Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Flavor physics - searches for rare decays at CMS (UC2.2.2)

Marco Buonsante, Federica Simone, Rosamaria Venditti (INFN e Università/Politecnico di Bari)

Spoke2 WP2 Meeting, 21 Maggio 2024

Motivations

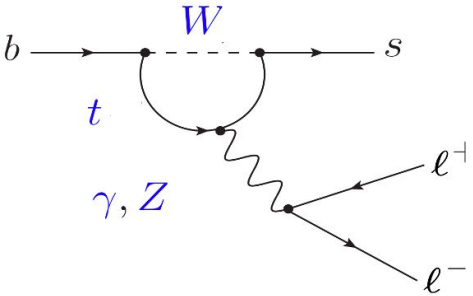
(Heavy) flavor physics: study B and D meson decays and look for new physics effects

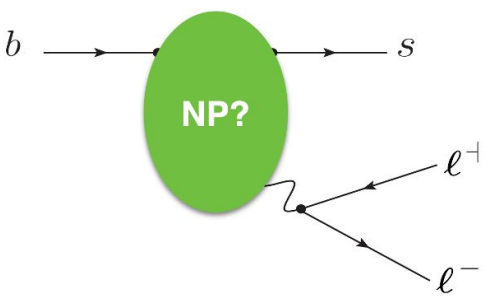
Rare processes - wish list:

- suppressed in the SM
- precise theoretical predictions
- sensitive to new physics
- experimentally accessible

Examples

- FCNC transitions
  e.g. $b \to s\ell^+\ell^-$, also helicity suppressed
- Decays forbidden by (accidental) symmetries of the SM: LUV, LFV and LNV



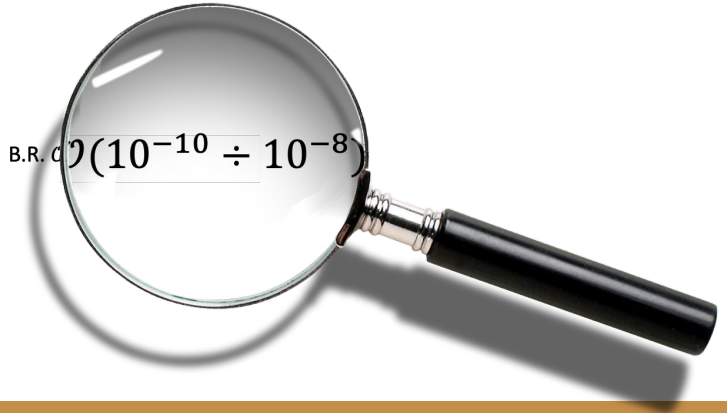


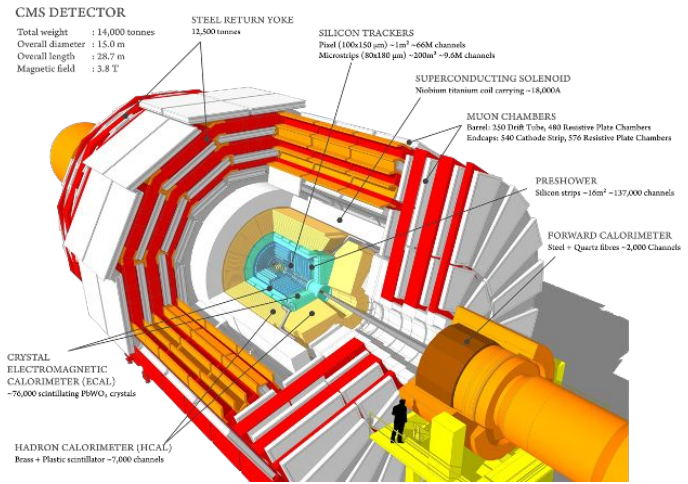
Motivations

(Heavy) flavor physics: study B and D meson decays and look for new physics effects

Rare processes - wish list:

- suppressed in the SM
- precise theoretical predictions
- sensitive to new physics
- **experimentally accessible**

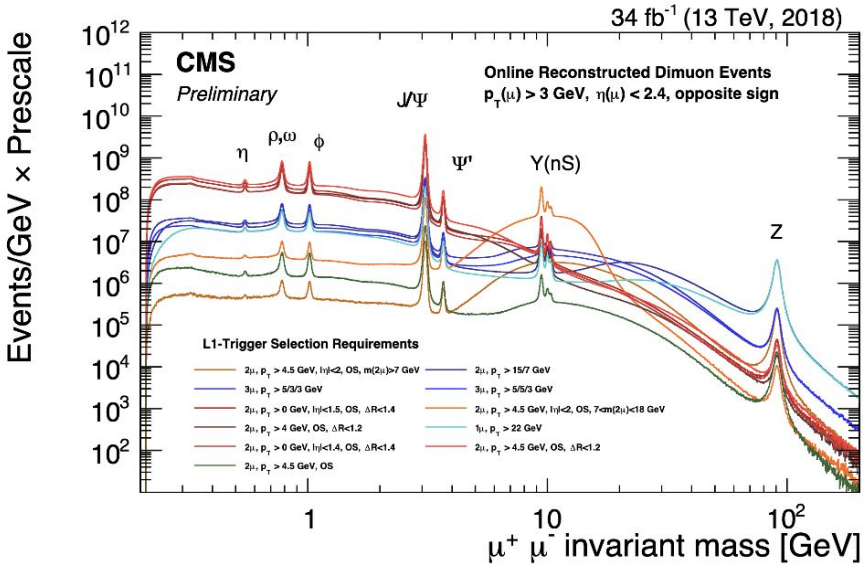




Triggers

Experimental signature: Low-pT muons → need careful trigger strategy

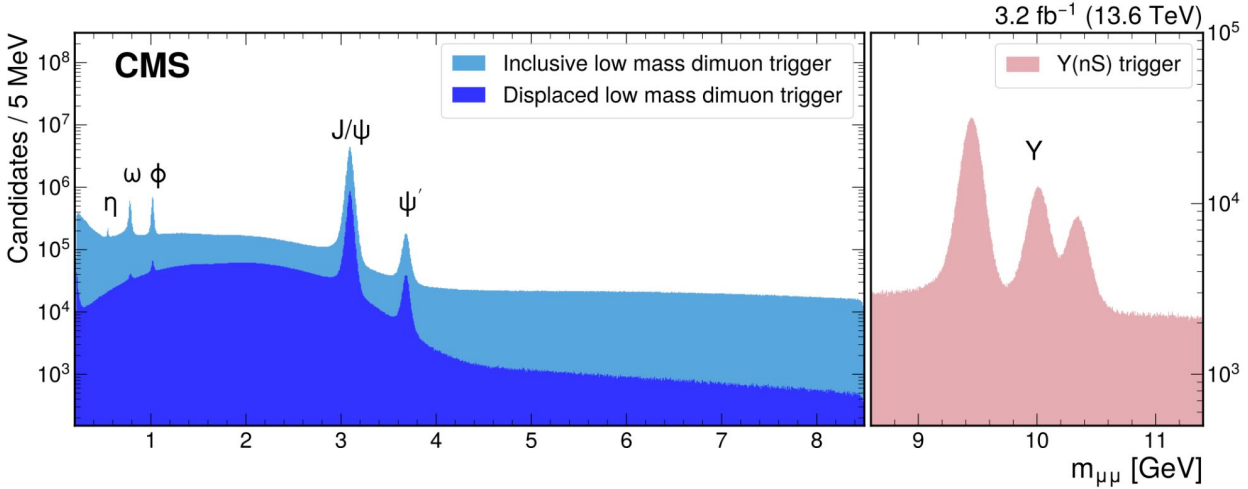
In **Run2**: combination of dimuon triggers with different complementary selections



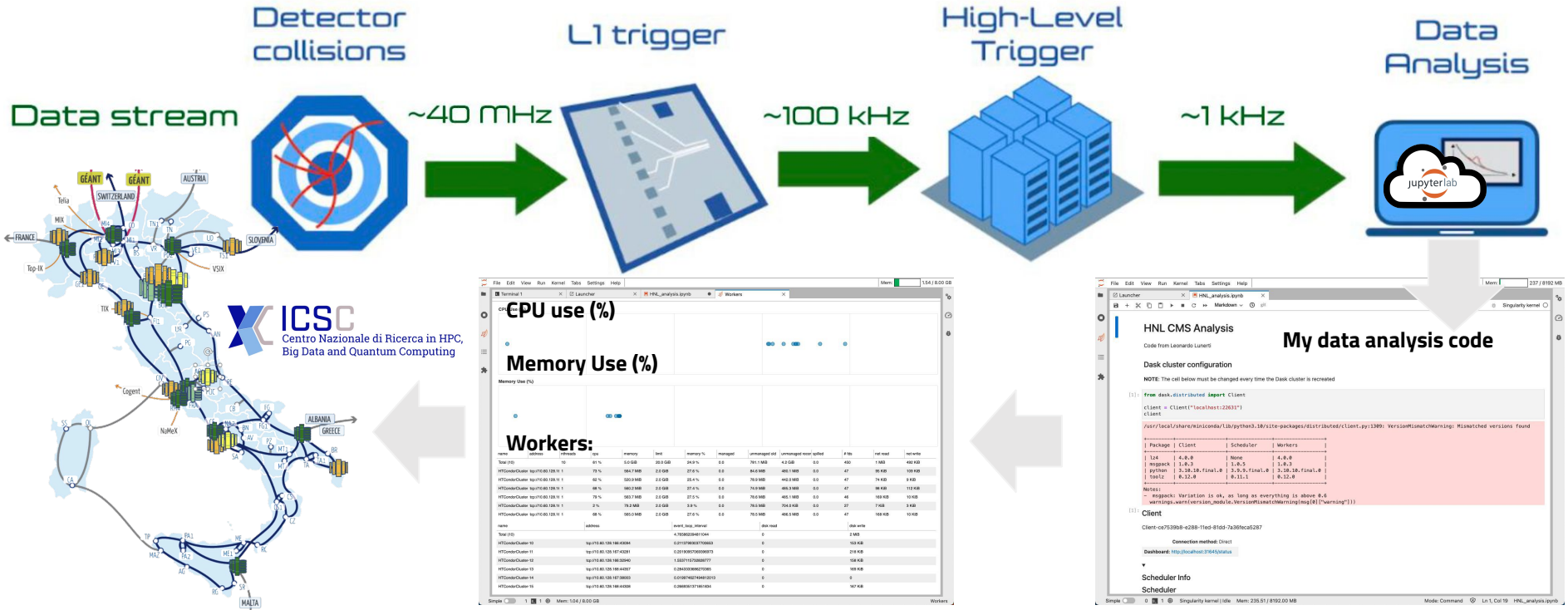
Triggers

Experimental signature: Low-pT muons → need careful trigger strategy

In **Run3**: "inclusive" dimuon trigger →dataset **5 times larger, parked for delayed reconstruction**



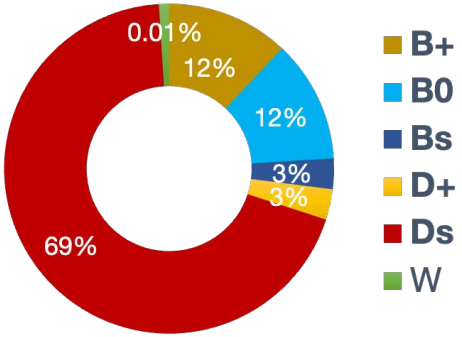
HEP data analysis with ICSC

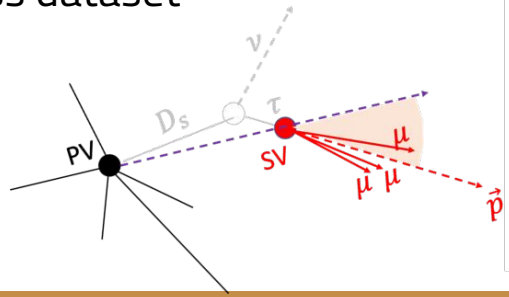


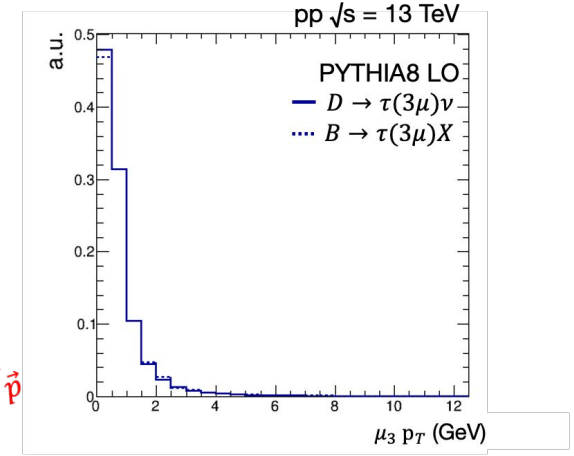
The physics use-case

Search for Lepton Flavor Violating decays tau -> 3mu

- Tau leptons produced by D and B decays with very low momenta
- Run2 analysis recently published
  - (Phys. Lett. B 853 (2024) 138633)
- Run3 analysis ongoing
  - Uses "inclusive" L1 dimuon triggers to seed custom HLT path
  - ParkingDoubleMuonLowMass dataset



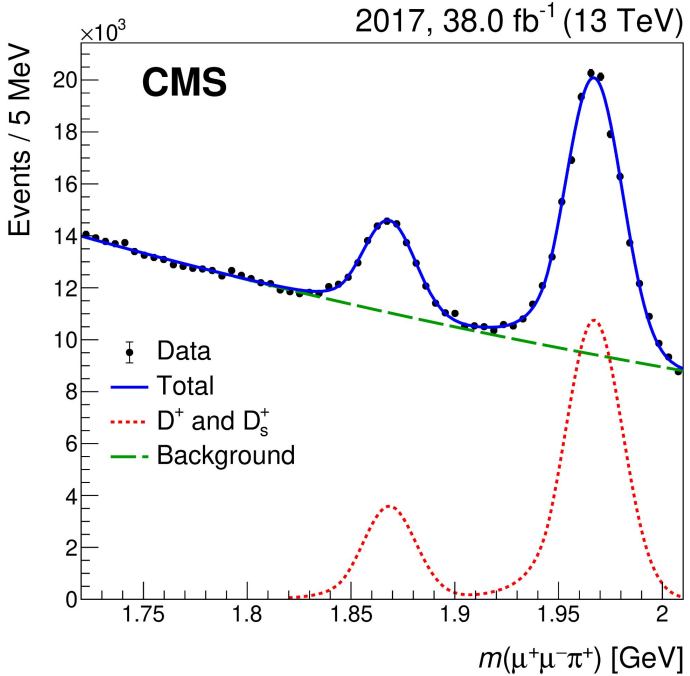




The physics use-case

Search for Lepton Flavor Violating decays tau -> 3mu

- Tau leptons produced by D and B decays with very low momenta
- Run2 analysis recently published
  - (Phys. Lett. B 853 (2024) 138633)
- Run3 analysis ongoing
  - Uses "inclusive" L1 dimuon triggers to seed custom HLT path
  - ParkingDoubleMuonLowMass dataset
- Normalisation channel of this analysis: Ds->phi(mumu)pi decays



Legacy analysis workflow

- Start from CMS MINIAOD data tier
- Data compression into ROOT ntuples
- Loop-based analysis implemented using ROOT TTree:MakeClass
  - **what:** define high-level variables, apply selections, apply scale factors, select best Ds candidate per event
  - **how:** split computation in batches of input files, run separately as HTCondor jobs, gather the output rootfiles

RECAS

ROOT
Data Analysis Framework

New analysis workflow

- Start from CMS MINIAOD data tier
- Data compression into ROOT ntuples
- Ntuples read as RDataFrame
  - Almost all operations "lazy" → no loop triggered till the end
  - One single output (e.g. histogram) produced

  - Option 1: use the MultiThreading capability of RDF
  - Option 2: go distributed using Dask

Dataset size

- CMS MINIAOD data tier
  - **Run2:** "DoubleMuonLowMass"
    - UL2018A⟶ 1.5 TB/ 14 fb-1 = ~ 500 GB / fb-1
  - **Run3:** "ParkingDoubleMuonLowMass"
    - Prompt2022C ⟶ 3.2 TB * 8 streams / 18 fb-1 = ~ 1.4 TB / fb-1

- Data compression into ROOT ntuples ⟶ ~ 3 GB / fb-1
  - to be taken with a grain of salt:
    - at ntuple/NANOAOD level we select events with 3 muons (or 2 muons + 1 track) within our mass region of interest, forming a good vertex etc
    - plain data format

Event selection	Apply event weighting	Selections on muons and tracks	Selections on Ds candidate	Mass plot
Apply per-event requirements (HLT, L1)	Pile-up reweighting, scale factors etc	Per-object cuts (ID, pT, eta etc) Topological cuts (collimation, displacement).	SV quality, displacement, invariant mass Select best candidate per event based on vertex chi2	Fit to extract Ds→phi(μμ)π yield

Event selection	Apply event weighting	Selections on muons and tracks	Selections on Ds candidate	Mass plot
Apply per-event requirements (HLT, L1)	Pile-up reweighting, scale factors etc	Per-object cuts (ID, pT, eta etc) Topological cuts (collimation, displacement).	SV quality, displacement, invariant mass Select best candidate per event based on vertex chi2	Fit to extract Ds→phi(μμ)π yield

Analysis Facility
(CMS INFN AF, ICSC AF under maintenance)

Setup

- CMS AF: https://cms-it-hub.cloud.cnaf.infn.it/ (ICSC highrate AF under maintenance)
- Image: unpacked.cern.ch/registry.hub.docker.com/dodasts/root-in-docker:ubuntu22-kernel-v1-monitoring

Code

https://github.com/fsimone91/BPH_interactive_analysis/tree/main

to be then included in the central repo

BPH_interactive_analysis

Porting typical BPH analysis cutflow to RDF for testing on highrate platforms. This activity is framed in the Spoke 2 of the ICSC – Italian Research Center on High Performance Computing, Big Data and Quantum Computing (https://www.supercomputing-icsc.it/en/), specifically under WP2, UC2.2.1

Use case

The selected use-case is the analysis of Ds->phi(mumu)pi decays, that serve as control and normalisation channel for the tau->3mu search (PLB 853 (2024) 138633).

CMS-BPH-21-005_Figure_001

The analysis plaftorms

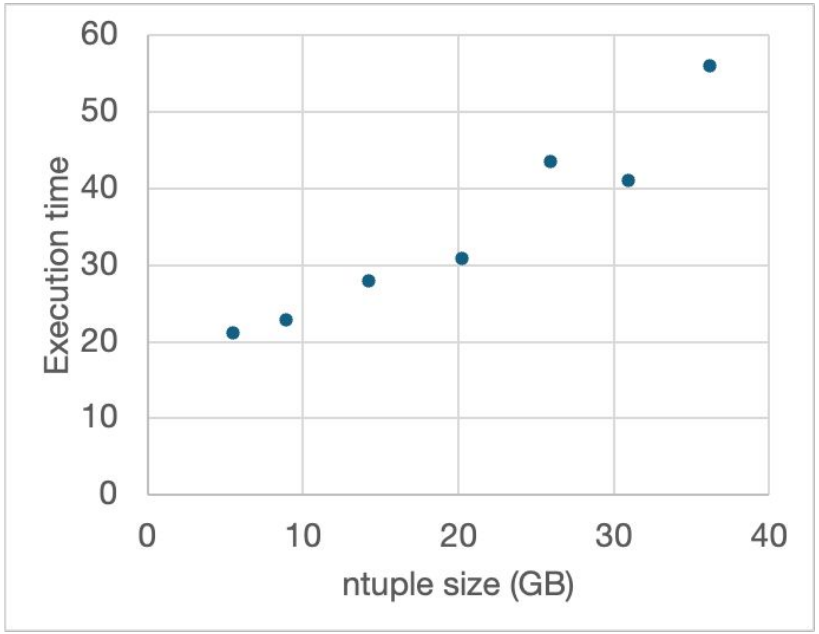
The starting point is a prototypal CMS high throughput analysis platform, offloaded on local Tier-2: https://infn-cms-analysisfacility.readthedocs.io/en/latest/

Within ICSC a new platform is being developed as documented here: https://icsc-spoke2-repo.github.io/HighRateAnalysis-WP5/sections/intro.html

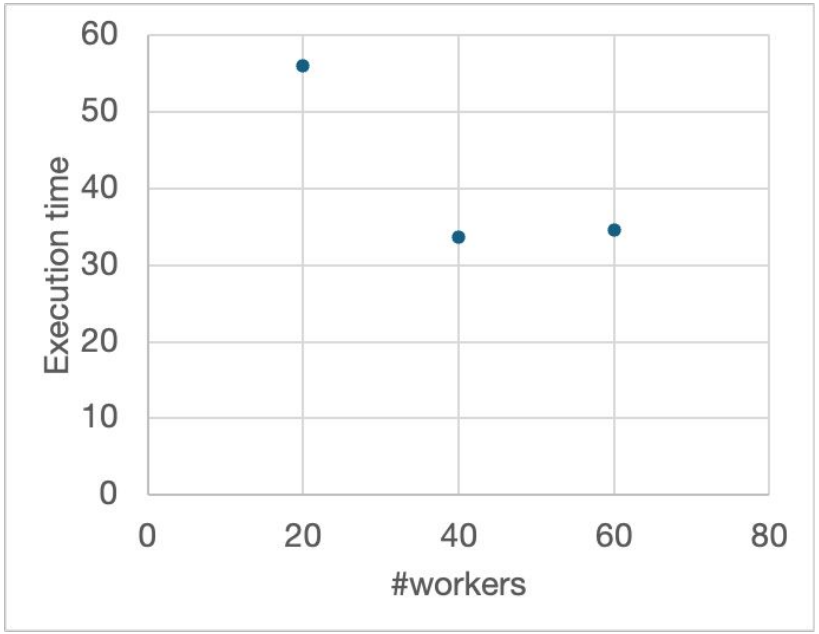
References

- https://doi.org/10.1016/j.physletb.2024.138633

Simple and preliminary tests

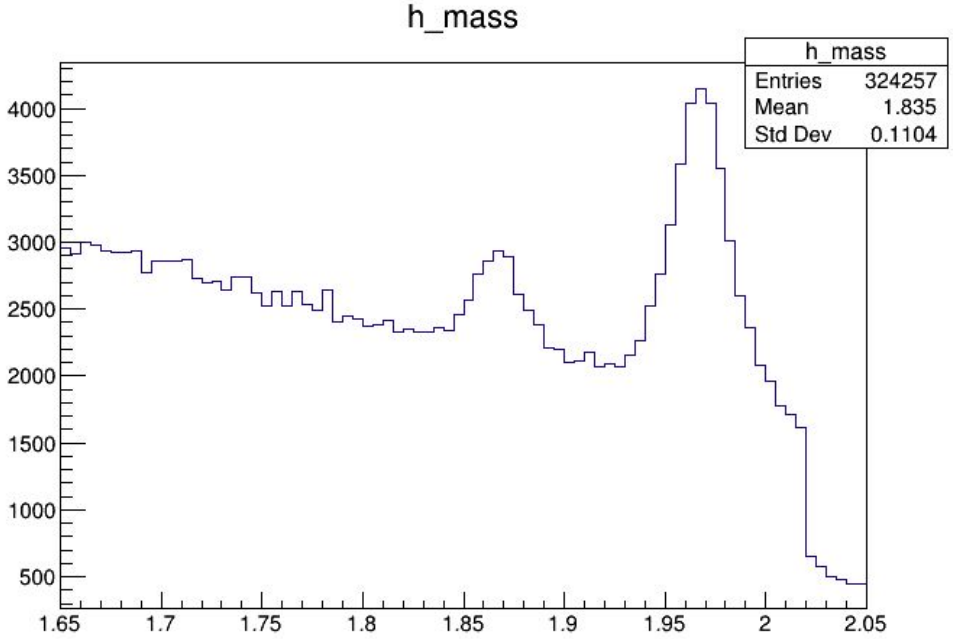


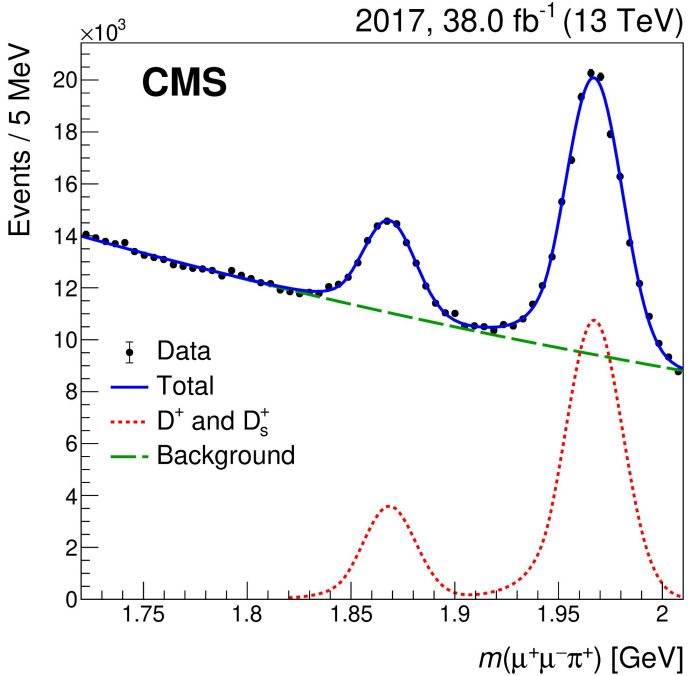
20 workers, 20 RDF partitions



RDF partitions = #workers, 36 GB input

Output


h_mass
h_mass
Entries 324257
Mean 1.835
Std Dev 0.1104
4000
3500
3000
2500
2000
1500
1000
500
1.65
1.7
1.75
1.8
1.85
1.9
1.95
2
2.05


2017, 38.0 fb^{-1} (13 TeV)
CMS
×10^3
Events / 5 MeV
20
18
16
14
12
10
8
6
4
2
0
Data
Total
D$^+$ and D$_s^+$
Background
1.75
1.8
1.85
1.9
1.95
2
$m(\mu^+\mu^-\pi^+)$ [GeV]

User-experience and summary

- User experience (CMS AF):
  - some instabilities when increasing # workers
  - strongly depends on the Site, difficult to debug

- Plans:
  - include fit step (need newer ROOT version with “pythonised” RooFit, NOT distributed though!)
  - reproduce approved plots for validation
  - scale to Run3

- Abstract submitted to CHEP2024 https://indico.cern.ch/event/1338689/abstracts/175102/