Highlights from CMS experiment and prospects for Run3

Maurizio Pierini
CERN

European Research Council

This talk in a nutshell

- *With Run3 starting, we are still in the process of analysing Run2 data*
  - *There are interesting (but still small) discrepancies to watch for)*
  - *More could come*
- *Pursuing the search process is a must*
- *With x2 more data coming at ~ same energy, maximal gain from exploring (also) new directions*
- *Alternative data taking could be crucial, to extend our search to the data that we normally throw away*
  - *Scouting*
  - *Parking*
  - *Anomaly detection*
  - …

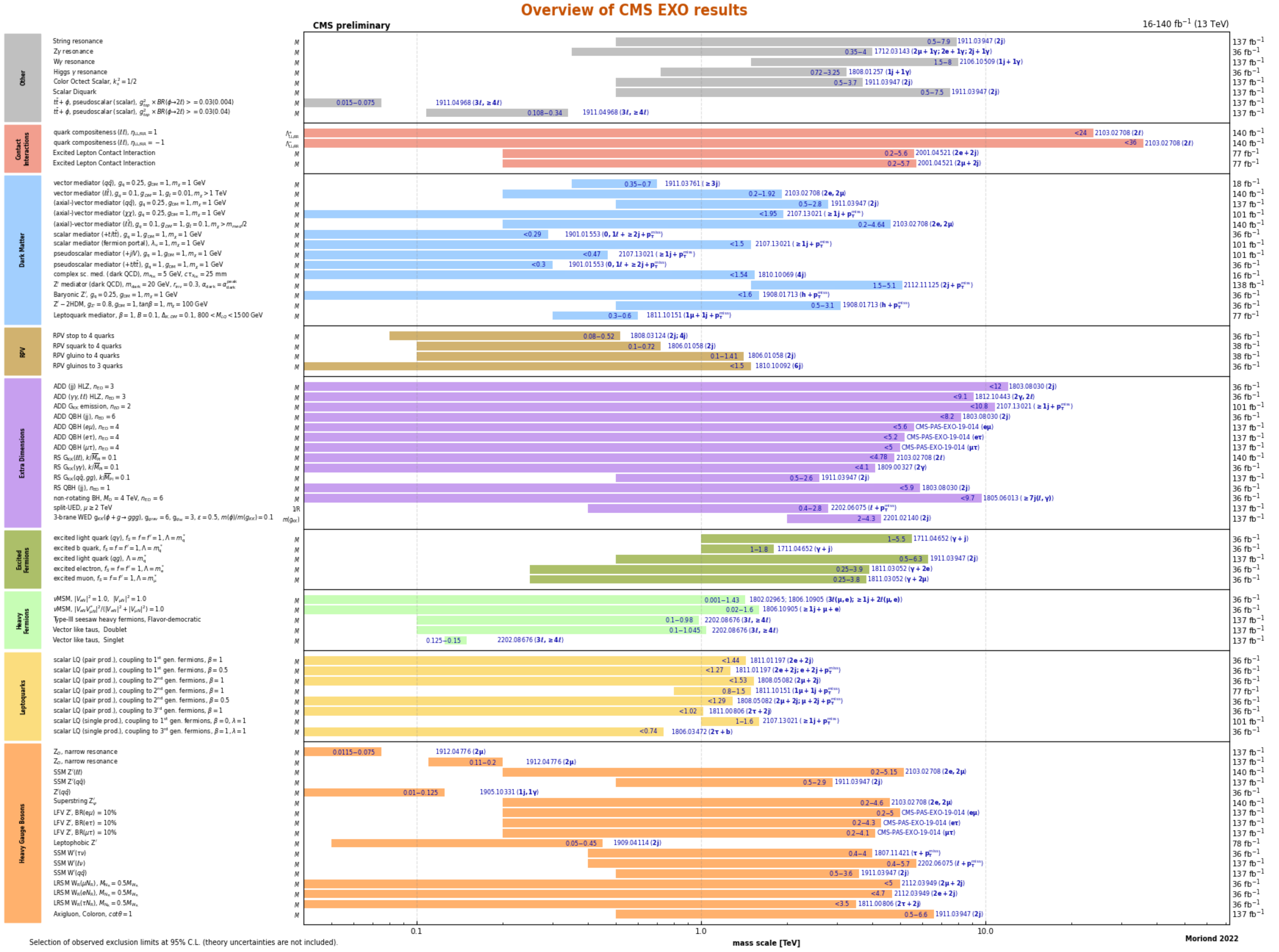
This talk in a nutshell

- With Run3 starting, we are still in the process of analysing Run2 data
  - There are interesting (but still small) discrepancies to watch for)
  - More could come
- Pursuing the search process is a must
- With ... lso) new ...
- Alte ... a that we normally throw away
  - Scouting
  - Parking
  - Anomaly detection
  - …

I am focusing on searches today, for lack of time
A lor more happening with Higgs, top, EW, QCD and Heavy Ions

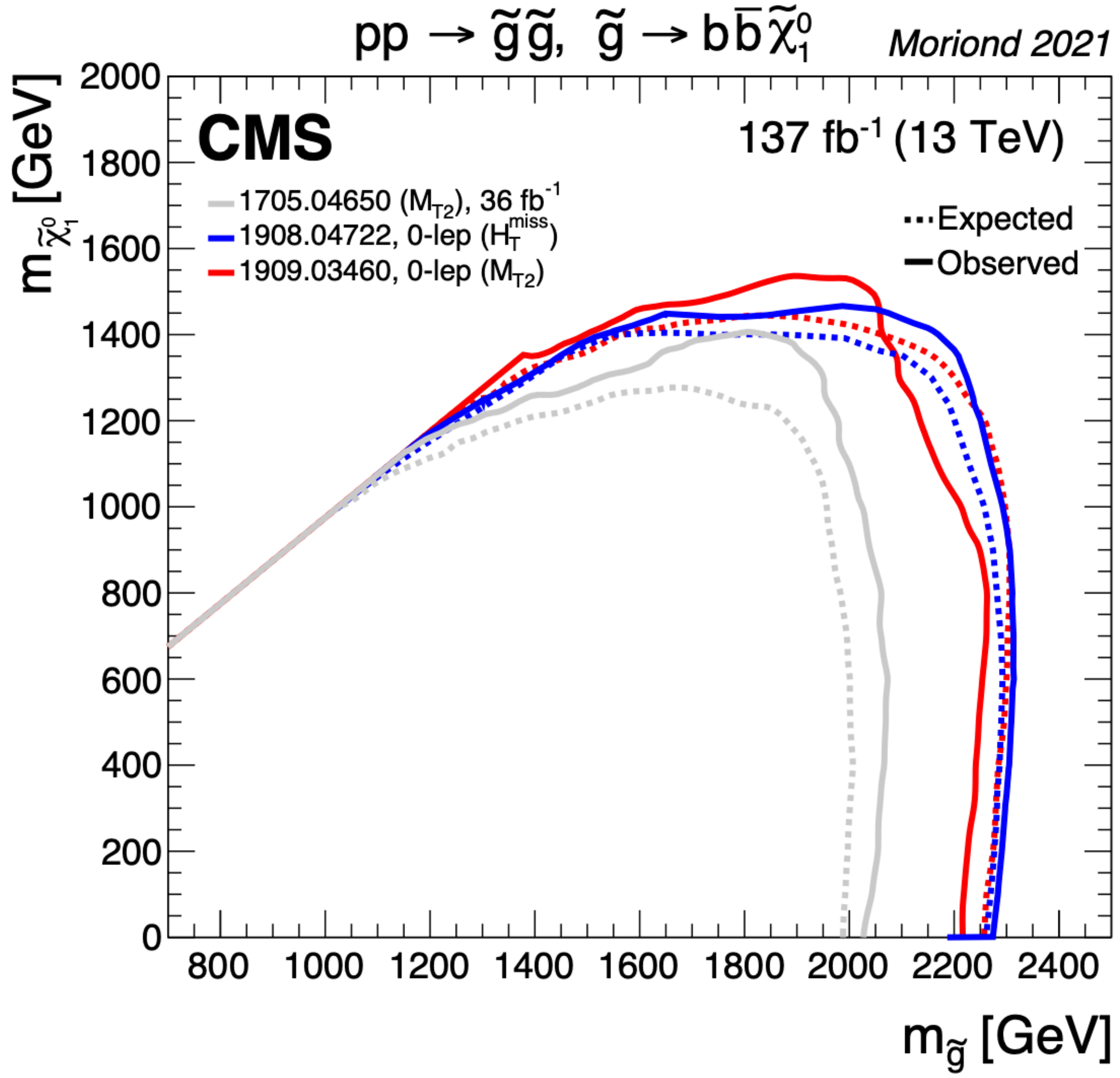
A little bit of history

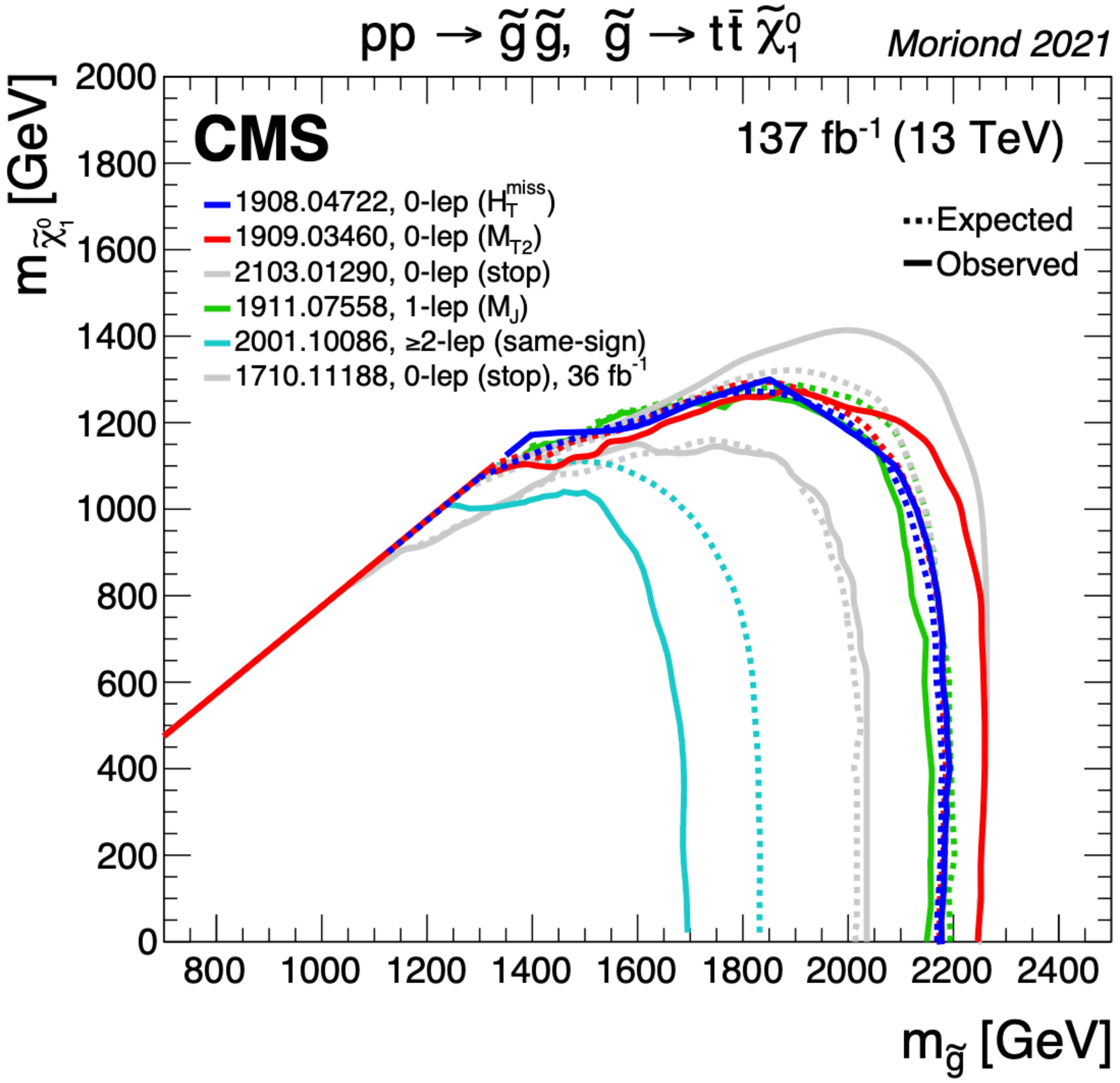
- Run 1 (7 TeV/8TeV) was a triumphant journey towards the Higgs discovery
- In Run 2 we shifted attention to new physics, thanks to energy increase (to 13 TeV)
- No discovery reported, but
  - several analyses reported 3-4 sigma (local) excesses that should be monitored with more data
  - many analyses are still ongoing



Let's keep in mind…

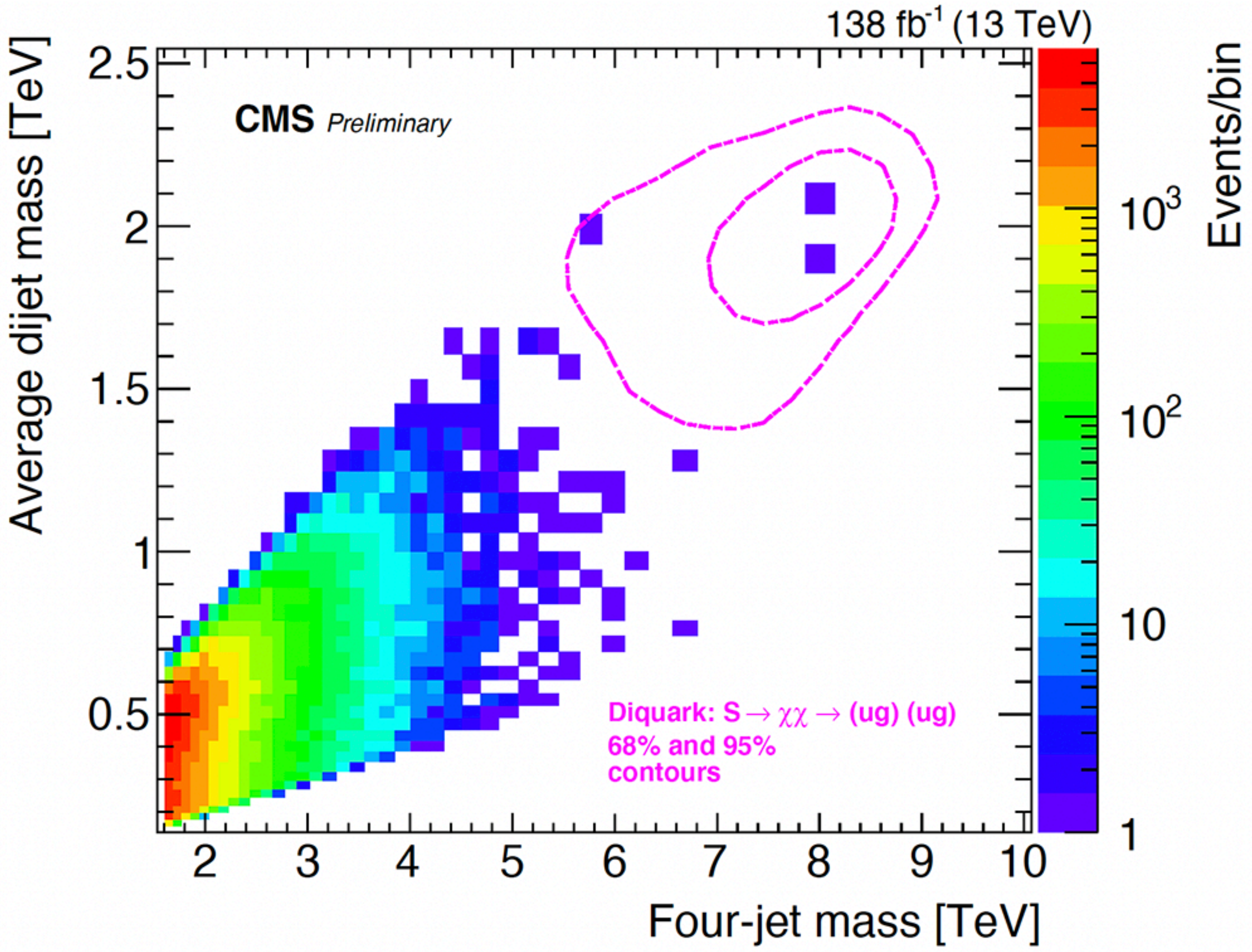
- **Exclusion plots come with assumptions and warnings**
- *For instance, SUSY limits use simplified models*
  - *Not a complete model, BR=100% typically assumed, Cross section computed assuming all other sparticles are decoupled*
- *The actual exclusion might be weaker*





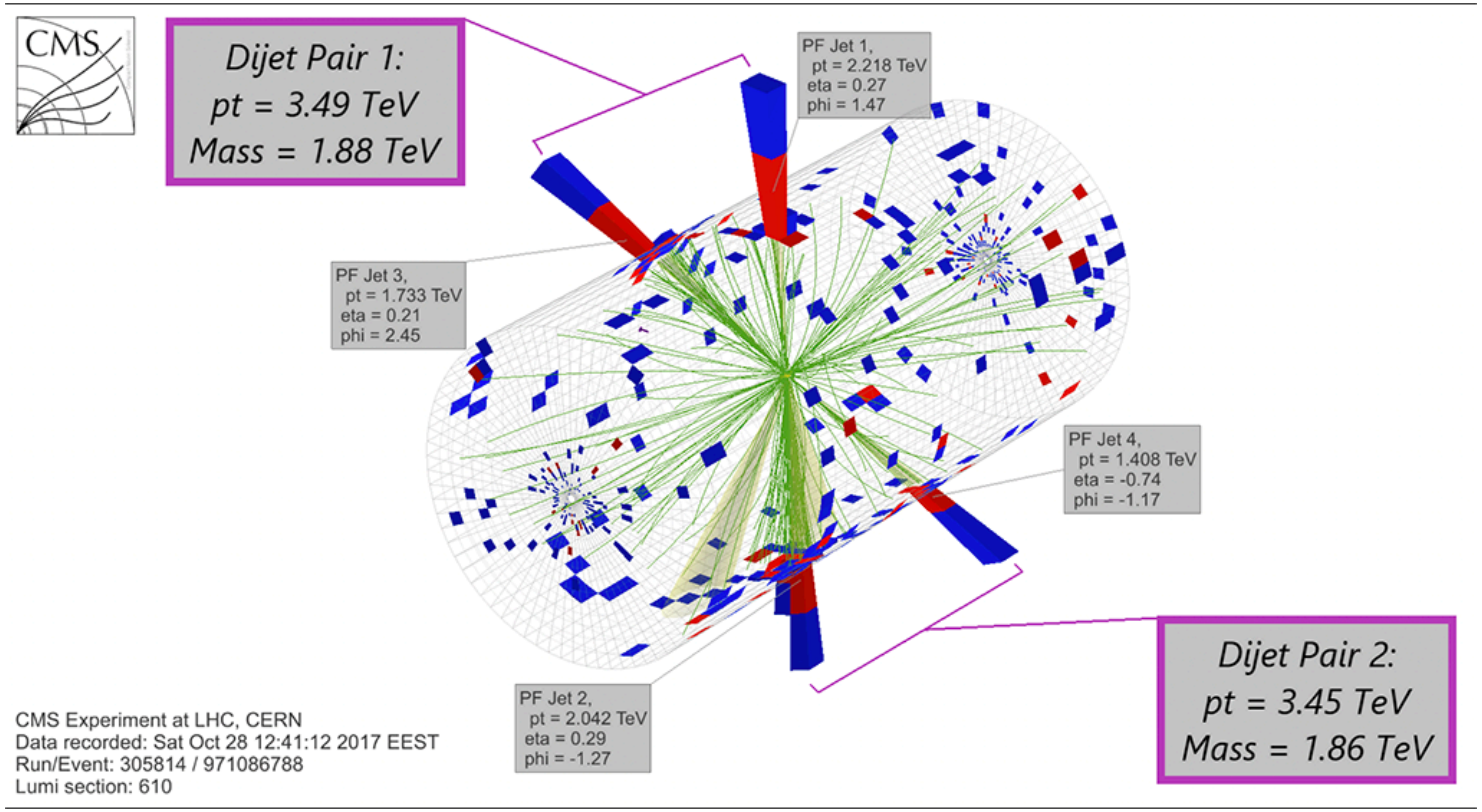
Some interesting excess

- The dijet resonance search identified an excess of a few events on the tail
- When inspected, these events revealed a common structure
  - two jet pairs, merged into single jets with M~2 TeV
  - The dijet system had mass above 6 TeV
- Nothing yet significant, but something to watch



Some interesting excess

- *The dijet resonance search identified an excess of a few events on the tail*
- *When inspected, these events revealed a common structure*
  - *two jet pairs, merged into single jets with M~2 TeV*
  - *The dijet system had mass above 6 TeV*
- *Nothing yet significant, but something to watch*



Some interesting excess

- *Excess seen in ATLAS dE/dx analysis. CMS still analysing Run2 data + will repeat the search in Run3*

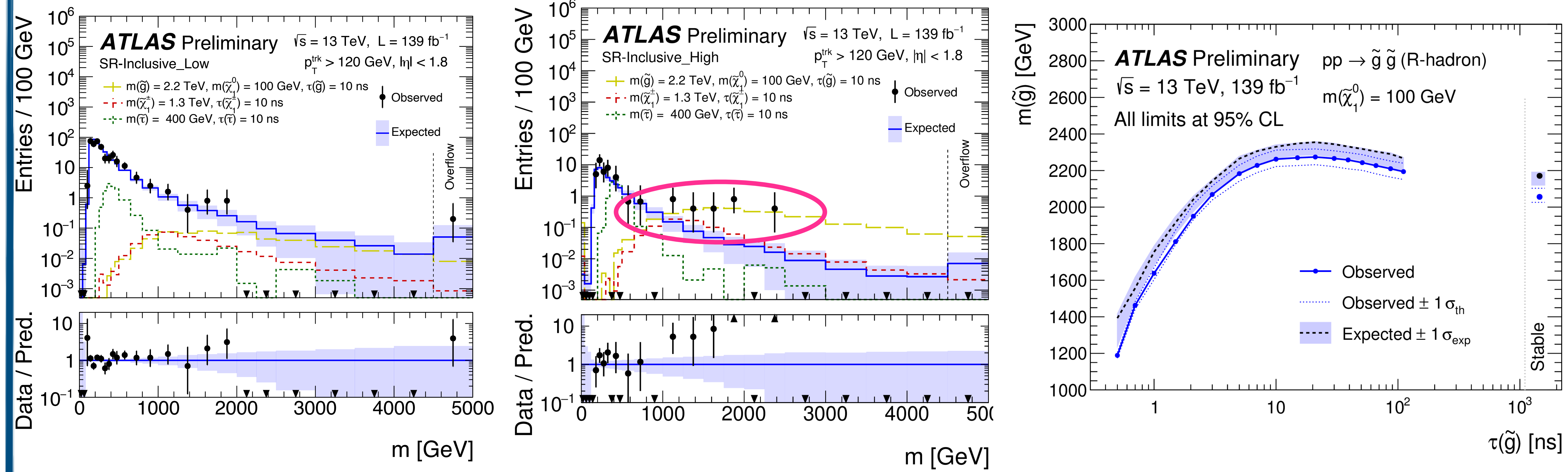
Talk by Julia Gonski @Moriond

Pixel dE/dx Result

SUSY-2018-42

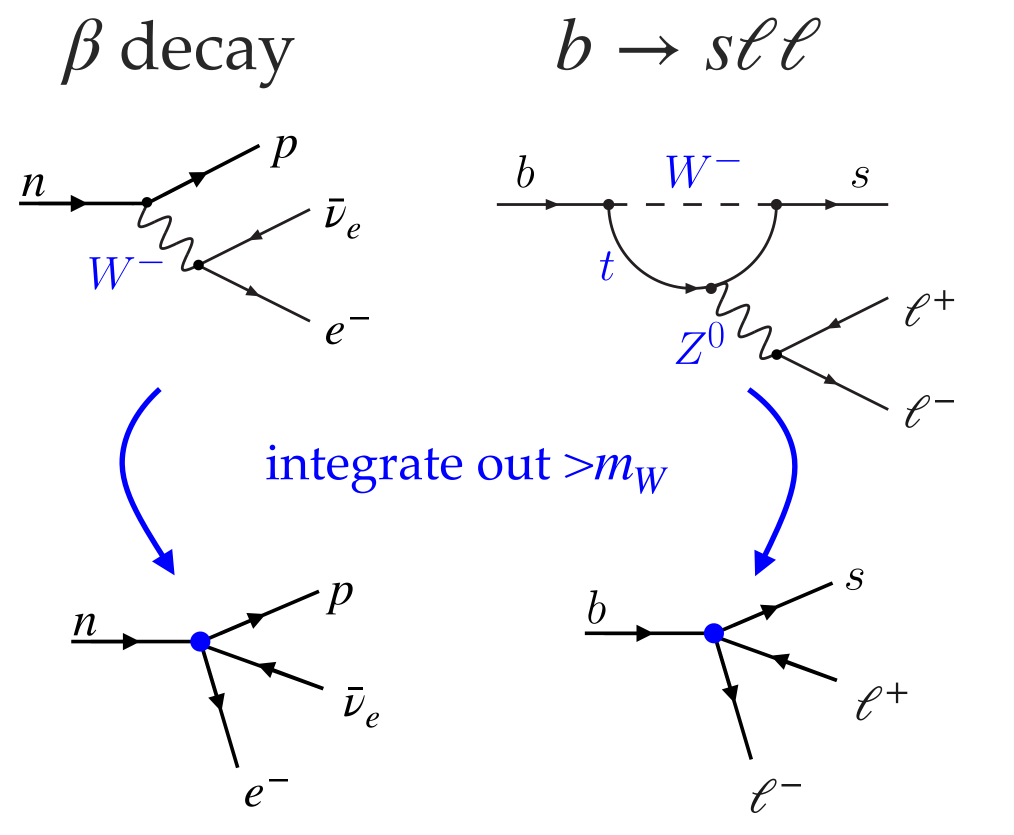
- **Excess (3.6σ local, 3.3σ global)** in high dE/dx SR (> 2.4) with for target mass hypothesis of 1.4 TeV
  - ❗ Cross check of candidate tracks with TileCal and MS time-of-flight variables was consistent with β = 1, therefore ***not* consistent with LLP hypothesis**
- Set limits on gluino, chargino, stau hypotheses
  - Max sensitivity for τ ~10-30 ns
  - Gluino R-hadrons with mass < 2.27 TeV excluded with τ = 20 ns and LSP mass = 100 GeV

SR Low-dE/dx [1.8, 2.4] | SR High-dE/dx [> 2.4] | Gluino Exclusion



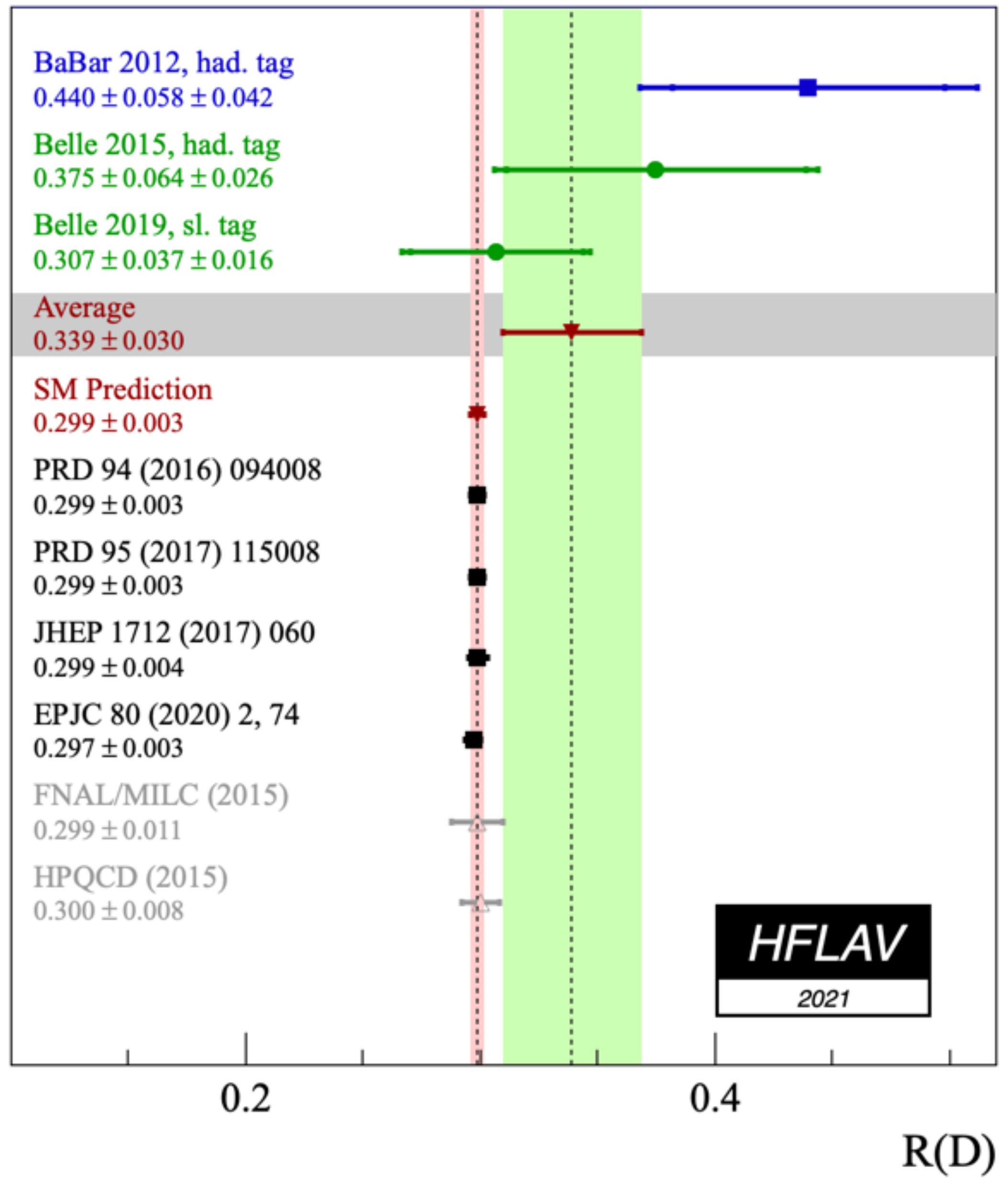
And the biggest excitement

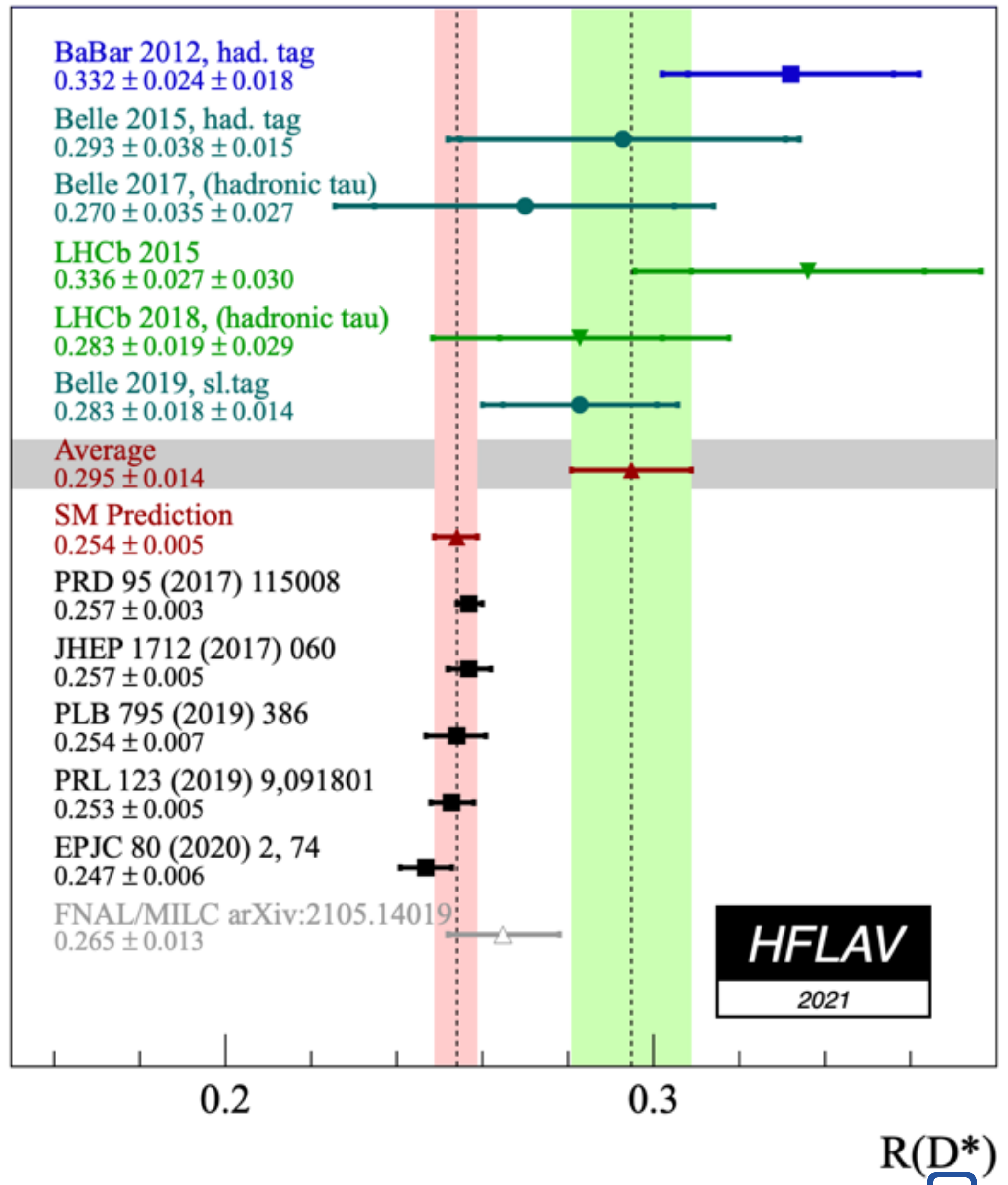
- LHCb anomalies point to breaking of Lepton Flavor Universality
  - In μ/e ratio of b-> sll
  - In τ/μ ration of b -> clv rates
- Confirmation from LHCb at larger statistics + independent experiment

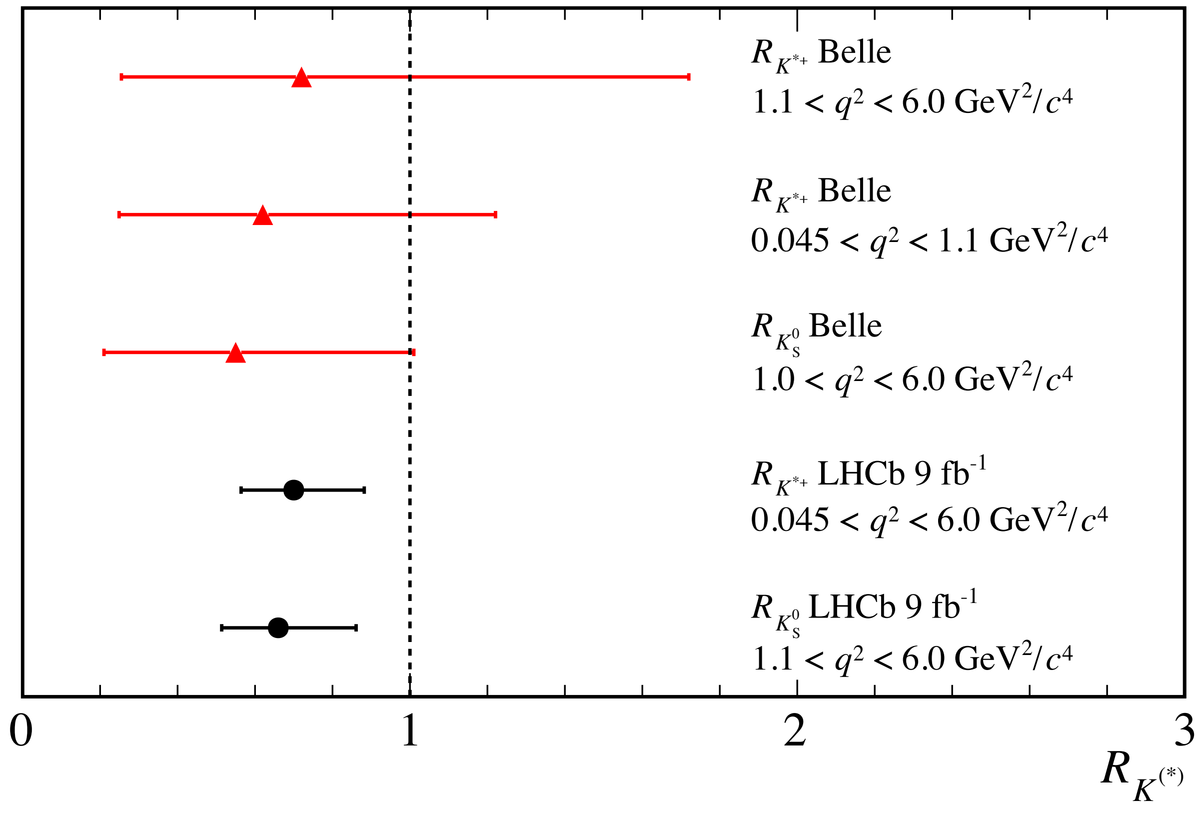


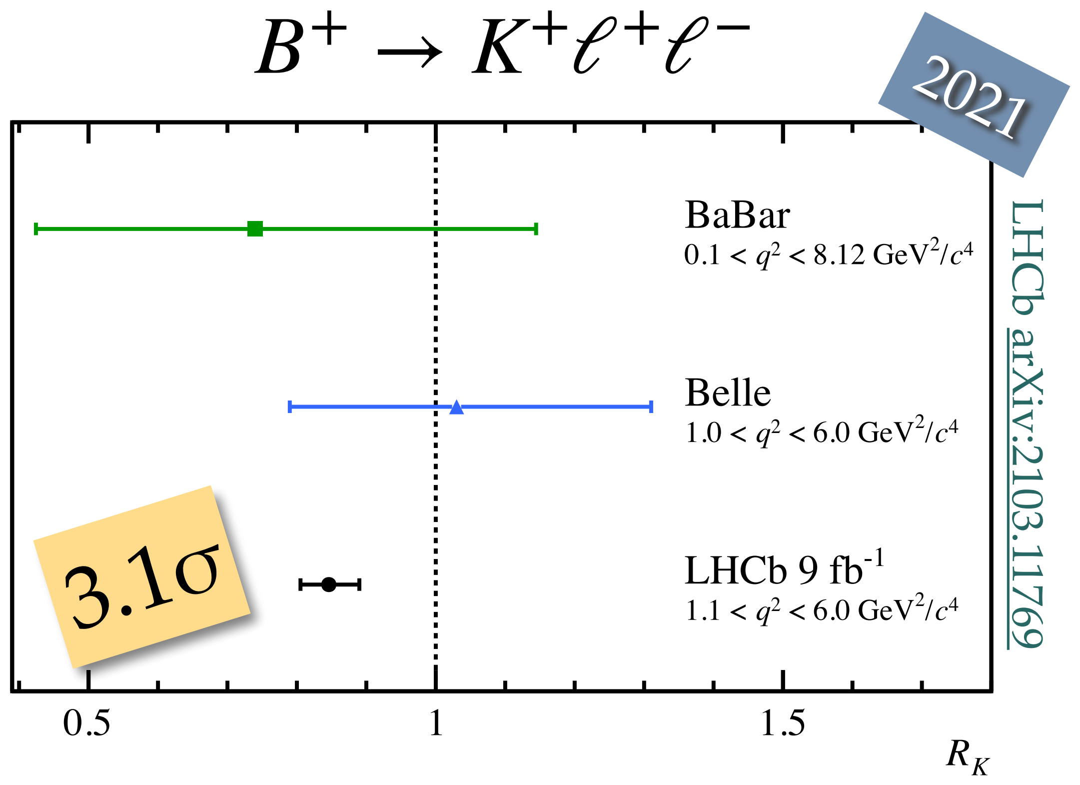
$$\mathscr{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_i C_i O_i$$

$$R_H = \frac{\int_{q^2_{\min}}^{q_{\max}} \frac{\mathrm{d}\mathscr{B}(B \to H\mu^+\mu^-)}{\mathrm{d}q^2}\,\mathrm{d}q^2}{\int_{q^2_{\min}}^{q_{\max}} \frac{\mathrm{d}\mathscr{B}(B \to He^+e^-)}{\mathrm{d}q^2}\,\mathrm{d}q^2} \overset{\text{SM}}{\cong} 1$$





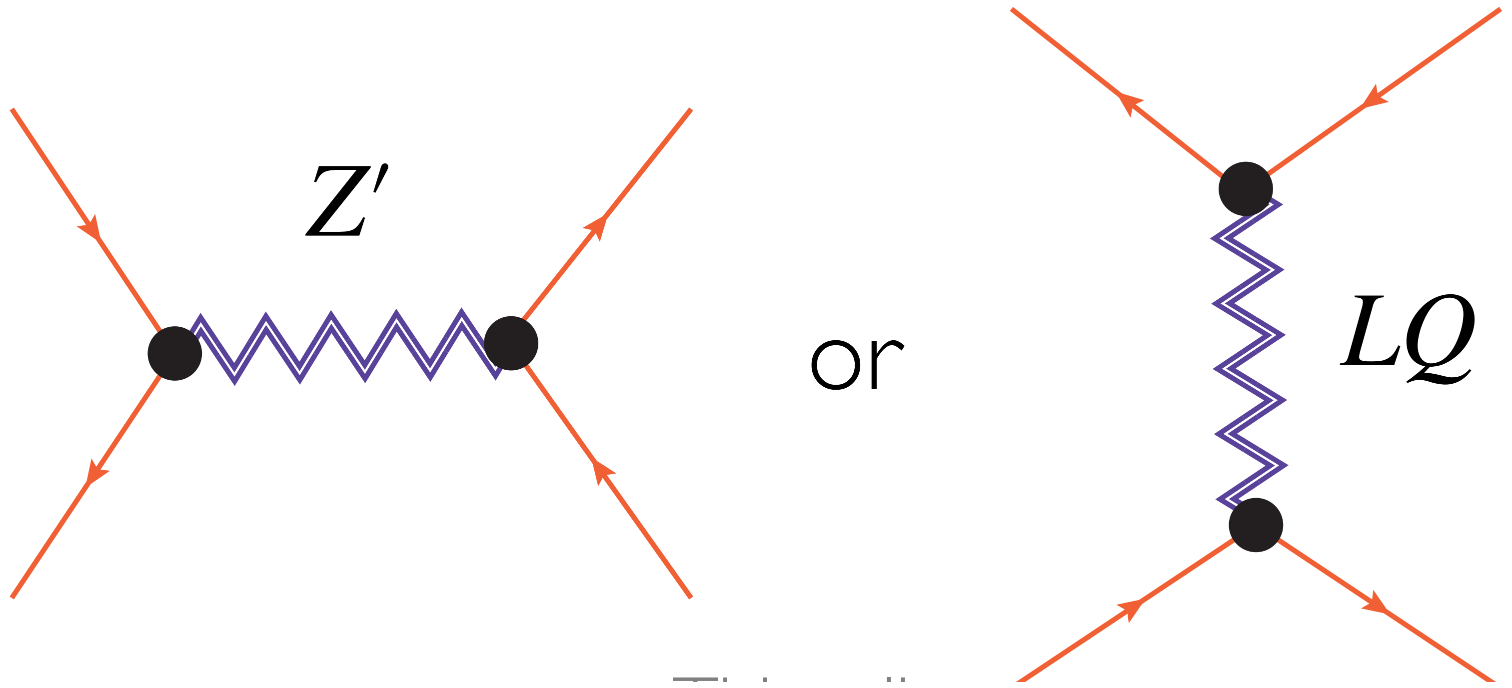


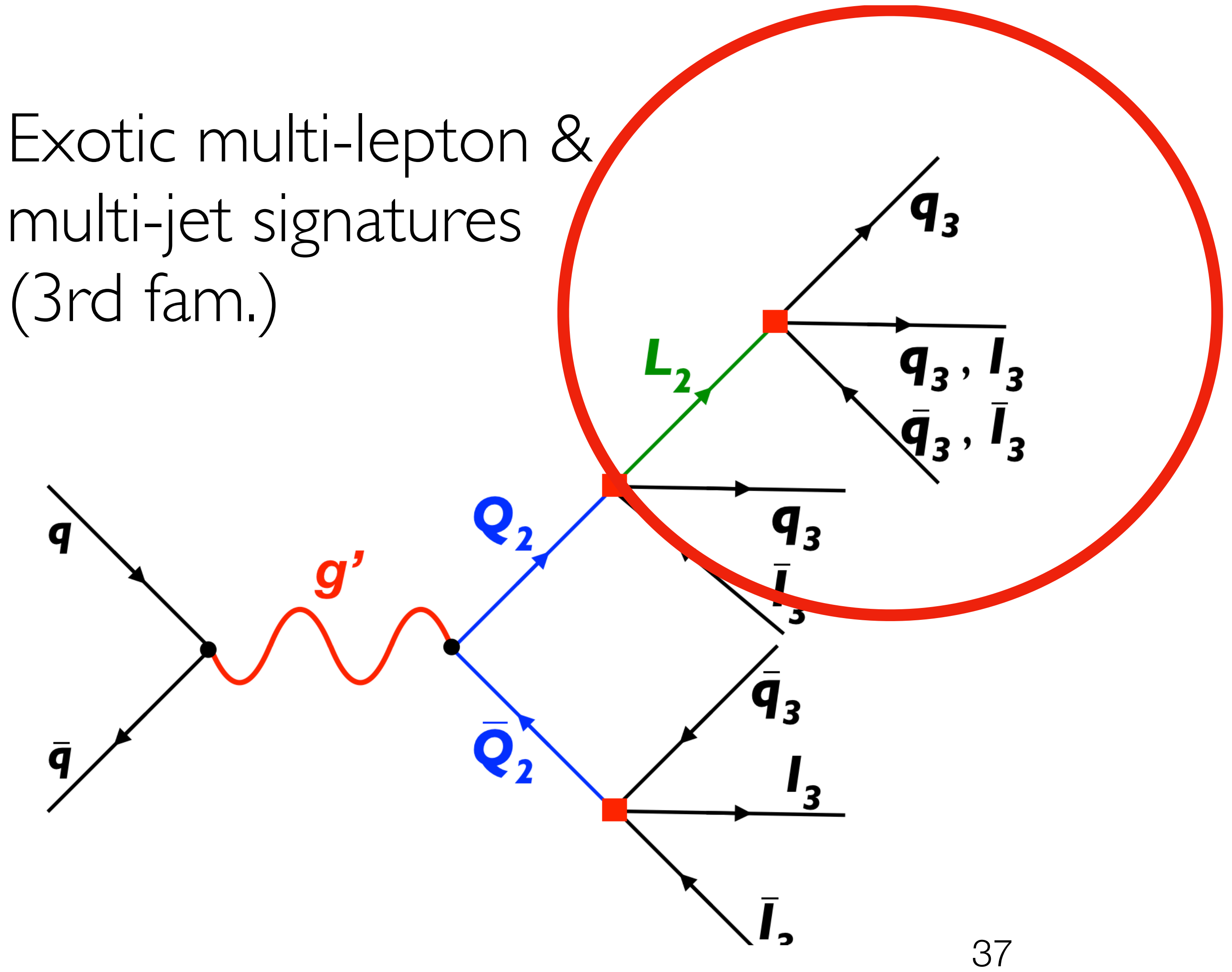


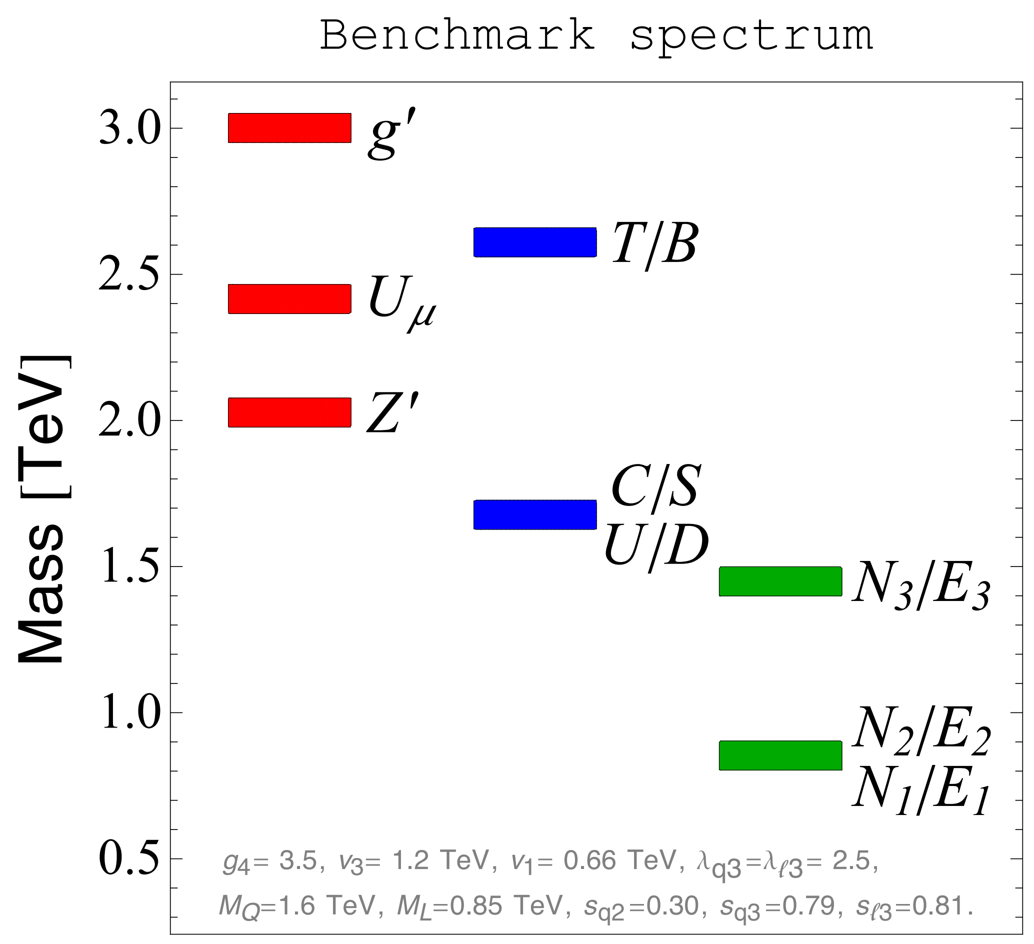
European Research Council

And the biggest excitement

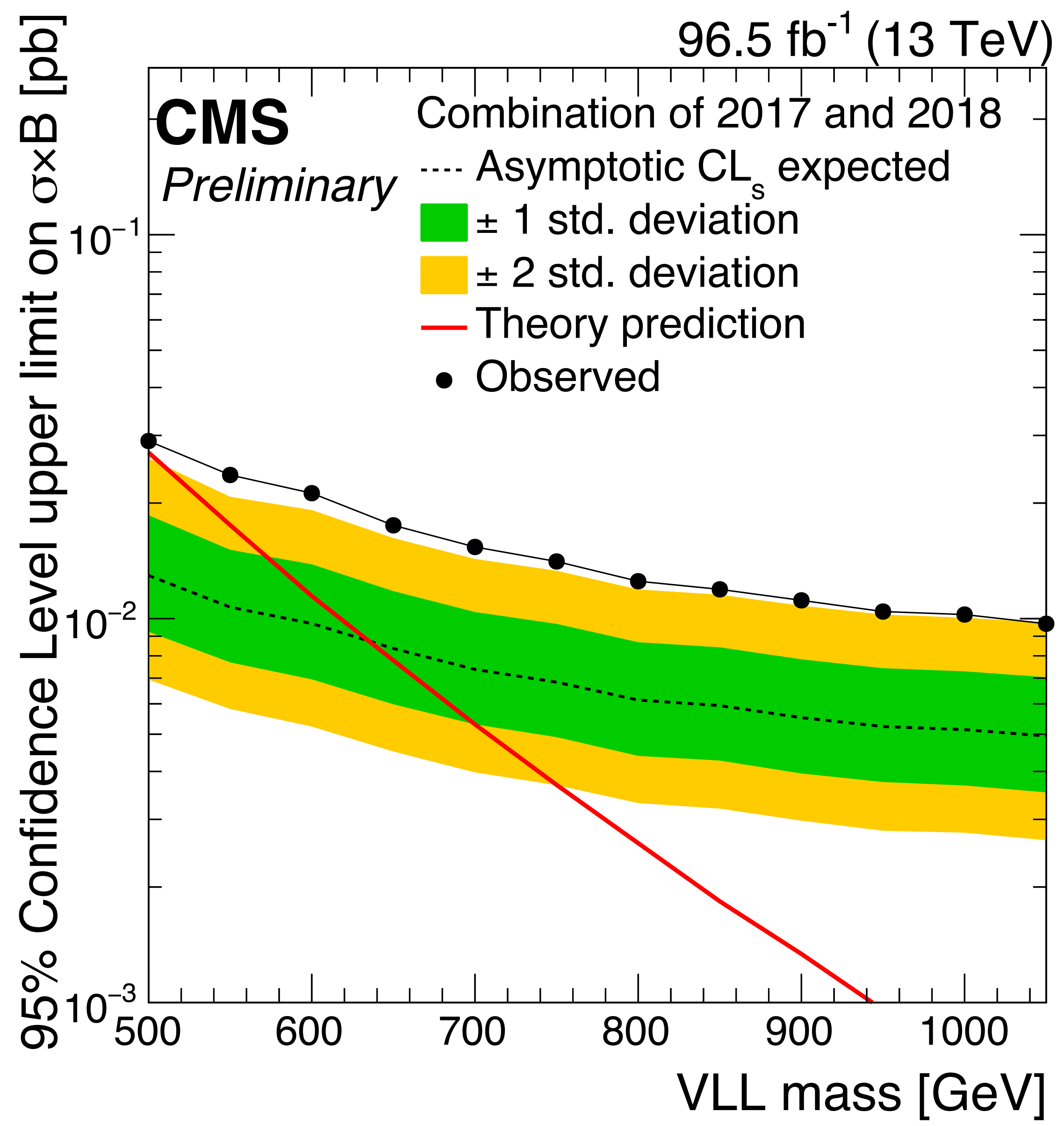
- *Several implications at high pT*
  - *Leptoquarks*
  - *Z'*
  - *Long decay chains to final states with 3rd generation particles*
  - *Mild excess observed (2.8σ of significance)*







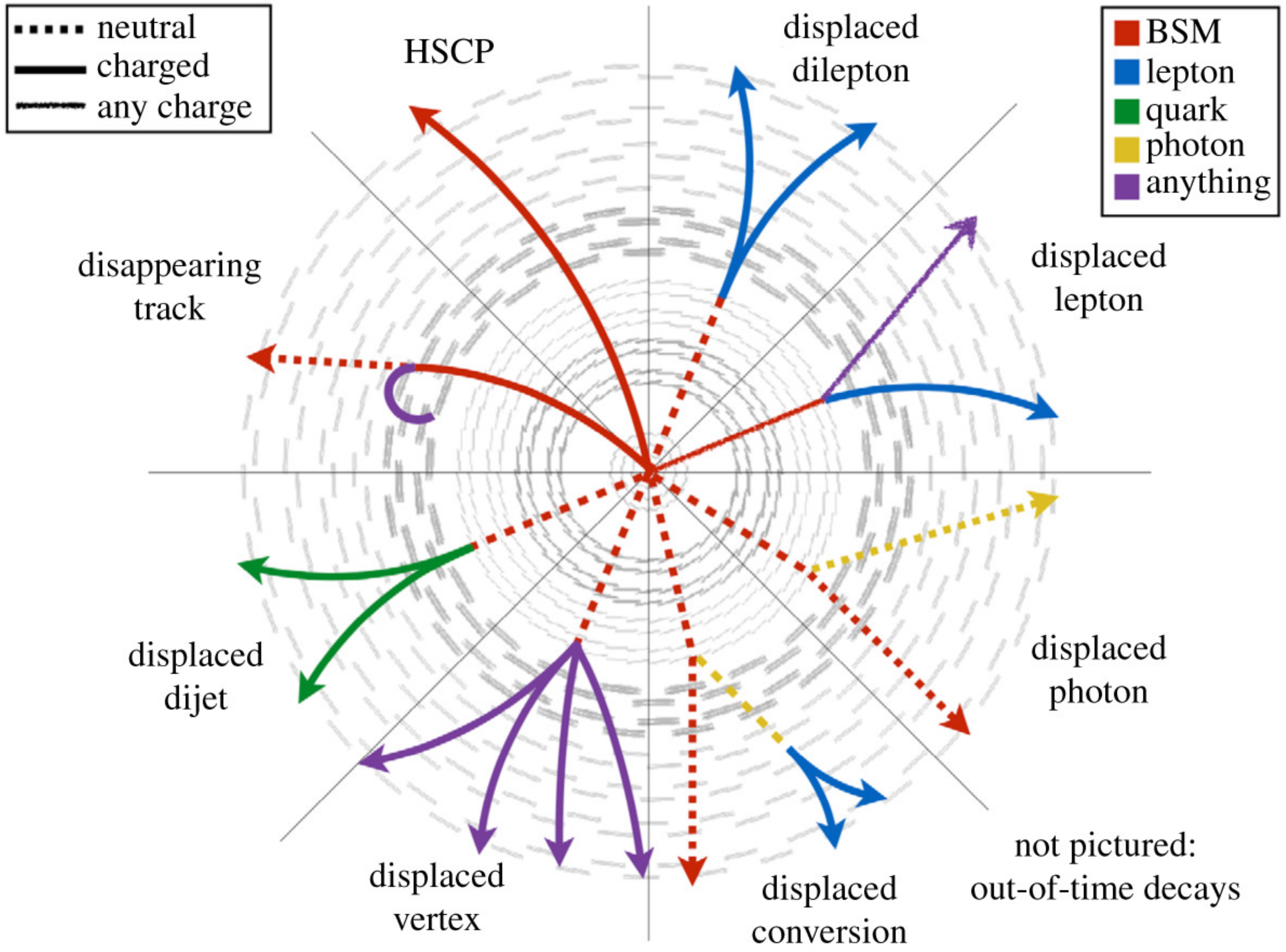
Di Luzio, Fuentes-Martin, AG, Nardecchia, Renner; 1808.00942

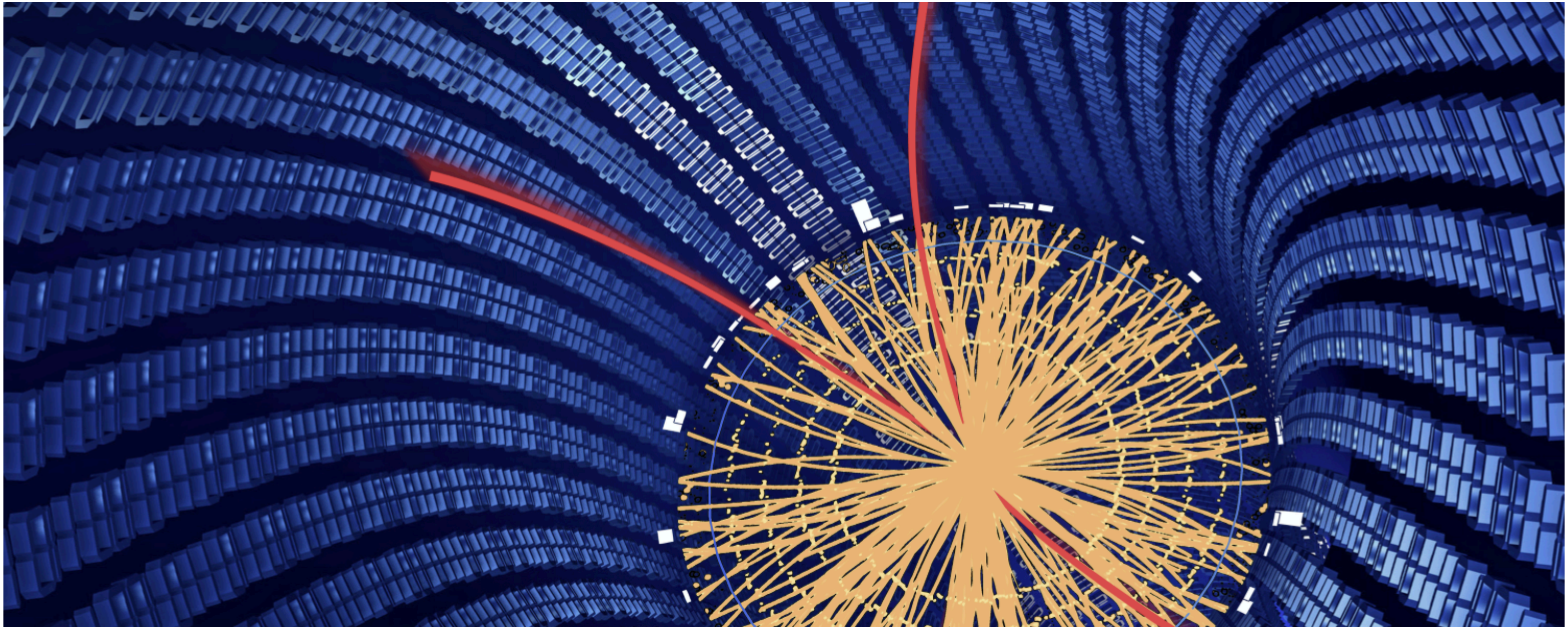


European Research Council

What should we expect from Run3

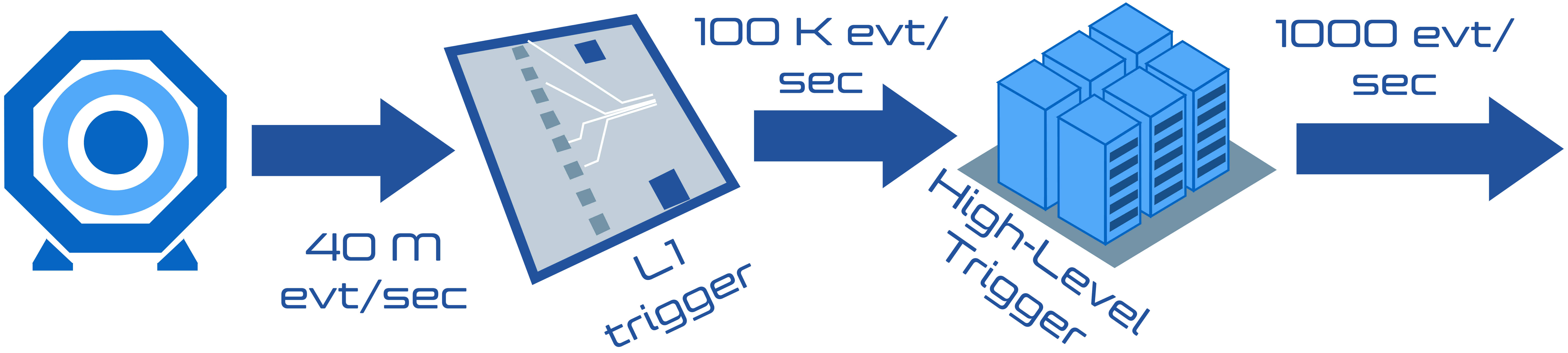
- Luminosity doubling
  - can improve current results by at least sqrt(2)
  - can investigate previous excesses with comparable datasets
- New trigger opportunities
  - Improved algorithms and hardware let us do more than in the past
  - New triggers imply new (i.e., unexplored) territory, basically a new experiment
  - For instance, a big push in this direction from long-lived particle searches





Alternative Data Taking Strategies: scouting

The LHC Big Data Problem

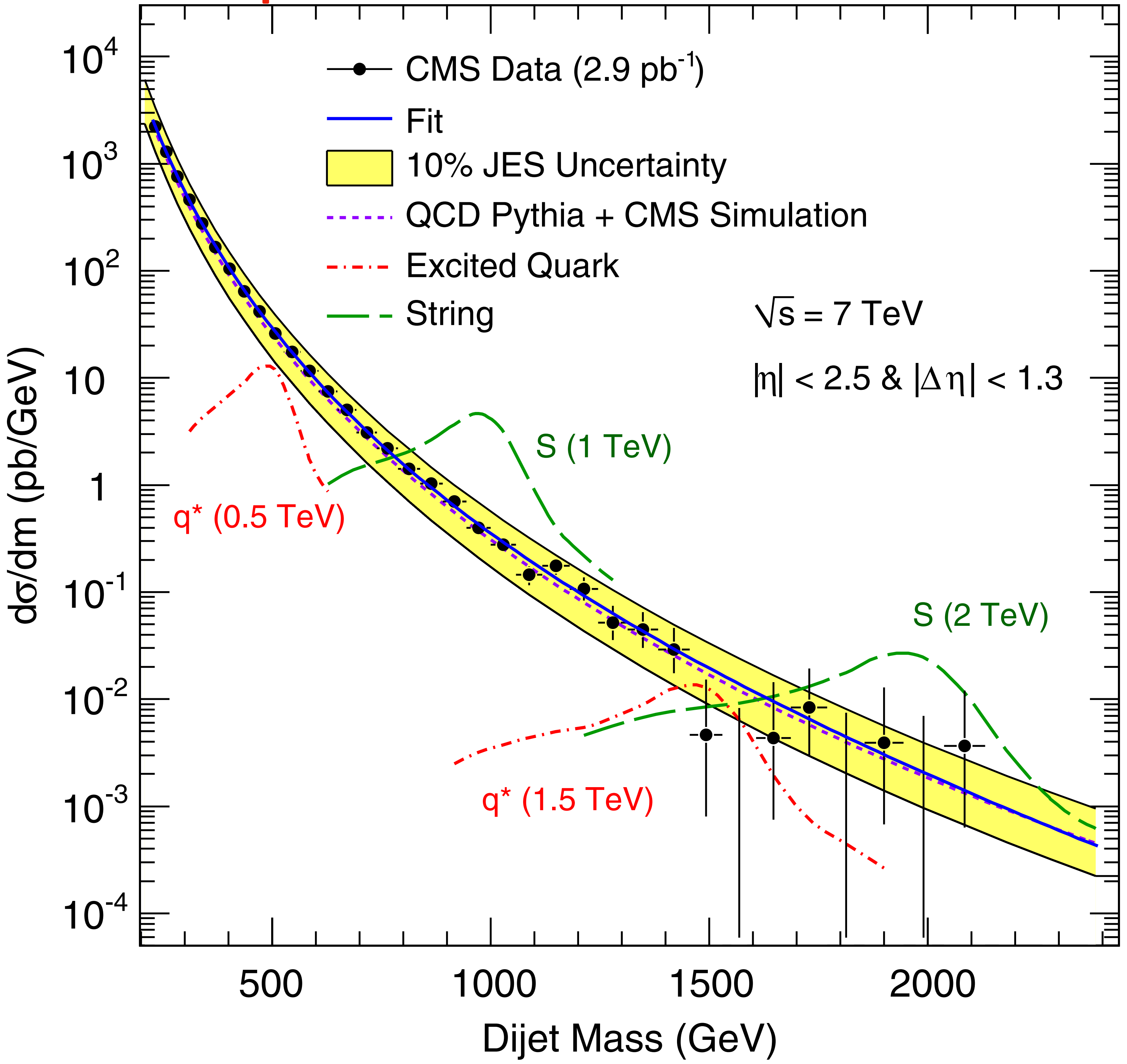


- LHC produces more events than what we can store
  - We then filter them using trigger
  - The menu is a negotiation between different physics topics and the rate of the corresponding processes
  - Many studies are trigger limited

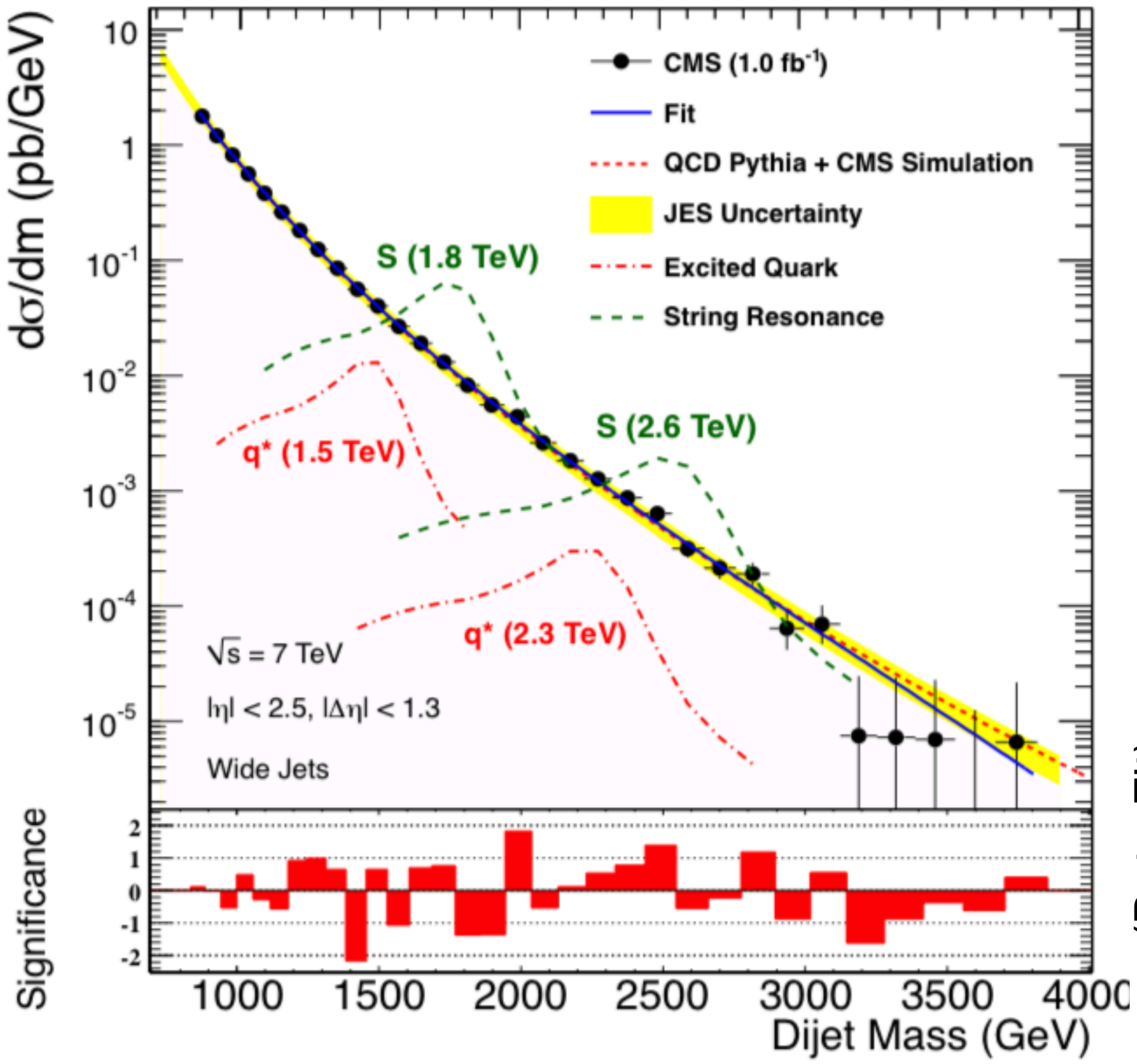
Example: Dijet resonance searches

- *Most of LHC events have two jets*
- *If one wants to look for dijet resonances, a hard cut on jet energy has to be applied*
- *For this reason, we cannot keep sensitivity to low-mass resonances*

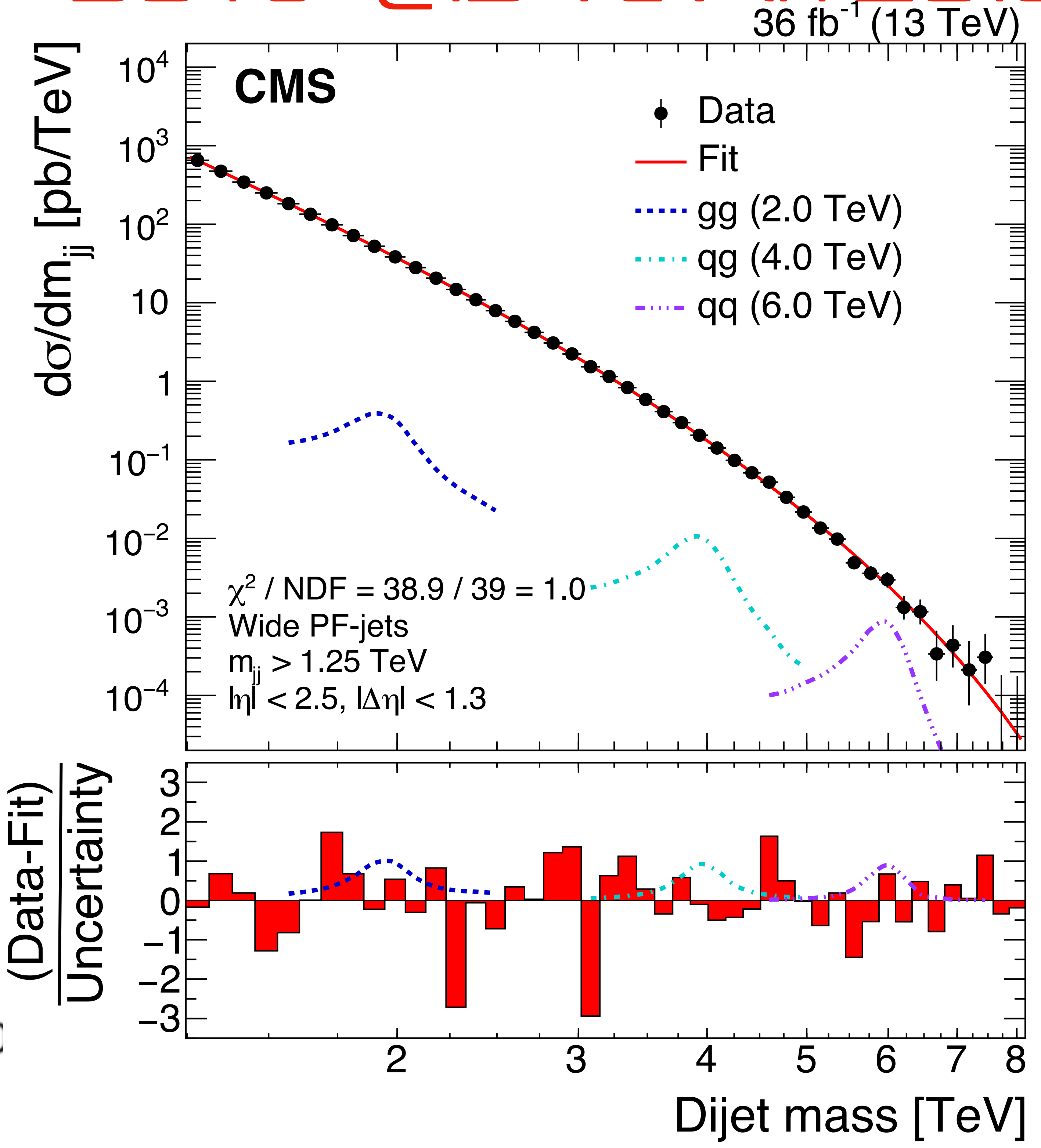
3 pb^{-1} @7 TeV in 2010



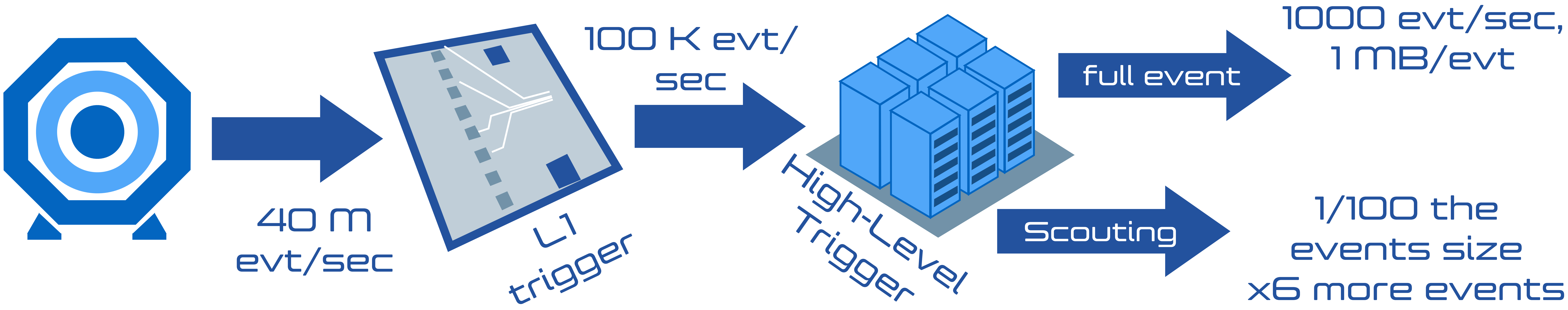
1 fb^{-1} @7 TeV in 2011



36 fb^{-1} @13 TeV in 2016



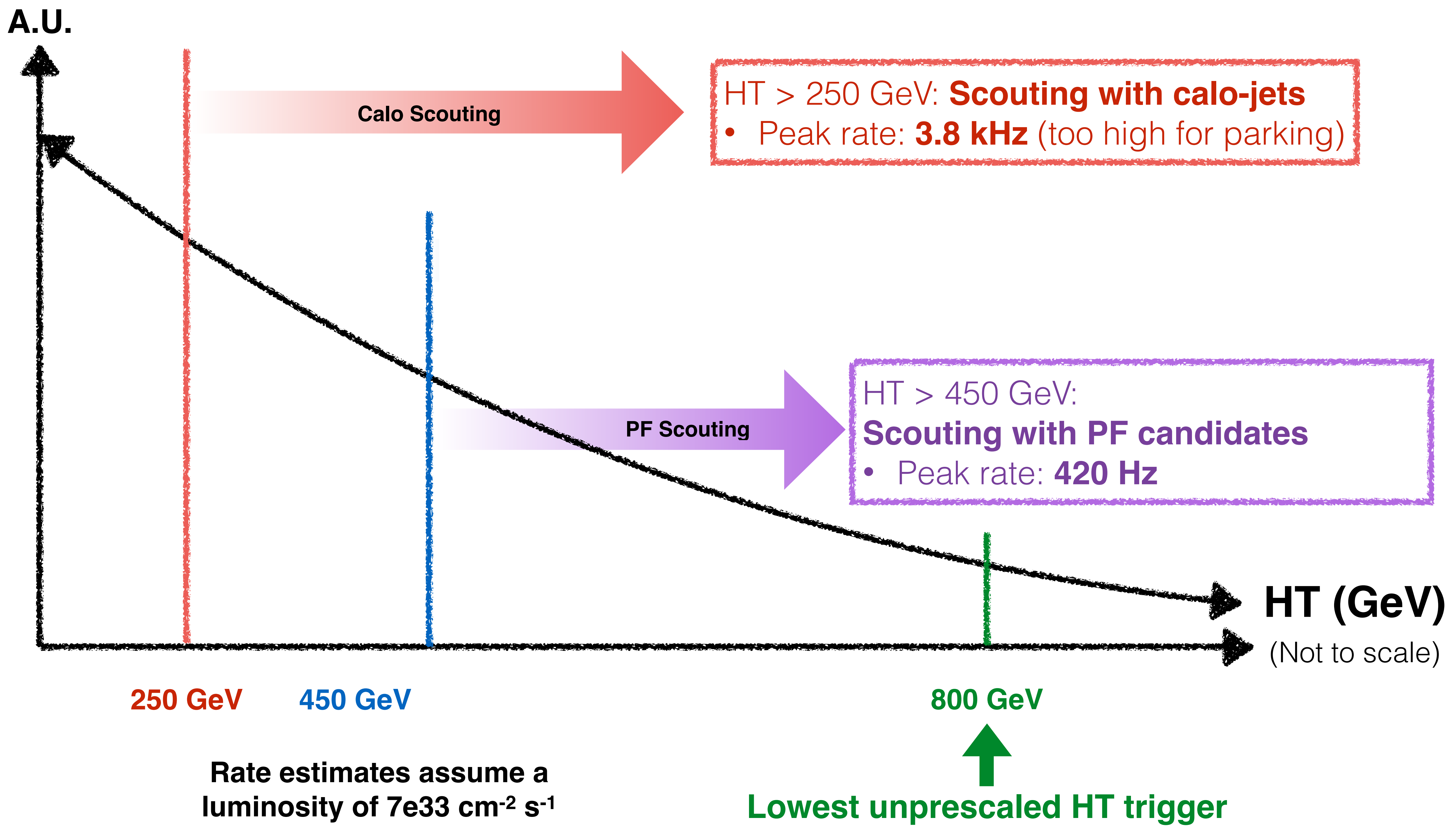
The Scouting Streams



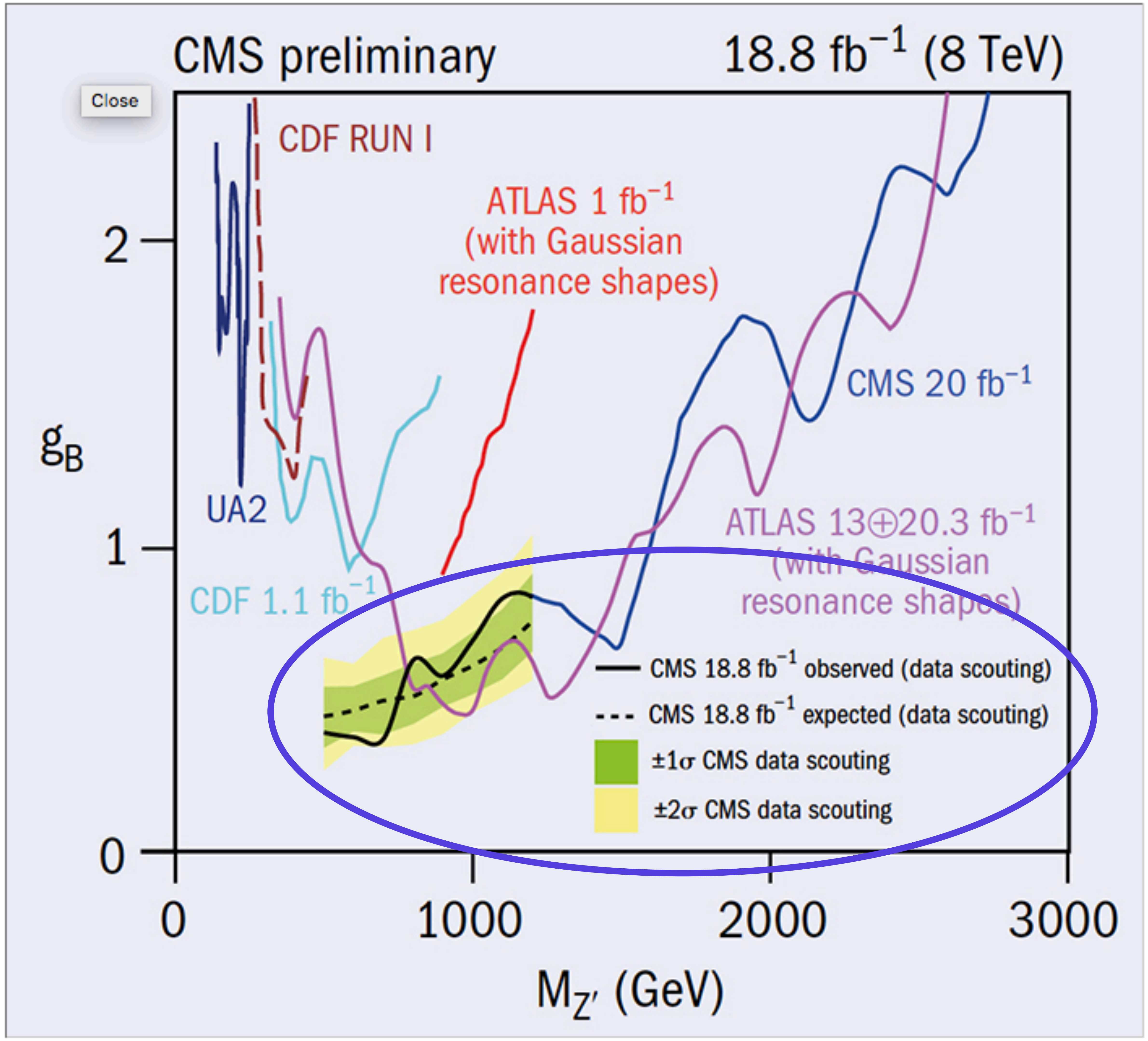
- *Trade-off between # of events & event size*
  - *write HLT objects (four-momenta, etc…) rather than full event content*
  - *write many more events (essentially HLT passthrough, limited by L1)*

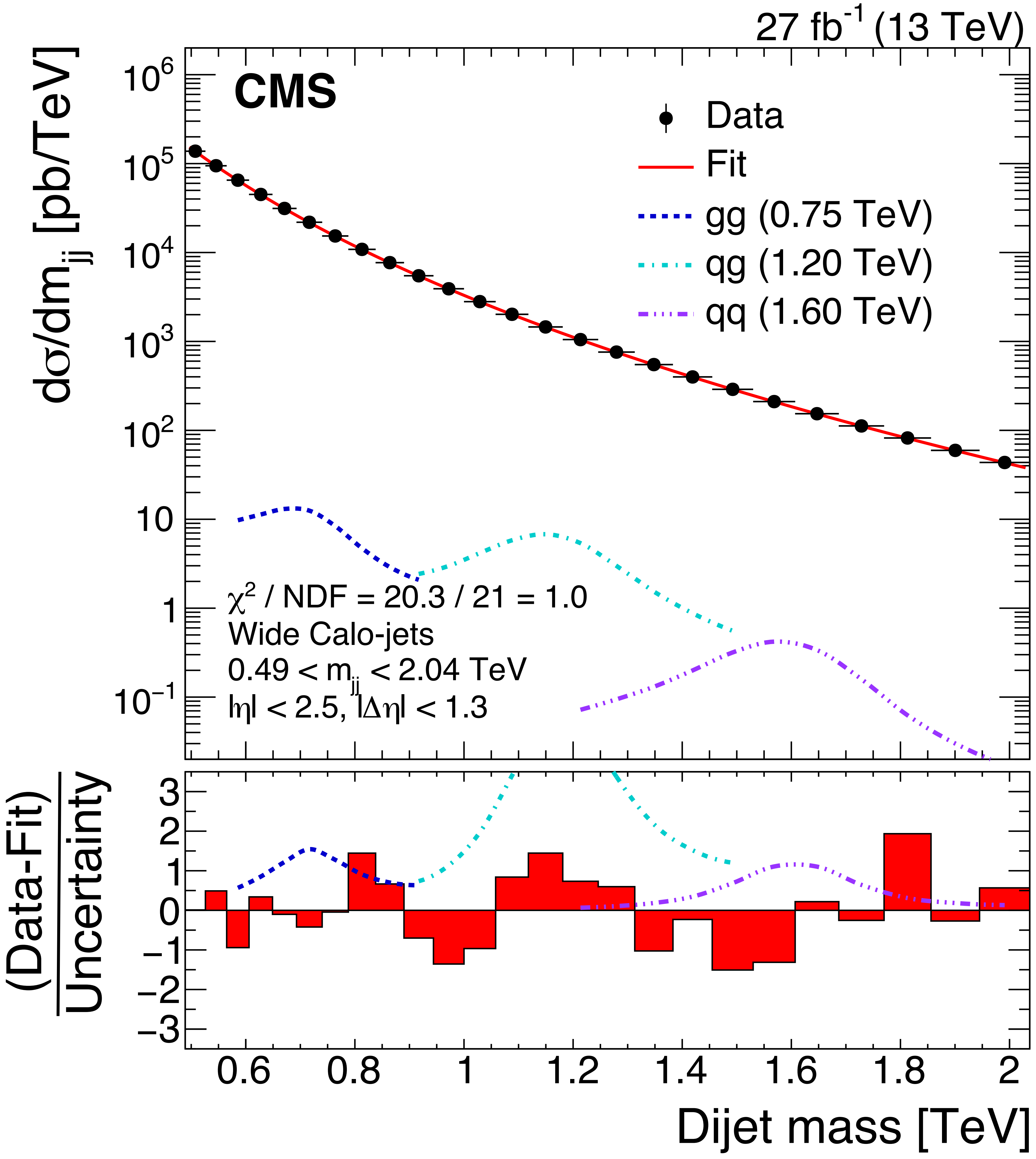
DiJet Scouting Streams

- *Dijet Data Scouting strategy in Run2 used two tiers*
  - *A calo-scouting stream with minimal event content, to go as down as possible (limited by L1 trigger)*
  - *A PF-scouting stream, going as down as HLT let us run the PF sequence. More complex event content and more flexibility (substructure, etc)*



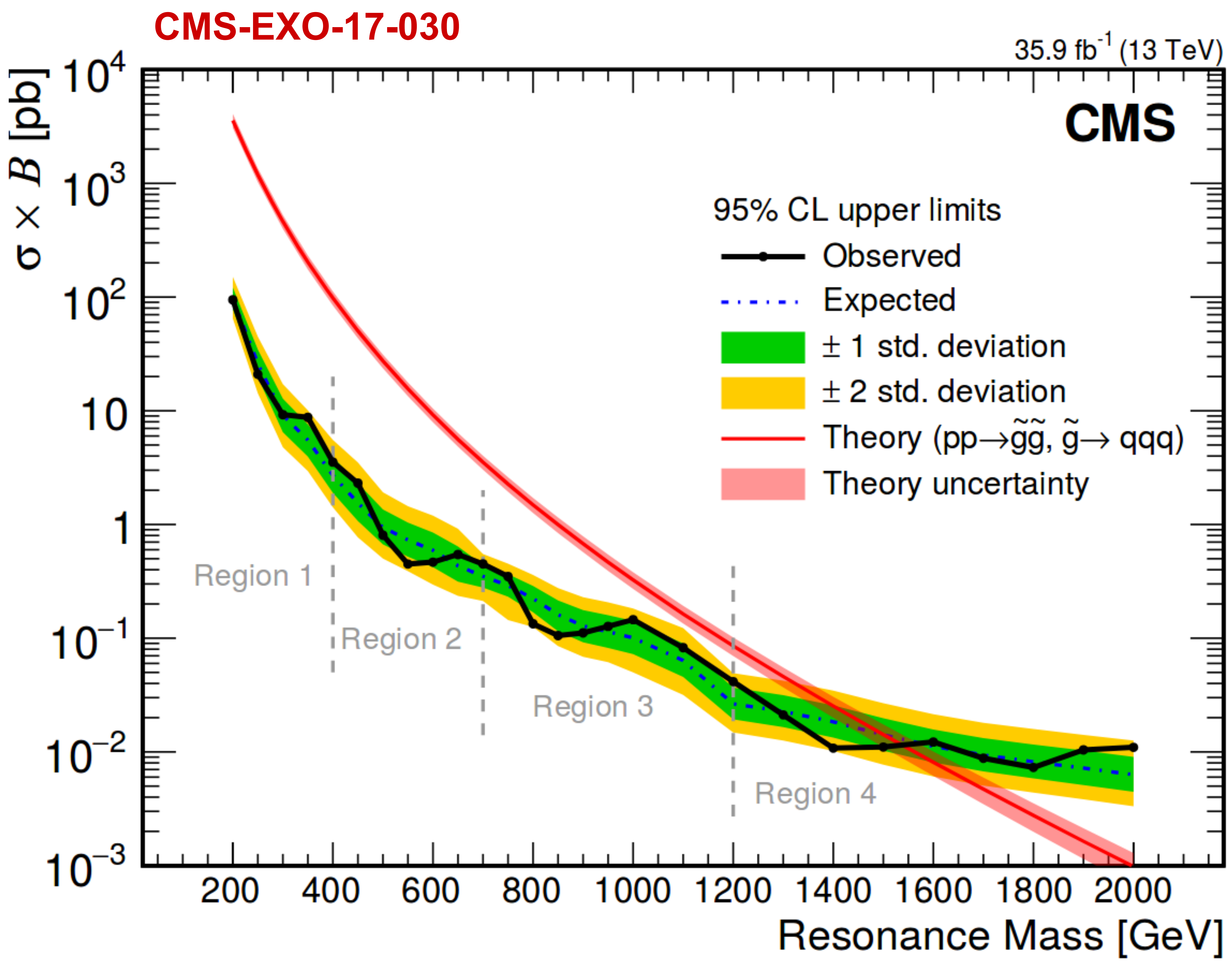
DiJet Scouting Streams





More than dijet

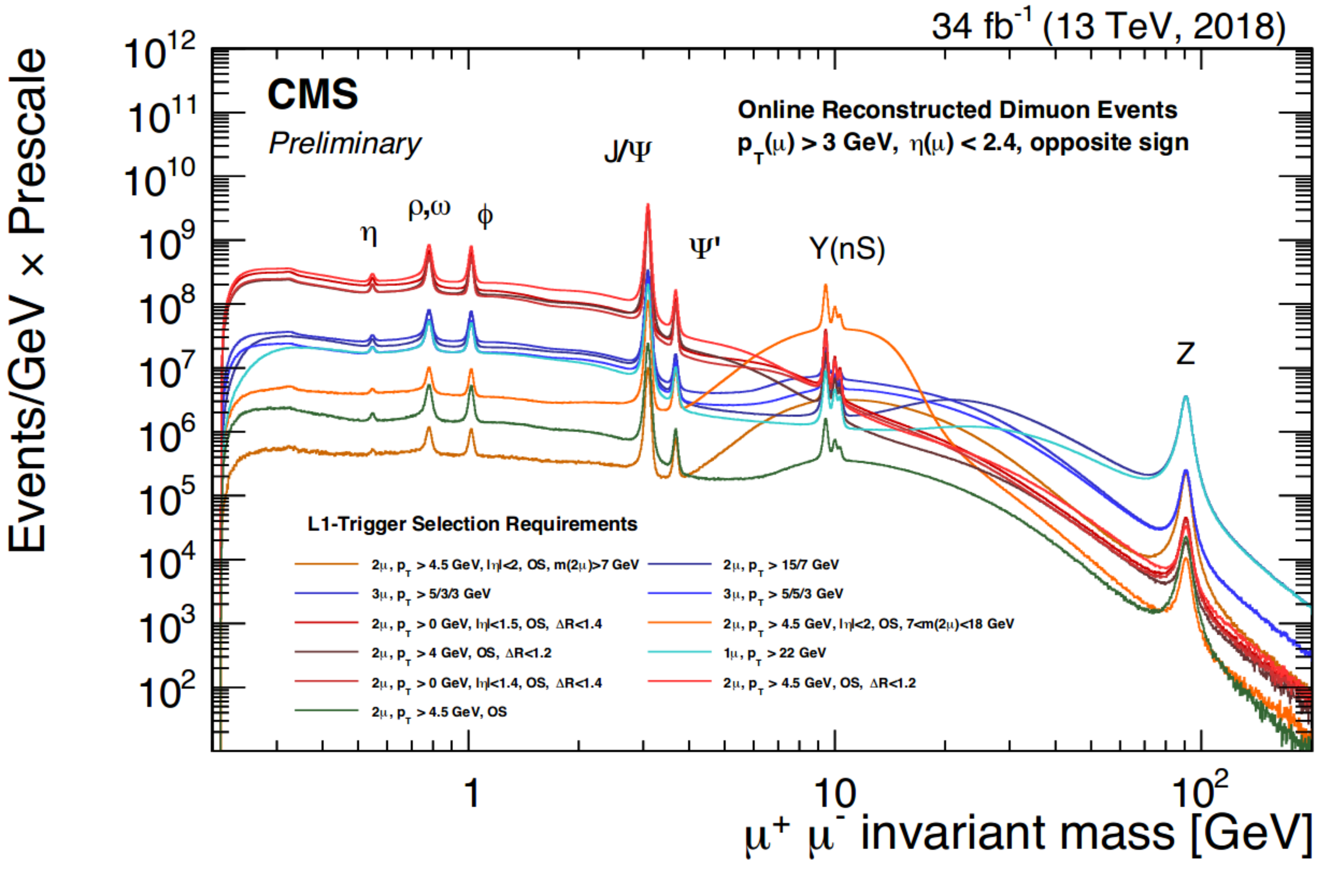
- *Same approached followed for other final states*
    - *multiple jet resonances (e.g., trijet in*
    - *jet substructure*
    - …
- *In particular, the use of PF (with full access to individual particles) showed a great potential, which so far was barely explored*

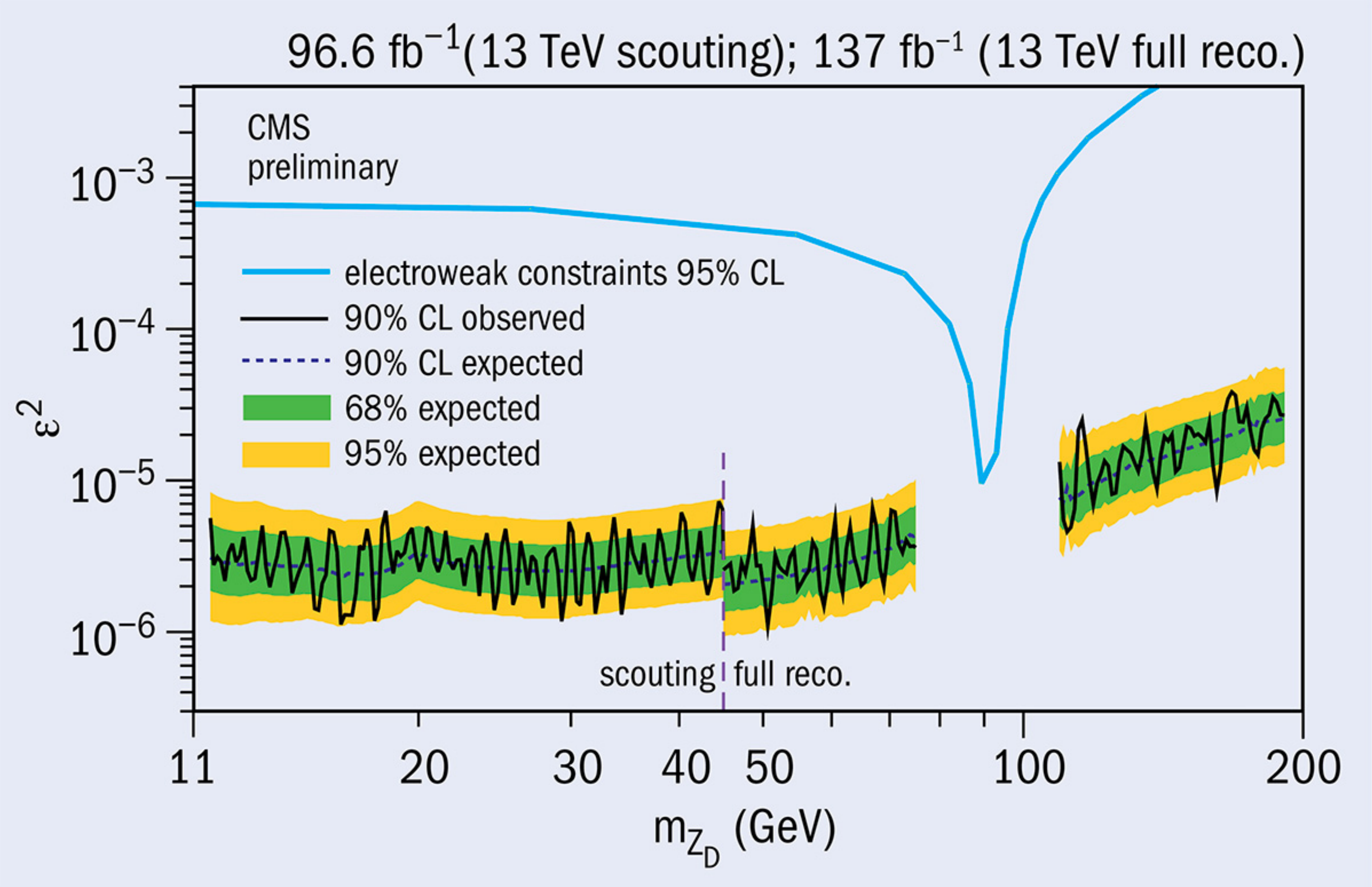


See D. Sperka's talk at HOW2019

DiMuon Scouting Stream

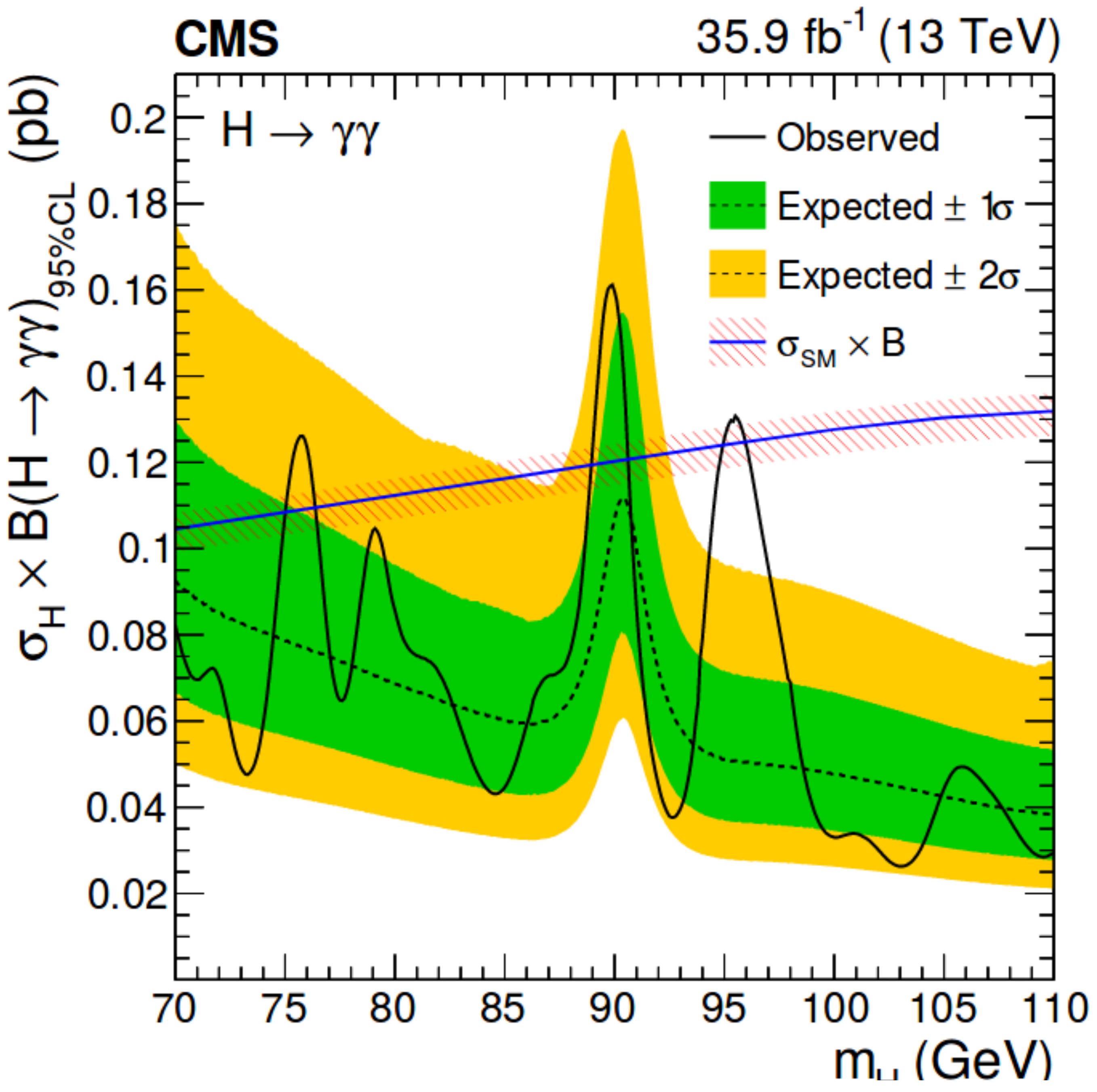
- *Starting in 2015, the same idea was applied to dimuon events, to probe light dimuon resonances (prompt and displayed)*
  - *Similar to LHCb turbostream analysis*
  - *Probed competitive Dark Photon parameter space*

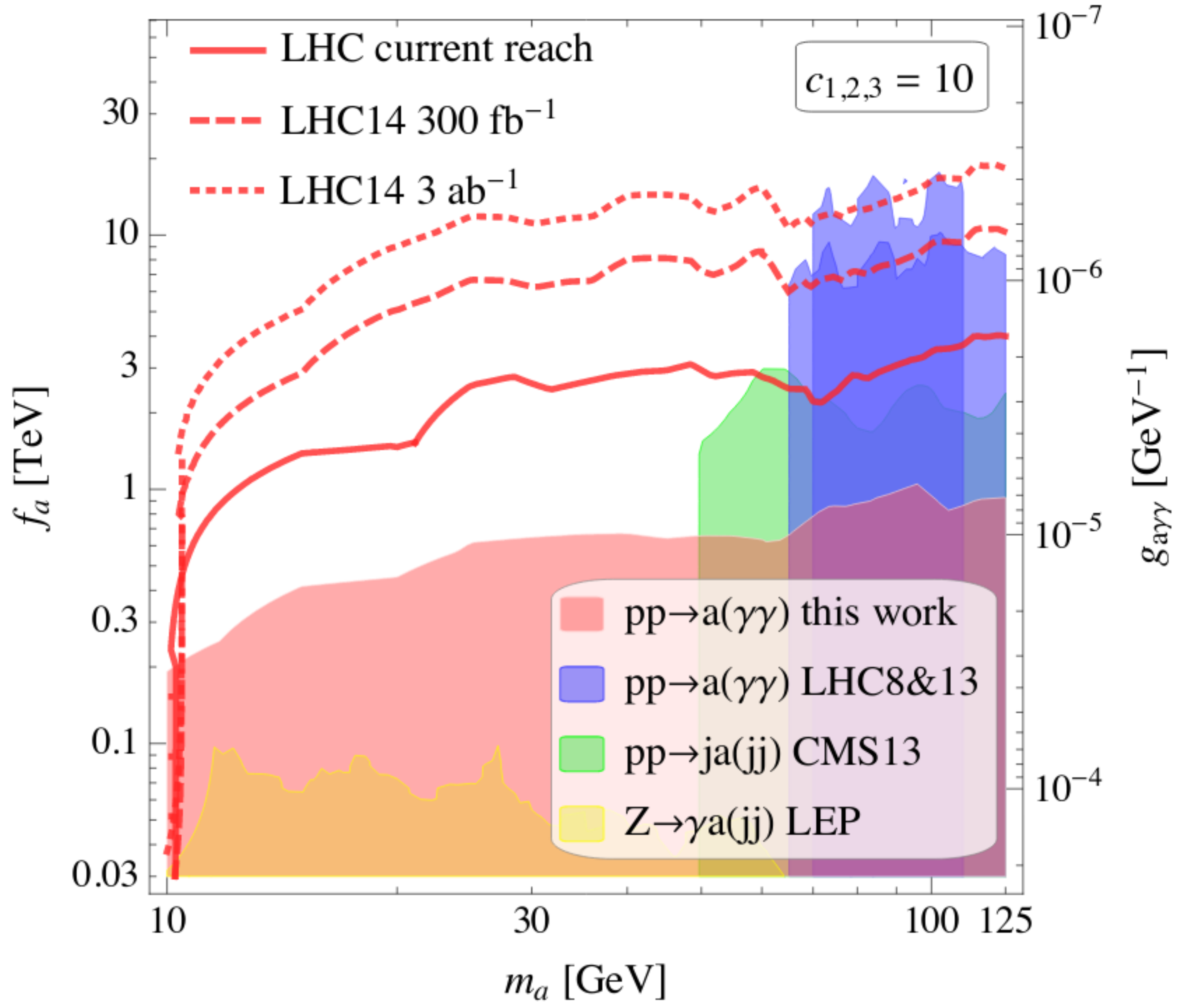




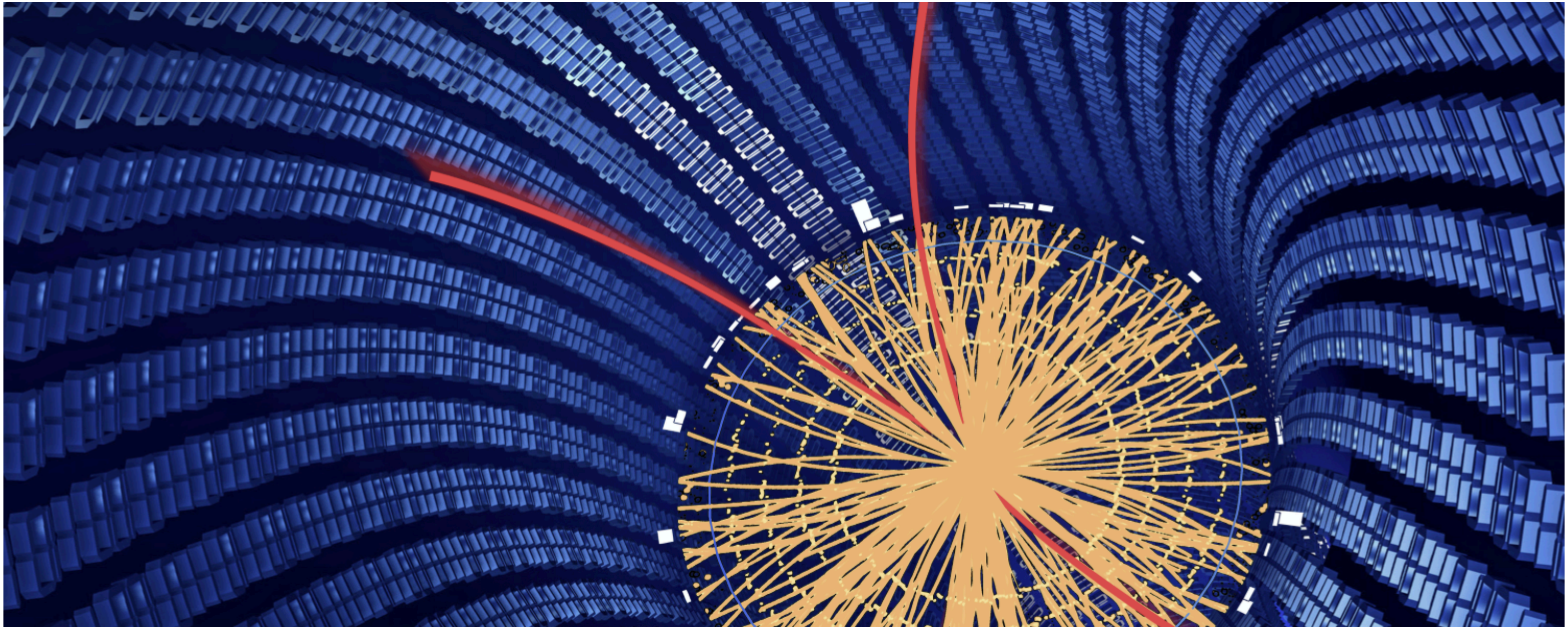
Scouting plans for Run3

- *Improved Trigger farm (heavy computing now on GPUs) makes it possible to run PF algorithm on a much larger dataset*
- *Can use PF scouting on O(10K) evt/sec, storing an event content including all PF particles*
- *Could boost the physics case of the scouting stream*



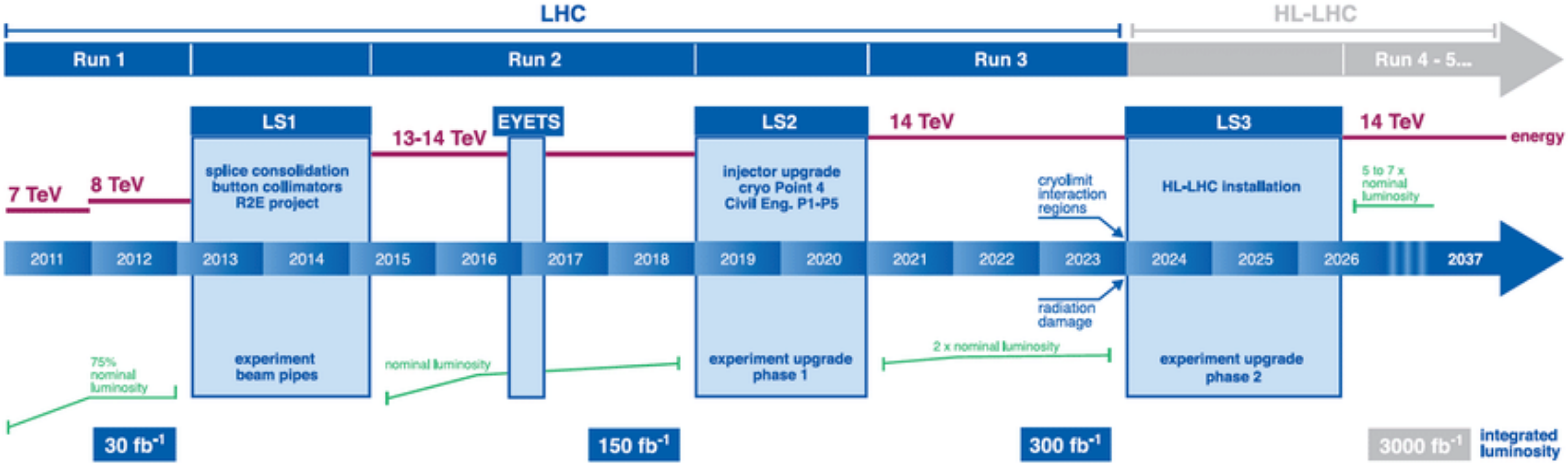


See D. Sperka's talk at HOW2019



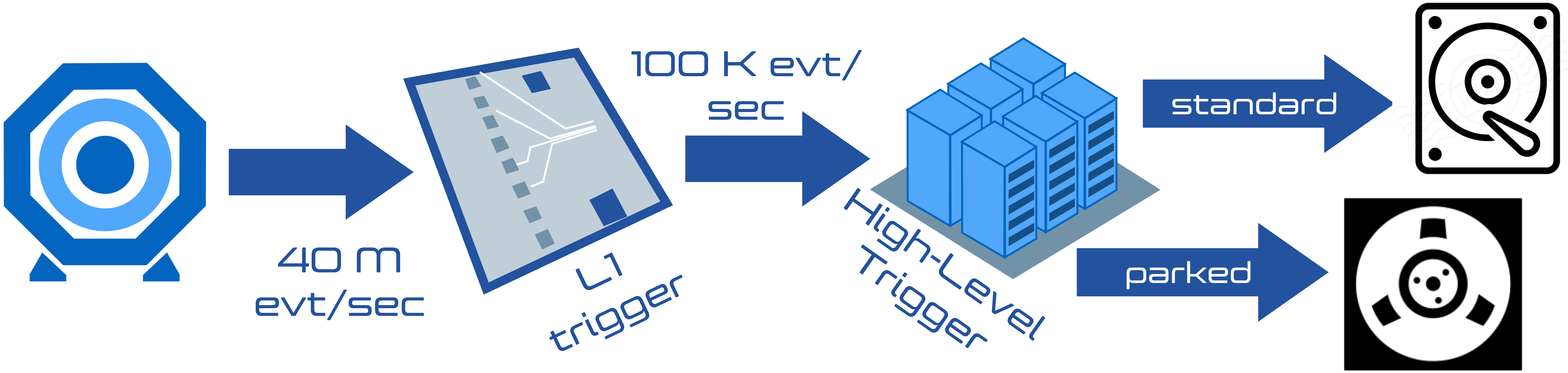
Alternative Data Taking Strategies: parking

LHC schedule: Runs and LeSs



- *When we take data, we saturate our computing facilities with the data we take*
- *We could take more, but then we don't have enough computing power*
- *During long shutdowns, our CPUs are under-utilised (e.g., for MC production)*

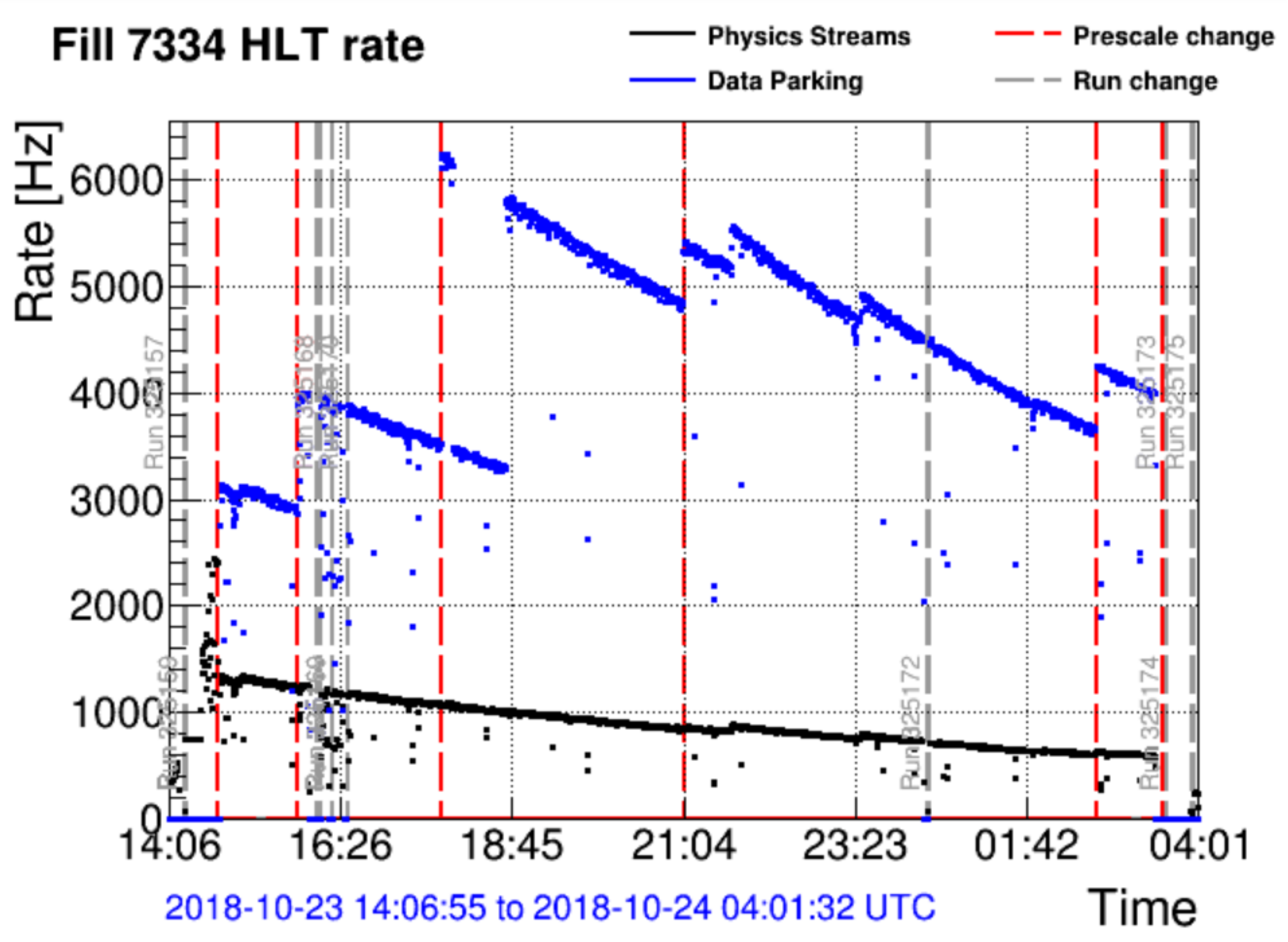
Data Parking



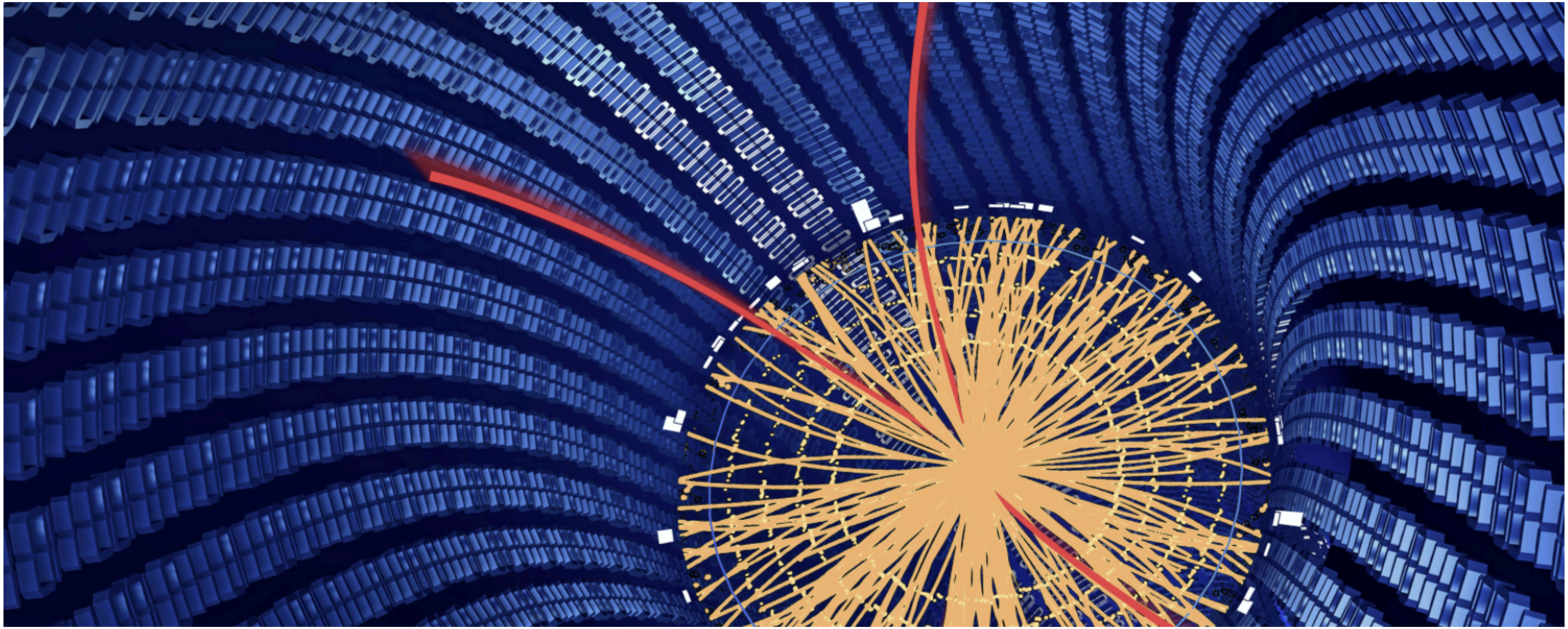
- Since Run 1, we open up the triggers during the last year before a LS
- We send the events on tape (cheaper than disk) and reconstruct them during the shutdown
- We target specific physics use cases
  - Inclusive/invisible Higgs in Run 1
  - B physics in Run 2

B-physics Parking

- Saturated available rate with displaced muon trigger
- Selected ~80% pure sample of BB mesons
    - one leg biased by trigger
    - one leg trigger-selection free
- Used to establish a physics program targeting the LHCb anomalies
    - Studies ongoing
    - More to come in Run3



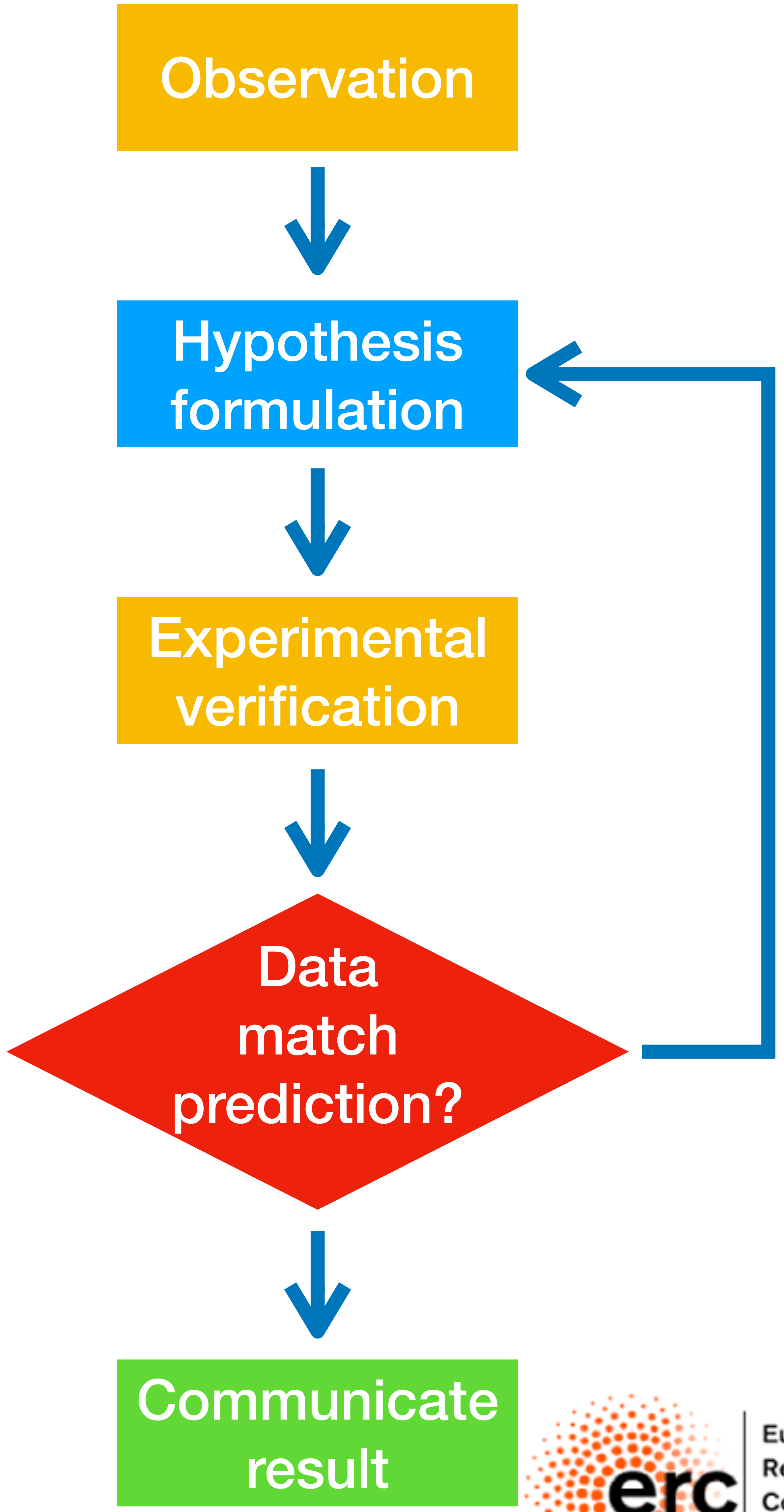
See D. Sperka's talk at HOW2019



Alternative Data Taking Strategies: anomaly detection

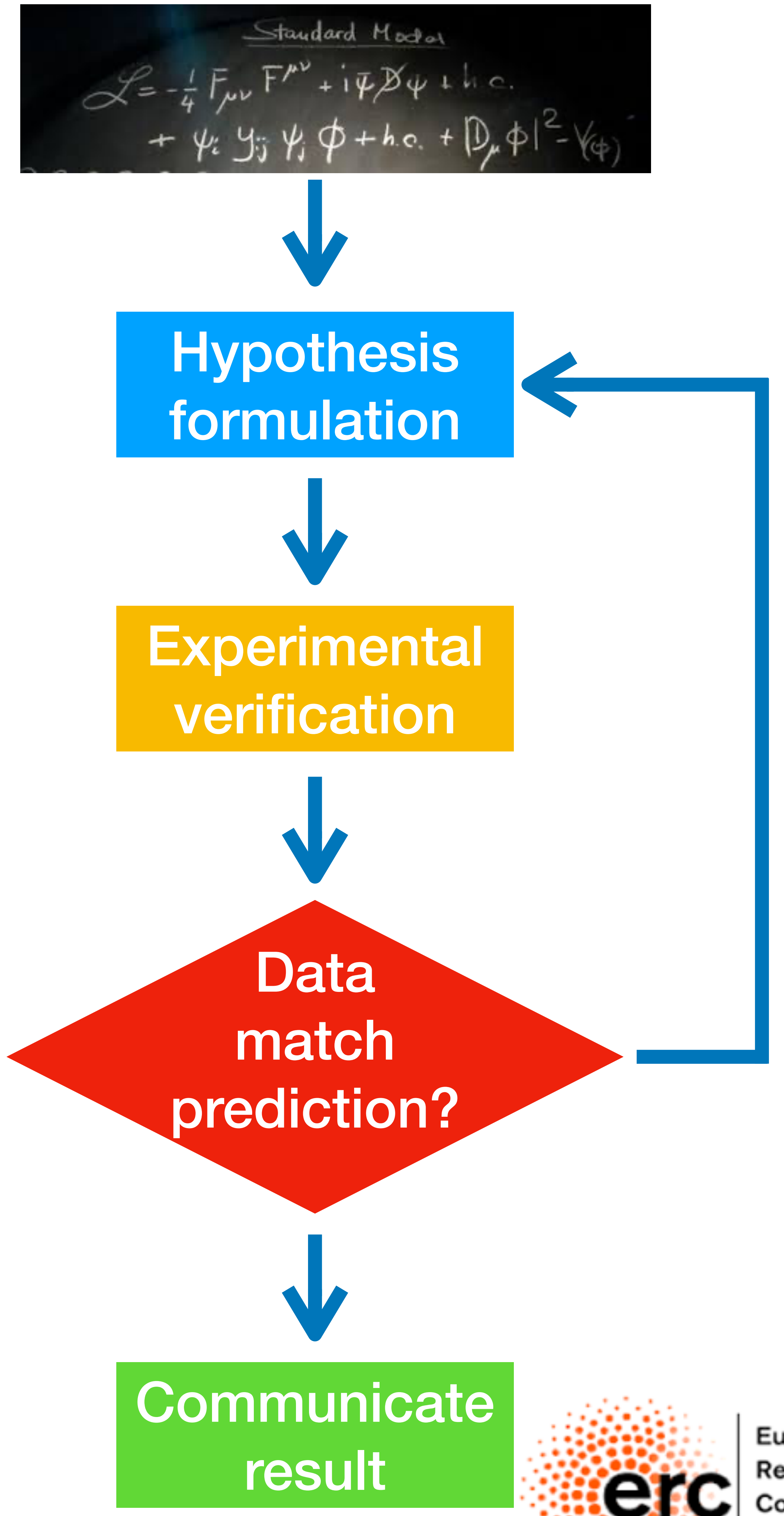
HEP searches in LHC era

- *The first step of the scientific method consists in observing nature*



HEP searches in LHC era

- The first step of the scientific method consists in observing nature
- In the last 40 years, our starting point has been the SM (which was put together from experimental information collected in the 70s)
- We are victim of our success:
  - We use data mainly to confirm our hypothesis
  - We have lost the value of learning from data
  - Not by chance, we totally endorsed blind analysis as the ONLY way to search

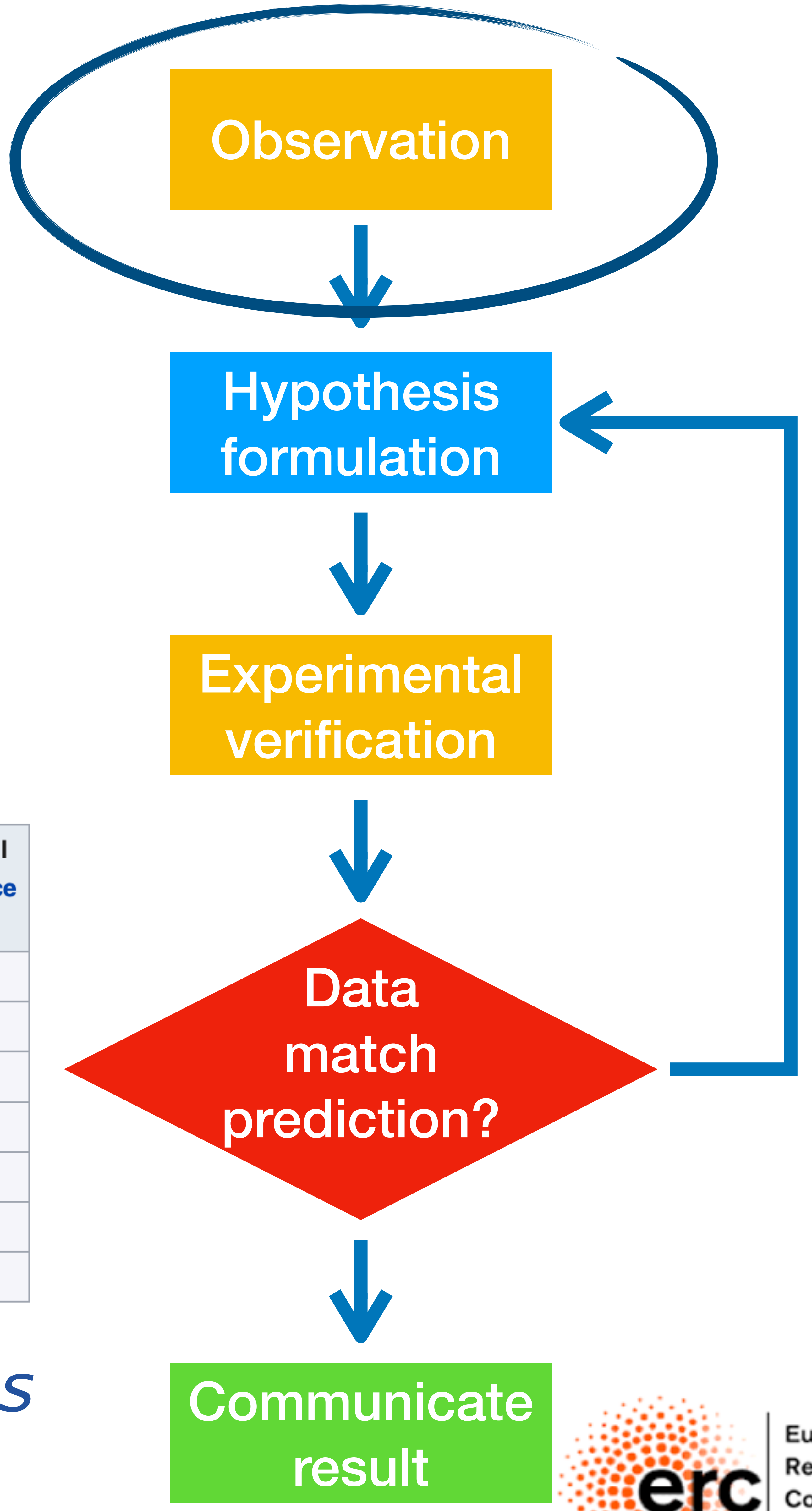


Learning from Data

- *Rather than specifying a signal hypothesis upfront, we could start looking at our data first*
- *Based on what we see (e.g., clustering alike objects) we could formulate a signal hypothesis*
- *EXAMPLE: star classification was formulated on empirical characteristics*

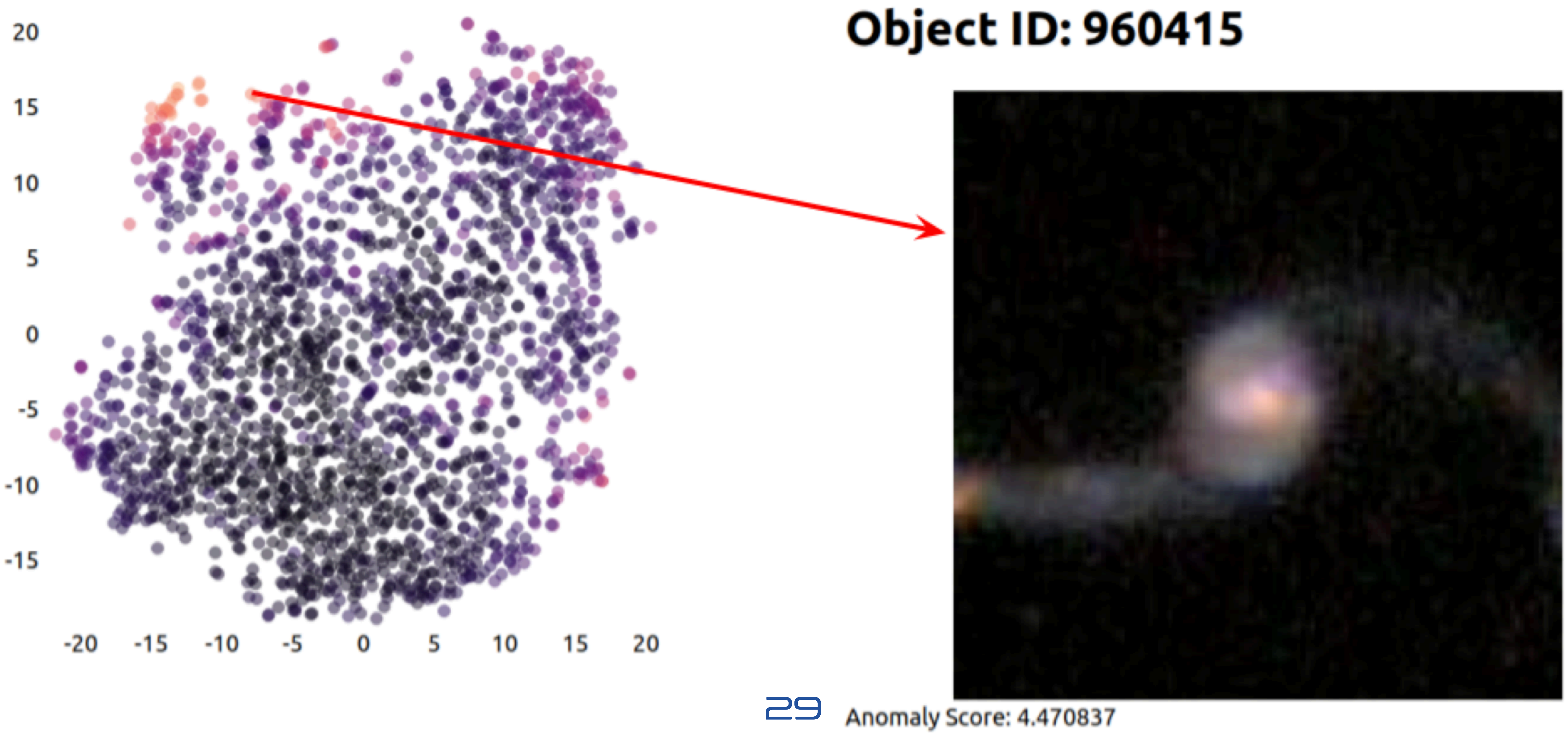
Class	Effective temperature[1][2]	Vega-relative chromaticity[3][4][a]	Chromaticity (D65)[5][6][3][b]	Main-sequence mass[1][7] (solar masses)	Main-sequence radius[1][7] (solar radii)	Main-sequence luminosity[1][7] (bolometric)	Hydrogen lines	Fraction of all main-sequence stars[8]
O	≥ 30,000 K	blue	blue	≥ 16 $M_\odot$	≥ 6.6 $R_\odot$	≥ 30,000 $L_\odot$	Weak	~0.00003%
B	10,000–30,000 K	blue white	deep blue white	2.1–16 $M_\odot$	1.8–6.6 $R_\odot$	25–30,000 $L_\odot$	Medium	0.13%
A	7,500–10,000 K	white	blue white	1.4–2.1 $M_\odot$	1.4–1.8 $R_\odot$	5–25 $L_\odot$	Strong	0.6%
F	6,000–7,500 K	yellow white	white	1.04–1.4 $M_\odot$	1.15–1.4 $R_\odot$	1.5–5 $L_\odot$	Medium	3%
G	5,200–6,000 K	yellow	yellowish white	0.8–1.04 $M_\odot$	0.96–1.15 $R_\odot$	0.6–1.5 $L_\odot$	Weak	7.6%
K	3,700–5,200 K	light orange	pale yellow orange	0.45–0.8 $M_\odot$	0.7–0.96 $R_\odot$	0.08–0.6 $L_\odot$	Very weak	12.1%
M	2,400–3,700 K	orange red	light orange red	0.08–0.45 $M_\odot$	≤ 0.7 $R_\odot$	≤ 0.08 $L_\odot$	Very weak	76.45%

- *Afterwords, the connection to physics properties (temperature) was understood*

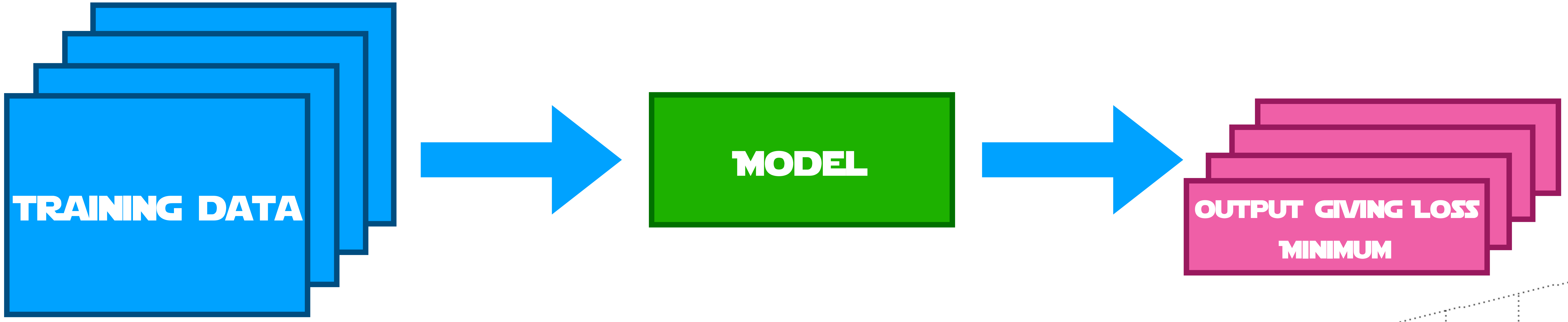


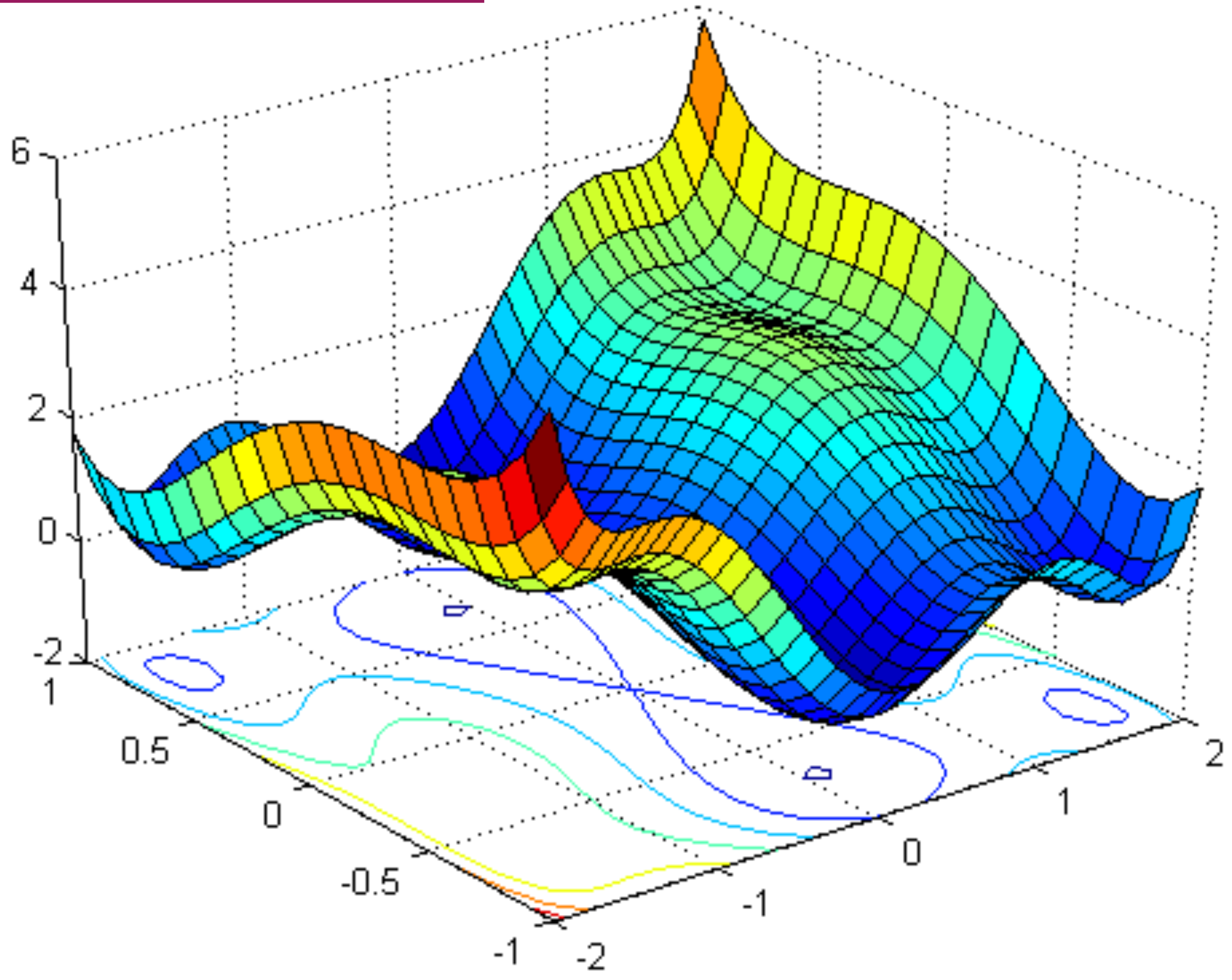
Learning from Anomalies

- Anomaly detection is one kind of data mining technique
  - One defines a metric of "typicality" to rank data samples
  - Based on this ranking, one can identify less typical events, tagging them as anomalies
  - By studying anomalies, one can make hypotheses on new physics mechanisms



Unsupervised learning





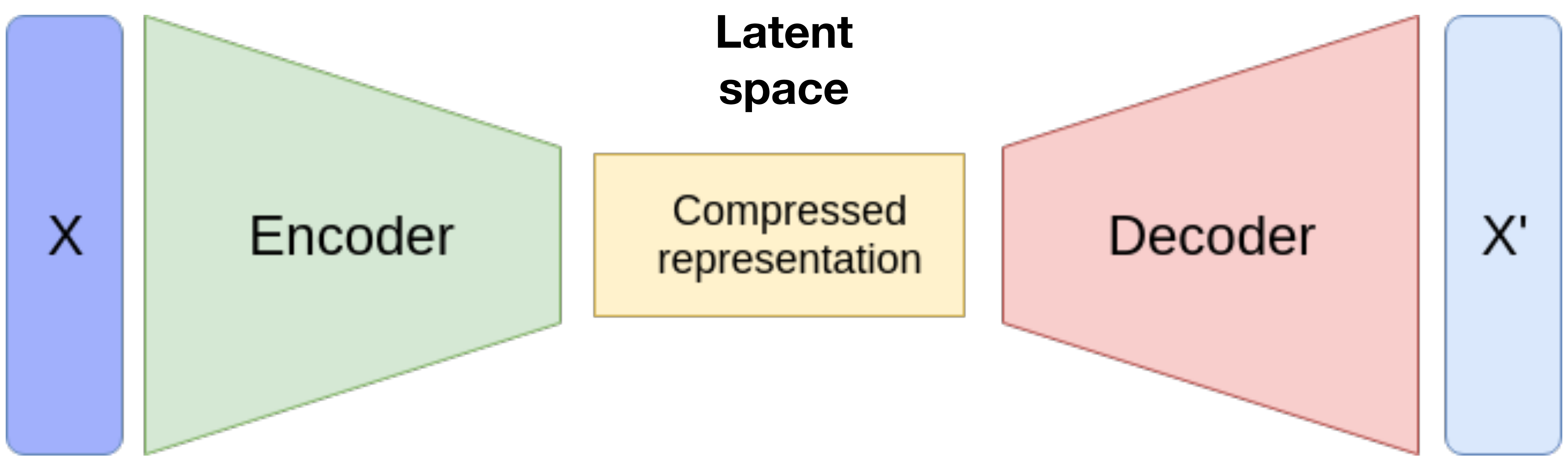
- A training dataset ***x***
- No target *y*
- A model providing an output *y* at the minimum of the loss
- A loss function of *x* and *y* specifying the task
  - e.g., clustering: group similar objects together

Autoencoders

- *Autoencoders are networks with a typical "bottleneck" structure, with a symmetric structure around it*
  - *They go from* $\mathbb{R}^n \rightarrow \mathbb{R}^n$
  - *They are used to learn the identity function as* $f^{-1}(f(x))$

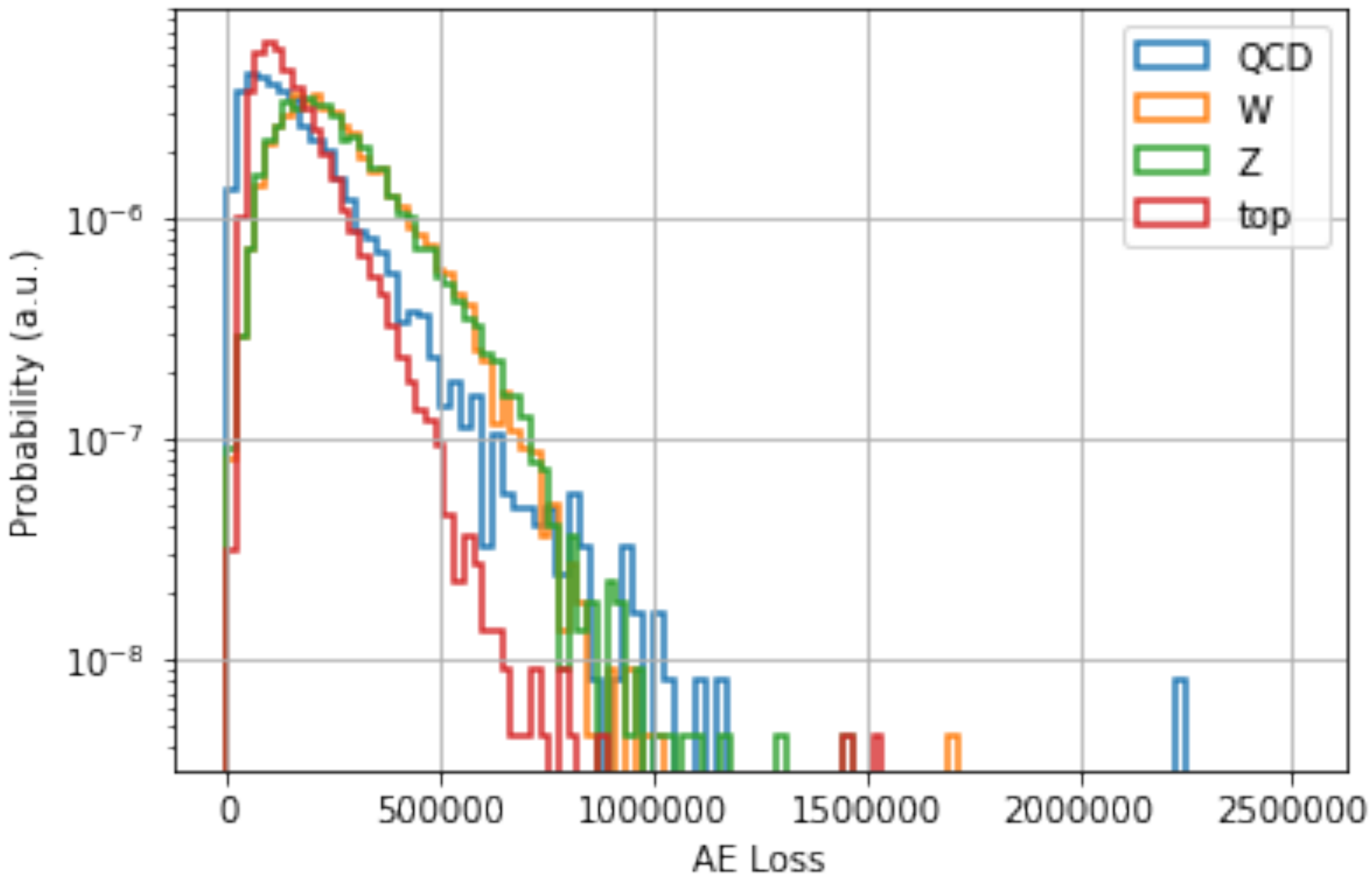
where $f: \mathbb{R}^n \rightarrow \mathbb{R}^k$ *and* $f^{-1}: \mathbb{R}^k \rightarrow \mathbb{R}^n$

- *Autoencoders are essential tools for unsupervised studies*



Anomaly detection

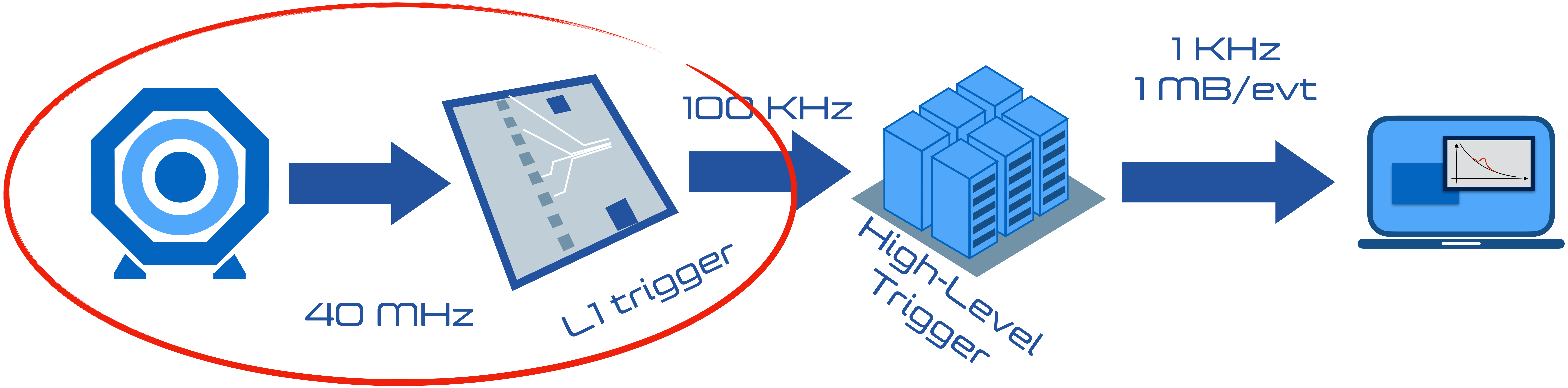
- *An autoencoder can reproduce new inputs of the same kind of the training dataset*
  - *The distance between the input and the output will be small*
- *If presented an event of some new kind (anomaly), the encoding-decoding will tend to fail*
  - *In this circumstance, the loss (=distance between input and output) will be bigger*



T. Heimel et al., https://arxiv.org/abs/1808.08979
M. Farina et al., https://arxiv.org/abs/1808.08992

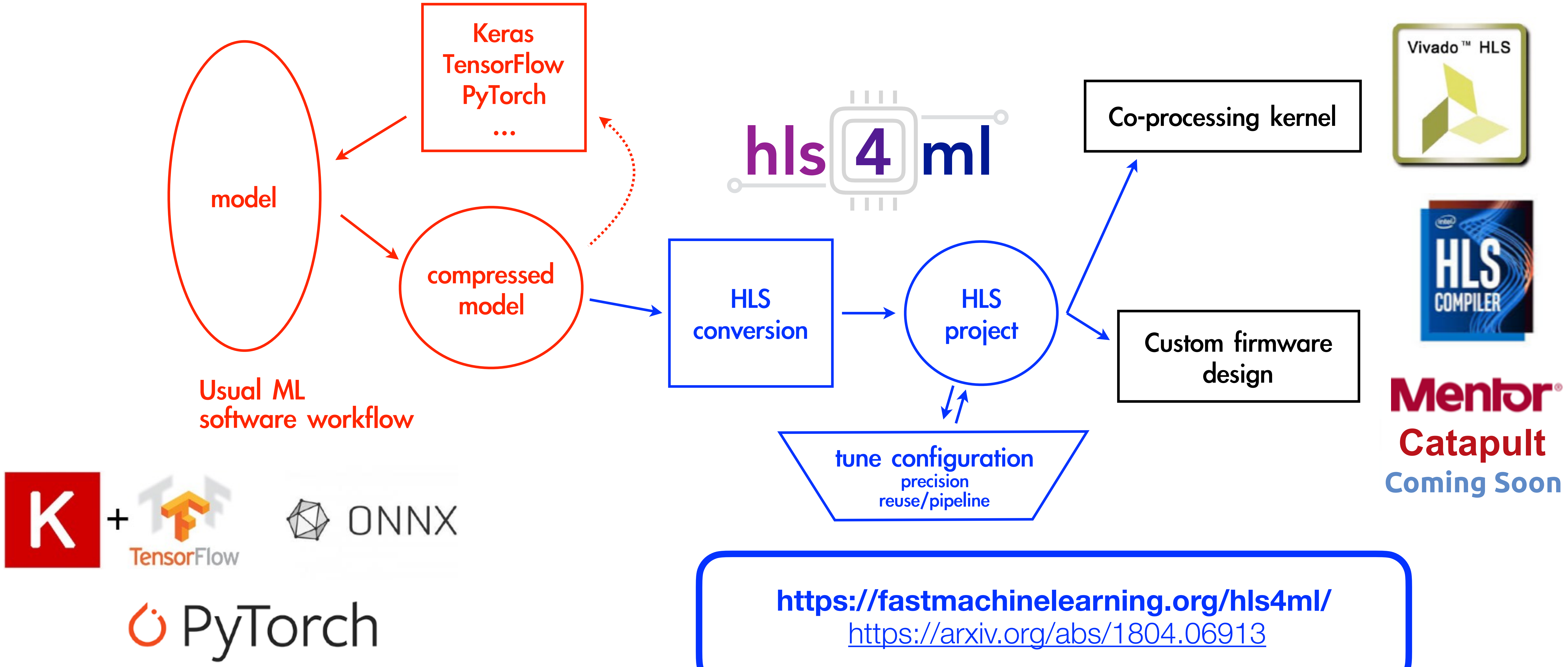
Offline might be too late

- *With 40M beam crossings/seconds and 1000 stored, we might just being writing the wrong events*
- *If we want to take action on a "plan B" path, this has to happen at the trigger*



HLS4ML: NN to FPGAs

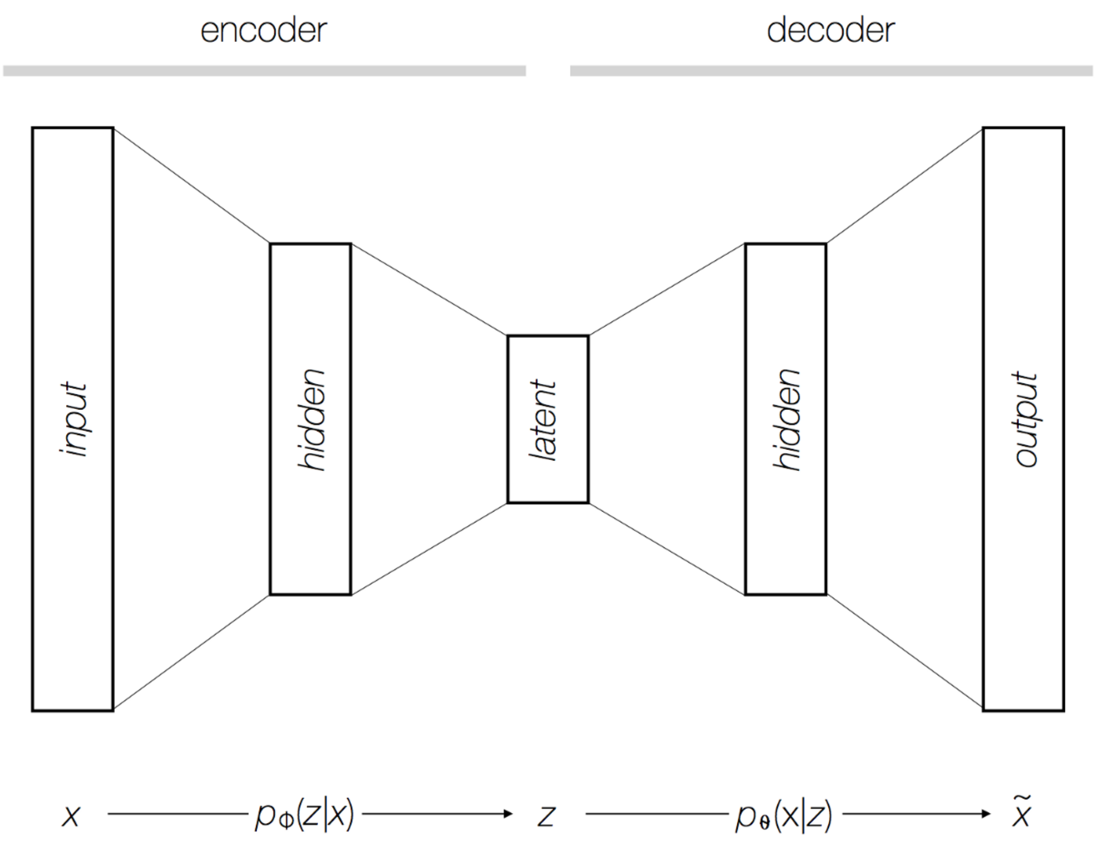
- HLS4ML aims to be this automatic tool
  - reads as input models trained on standard DeepLearning libraries
  - comes with implementation of common ingredients (layers, activation functions, etc)
  - Uses HLS softwares to provide a firmware implementation of a given network
  - Could also be used to create co-processing kernels for HLT environments

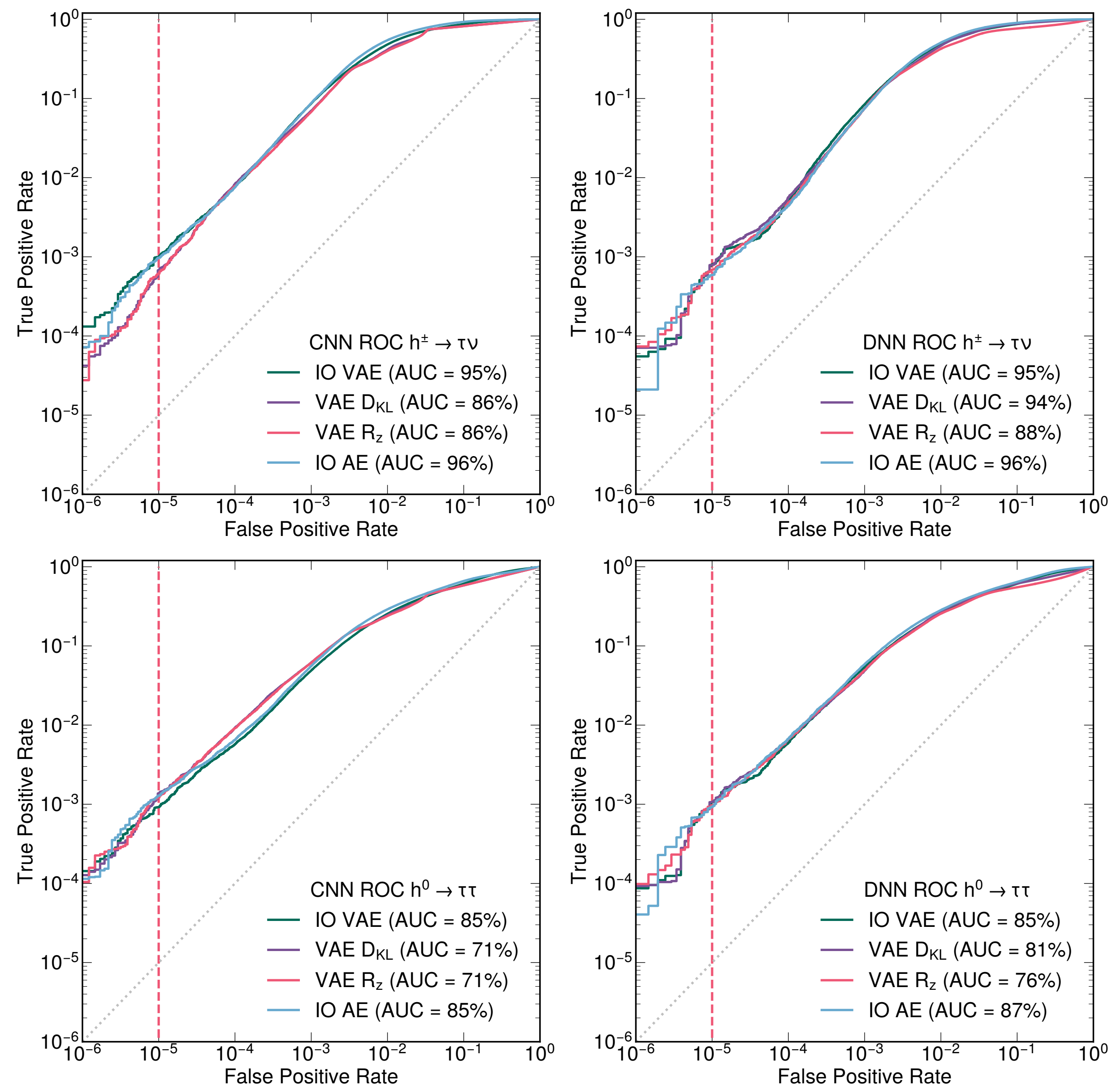


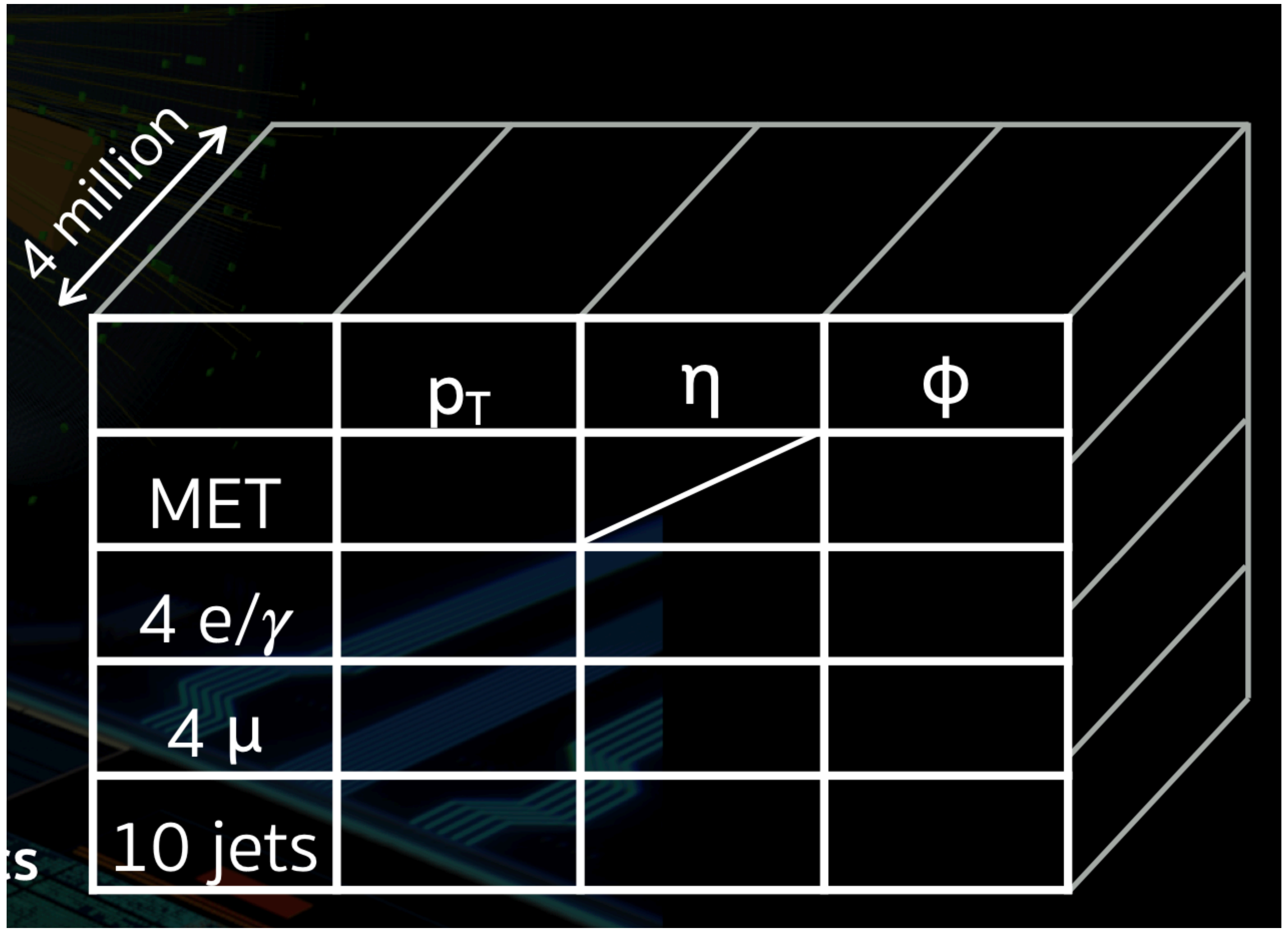
https://fastmachinelearning.org/hls4ml/
https://arxiv.org/abs/1804.06913

Anomaly Detection on FPGA

- *Autoencoders can provide event discrimination for anomaly detection, even with the little information available at L1*



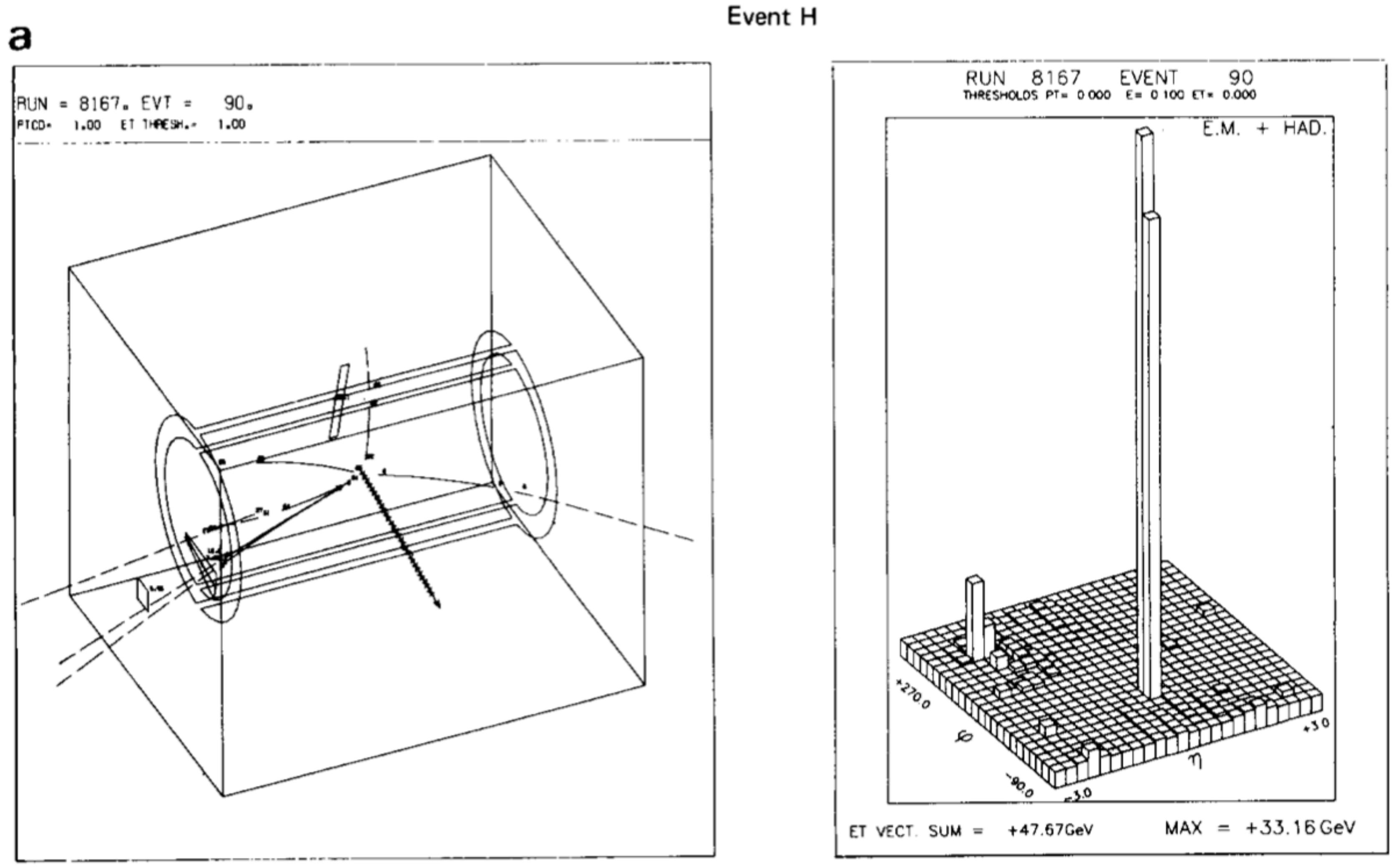




Model	DSP [%]	LUT [%]	FF [%]	BRAM [%]	Latency [ns]	II [ns]
DNN AE QAT 8 bits	2	5	1	0.5	130	5
CNN AE QAT 4 bits	8	47	5	6	1480	895
DNN VAE PTQ 8 bits	1	3	0.5	0.3	80	5
CNN VAE PTQ 8 bits	10	12	4	2	365	115

Anomaly Dataset

- The minimal final deliverable is a dataset of anomalous events
  - Could be used to look for recurrent topologies
  - By visual inspections, these "clusters" could inspire new searches on other (future) datasets
- Not very far away from how our field used to operate before big-data computing solutions took over



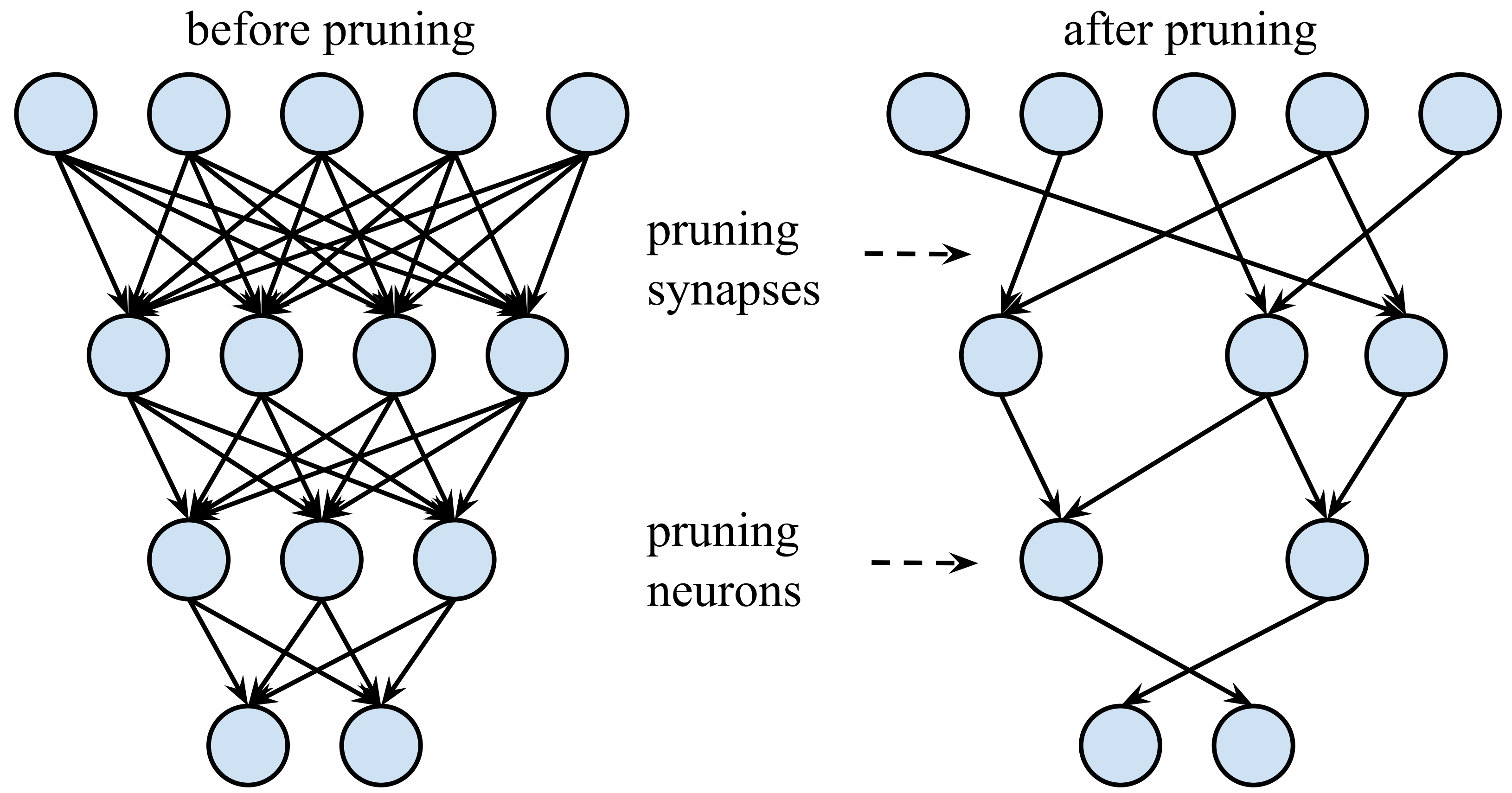
Much more than BSM program

- LHC experiments will deliver much more than this
  - More precise W and top mass measurements
  - EFT operator analyses from differential cross section measurements in Higgs, top, and EW processes
  - Broader reach to Flavor physics
  - More exiting Heavy Ion physics
- Stay tuned…

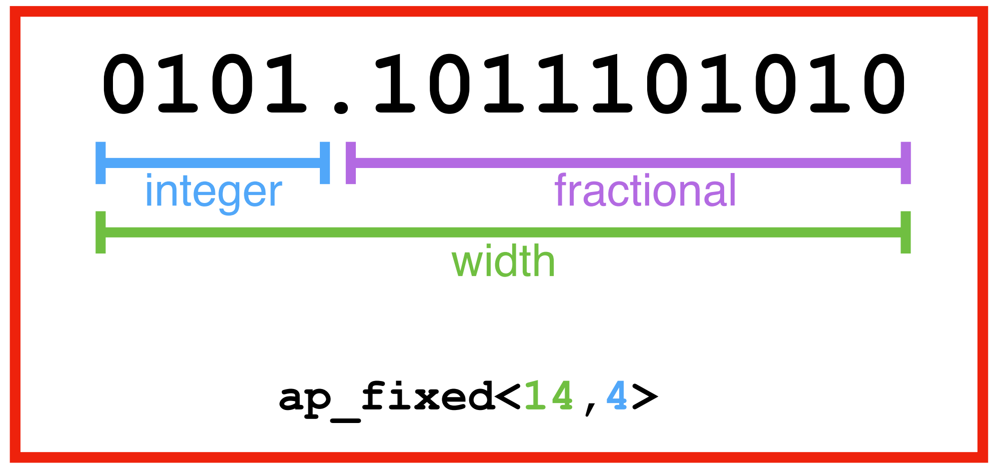
Backup Slides

Network compression techniques

- *Pruning: identify which part of the network is really relevant and remove the rest (makes network smaller)*



- *Quantization: use limited precision for numerical representations (save resources)*



- *Reuse: dilute the network inference on multiple clock cycles, trading off resource needs for latency*