

Reduced Formats VS Exotic Signatures

Bryan Cardwell

*Thanks to Andre Frankenthal, Andrew Hart, and Juliette Alimena for input

Analysis Ecosystems Workshop II

May 24, 2022

CMS data formats

| | | |
|---------|-------------|-----------------------|
| RAW | ~1 MB/evt | |
| RECO | ~3 MB/evt | |
| AOD | ~500 kB/evt | ← most run 1 analyses |
| MiniAOD | ~50 kB/evt | ← most run 2 analyses |
| NanoAOD | ~2 kB/evt | ← most run 3 analyses |

But what about analyses that aren't *most analyses*?

Exotic signatures present additional challenges when trying to move to smaller data tiers

- Focus on long-lived particles (LLPs)
- Similar arguments generally apply to B physics and heavy ions

LLP analyses

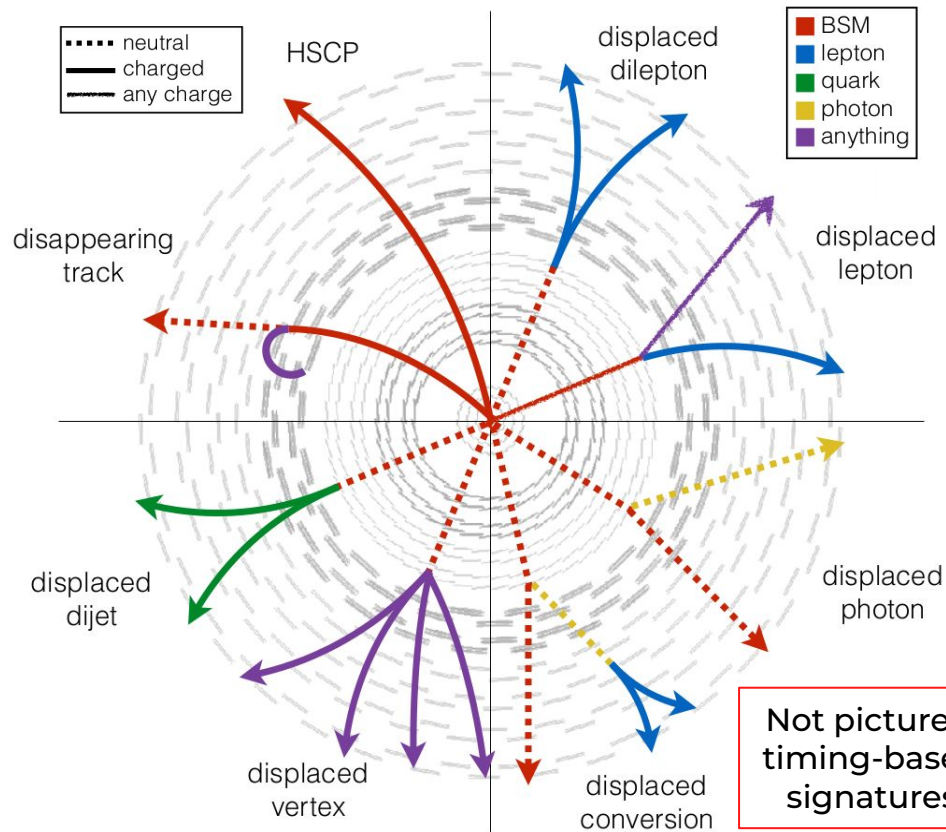
SM particles come with a wide range of lifetimes, BSM particles may do the same

BSM LLPs with detector-scale decay lengths could produce striking signatures

But unusual signatures \rightarrow dedicated triggers, atypical backgrounds, unique reco, unusual discriminating variables

Large majority (~90%) of Run 2 CMS LLP searches require AOD (or RECO or RAW...)

AOD will not be easily available in Run 3!



What to do?

Add info to Mini/NanoAOD

- Difficult: many analyses require inherently not-small detector-level info
- Try to identify collections that are useful to many analyses

Make custom skims that include AOD, RECO, or RAW info

- Again, best to coordinate effort to minimize duplicated info
- Some analyses require lots of information from a smaller number of events, others require a little extra information from many events

This talk is supposed to be about NanoAOD, but we have to get to MiniAOD first...

Standard muon reco results in low efficiency and poor momentum resolution for displaced muons

Dedicated muon reconstruction using only the muon system or displaced tracker tracks + muon system improves reco efficiency and momentum resolution

Add these displaced muon collections to miniAOD

Minimize size by filtering out muons that overlap already-available prompt muons

Useful for many LLP analyses, increases miniAOD size by ~2.5%

Disappearing tracks

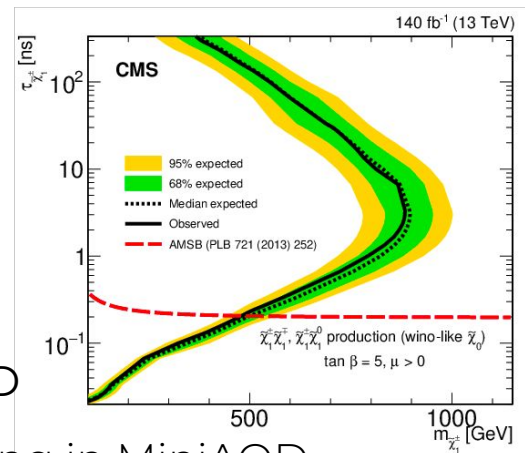
arxiv.org/abs/2004.05153

Search for charged LLPs that decay to neutral and/or unreconstructed particles in the tracker volume

- ≥ 1 isolated track with missing hits in the outermost tracker layers < 10 GeV of associated calorimeter deposits

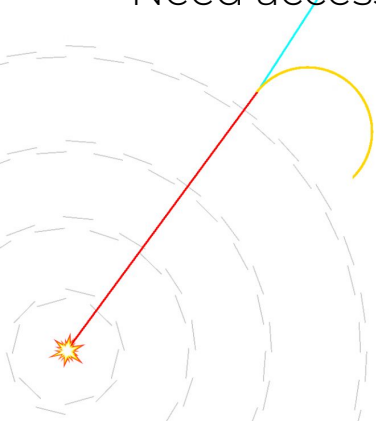
Run 2 performed in miniAOD + a few extra collections from AOD

- Need access to tracks and calorimeter rechits that are missing in MiniAOD



Collections are too big for Mini/NanoAOD, so prepare custom skim

- Store all ECAL and HCAL rechits from RECO
- Only keep events with ≥ 1 track with $p_T > 25$ GeV, $|\eta| < 2.1$



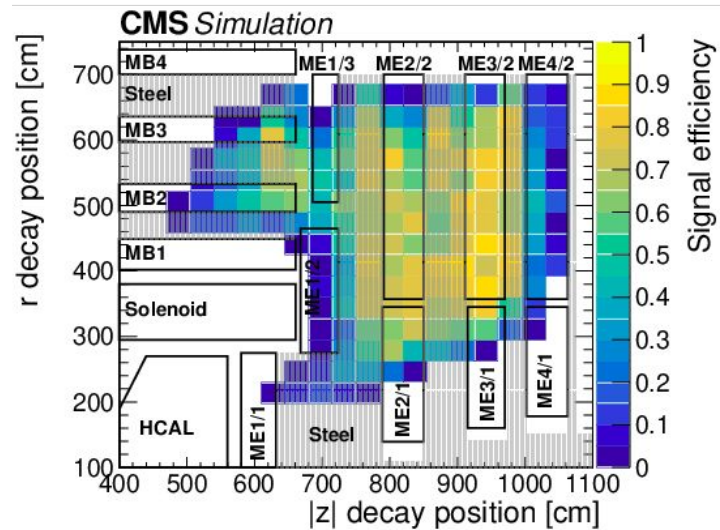
Search for LLPs that decay in the CMS muon system

- Look for clusters of > 50 hits in muon system

Use existing high-MET and $Z \rightarrow \mu\mu$ skims to validate signal simulation

New skims with analysis-specific triggers

- Save AOD + Reco-level muon-system rechits for events that pass DT cluster triggers ($\sim 1\%$ skim rate)
- Save RECO for events that pass CSC cluster triggers ($< 1\%$ skim rate)



Magnetic monopoles

Magnetic monopoles would bend in ρ - z (instead of ρ - ϕ) and be highly ionizing

- Unique tracks, unique ECAL shower shapes

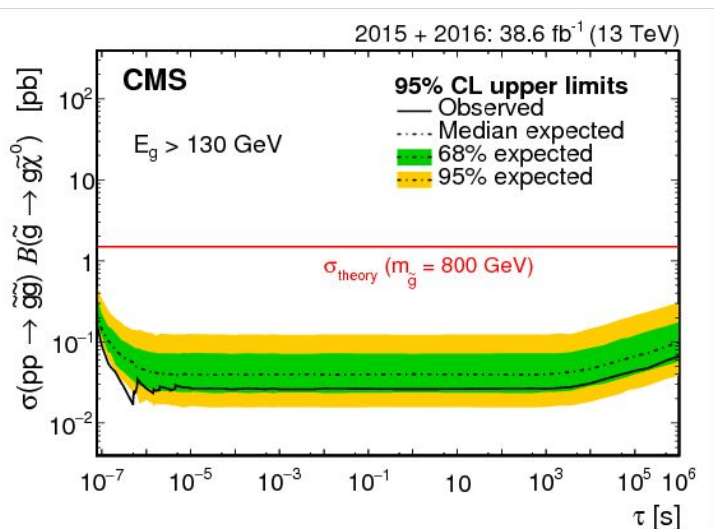
Dedicated reconstruction requires extra tracker and ECAL info

- Run 2 skim kept AOD + additional collections for photon-triggered events
 - Track segments for track reco
 - Uncleaned ECAL cluster collections to avoid spike rejection
 - Reasonable rate, but large per-event size
- Run 3 skim will reduce per-event size by only keeping needed collections

Search for LLPs that stop in the detector and decay at later times

- Look for decay products in calorimeters and muon system

Runs 2 & 3: RECO-level skim of events at least ± 2 bunch crossings away from collisions



No event selection, but keep only RECO-level information necessary for analysis

- Calo jets, HCAL noise info, muon-system hits/segments, displaced muons
- Reduces RECO event size by 70%
- Overall low rate (cosmics, beam halo, noise)

Use tracker dE/dx to identify charged, heavy LLPs

MiniAOD likely sufficient for main Run 3 analysis

...but need extra AOD-level info for calibrations:

- dE/dx for low- p_T tracks
- DT and CSC segments for $1/\beta$ calibration
- Small subset of events is sufficient as long as all running periods are represented

High-MET skim for many exotic-signature analyses

- Keep RAW and RECO for events with $\text{MET} > 200 \text{ GeV}$

Delayed jets (arxiv.org/abs/1906.06441)

- Keep RECO for analysis-specific triggers (based on jet timing, MET, and HT)

Displaced jets (arxiv.org/abs/2012.01581)

- Keep RAW+AOD for a small fraction of large HT events to assess systematics

Exotic signatures and reduced formats tend not to get along

- Reducing event size == discarding info *most* analyses don't need
- Exotic signatures != *most* analyses

Most Run 2 CMS LLP searches rely on AOD (and some RAW or RECO)

Focus now is on making sure Run 3 searches can use MiniAOD + dedicated skims

- Displaced muons in MiniAOD
- Many dedicated skims; some multi purpose, some analysis specific

This talk mostly focused on the challenges of going to reduced formats, hopefully the discussion will include some thoughts on solutions :)