GEN Tutorial Part. II

Sitian Qian (PKU)

CMS China Winter Camp 2025

2025/01/18

So, You are an Expert of MG5 now!

- And your analysis team is going to ask you to take care of the MC :)
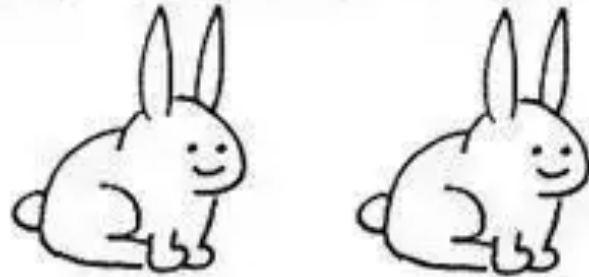
  - How to do that?

FIRST OF ALL, MG5 IS DEFINITELY NOT THE END OF THE DAY

VBF production of higgs boson and decaying into 2 muons
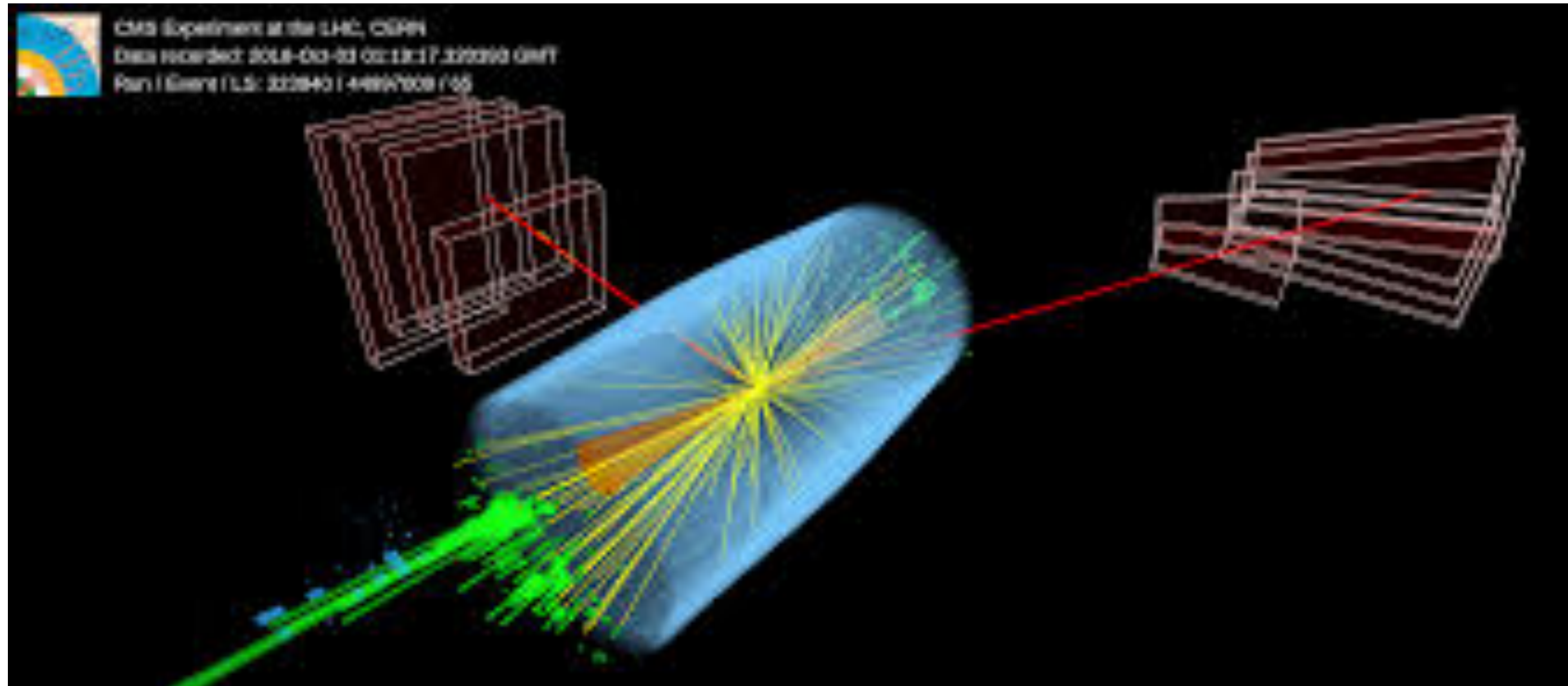
MG5 syntax is:

import model sm-lepton_masses
generate p p > h > mu+ mu- j j
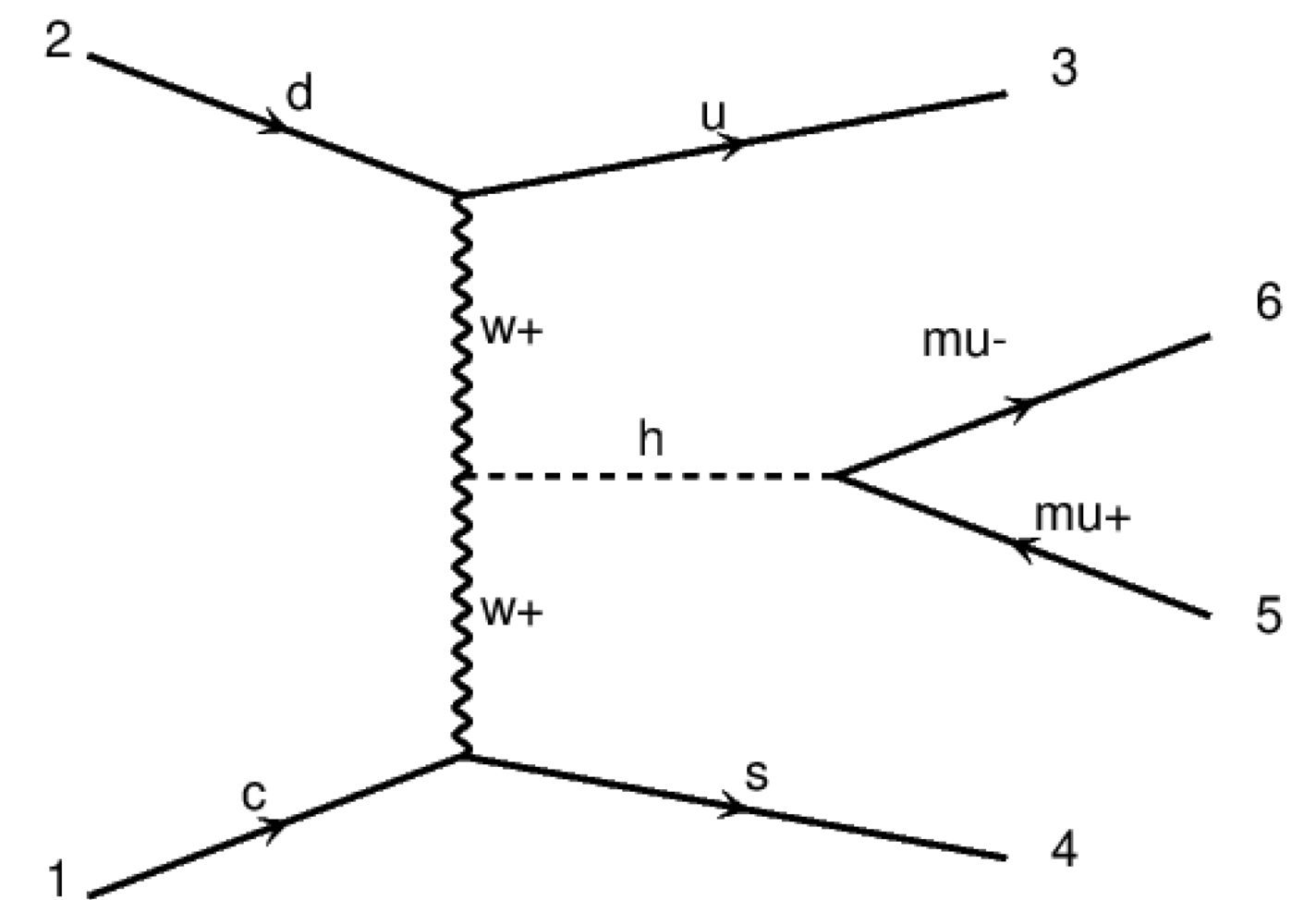
Try generate events and
count **how many particles**
are there in the LHE file :)

Hadronization of parton shower particles and further decay

Parton shower: the evolution of the particles from Hard process

Double parton scattering
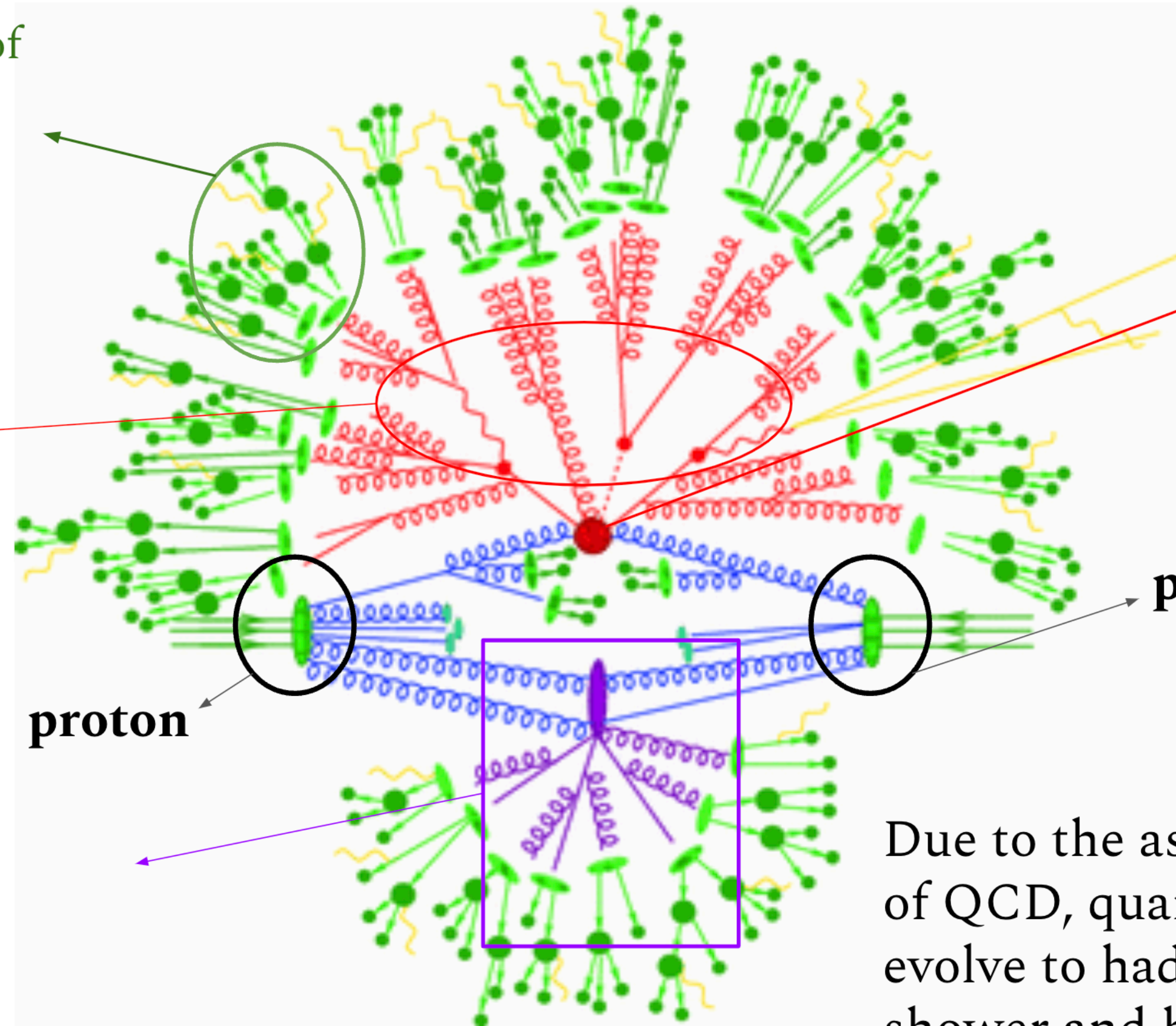
Hard process, this is what people usually interested in.

MadGraph gives you this only!

proton

proton

Due to the asymptotic freedom of QCD, quark and gluon will evolve to hadron, i.e., the parton shower and hadronization.

Generator (***Matrix Element*** modeling) usage breakdown based on legacy Run2 dataset
Pythia8 mostly chosen for ***parton shower*** and ***hadronization***

Num. of **Requests**

■	madgraph
■	pythiaOnly
■	powheg
■	amcatnlo
■	madgraphMLM
■	evtgen+pythia
■	amcatnloFXFX
■	other generators
■	sherpa
■	mcfm
■	powhegMiNNLO
■	herwig

11%
6.47%
3.65%
3.61%
2.84%
2.38%
1.27%
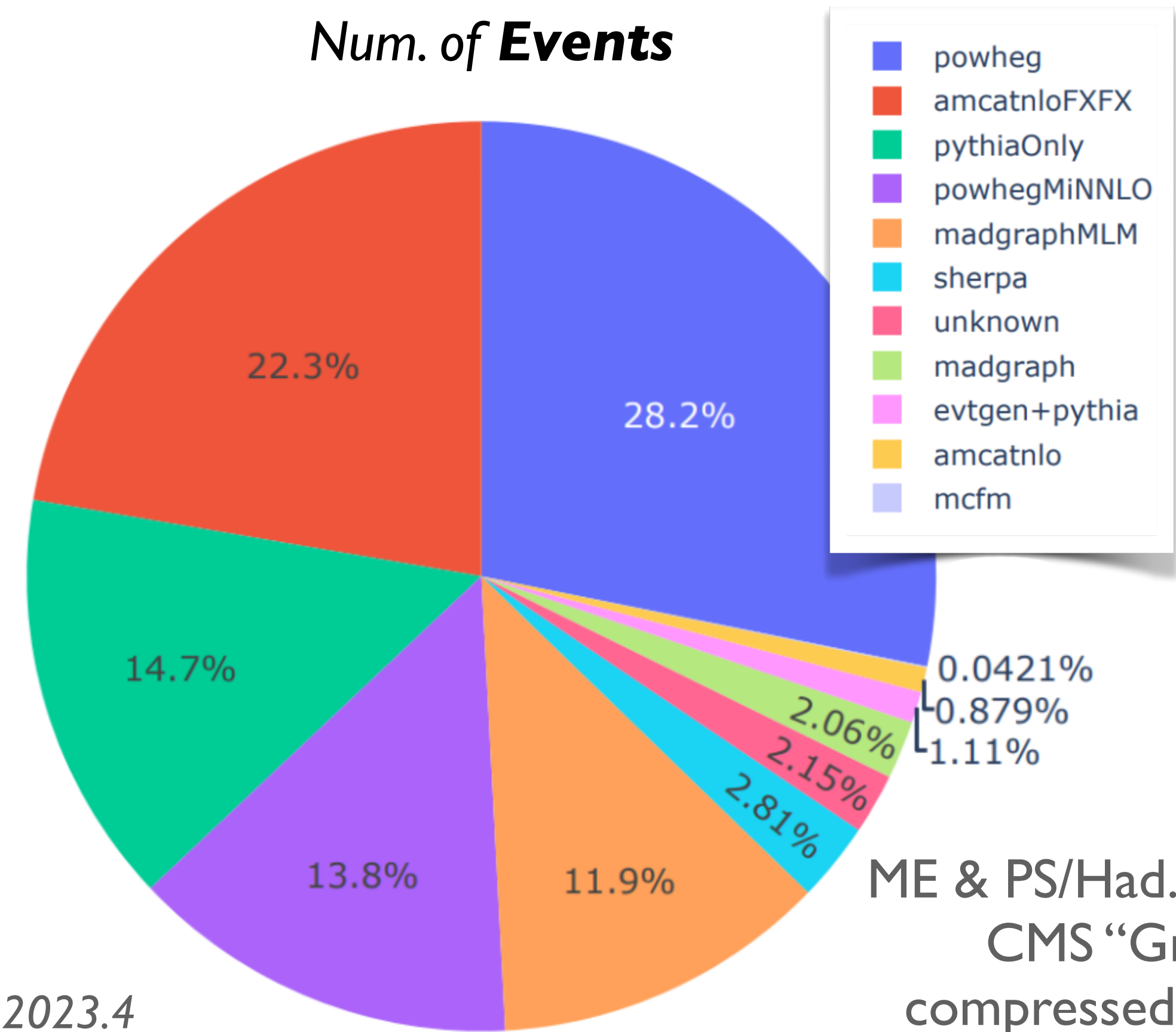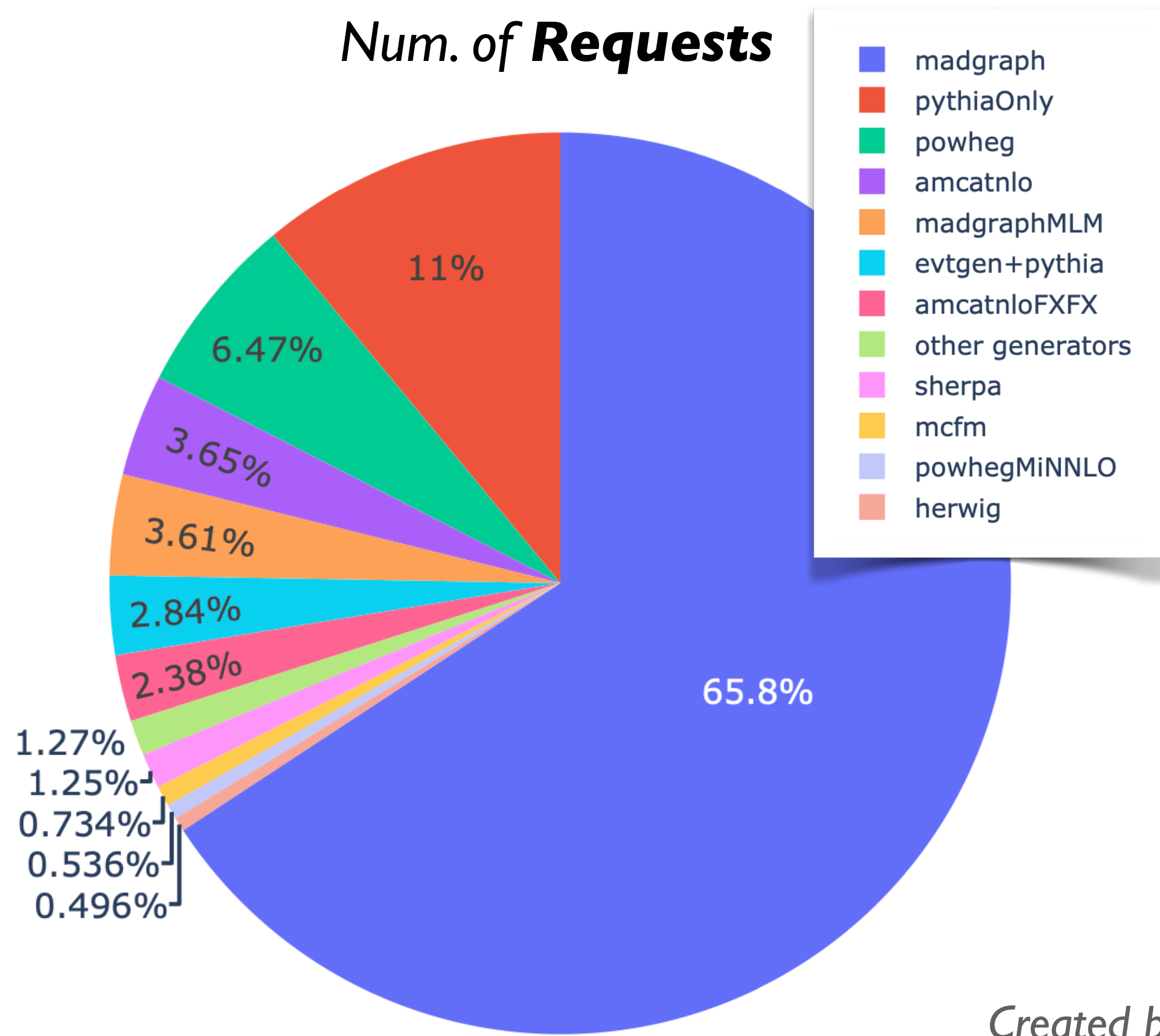1.25%
0.734%
0.536%
0.496%
65.8%

Created by 2023.4

Num. of **Events**

■	powheg
■	amcatnloFXFX
■	pythiaOnly
■	powhegMiNNLO
■	madgraphMLM
■	sherpa
■	unknown
■	madgraph
■	evtgen+pythia
■	amcatnlo
■	mcfm

28.2%
22.3%
14.7%
13.8%
11.9%
2.81%
2.15%
2.06%
0.0421%
0.879%
1.11%

ME & PS/Had. factorization:
CMS "Gridpack":
compressed tarball with
precompiled ME grids
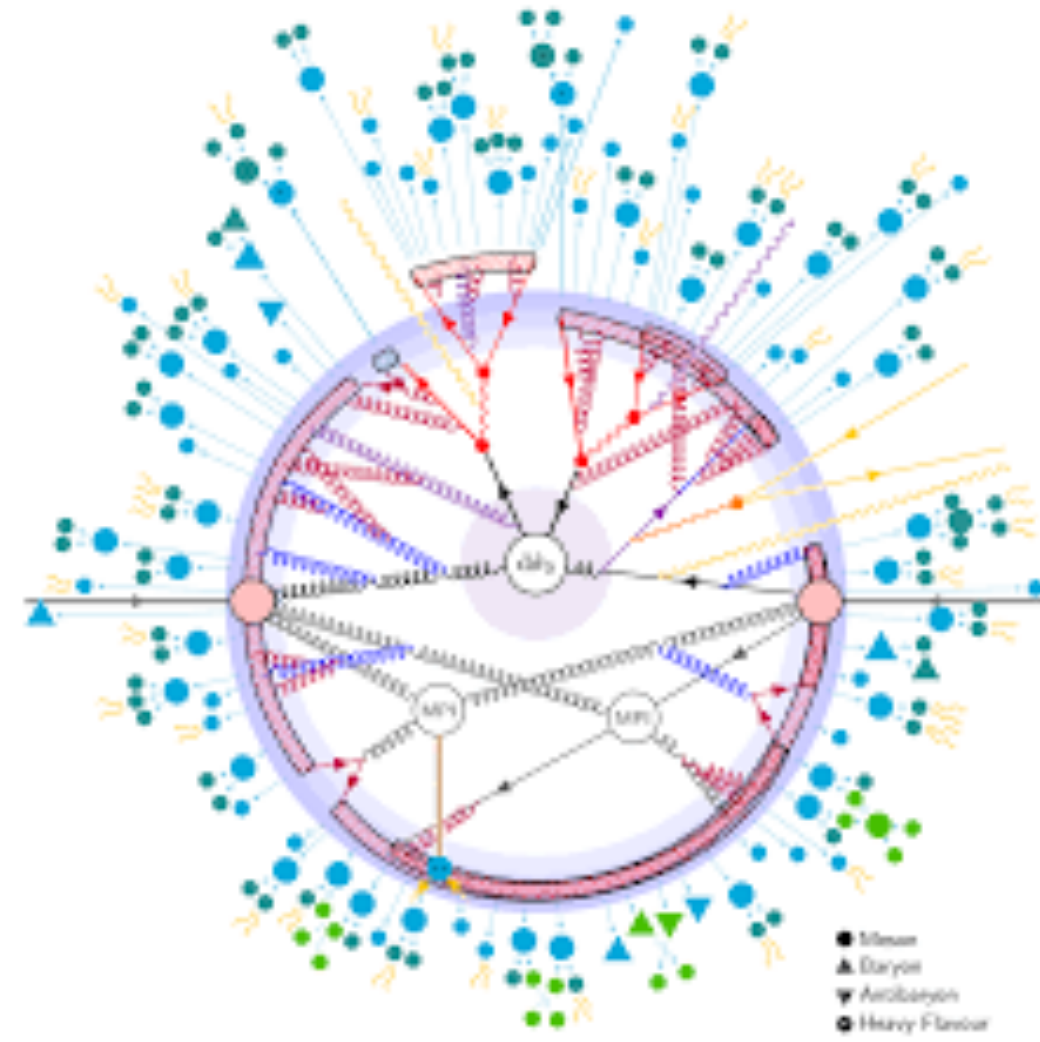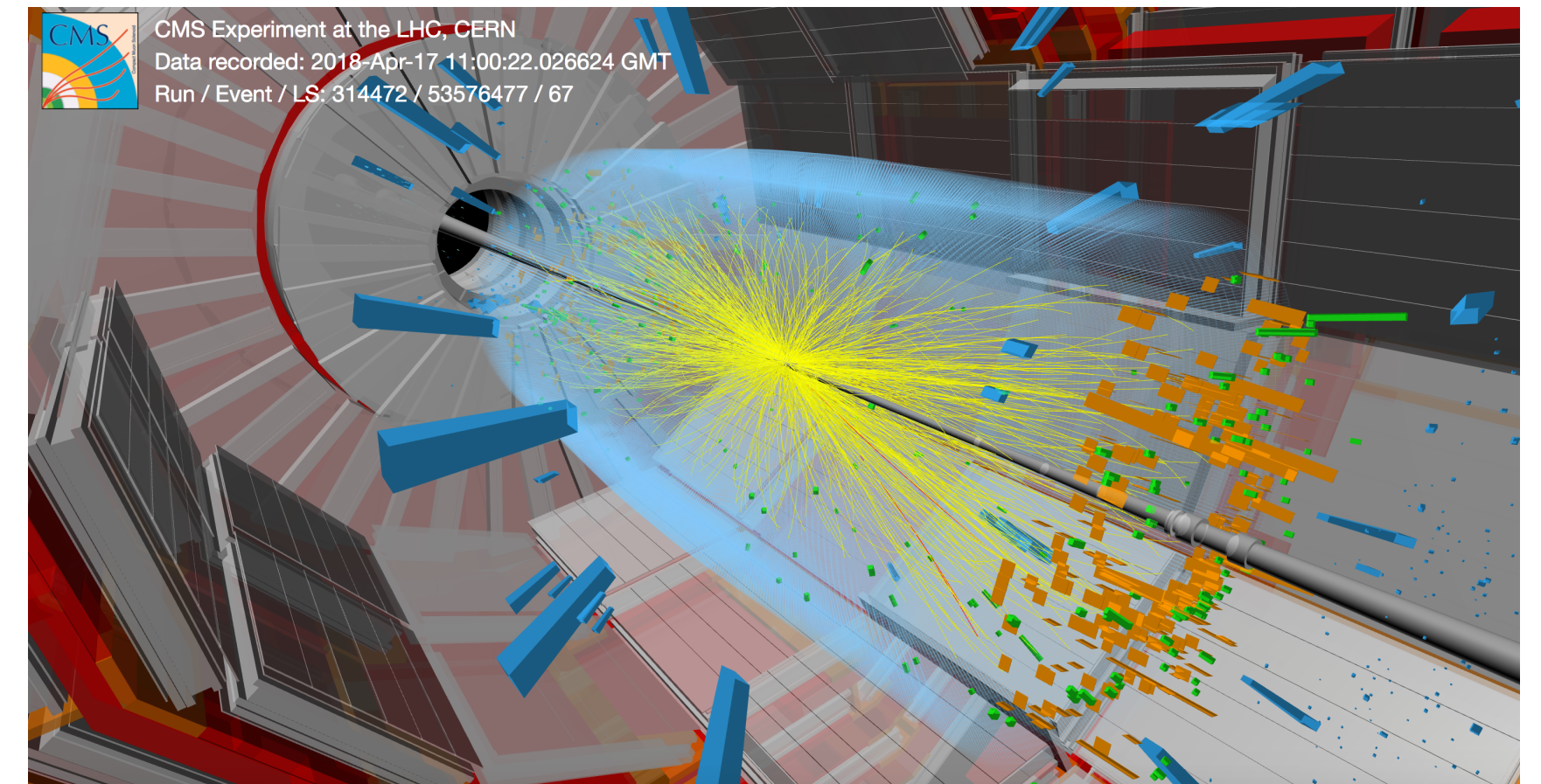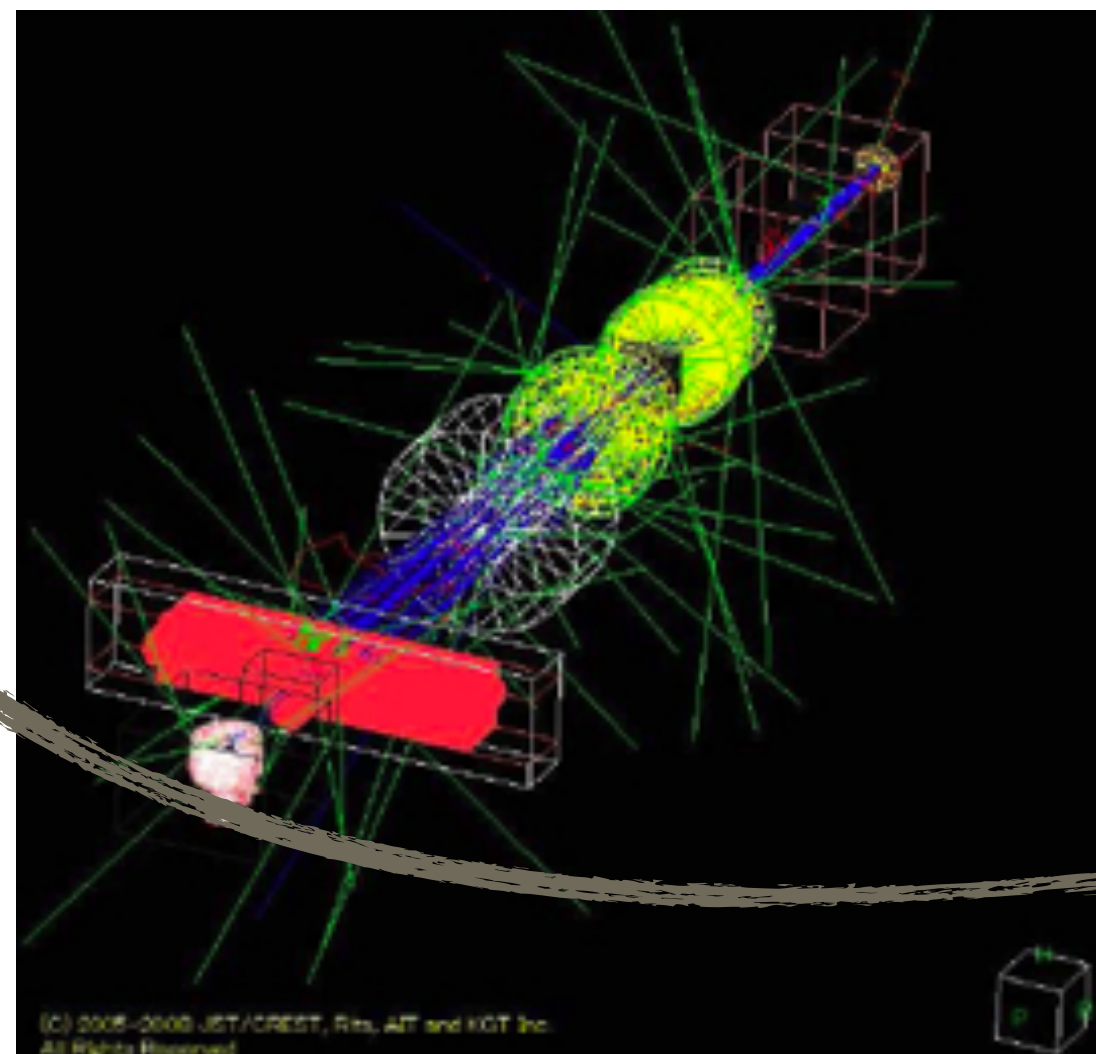
You need to simulate the detector response
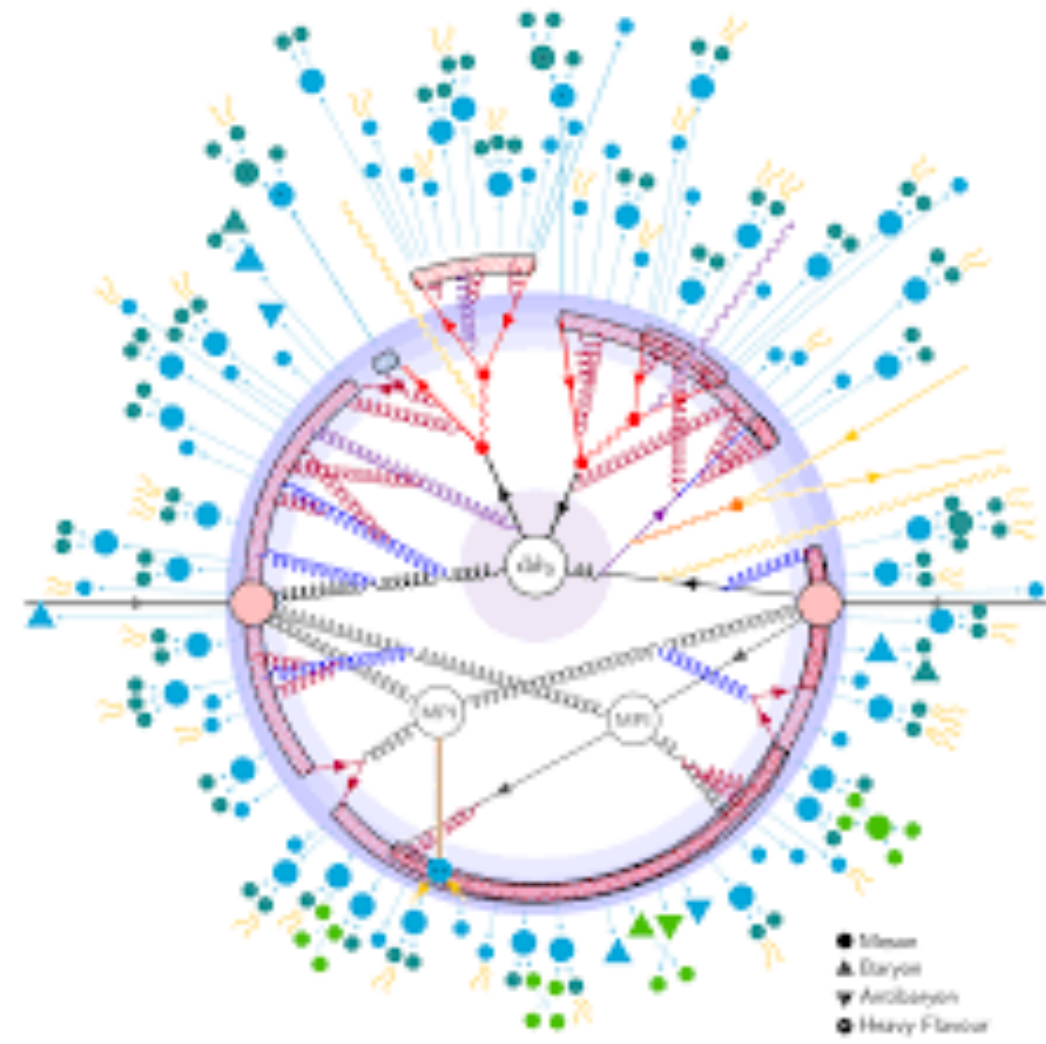
You need to simulate the detector response
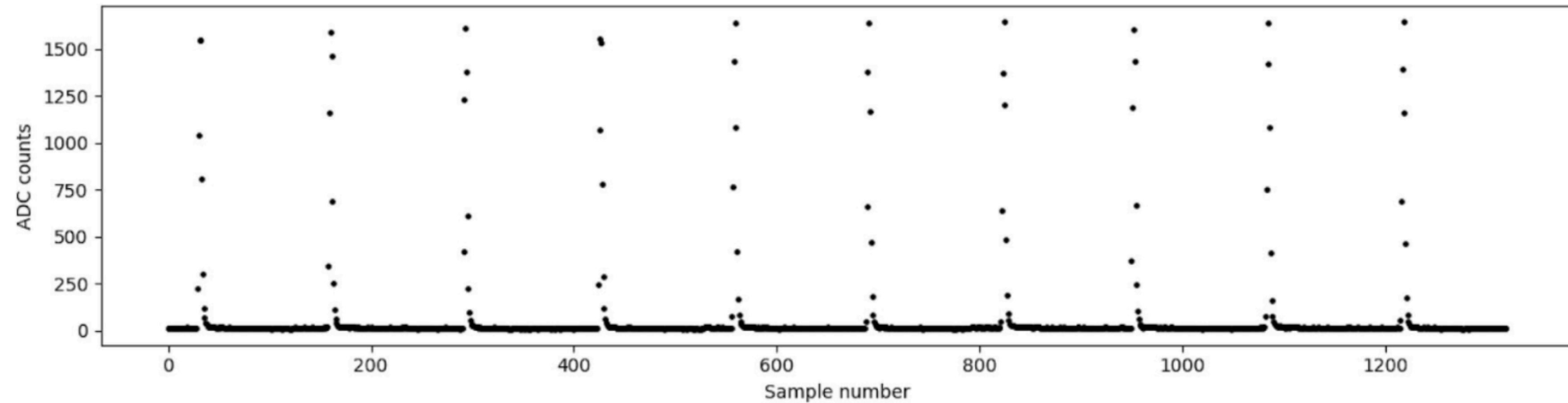
And detectors' output are digital signals

Figure 5: ECAL signals observed during the Nov 2022 test beam campaign, during which the full Phase-2 readout chain was tested from crystals to backend readout board.

- *L1 trigger: local, hardware based, on FPGA, @experiment site*
- *HLT: local/global, software based, on CPU, @experiment site*
- *Offline: global, software based, on CPU, @CERN T0*
- *Analysis: user-specific applications running on the grid*

40 MHz → L1 trigger → 100 KHz → High-Level Trigger → 1 KHz / 1 MB/evt →
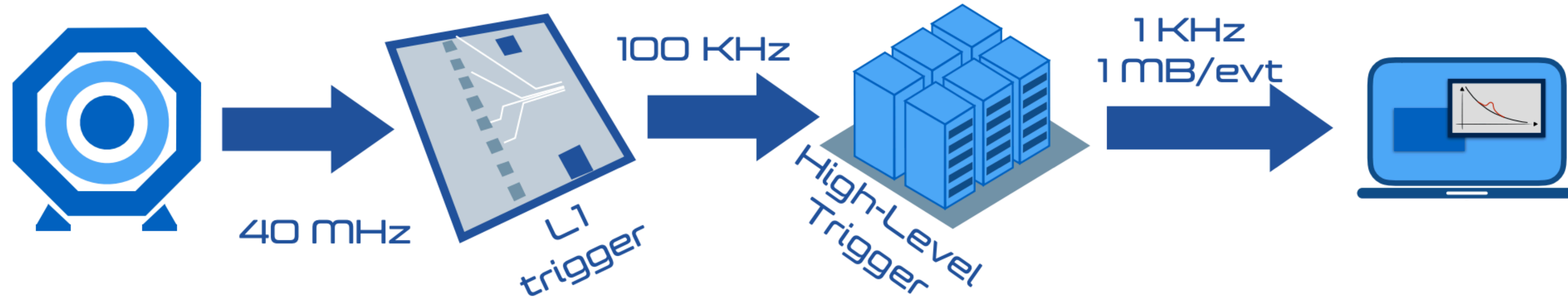
When you select (trigger on) a signature, you could collect everything happening in that short time window

"Pile Up"

How to model?

Hint

*You trigger on a relatively rare signature to **select** events of interest*

▸ *L1 trigger: local, hardware based, on FPGA, @experiment site*

▸ *HLT: local/global, software based, on CPU, @experiment site*

▸ *Offline: global, software based, on CPU, @CERN T0*

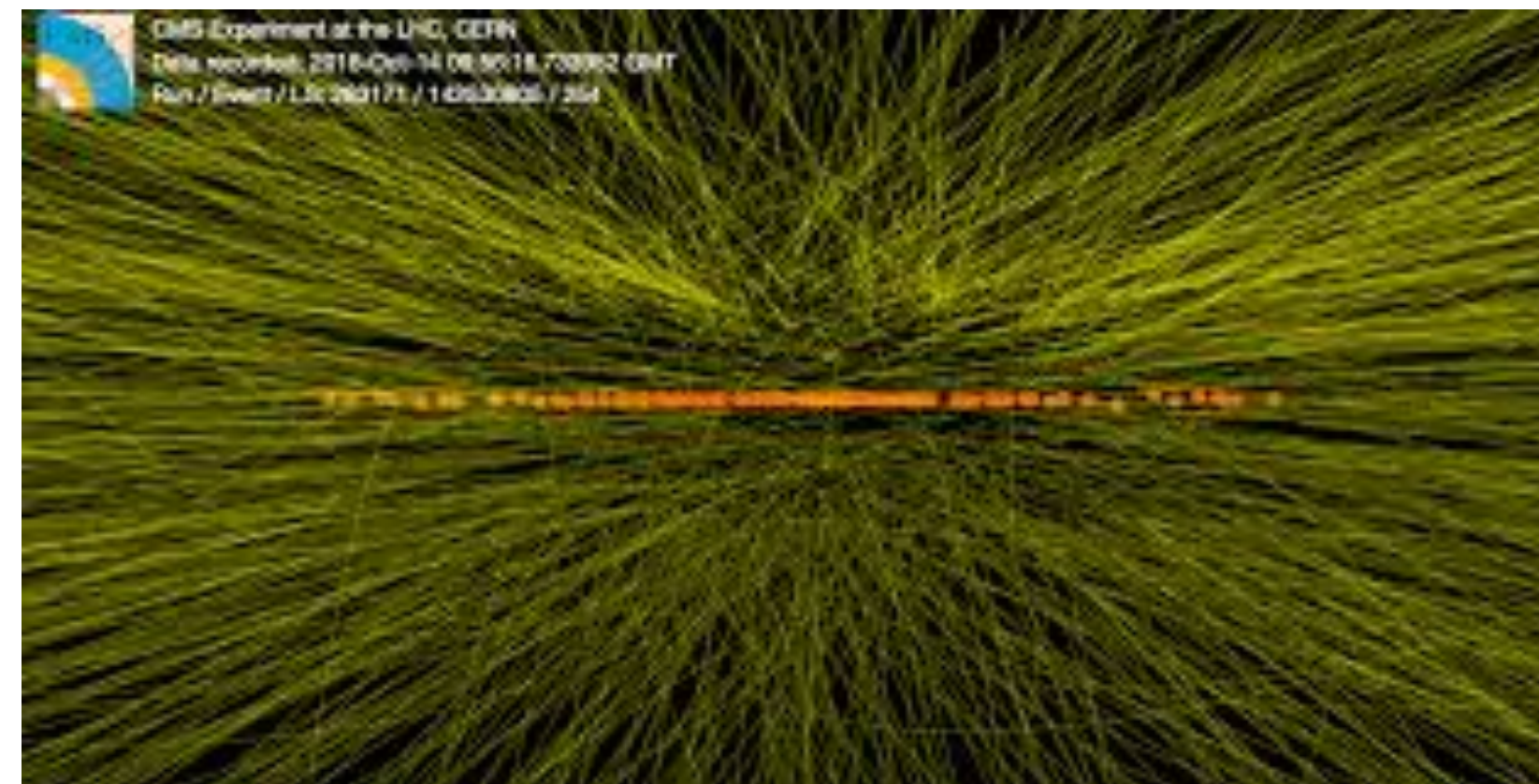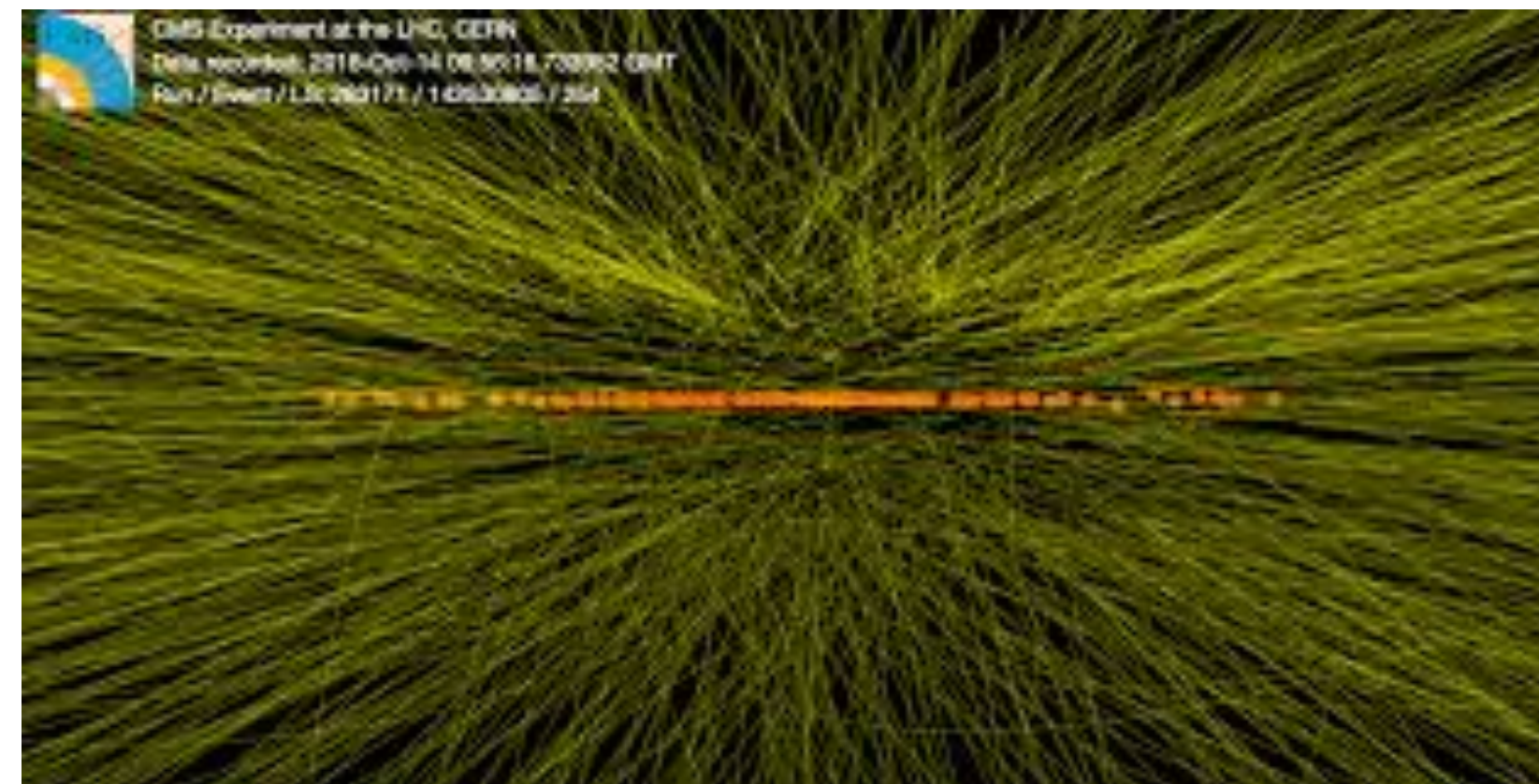▸ *Analysis: user-specific applications running on the grid*

40 MHz   L1 trigger   100 KHz   High-Level Trigger   1 KHz 1 MB/evt
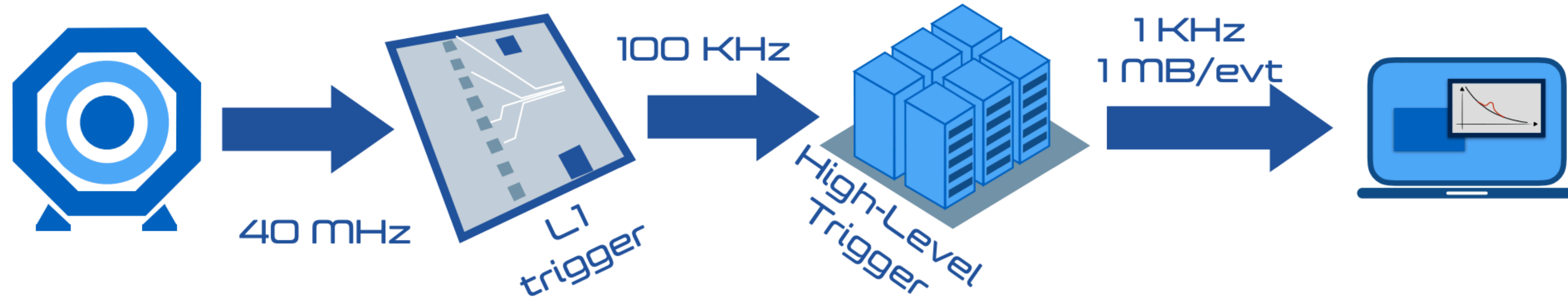
When you select (trigger on) a signature,
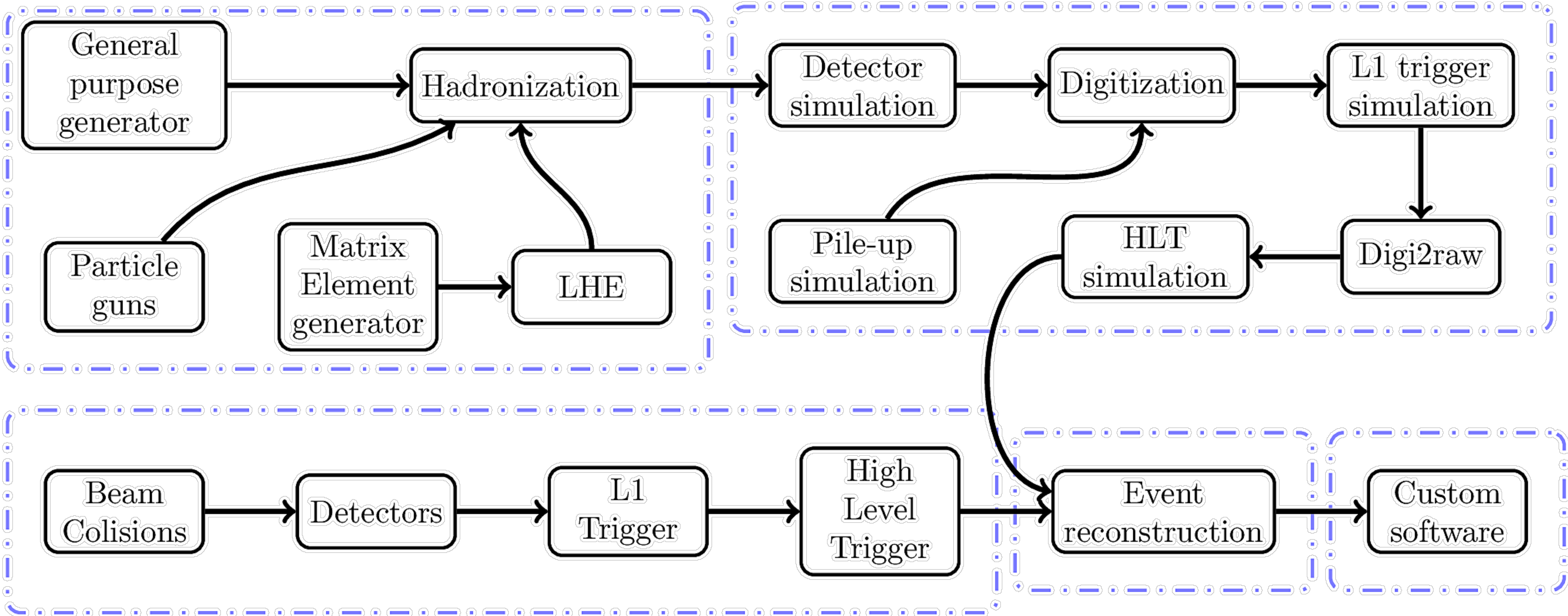you could collect everything happening in that short time window

"Pile Up"

How to model?

Model:

*Through a **mixing** of "most probably happening" events*

9

From CMS open data

So a complete chain of CMS simulation contains

From CMS Physics datasets and monte carlo validation (PdmV)
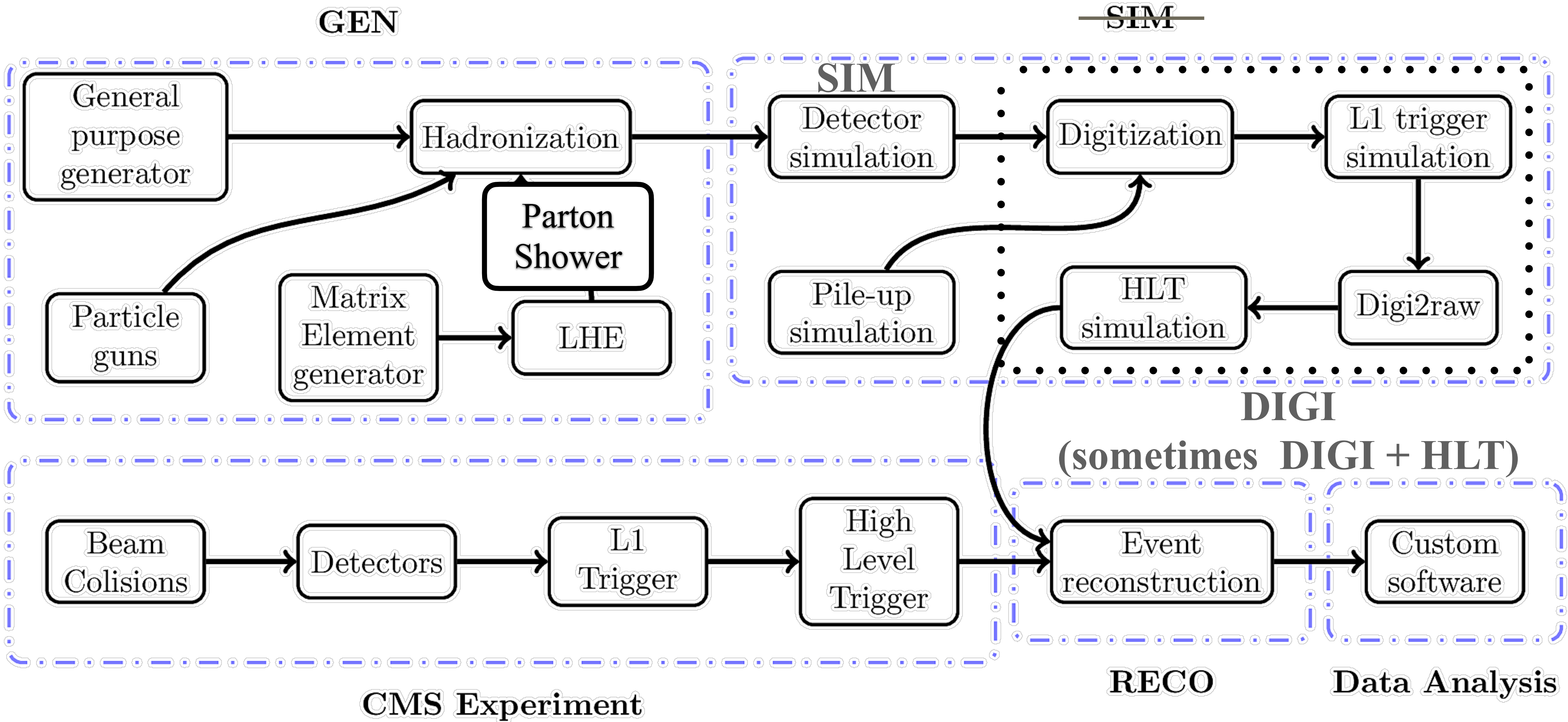
ONE MORE THING: WHAT IS GRIDPACK

- Did you recall all the "compilation messages"?
  - That means if you just use a "script" to run MG5, it needs compilation
  - Hence, CMS (semi-)invented the concept of gridpacks
    - Packing the pre-compiled MG5 output directory to grid

SOUNDS FAMILIAR...?

NOW HANDS ON:

- Small examples on how to generate gridpacks

- Small examples on how to use gridpacks

- Bonus:

  - Event display examples

  - McM/Grasp/...

THANKS!