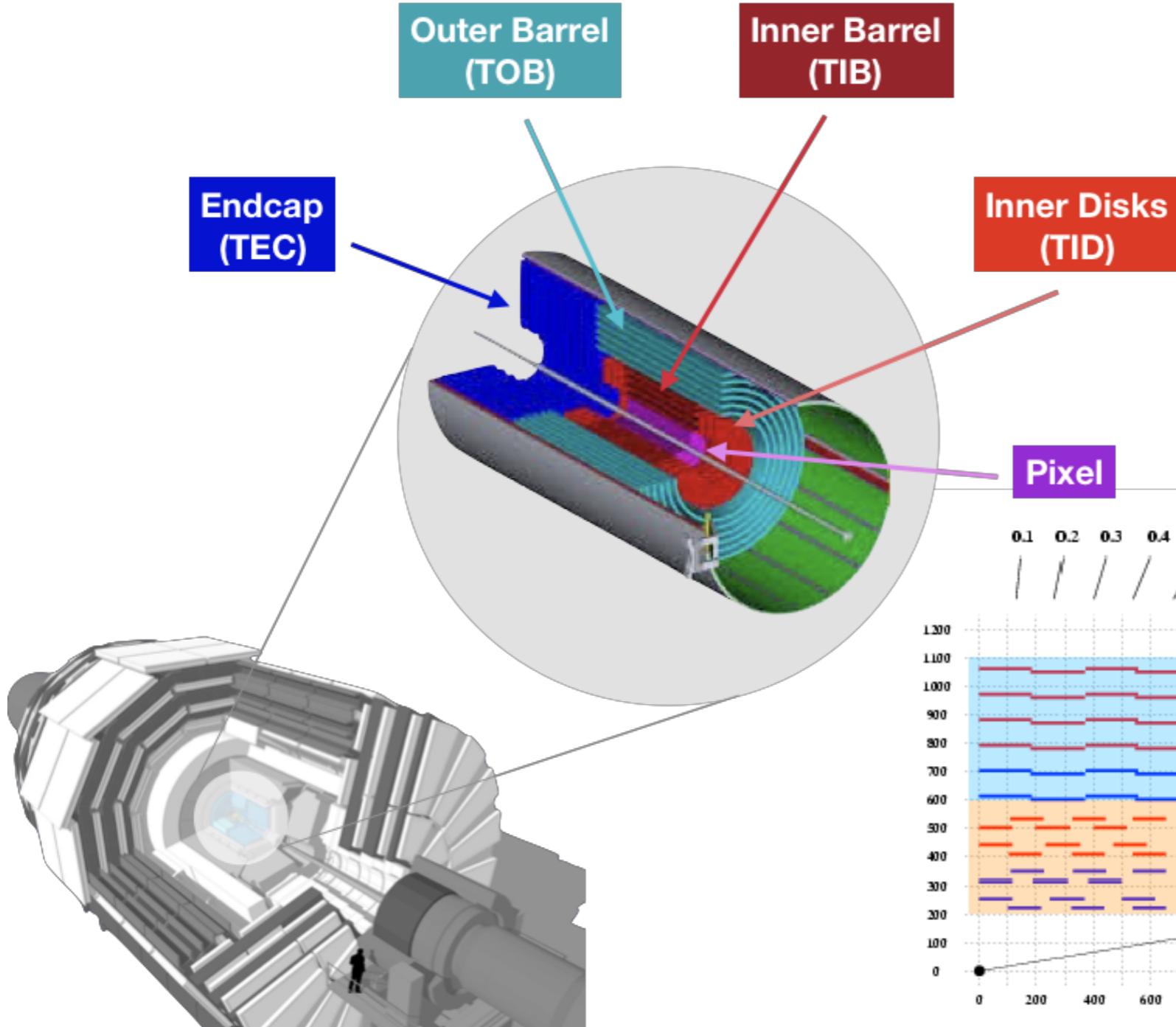
Requirements:

- High resolution & low occupancy: Resolve and isolate individual tracks, reconstruct vertices. Finer granularity is needed closer to the IP: High particle density, small tracking volume
- High rate capability: Fast charge collection time and read-out electronics to keep up with the expected event rates
- Low material budget: Minimal multiple scattering
- Radiation hardness: Innermost subdetectors, receive highest particle fluence

# CMS tracker: Silicon strips

10 million strips

8 - 14 hits per track



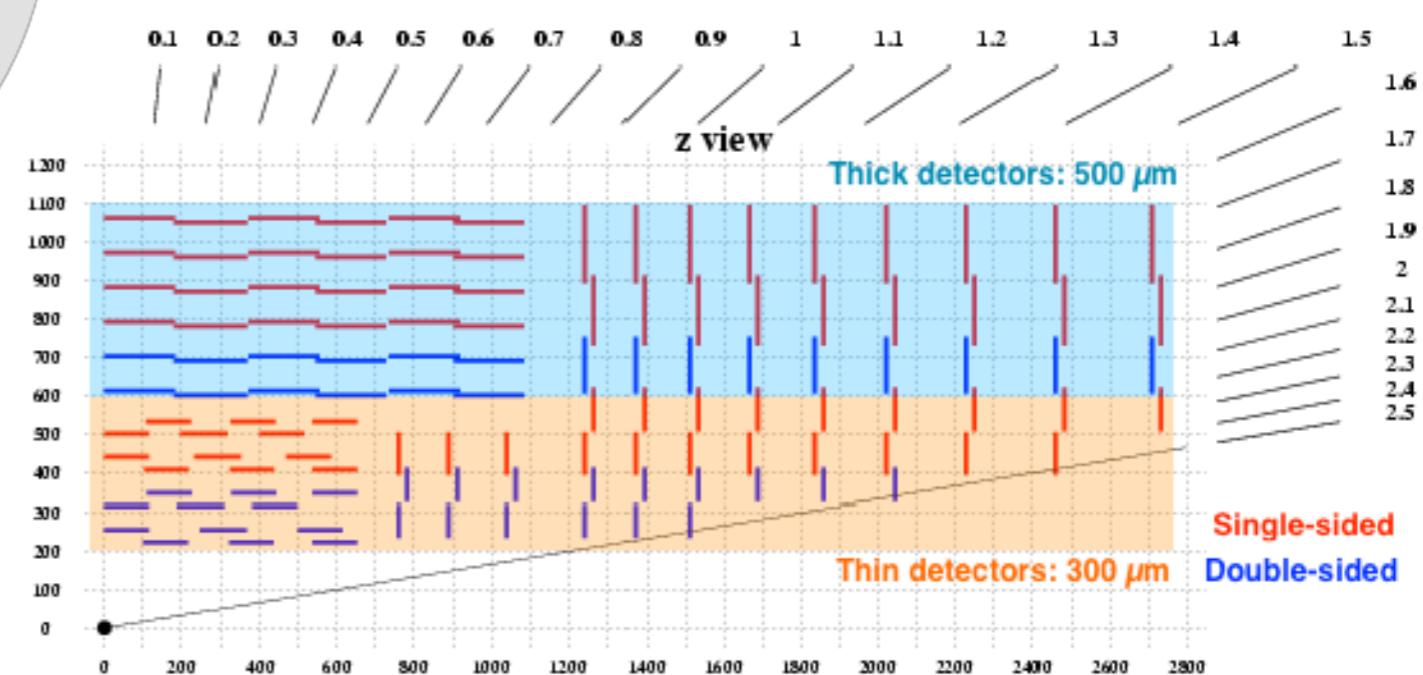
Layers:

- inner Barrel (TIB): 4
- inner Disks (TID): 3
- outer Barrel (TOB): 6
- endcap (TEC): 9

Many layers: Redundancy

Occupancy: 1 - 3%

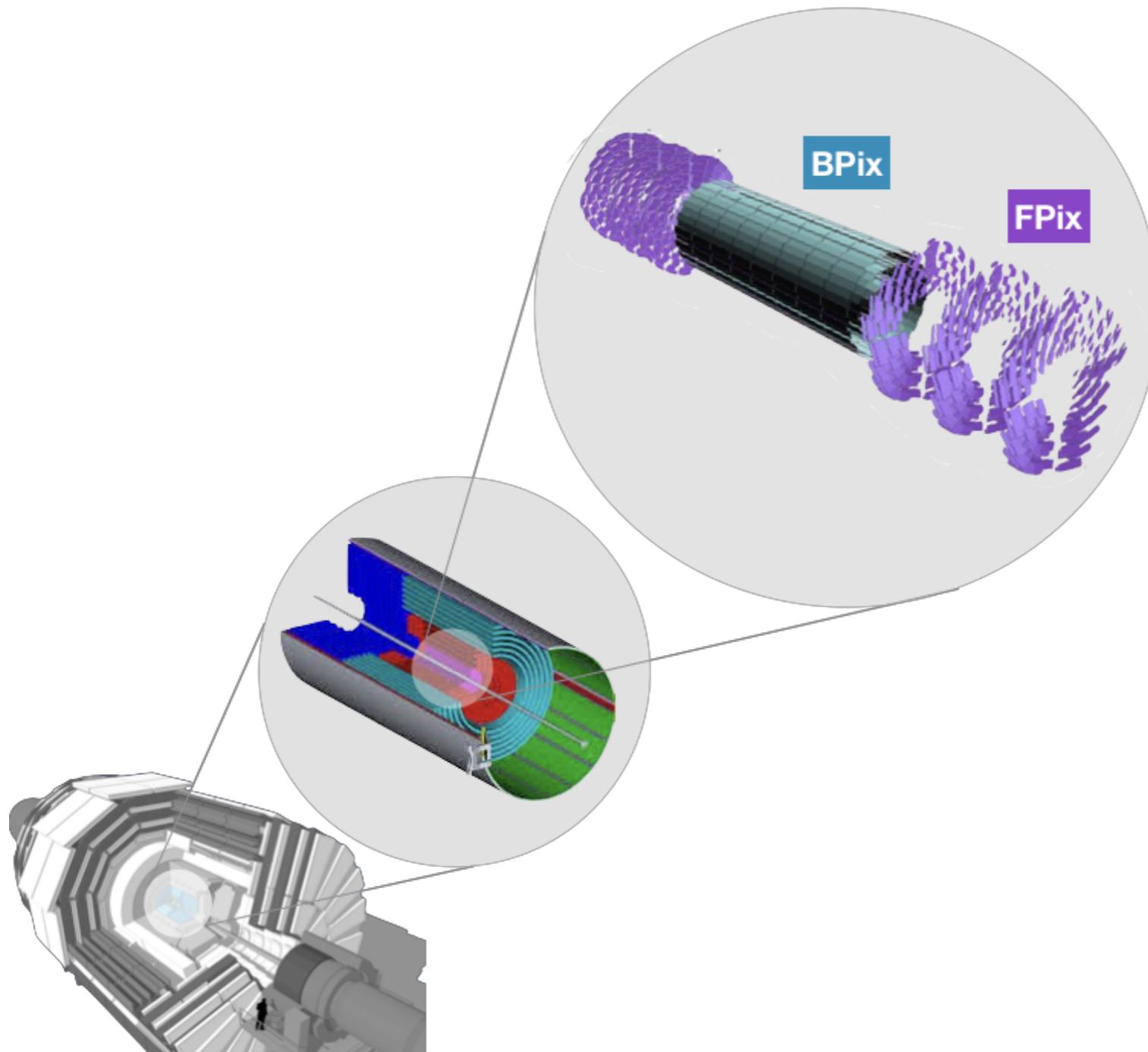
Resolve and isolate tracks



# CMS tracker: Silicon pixel

127 million pixels,  $100 \times 150 \mu\text{m}^2$  in size

4-hit coverage up to  $|\eta| = 2.5$



Layer position [cm]:

- BPix: 2.9, 6.8, 10.9, 16.0
- FPix: 29.1, 39.6, 51.6

Low occupancy:  $10^{-3}$

High segmentation:

- High quality seeds for offline tracking ( $>95\%$  efficiency)

Fast online tracking (HLT):

- vertex reconstruction
- $e / \gamma$  identification
- muon reconstruction
- tau identification
- b-tagging

# CMS tracker: pixel upgrade

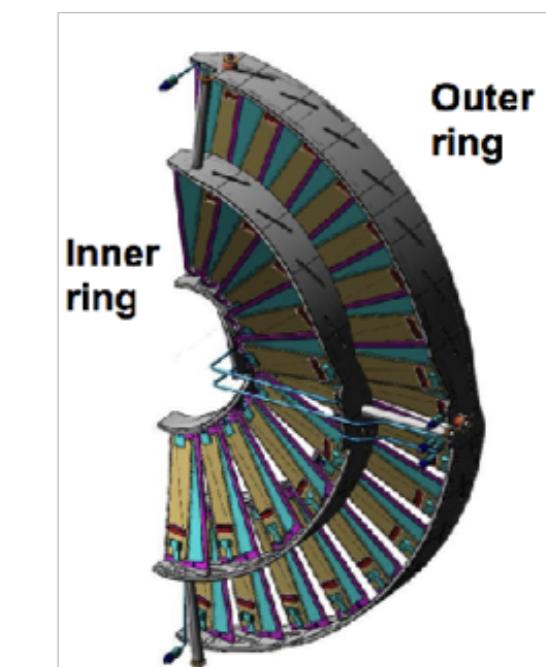
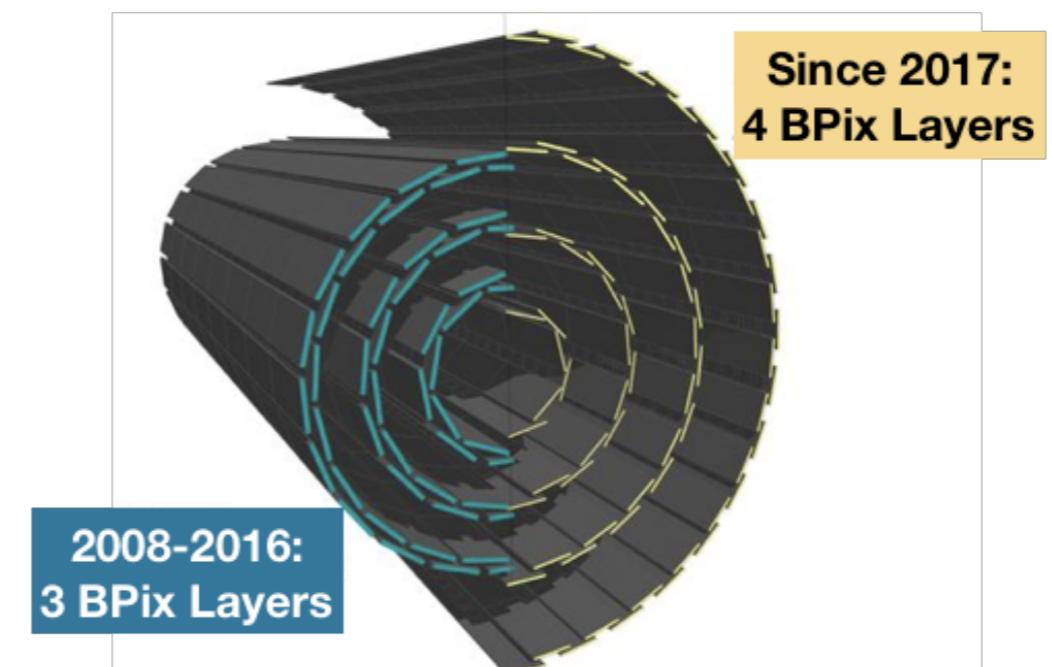
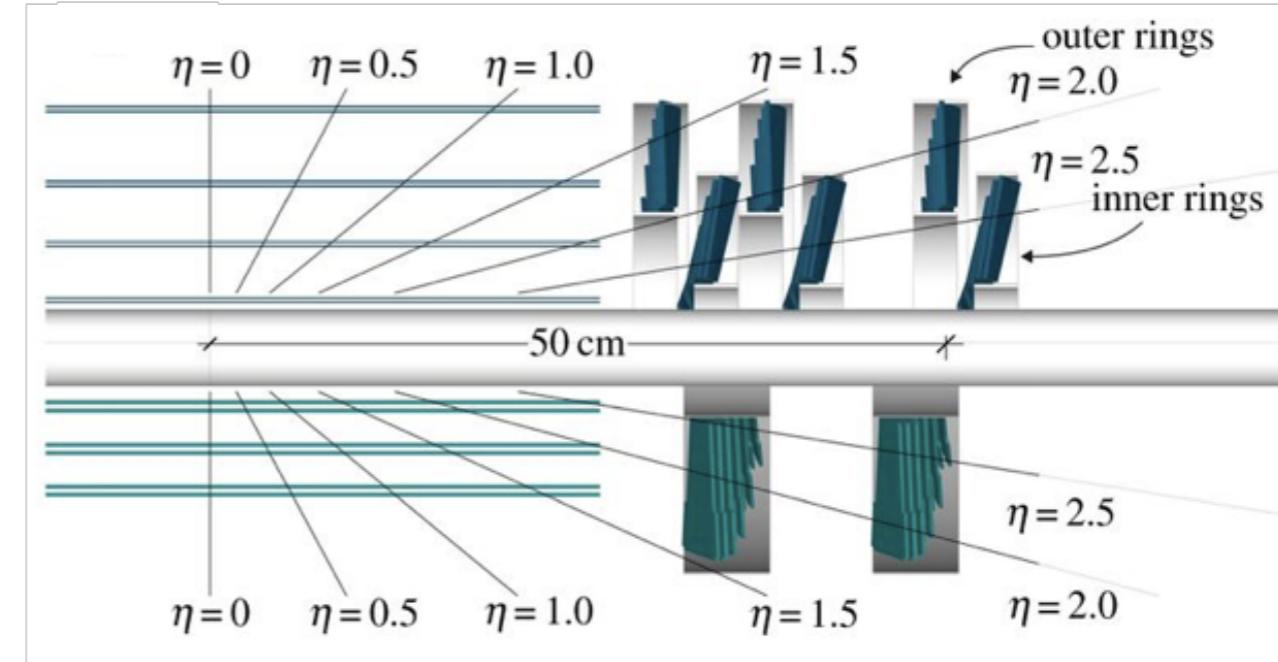
one additional tracking point,  
in both barrel and forward  
regions:  
 ⇒ 4-hit seeds  
 ⇒ lower fake rate!

smaller radius of the innermost  
pixel layer  
 ⇒ closer to the interaction  
region  
 ⇒ Improved track  
reconstruction

reduced material budget  
 ⇒ reduce multiple scattering

**Since 2017:**  
**4 BPix Layers**  
**3 FPix Layers**  
**with 2 Disks**

**2008-2016:**  
**3 BPix Layers**  
**2 FPix Layers**

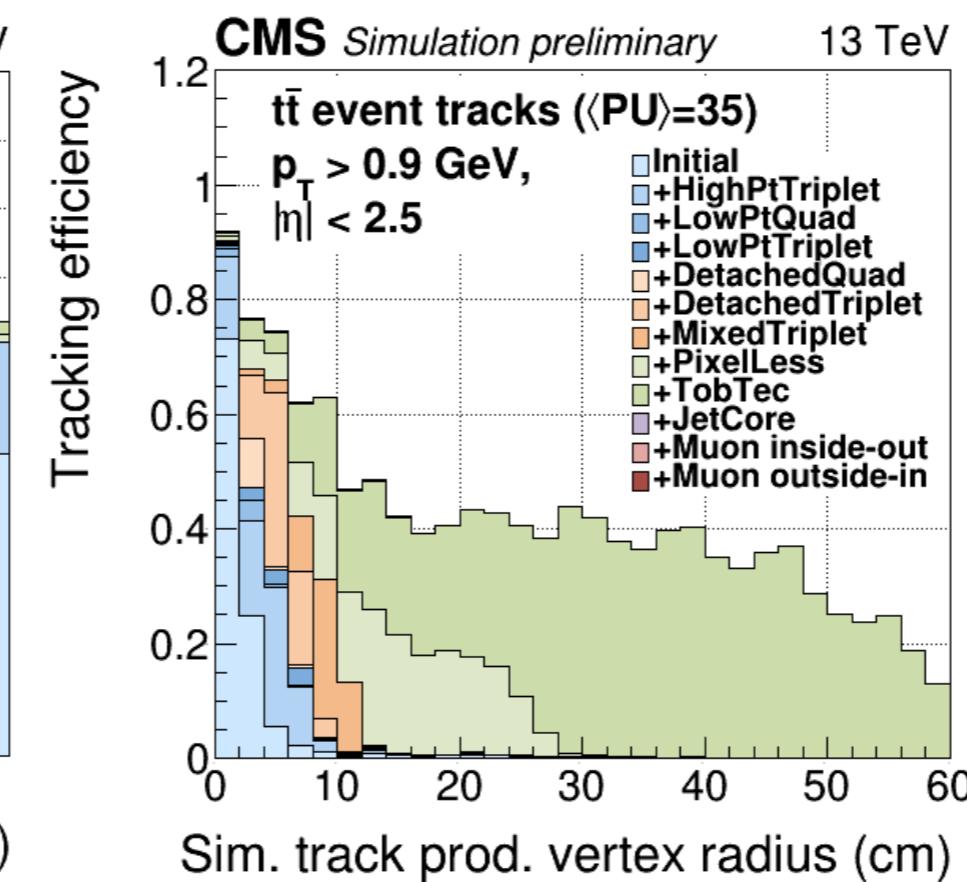
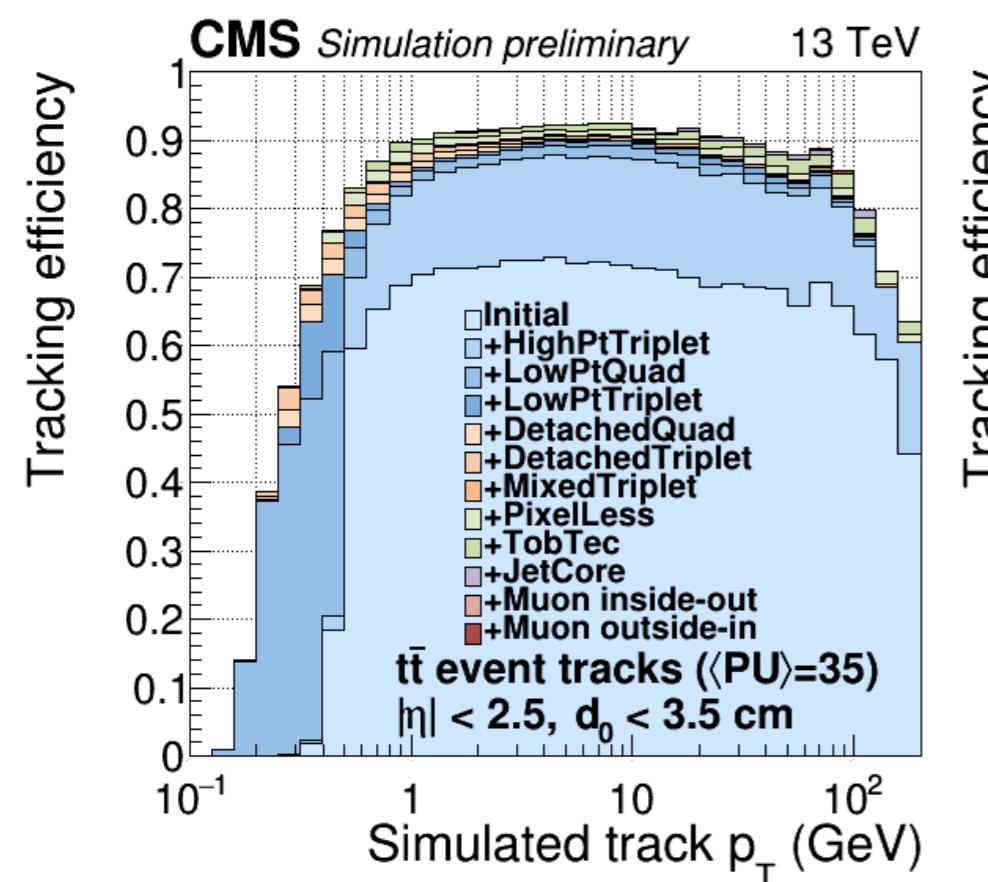


# Iterative tracking

In CMS, tracks reconstruction is an iterative procedure: High-quality tracks are reconstructed, their hits are removed, and other tracks are reconstructed from the remaining hits.

In the InitialStep, high- $p_T$  quadruplets coming from the beam spot region are used.

Subsequent steps use triplets, or improve the acceptance either in  $p_T$  or displacement.

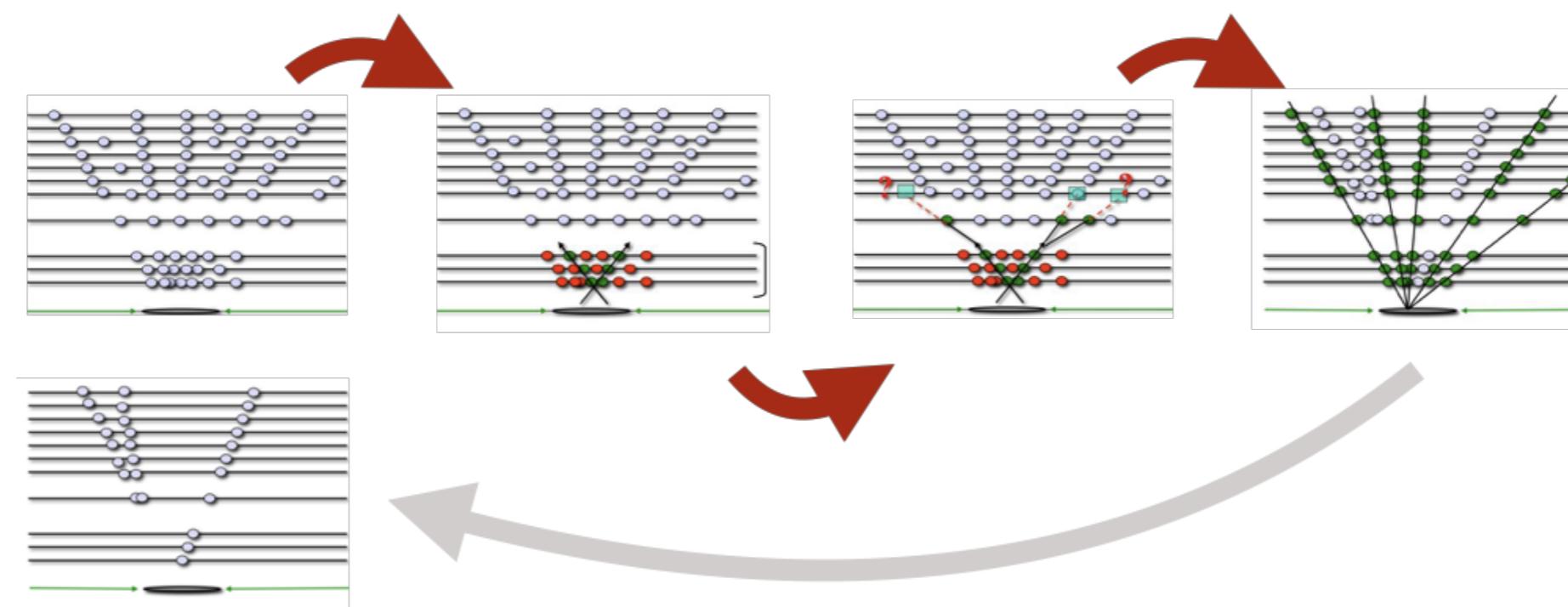


The later steps use seeds with hits from the strip detector to find detached tracks, e.g. from LLPs.

# Tracking at CMS

In each iteration, tracks are reconstructed in four steps:

1. **Seeding:** Provides track candidates, with an initial estimate of the trajectory parameters and their uncertainties.
2. **Pattern recognition:** Track candidates are propagated to find new compatible hits. Track parameters are updated.
3. **Final fitting:** Provides the best estimate of the parameters of each smooth trajectory, after combining all associated hits.
4. **Selection:** Sets quality flags based on the fit  $\chi^2$  and the track compatibility with interaction region. Aims to reject fake tracks.



# Tracking at CMS

First part of the exercise:

- extract basic track parameters
- reconstruct invariant masses from tracks

## 1. Introduction to tracking

## 2. Search for disappearing tracks

1. track tag
2. signal region selection
3. prompt background estimation
4. non-prompt background estimation
5. limit estimation

### XVIII CMS Data Analysis School

Jan 14-18, 2019 at the Fermilab LPC



“Run2 Physics”



A school designed to teach CMS members how to perform data analyses with the CMS analysis software.

Hands-on exercises will cover all physics objects, trigger, visualization, statistics and participants will engage in full-fledged physics analyses with CMS data collected during LHC Run2, ranging from standard model measurements to searches for physics beyond the standard model.

<https://indico.cern.ch/e/cmsdas2019>

#### Organizing Committee:

Gabriele Benelli (Brown)  
Cecilia Gerber (UIC)  
Sergo Jindariani (Fermilab)  
Marguerite Tonje (UIC)



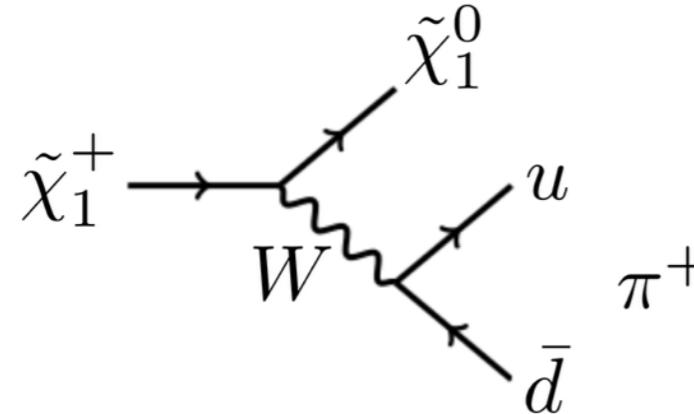
CMS Schools Committee:  
Lothar Bauerdick (Fermilab)  
Jack Chen (NTU)  
Nicola De Filippis (Bari, co-Chair)  
Elisabetta Gallo (DESY)  
Cecilia Gerber (UIC)  
Sudhir Malik (UPRM, co-Chair)  
Martijn Mulders (CERN)  
Fabrizio Palla (Pisa)  
Gigi Rolandi (Pisa)  
Phat Srimanobhas (Chulalongkorn)

#### Administrative Support:

Carrie Farver, Terry Grozis, Terry Read, Sonya Wright



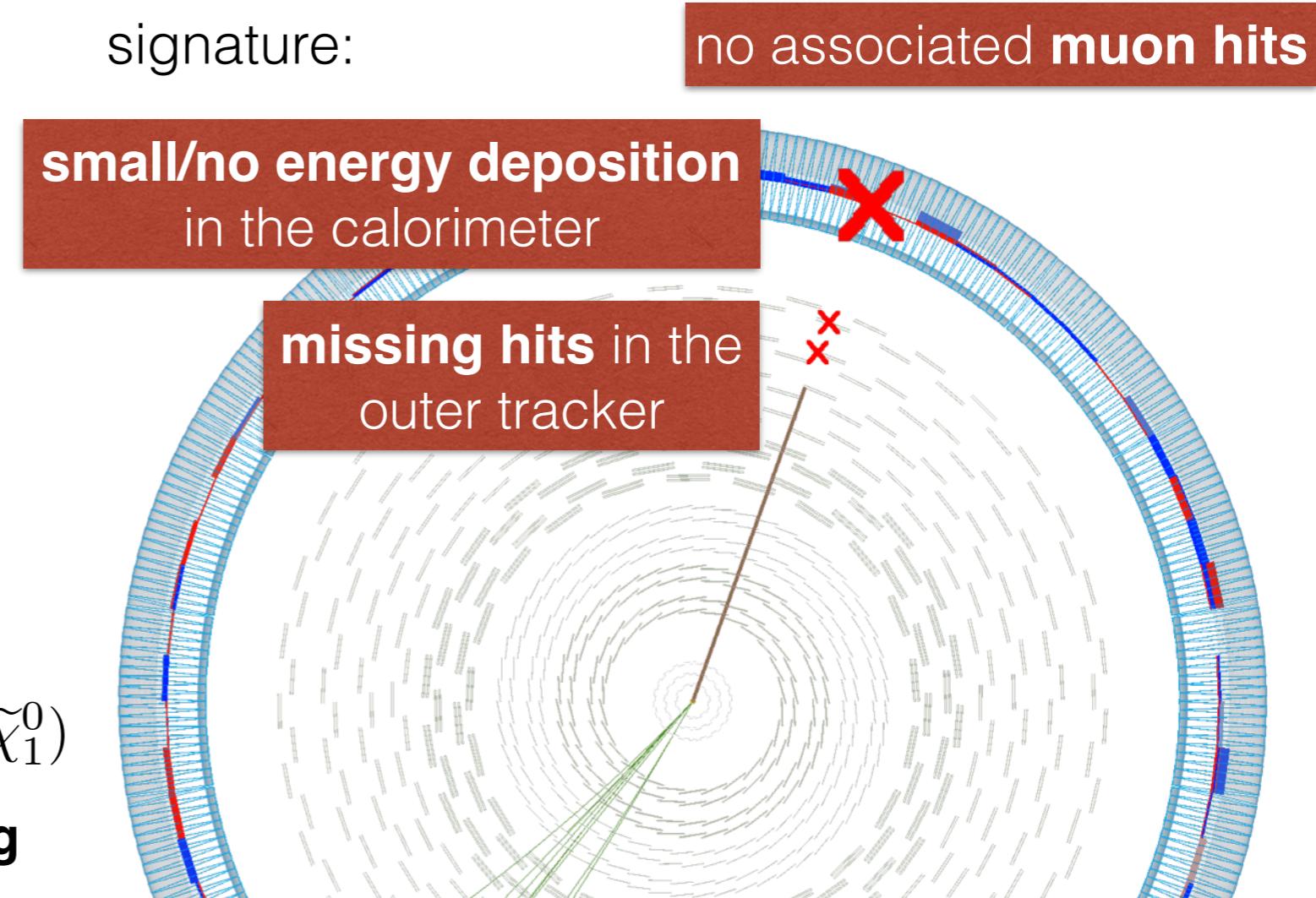
# Search strategy



- pion: too soft to be reconstructed

mass splitting:  $\Delta m = m(\tilde{\chi}_1^\pm) - m(\tilde{\chi}_1^0)$

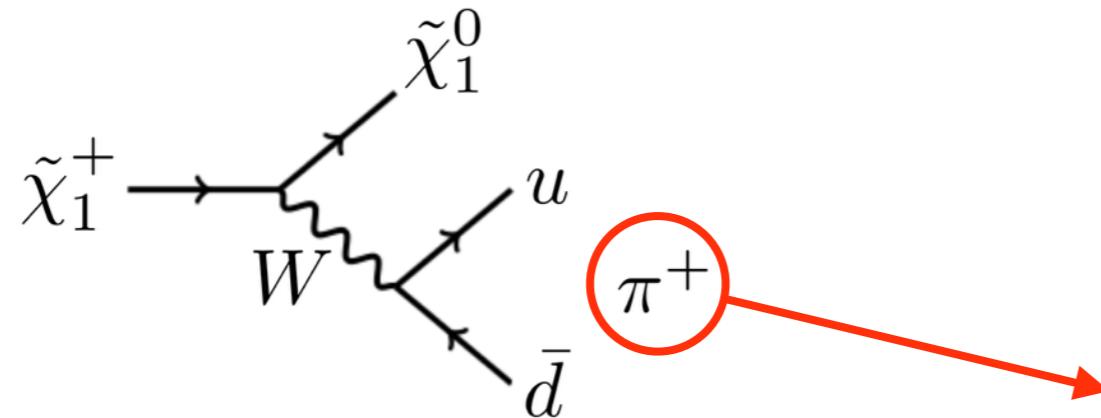
- $m_\pi < \Delta m \lesssim 200$  MeV: **disappearing track visible in tracker**
- $\Delta m \gtrsim 200$  MeV: decay before first tracking layer



need to suppress background from

- fake tracks
- tracks from failed lepton reconstruction

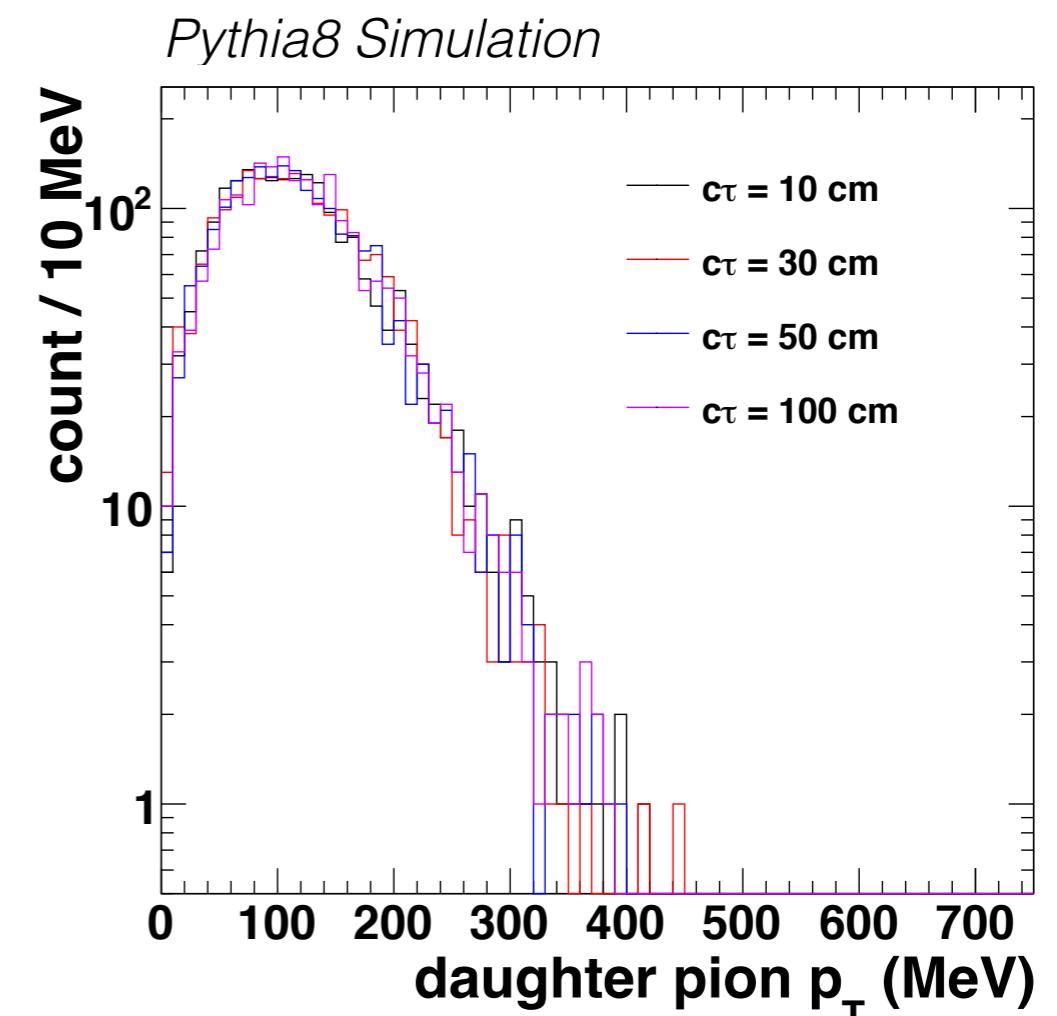
# Search strategy



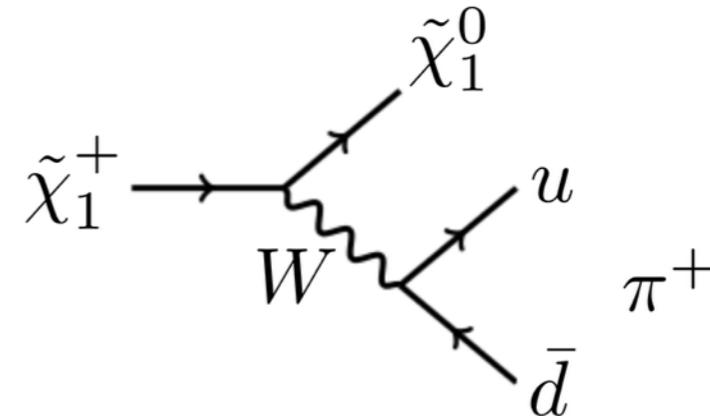
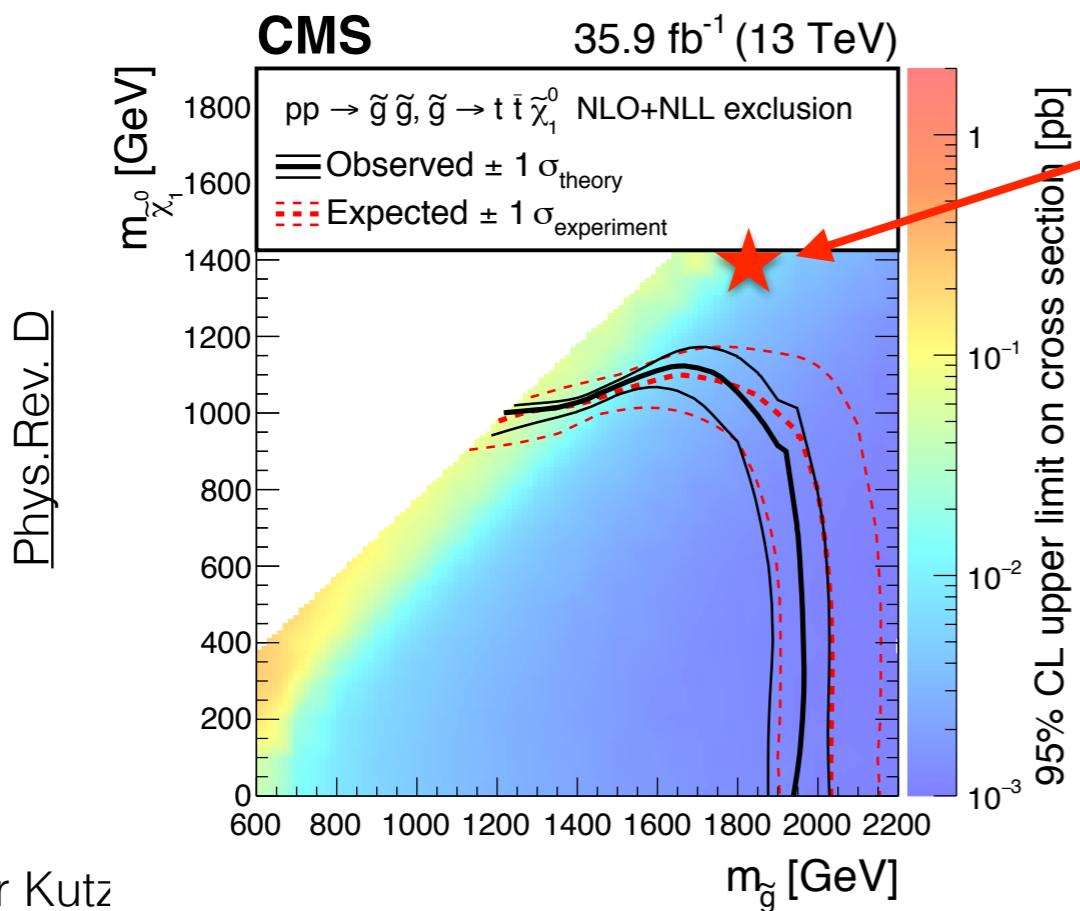
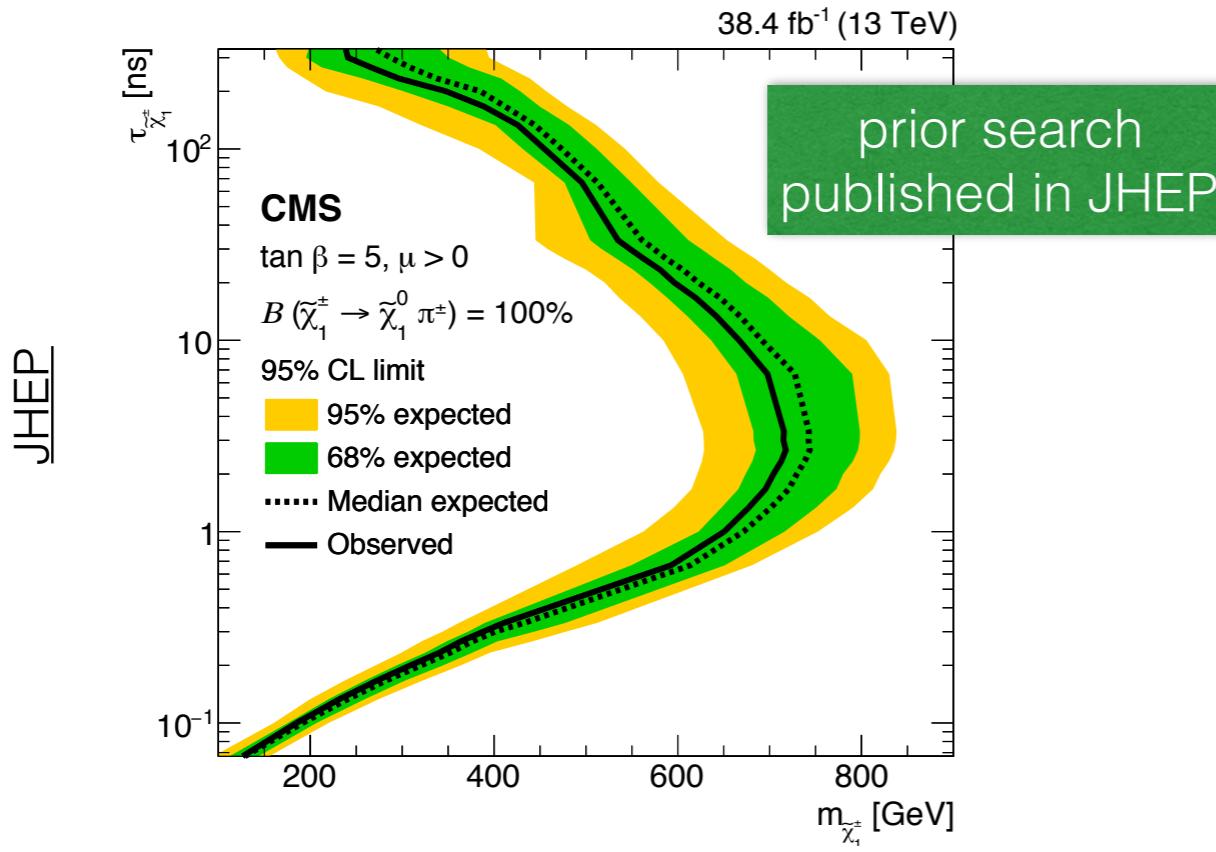
- pion: too soft to be reconstructed

mass splitting:  $\Delta m = m(\tilde{\chi}_1^\pm) - m(\tilde{\chi}_1^0)$

- $m_\pi < \Delta m \lesssim 200$  MeV: **disappearing track visible in tracker**
- $\Delta m \gtrsim 200$  MeV: decay before first tracking layer



# Disappearing track signal



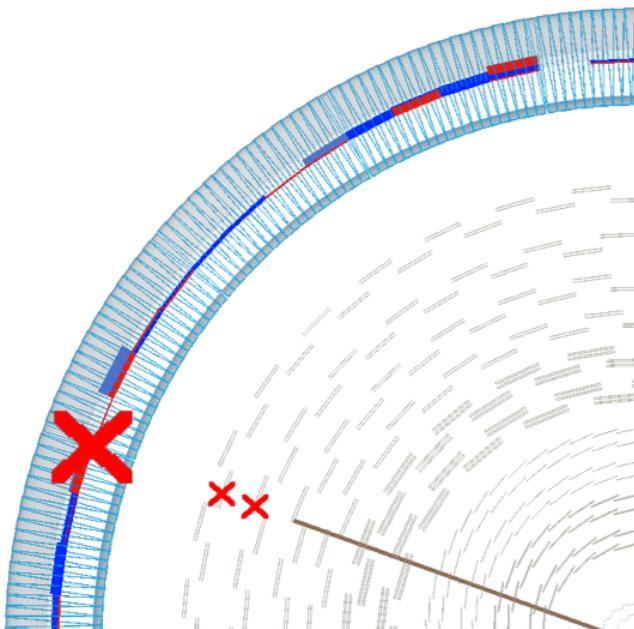
**signal generation** with Pythia8:  
strong chargino production through  
 $gg \rightarrow \tilde{g} \tilde{g}$

- $m(\tilde{g}) = 1800 \text{ GeV}$
- $m(\text{chargino}) = 1400 \text{ GeV}$
- $\Delta m = 180 \text{ MeV}$
- $c\tau = 10, 30, 50, 100 \text{ cm}$

include models from pMSSM scan for event selection

# Tagging disappearing tracks

## relevant track properties:

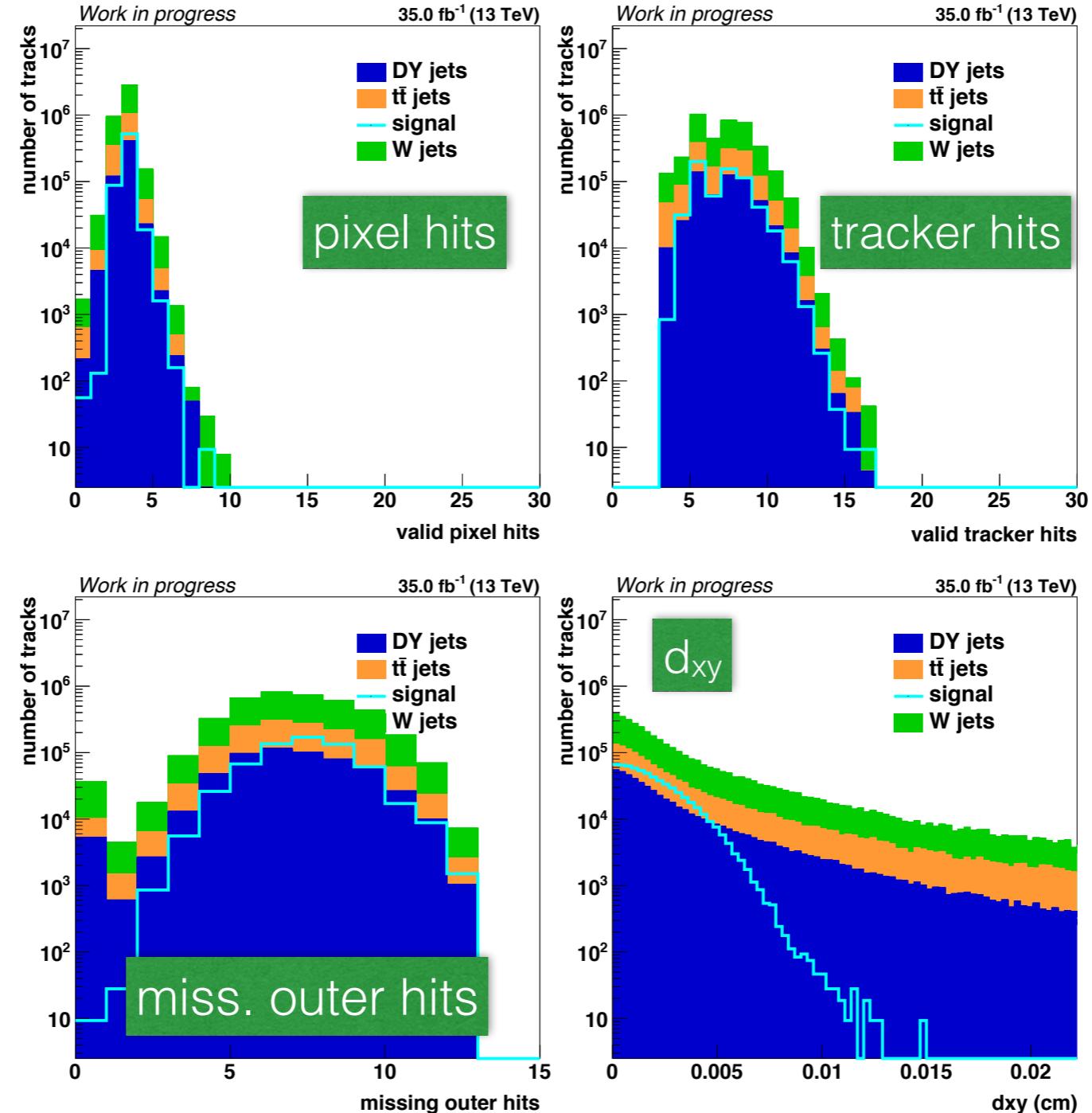


- number of pixel / tracker hits
- missing inner / middle / outer hits
- distance to primary vertex ( $d_{xy}$ ,  $d_z$ )
- $\chi^2 / \text{ndof}$
- $\Delta p_T / p_T^2$
- high purity track quality
- deposited energy in calorimeter within cone of  $\Delta R < 0.5$
- avoid **model dependence** by too tight cuts on  $p_T$ ,  $\eta$ , or isolation

# (Some) tracking variables

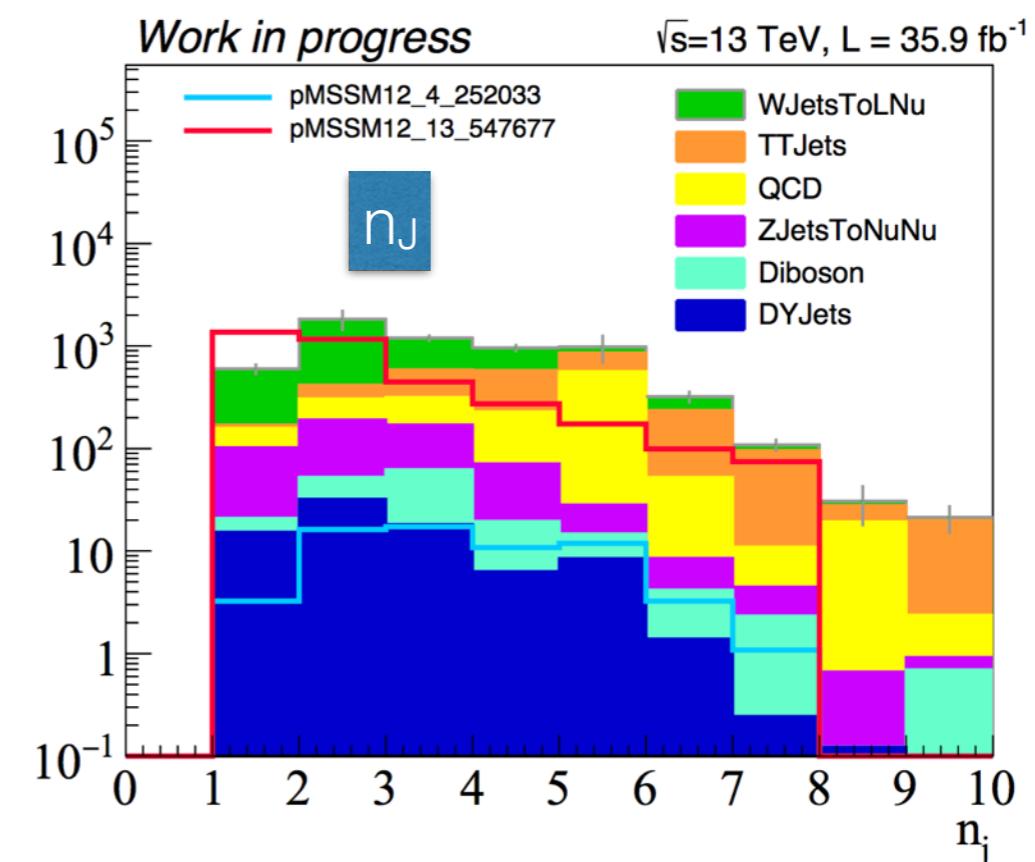
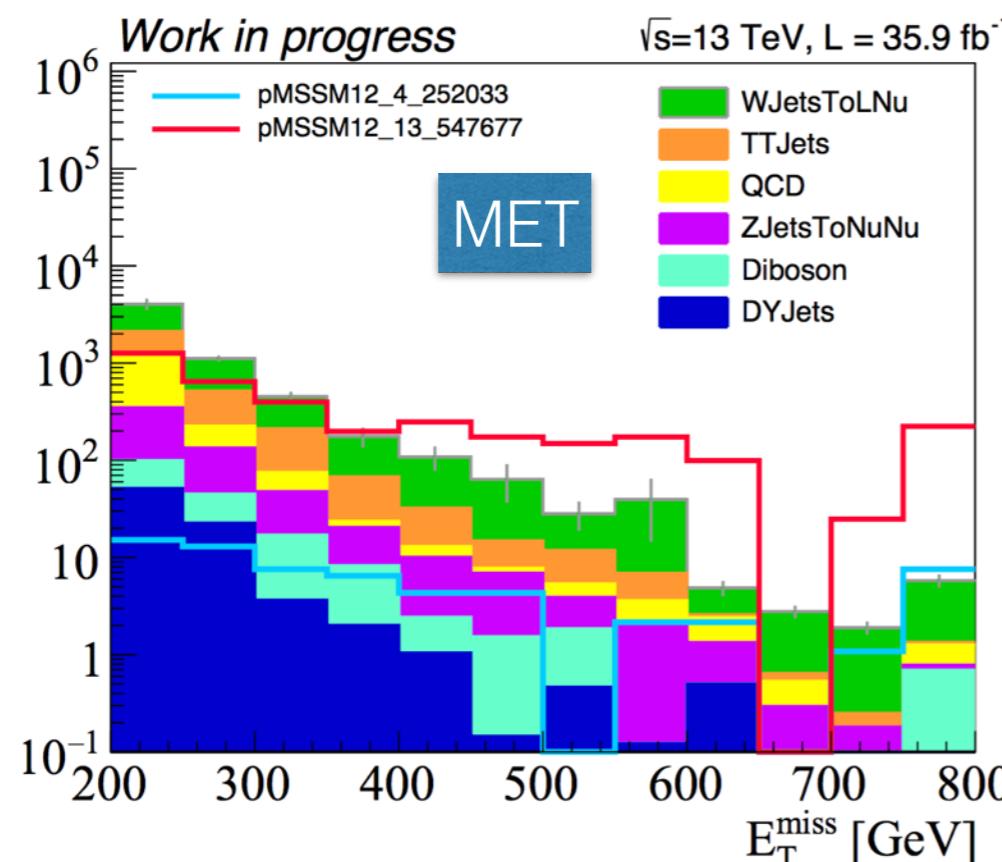
track variables after applying a basic preselection including  
**a reconstructed lepton veto & no missing inner hits:**

medium tracks,  
 signal scaled to 100 pb



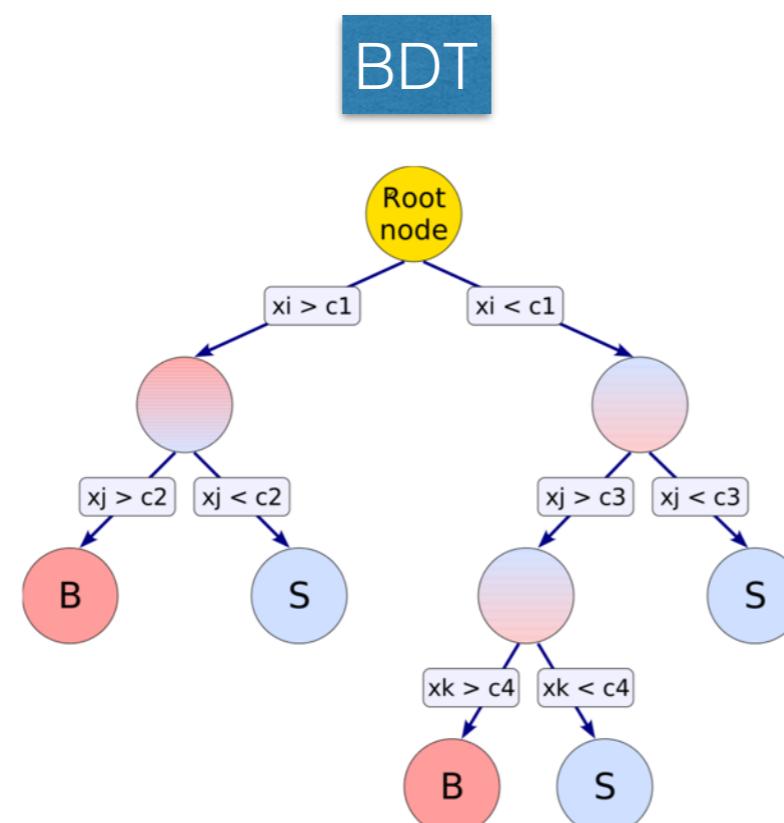
# Event topology

- use **disappearing track tag** for event selection
- look at **event topology** variables



# Optimization methods

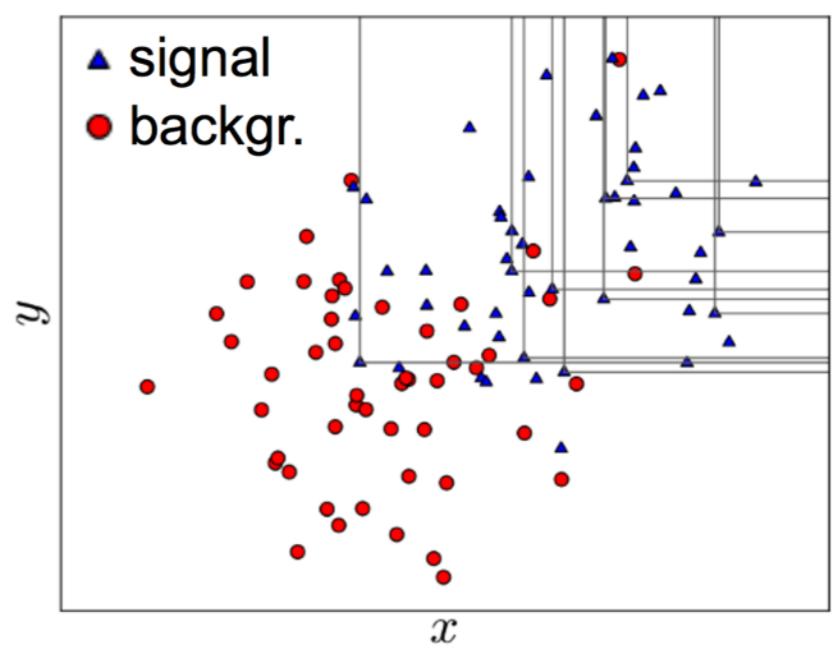
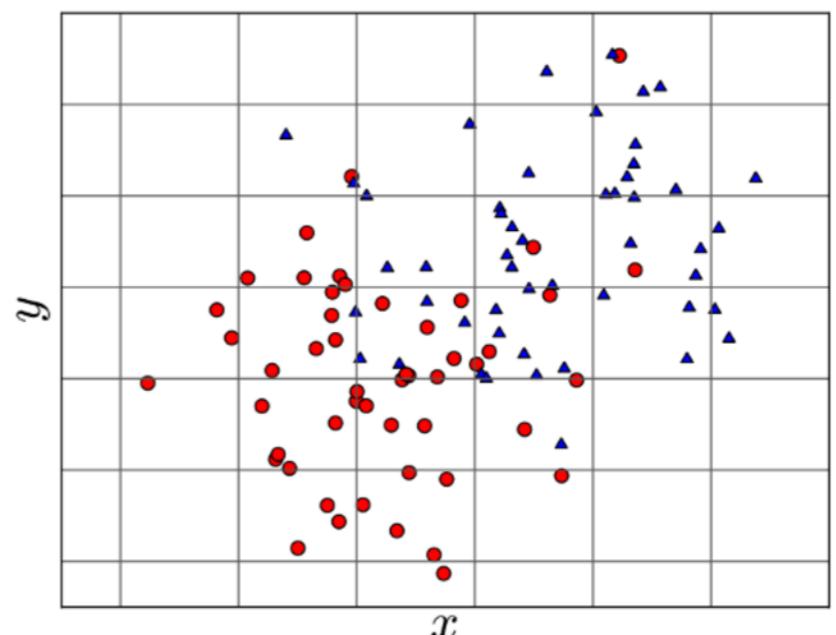
- **random grid search (RGS)**: find optimal cuts via importance sampling, focus on signal-like regions rather than entire phase space
- **boosted decision tree (BDT)** using TMVA



both methods:

select best cuts by maximizing significance  $s/\sqrt{s + b}$

random grid search



BDT illustration by Böser, Fink, Röcker (KIT)

# Today's exercise

## **today:**

- introduction to tracking in CMS
- construct disappearing track tag using a BDT
- define signal regions with RGS for event selection

## **tomorrow:**

- background estimation methods for
  - prompt background
  - fake track background
- systematic uncertainties, derivation of expected limit



# Today's exercise

- start here: <https://github.com/ShortTrackSusy/cmsdas>

The screenshot shows the GitHub repository page for 'LongLivedSusy / cmsdas'. The repository has 118 commits, 1 branch, 0 releases, and 1 contributor. The latest commit was 2816a21, 2 days ago. The repository description is: 'No description, website, or topics provided.' The README.md file contains the following content:

## CMSDAS @ FNAL 2019

Welcome to the 2019 FNAL CMSDAS exercise on disappearing tracks! This long exercise will walk students through a number of steps needed to set up and implement a search for new physics at CMS. Enjoy :)

If you're doing teh exercise at the school, please send an email to me so I can sign you up for Mattermost ([samuel.bein@gmail.com](mailto:samuel.bein@gmail.com))

<https://mattermost.web.cern.ch/cmsdaslpc2019/channels/shorttrackteam>

Note about the samples: This exercise is built largely on pre-made ntuples, and is thus mostly independent of CMSSW. The code that generated the ntuples is contained in the repo: <https://github.com/longlivedsusy/treemaker>