

A Shot in the Dark

Searching for a Dark Sector with the CMS
Experiment

Benedikt Maier, KIT
SLAC, Sept 5, 2023



No Dark Matter candidate in SM

Dark matter about **5x more abundant** than visible matter

Inferred from **astrophysical observations**, like gravitational lensing or rotational curves

u	c	t	g
d	s	b	W
e	μ	τ	Z
ν_e	ν_μ	ν_τ	γ
			H



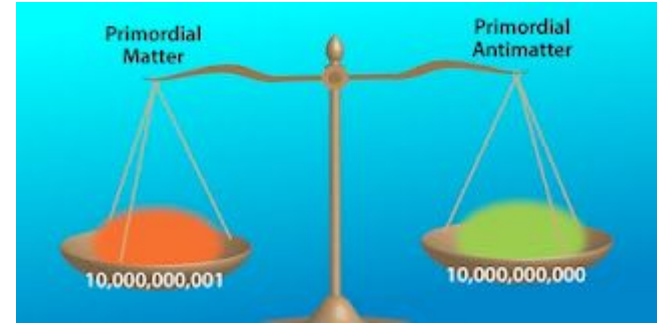
Many other open questions in cosmology and particle physics, like

Matter-antimatter asymmetry

Hierarchy problem (weakness of gravity, fine tuning at level 10^{16})

Dark sector could hold **the answer** to some of these, depending on the theory

Probing it at the LHC!



Known world

“Hidden valley”

Multitude of particles?
Multitude of interactions?

Standard Model



$SU(3) \times SU(2) \times U(1)$

Communicator

Dark sector



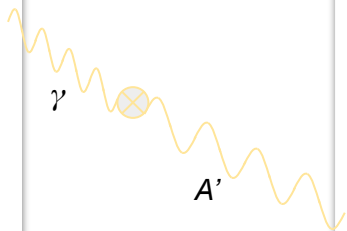
?

The “simplest” case: a dark photon A'



Standard Model

$SU(3) \times SU(2) \times U(1)$



Dark sector

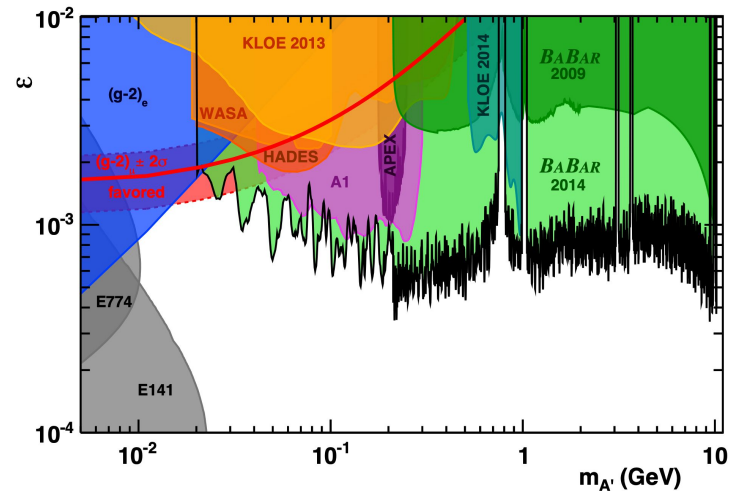
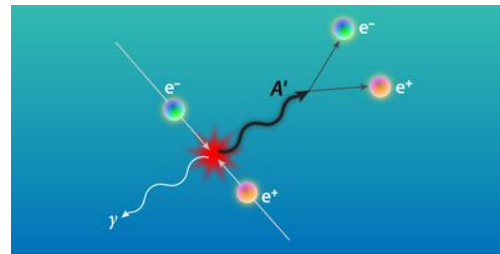
$$\mathcal{L}_{A'} = \underbrace{-\frac{1}{4} F'^{\mu\nu} F'_{\mu\nu}}_{\text{Its own field}} + \underbrace{\frac{1}{2} \frac{\epsilon}{\cos \theta_W} B^{\mu\nu} F'_{\mu\nu}}_{\text{Mixing with SM photon}} - \underbrace{\frac{1}{2} m_{A'}^2 A'^{\mu} A'_{\mu}}_{\text{Mass created by (dark?) Higgs}}$$

New broken $U_D(1)$ symmetry (with dark fermions charged under $U_D(1)$)

Free parameters: mixing ϵ , mass $m_{A'}$

$$e^+e^- \rightarrow \gamma A' (\rightarrow e^+e^-, \mu^+\mu^-)$$

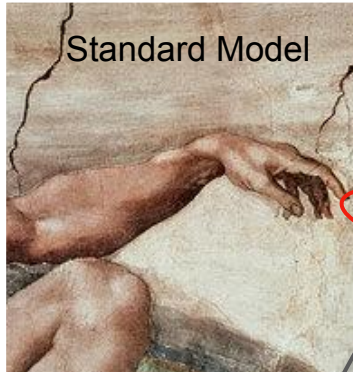
Dark photons searches at **BaBar**



The LHC can provide:

- Highest energies
- Unprecedented luminosities
- General purpose detectors and specialized experiments

Standard Model



$SU(3) \times SU(2) \times U(1)$

LHC

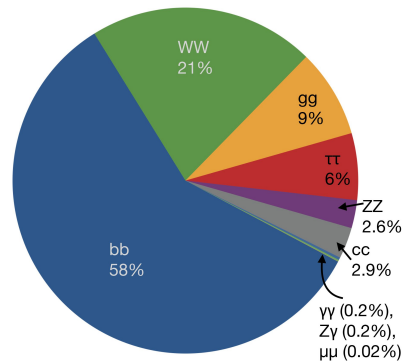
Dark sector

?



Higgs boson discovery, CERN, 2012

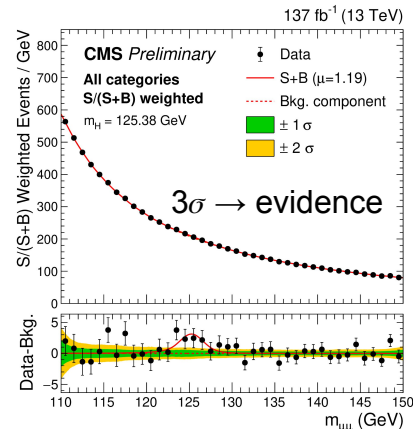
How do Higgs bosons decay according to the Standard Model?



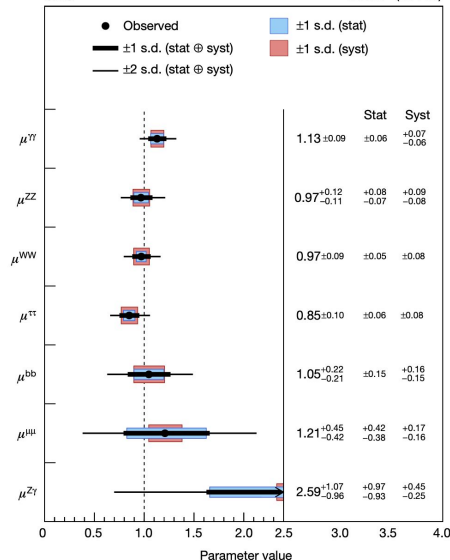
Higgs boson gives rest mass to elementary particles.
 Predicted in 60s, discovered 2012

Combination of all Higgs analyses currently allows an **upper limit of 15%** for new, unknown particles as decay products

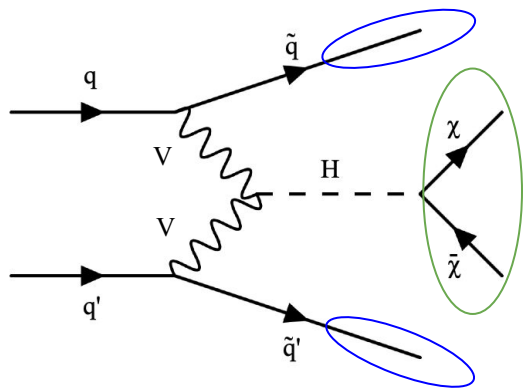
→ **Motivates search for decay to dark matter particles**



CMS 138 fb⁻¹ (13 TeV)



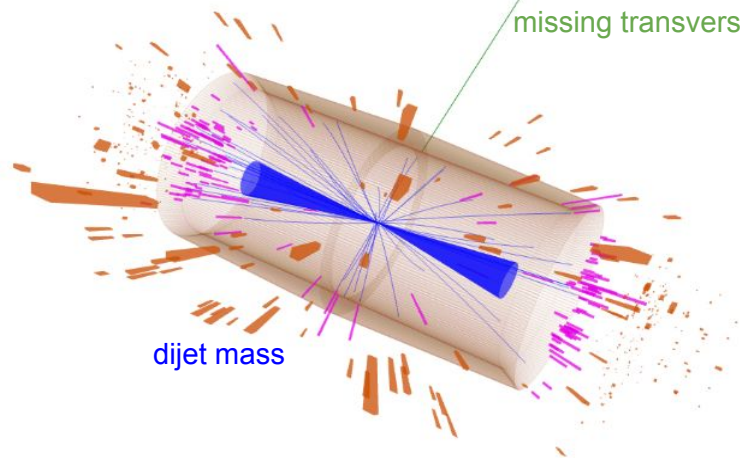
Disappearing Higgses (“Higgs portal”)



Dark matter, by definition, does not interact with the detector

- Sizable missing transverse energy (> 160 GeV)
- Requiring two jets with large pseudorapidity difference (“forward”)

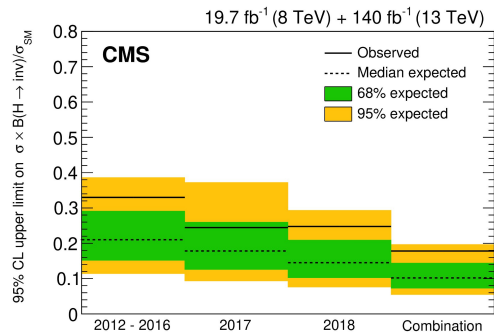
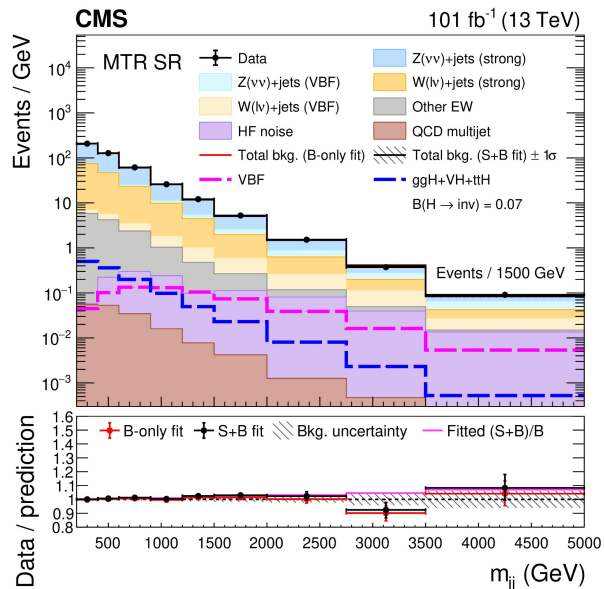
missing transverse energy



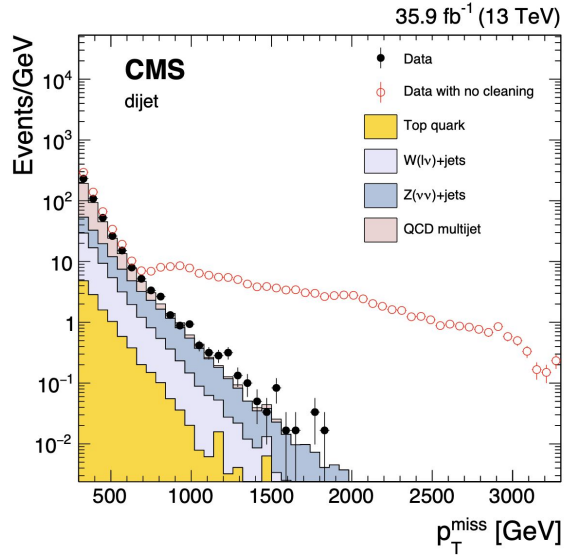
dijet mass

Statistical analysis of observed m_{jj} spectra gives:

18% (10%) observed (expected) upper limit for branching ratio $H \rightarrow$ invisible

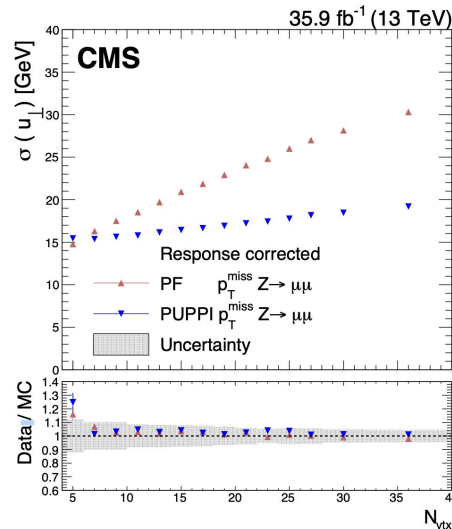
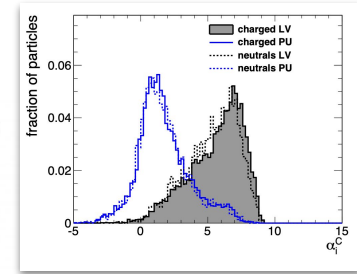
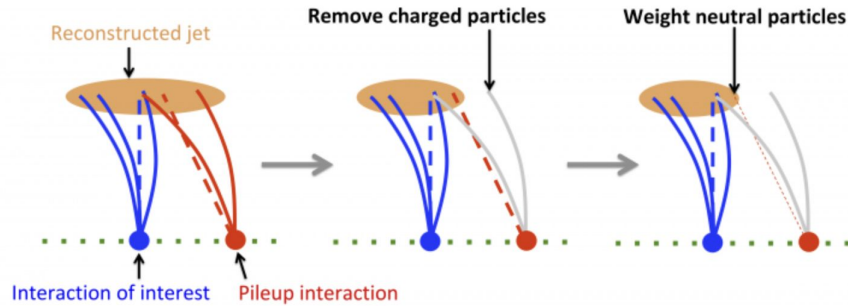


Missing transverse energy



Dedicated filters reject noise events that are not modelled with MC, e.g.:

- beam halo
- dead ECAL crystals
- ...



Pileup per particle identification (“PUPPI”) algorithm identifies particles from subordinate proton-proton collisions

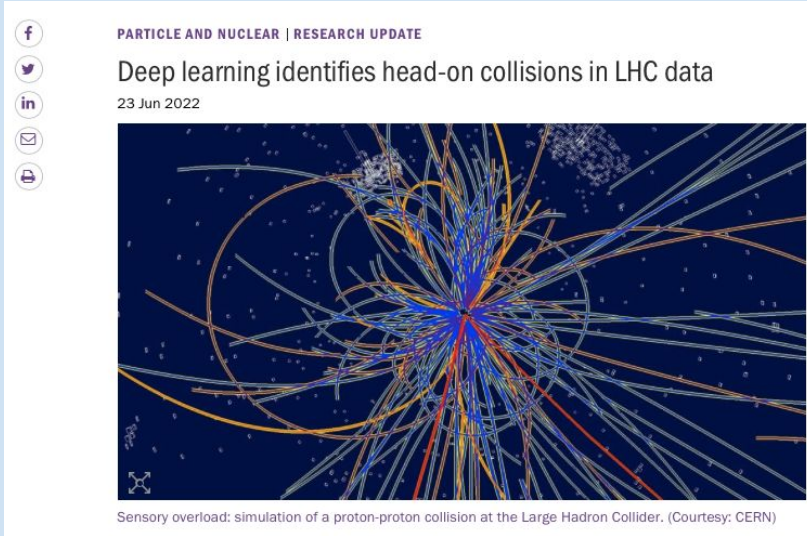
→ Dramatic improvement in resolution of missing transverse energy

What's next in Run-3?

AI-based pileup mitigation

“Pileup is never a dominating systematic uncertainty”

Disentangling up to 10,000 particles

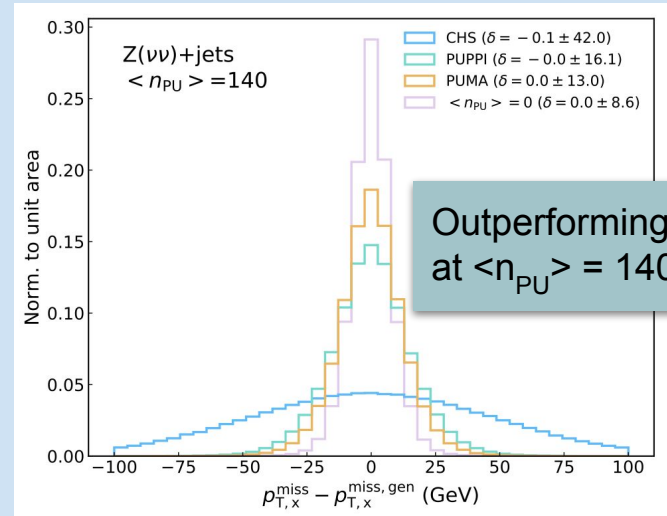


Relying on **transformers** (same as ChatGPT) to “translate” a sequence with up to 10,000 particles (a sentence with up to 10,000 words)


→ Need to use **sparse attention**

Developing a **novel pileup mitigation** algorithm that predicts if a neutral particle comes from primary vertex or not

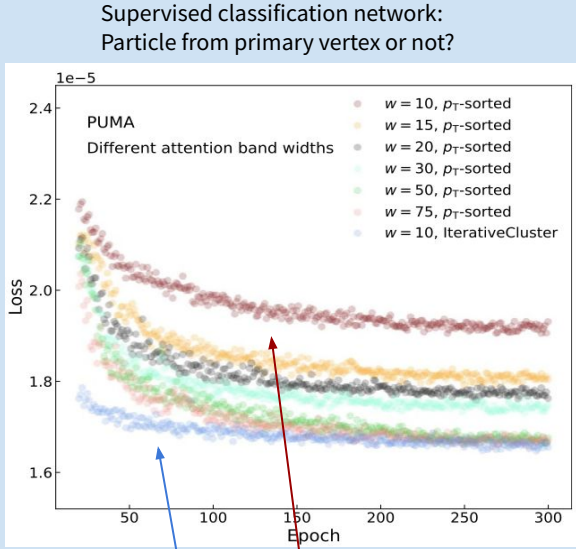
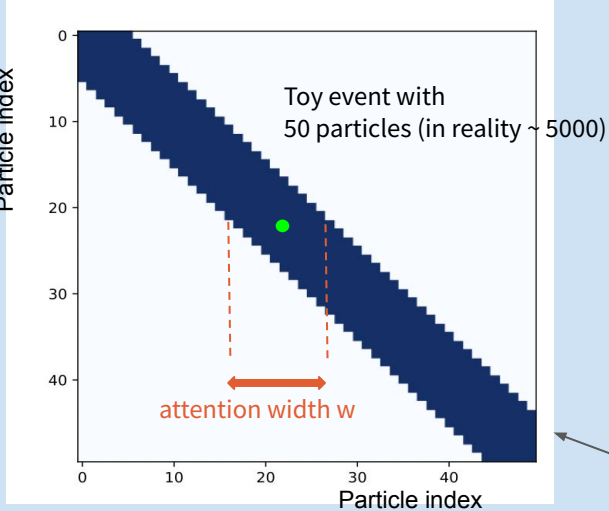
→ Eventually replacing PUPPI as standard PU mitigation



Attention mechanisms at work

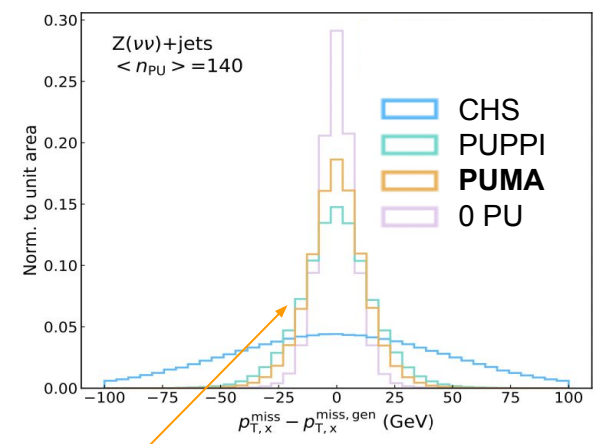
Problem: Cannot perform 9000 x 9000 matrix multiplications to get attention scores 

→ Introduce *sliding attention window*: a particle only “talks” to the particles before and after.



Particles sorted by p_T not really useful

Perform k-Means clustering in (η, ϕ) to embed sense of locality into particle sequence



25% improvement in MET resolution over rule-based algorithm (PUPPI)

Implementation in CMS ongoing!

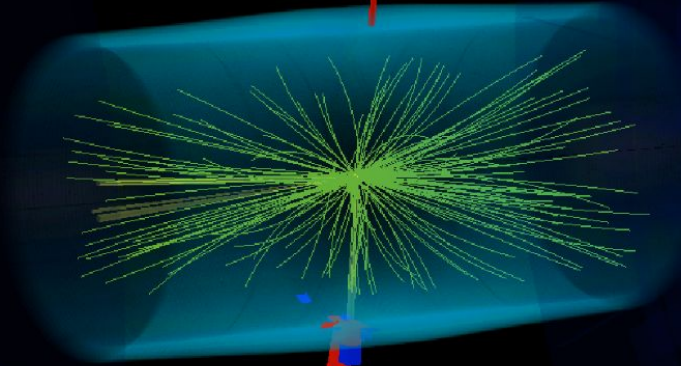
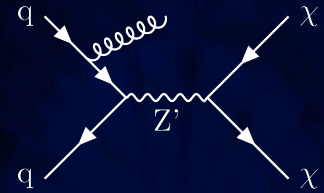


CMS Experiment at the LHC, CERN

Data recorded: 2018-Jul-14 21:03:24 EDT

Run / Event / LS: 319639 / 1418428259 / 986

MET,
pt = 1691.82 GeV
eta = 0
phi = 1.726

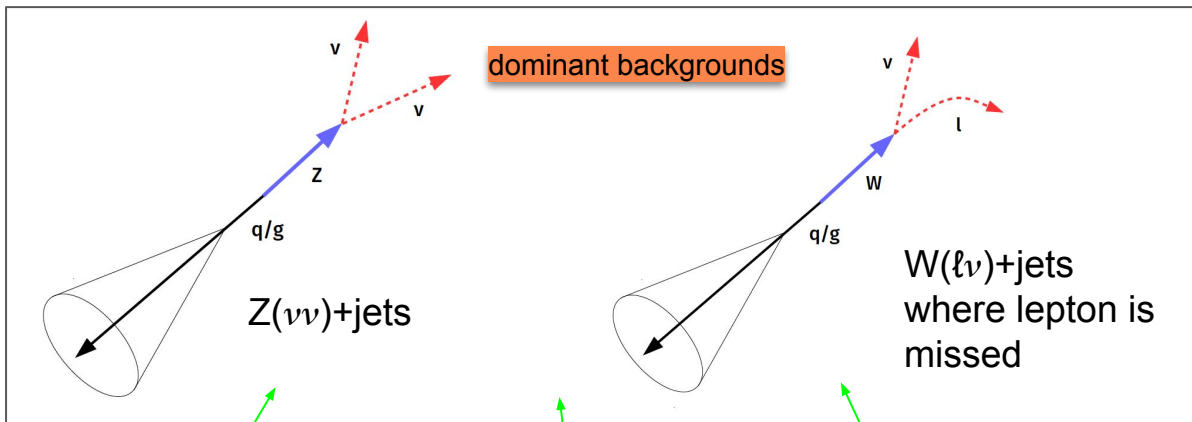


Jet,
pt = 1665.5 GeV
eta = 0.081
phi = -1.377

Energy frontier

Probing the heaviest communicators between the visible and dark sector

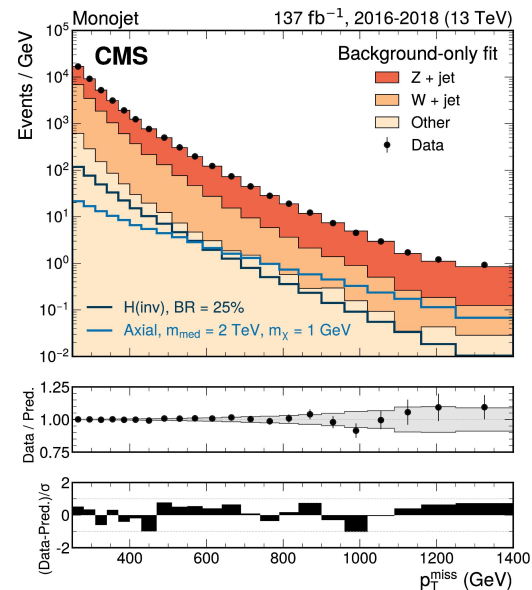
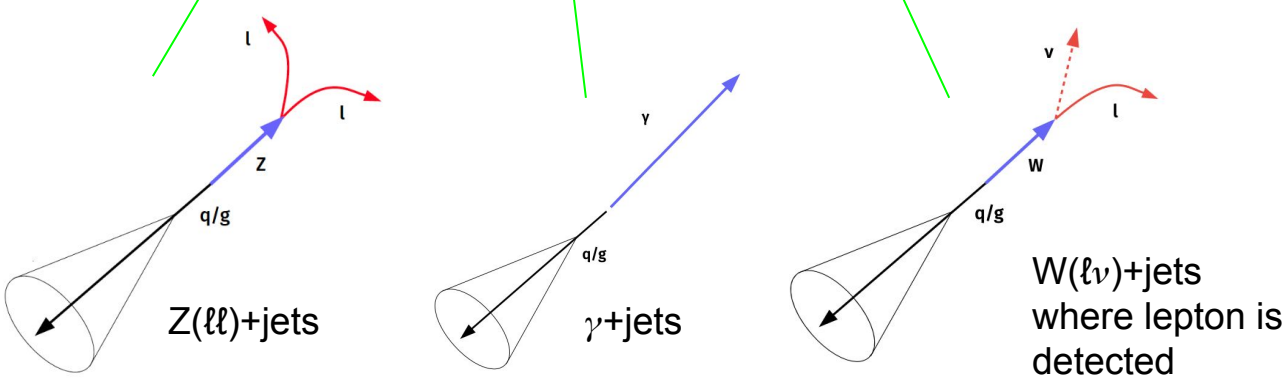
Standard Model backgrounds in dark matter searches

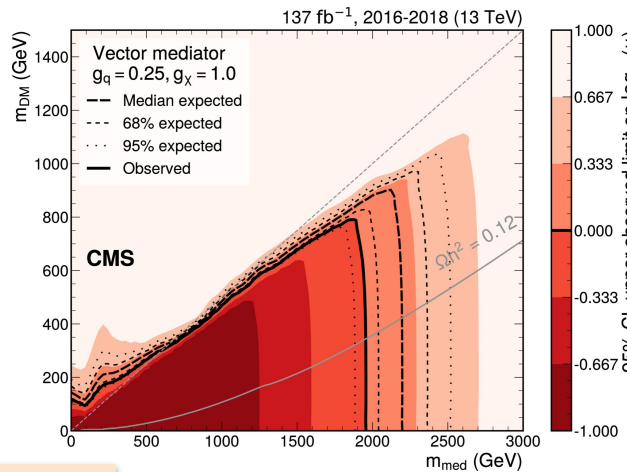


Use apparent symmetry between processes to constrain backgrounds

Pretending to not see photons & leptons to create a **proxy** for MET

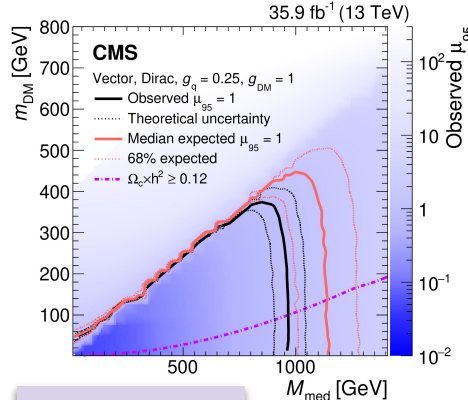
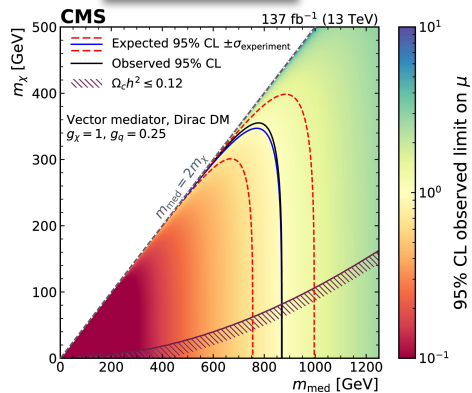
constrain





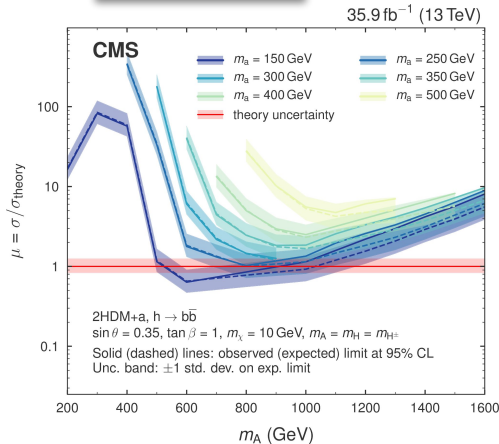
mono-jet

mono-Z(II)

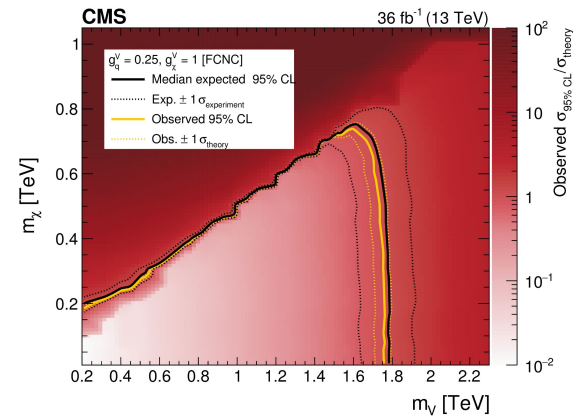
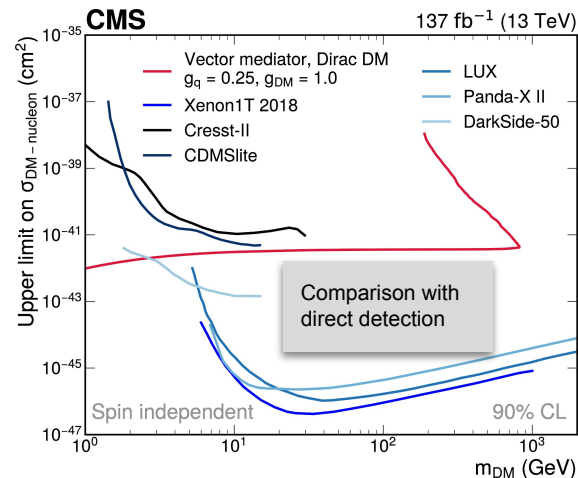


mono-photon

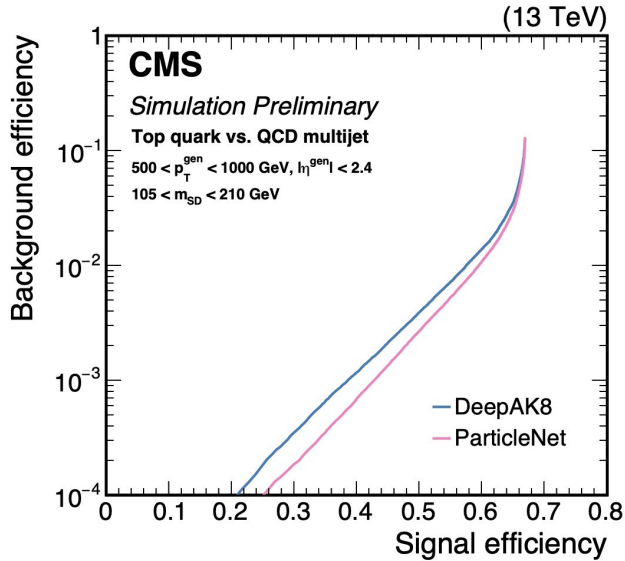
mono-H(bb)



mono-top



A particle net to tag heavy resonances



CMS at the forefront of accelerating machine learning-based solution in high energy physics

ParticleNet = current **state-of-the-art** in jet tagging

Jet constituents as graph

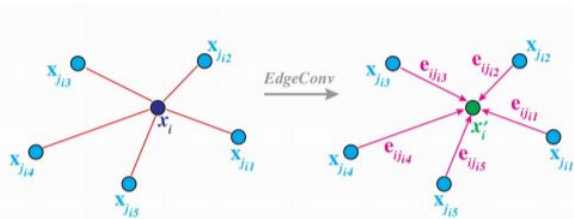
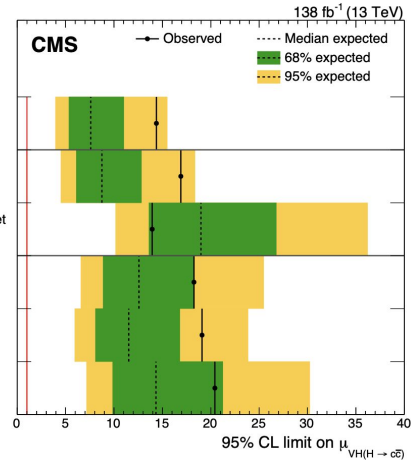
Edges dynamically determined in the latent space

Edge over ATLAS

CMS expected upper limit on $VH(cc)$: 7.8

ATLAS: 31 !!!

Is as if Run-2 lasted 16 times longer for CMS



What's next in Run-3?

Contrastive self-supervised jet taggers

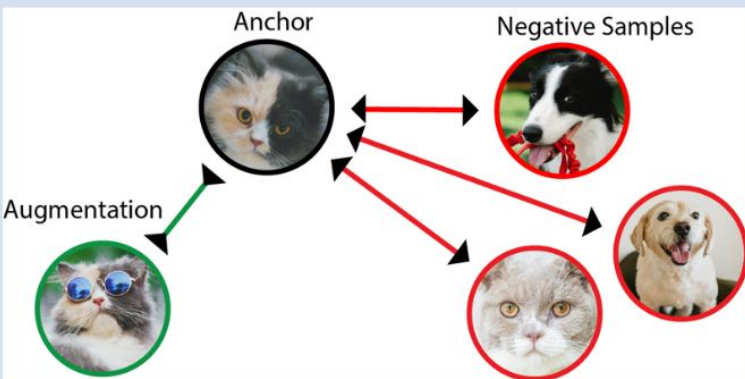
“Can we do better than ParticleNet?”

Contrastive learning

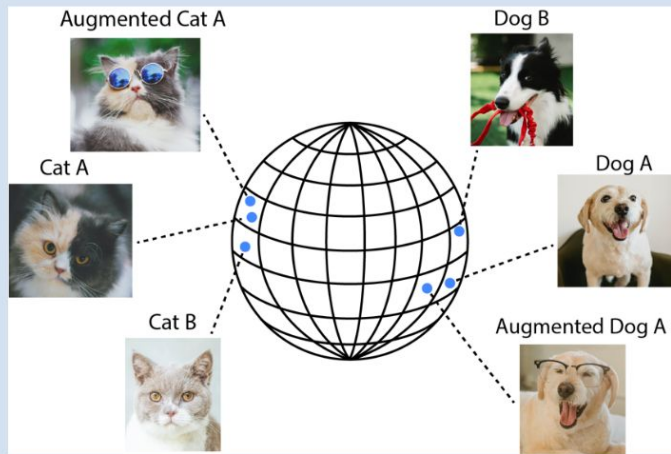
$$-\log \frac{\exp(\text{sim}(\mathbf{J}_i, \mathbf{J}'_i)/\tau)}{\sum_{k \in \mathcal{N}} \exp(\text{sim}(\mathbf{J}_i, \mathbf{J}_k)/\tau)}$$

- contrastive loss** maps together positive (anchor, augmentation) pair, pushes apart the rest

Large success in other AI domains!

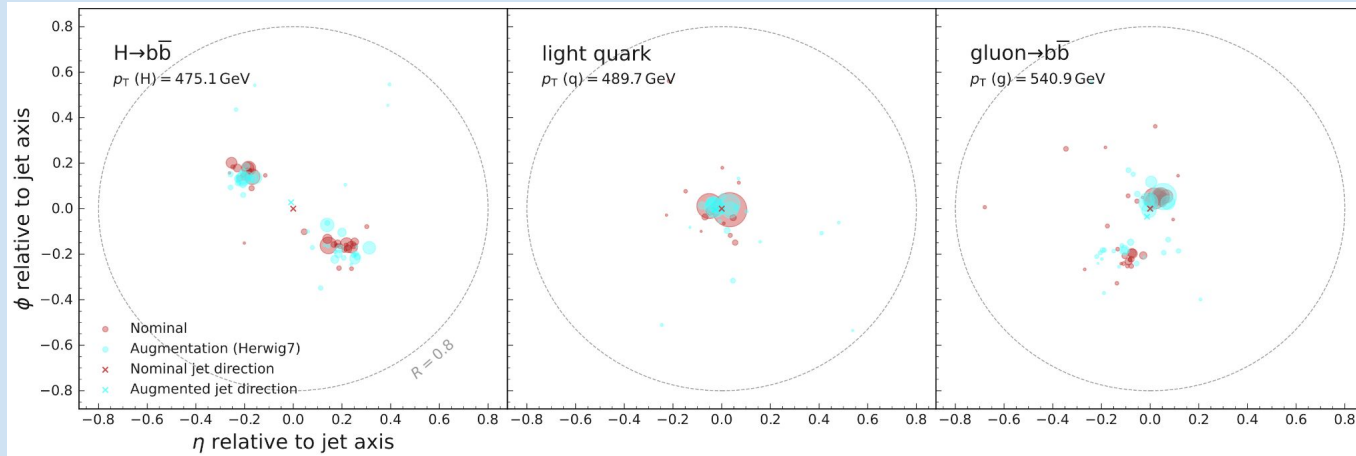
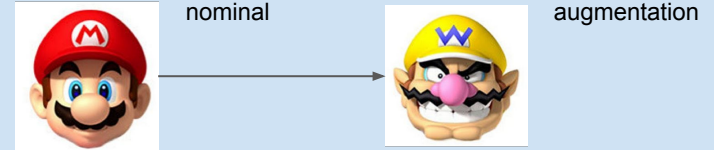


SimCLR loss (2002.05709)



Contrastive learning: Augmentations from re-simulation

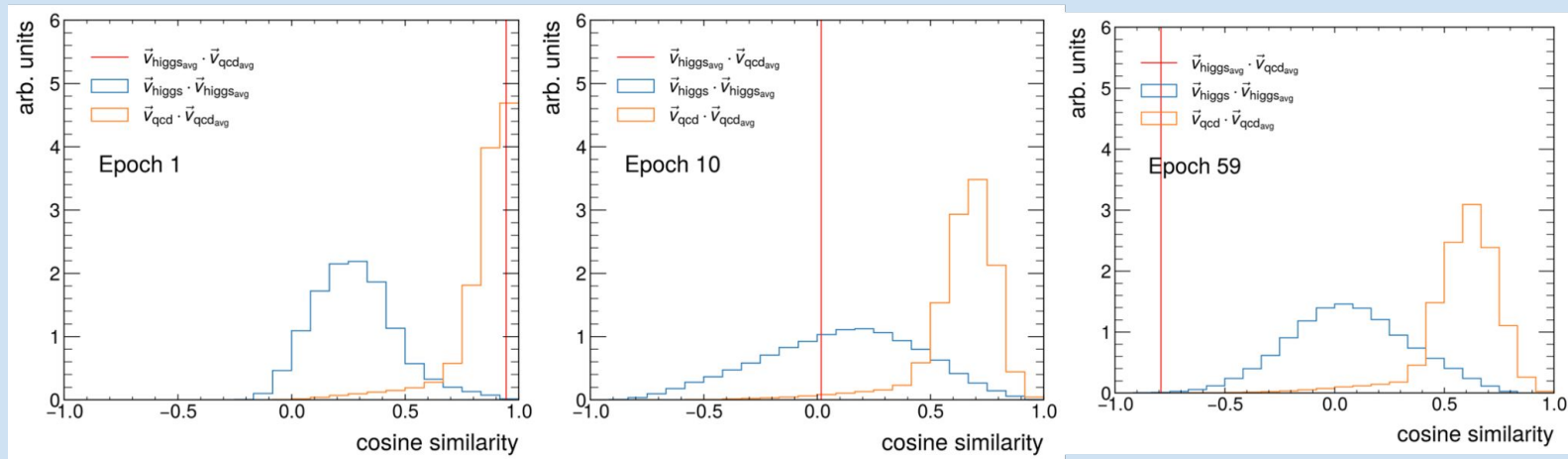
- What is a robust way to generate augmentations for particle physics data? How can we learn the full space of jets (at least what is available in simulator)?
- augmentations: re-simulated versions of the same generated partons
- We can re-shower with different configurations:
 - Pythia8 with different seed
 - Pythia8 with FSR scale varied up and down
 - Herwig7
- Essentially sampling configurations with all physical variations available in the simulator



Learning the contrastive representation

$$\cos(\theta) = \frac{\mathbf{J}_i \cdot \mathbf{J}_k}{\|\mathbf{J}_i\| \|\mathbf{J}_k\|}$$

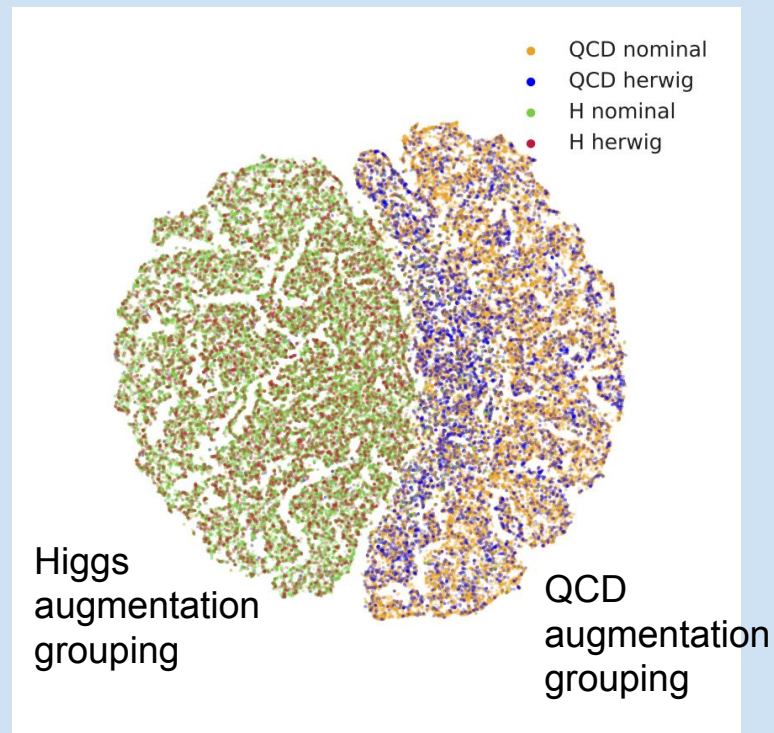
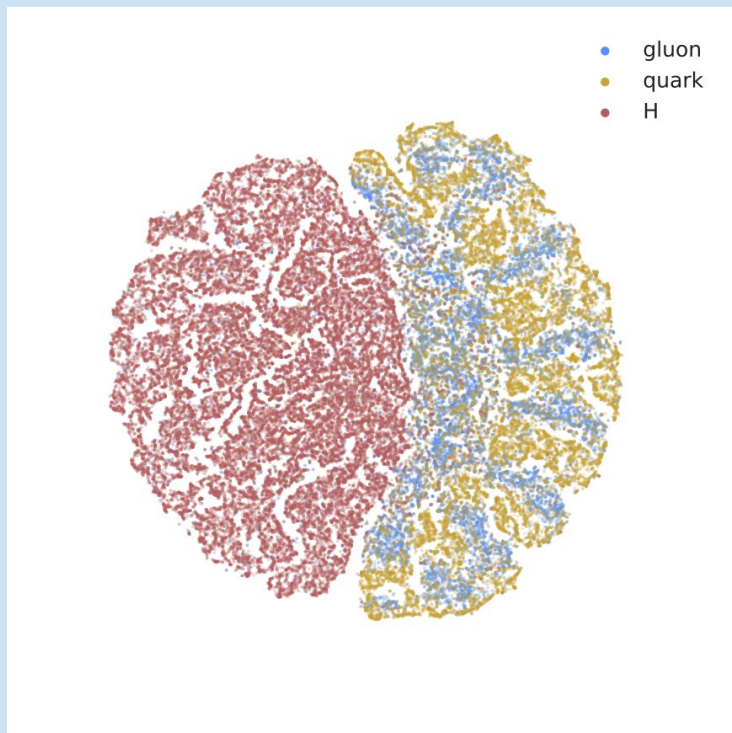
- Cosine similarity=primary distance metric used in simCLR loss
- Average QCD and Higgs vectors move back-to-back in 8D space (red line)



Training length

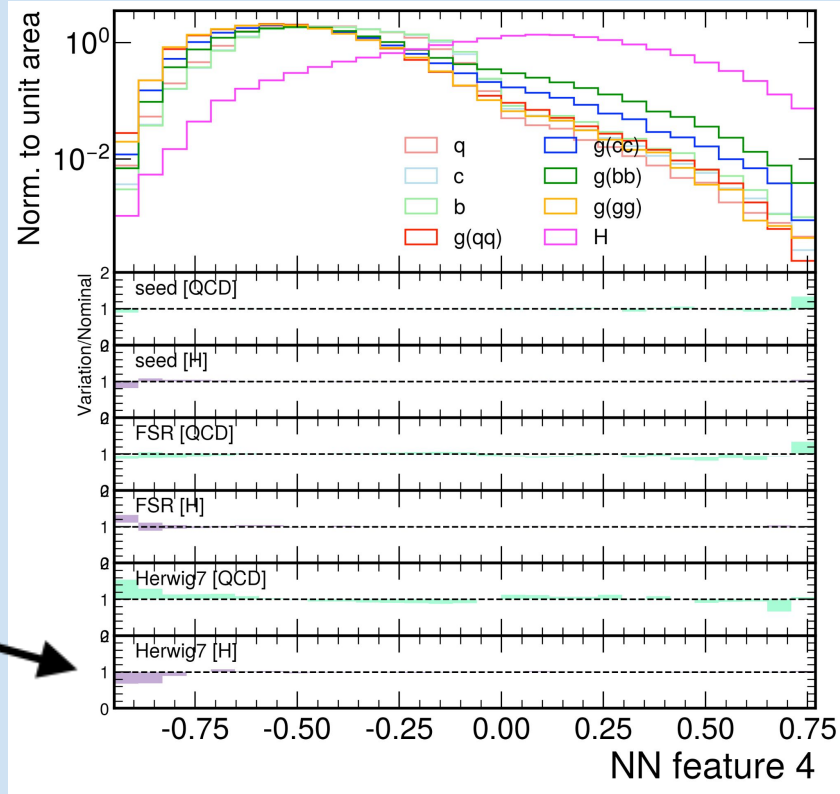
Investigating the contrastive space

The 8D representation space projected into 2D tSNE...



Contrastive space feature

Contrastive features distinguish Higgs and QCD



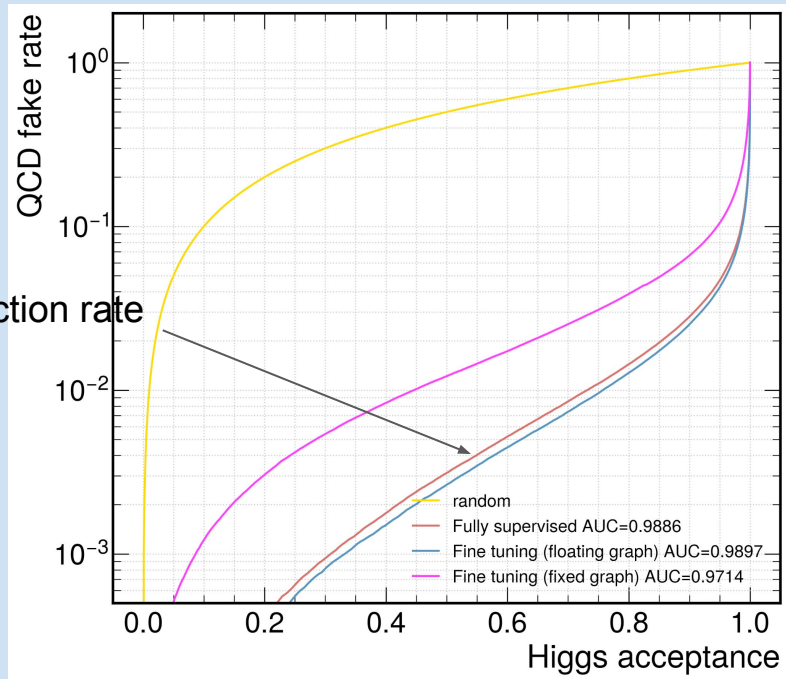
...while mitigating variations with simulator



Classification performance

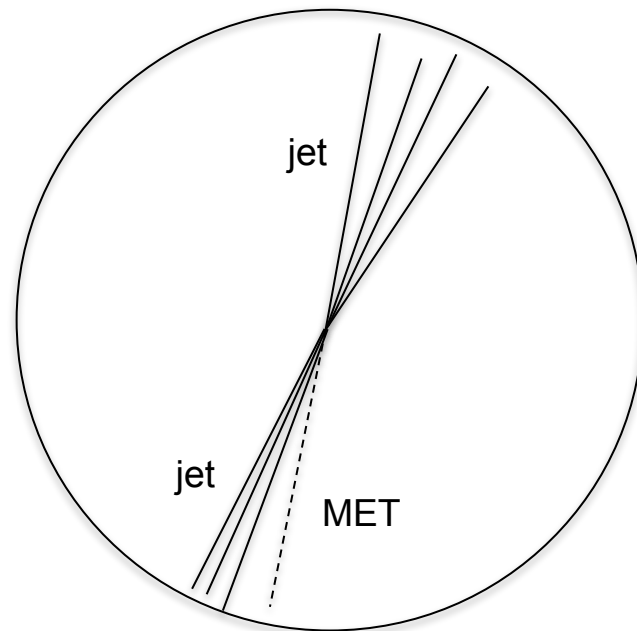
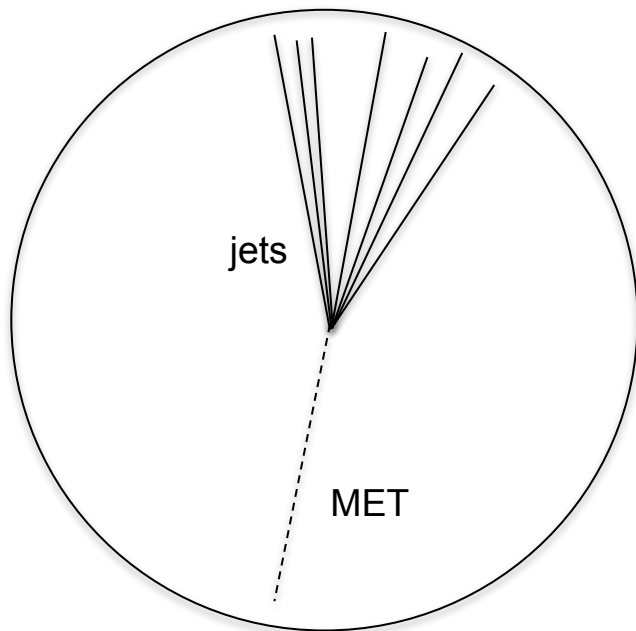
- Contrastive space as a backbone for classification (e.g. H vs QCD)
- Using graph weights pre-trained on SSL space, it finds a better minimum than fully supervised network

At 50% signal efficiency:
20% higher background rejection rate



We will see a lot of this during Run-3 ...

Back-to-back topology between MET and jets
Interesting, therefore kept
Easy to analyze

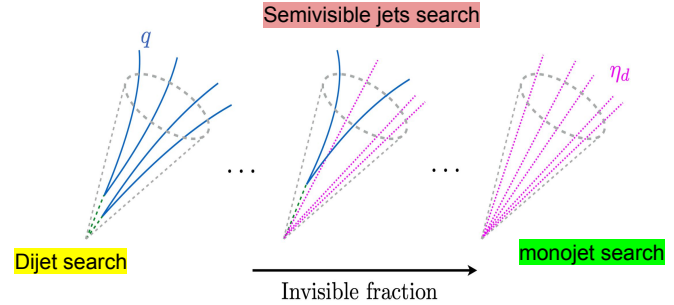
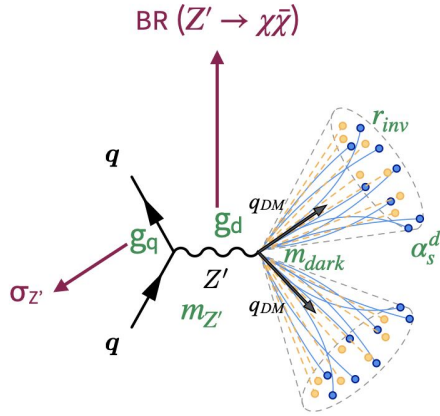


MET and jet **aligned**
Uninteresting (multijet background), thus discarded
Coming from mismeasurement of one jet ... or does it?

Semivisible jets – a dark version of the strong force

$$N_{\text{dark}}^c = 2, N_{\text{dark}}^f = 2$$

4 parameters: $m_{Z'}$, m_{dark} , r_{inv} , α_{dark}



$$r_{\text{inv}} \equiv \left\langle \frac{\# \text{ of stable hadrons}}{\# \text{ of hadrons}} \right\rangle$$

Dark showering process after leptophobic $Z' \rightarrow q_D q_D$

Some dark hadrons stable \rightarrow Visible states **interspersed** with dark hadrons

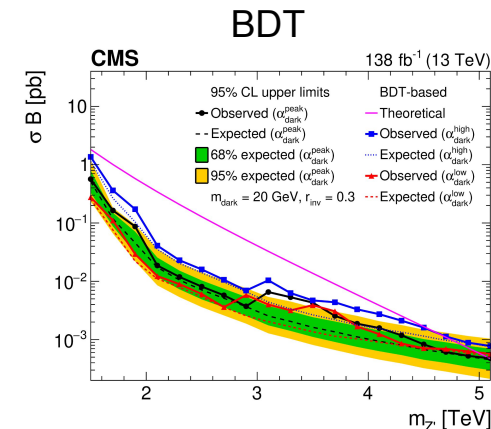
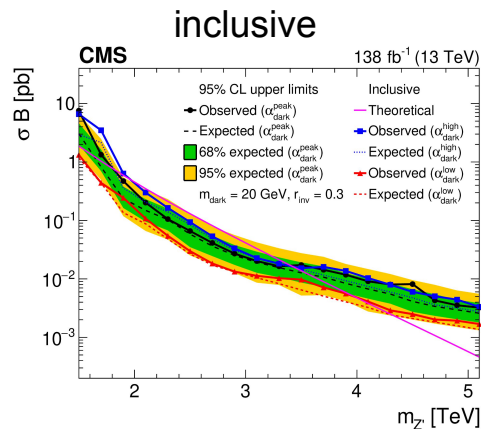
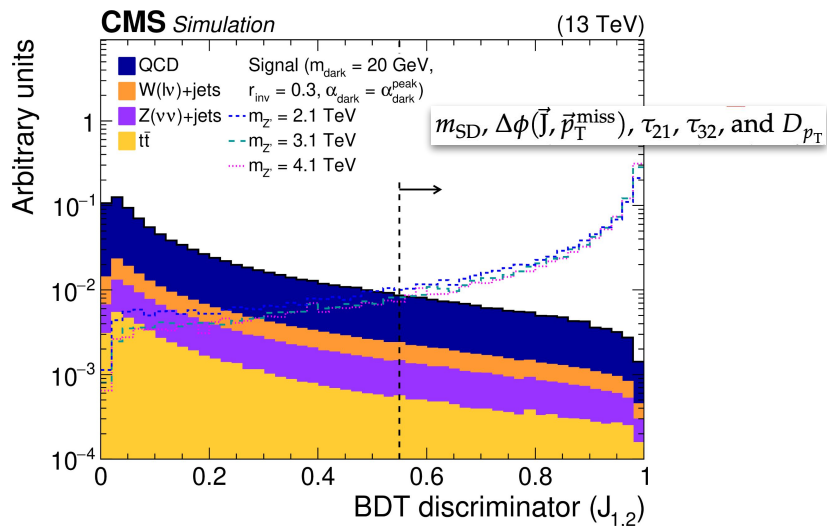
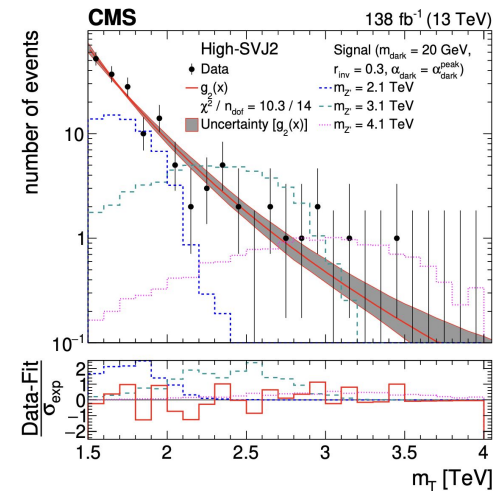
E_T^{miss} aligned with jets, not back-to-back \rightarrow mono-X searches not sensitive

Traditional jet quality cuts not sufficient; custom filters to eliminate dead cells that create fake MET

Semivisible jets – a dark version of the strong force

Sensitive variable: transverse mass m_T of dijet system and MET

A jet-level **boosted decision tree** is used to tag SVJs and define a high-purity category

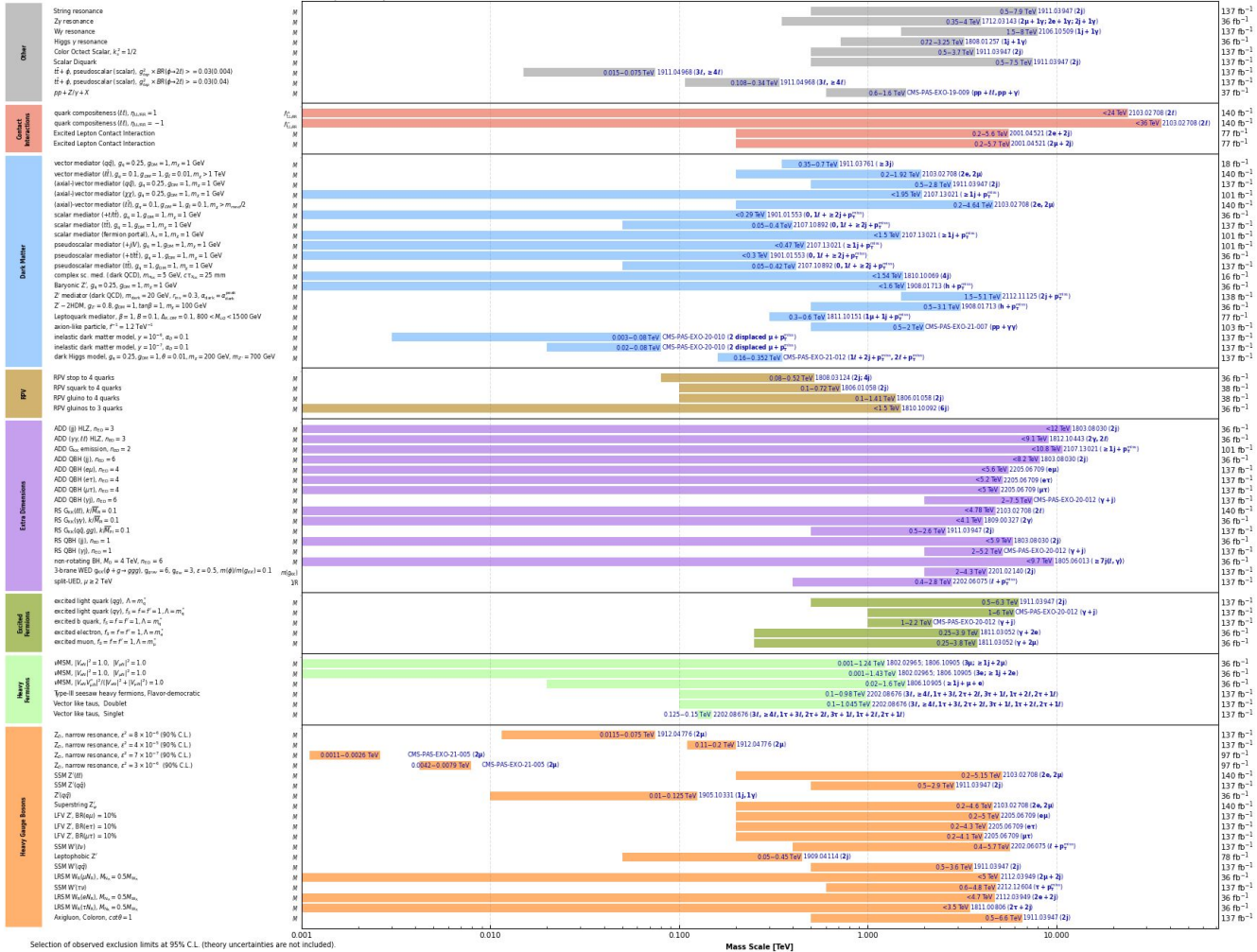


Slight excess around 3 TeV → Remains interesting for Run-3!

Overview of CMS EXO results

CMS preliminary

March 2023



Selection of observed exclusion limits at 95% CL. (theory uncertainties are not included).

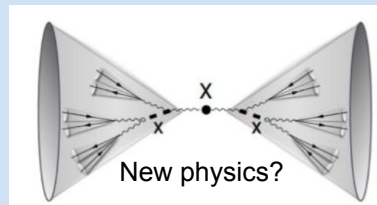
What's next in Run-3?

Model-independent searches

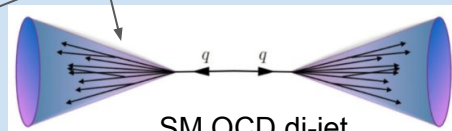
“Can we perform a catch-all search for new physics?”

Anomaly detection for model-independent searches

- Example: dijet analysis
- Exotic substructure or radiation patterns are common for BSM physics

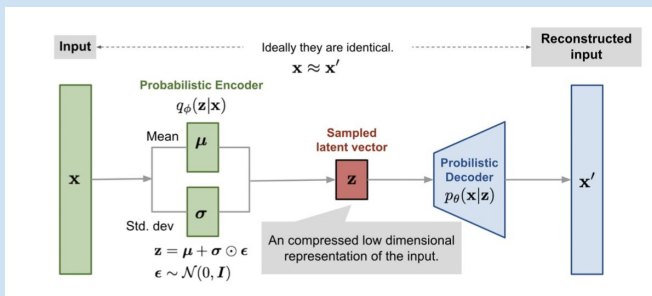


Large-radius jets (AK8)



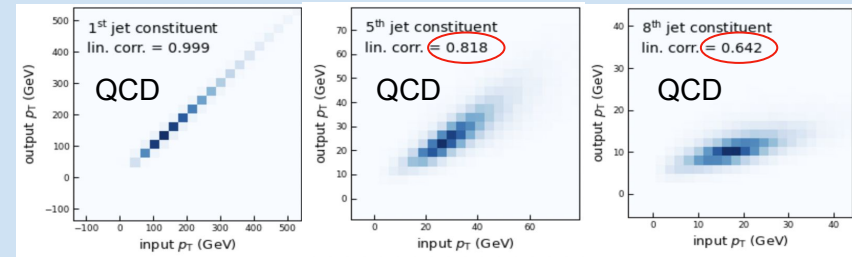
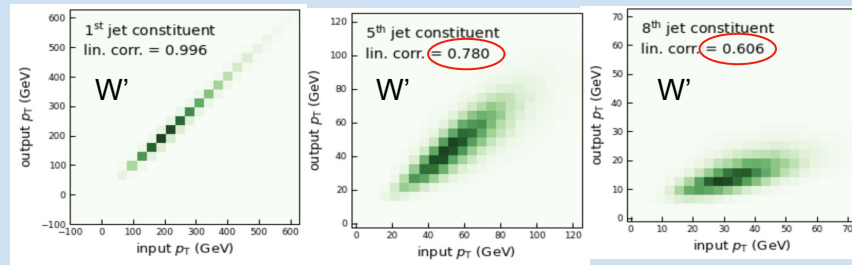
Challenge: Risk of being (mis-)guided by specific signal; need to be model-agnostic & data-driven.

Solution: Base selection on *SM jets veto* → anomaly detection based on ML to learn what SM jets look like



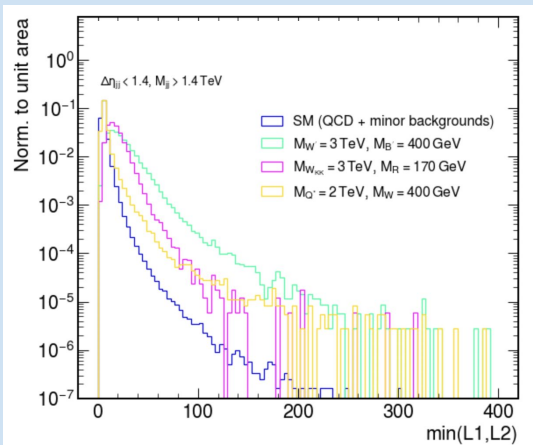
Train variational autoencoder (VAE) on QCD jets obtained from orthogonal control region

$M_{W'} = 3 \text{ TeV}, M_{B'} = 400 \text{ GeV } J_1$

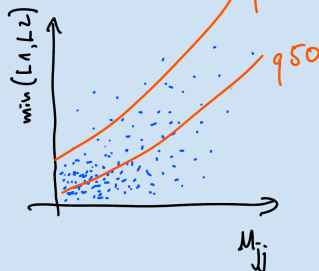
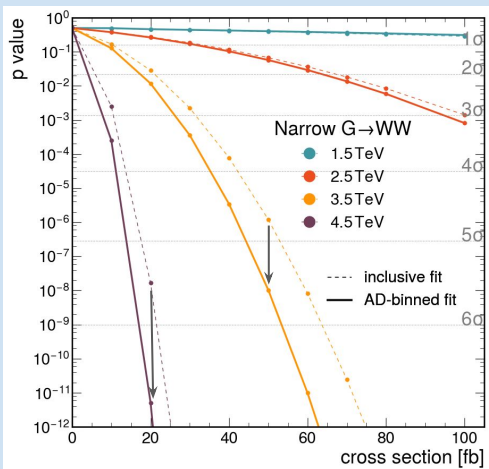
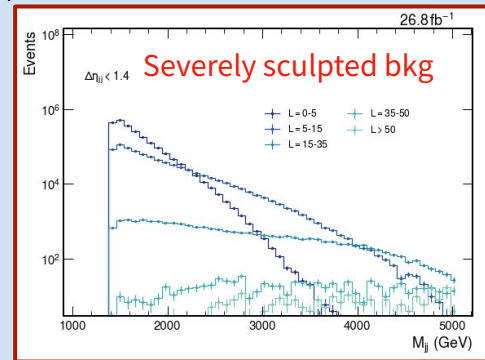
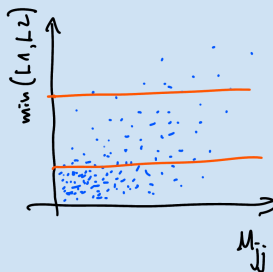


Worse reconstruction for BSM processes

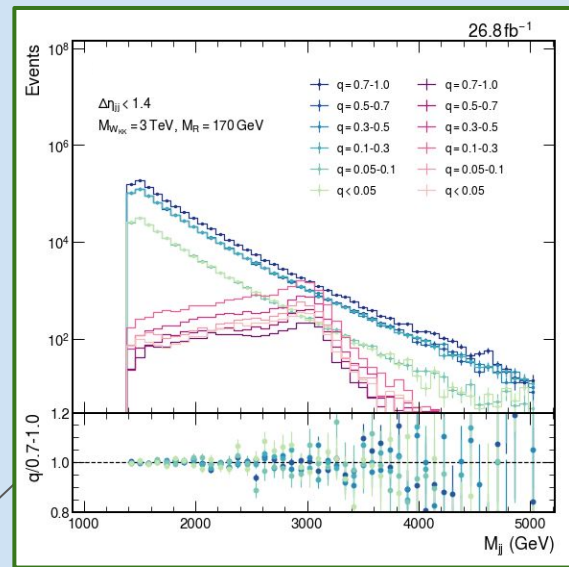
Search for clustered exotic hadronic activity, cont.



Loss (“Anomalousness”) is powerful event discriminator



Simultaneous fit of all categories gives dramatic improvement compared to inclusive fit

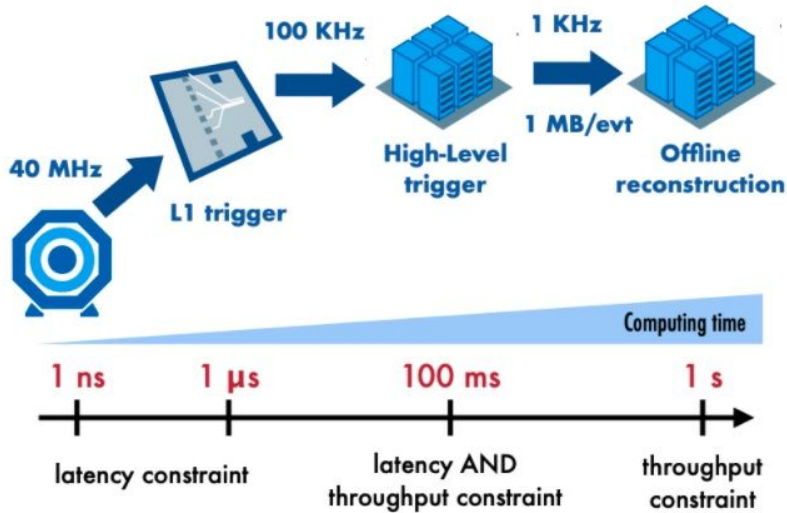




Intensity frontier

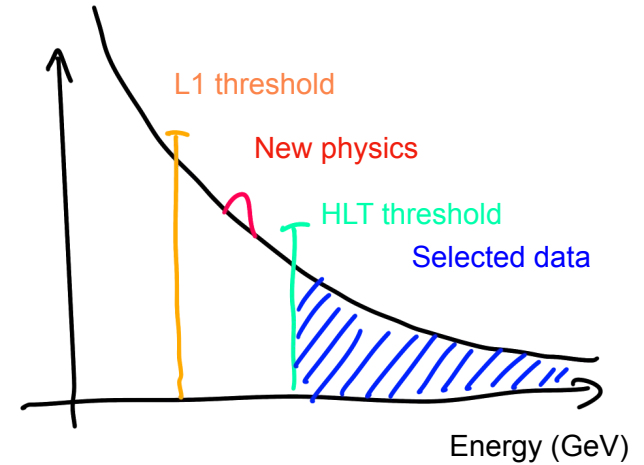
Look at smaller masses and higher rates

Bandwidth $\hat{=}$ rate \otimes size/evt



Limited bandwidth and disk space \rightarrow cannot afford to save the full detector information for every event to disk

Two-tiered **trigger system** in place to discard events in real-time (L1, on dedicated hardware) or almost-real-time (HLT, in software)



What if new physics is hiding below trigger threshold?

\rightarrow Change event content to accommodate lower thresholds
 \rightarrow Save only coarse HLT information in "Scouting" datastream

Only Hardware L1 trigger constraints; HLT pass-through; 100 fb⁻¹

Scouting at CMS during Run-2

L1-Trigger

Scouting Path	L1 trigger requirements
Muon CaloScouting	$2\mu, p_T > 15/7$ GeV $3\mu, p_T > 5/3/3$ GeV $2\mu, \text{OS}, p_T > 4.5$ GeV, $ \eta < 2, 7 < m(2\mu) < 18$ GeV $2\mu, \text{OS}, p_T > 0$ GeV, $ \eta < 1.5, \Delta R < 1.4$ $2\mu, \text{OS}, p_T > 4.5$ GeV, $ \eta < 2.5, \Delta R < 1.2$
H_T Calo/PF-Scouting	1 jet, $p_T > 180$ GeV 2 jets, $p_T > 30$ GeV, $ \eta < 2.5, m(2j) > 300$ GeV, $\Delta\eta < 1.5$ $H_T > 360$ GeV

Loose Cuts at HLT

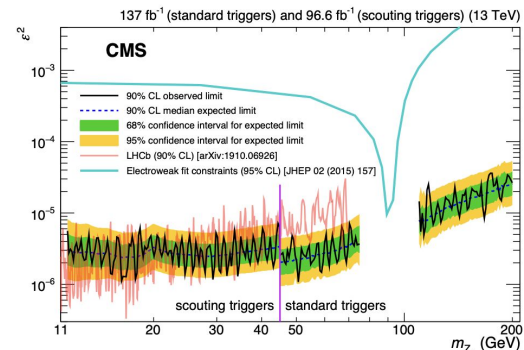
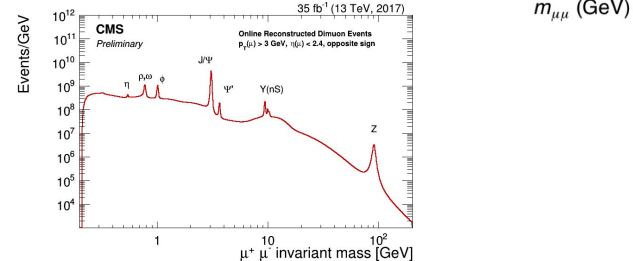
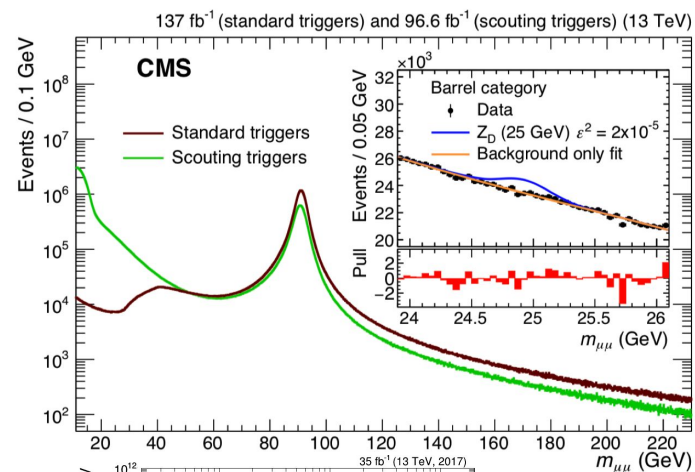
Scouting Path	Selection	Rate [kHz]	Proc. Time [ms]
Muon Calo-Scouting (2017–2018)	$(2\mu, p_T > 3$ GeV)	2.7	350
H_T Calo-Scouting (2016–2018)	$(H_T > 250$ GeV)	3	160
H_T PF-Scouting (2016–2018)	$(H_T > 410$ GeV)	0.7	1200

Scouting Datasets

Scouting Dataset	# events	dataset size	average size per event
Muon Calo-Scouting (2017–2018)	14.4 B	56 TB	3.9 kB
H_T Calo-Scouting (2016–2018)	37.7 B	78 TB	2.1 kB
H_T PF-Scouting (2016–2018)	7.7 B	66 TB	8.6 kB

Dimuon mass resolution slightly worse than LHCb (1-1.5%), not much worse than offline

Resolution for 2 TeV resonance using calo jets -20% compared to offline



Taking it to the extreme (i.e., below the Y)

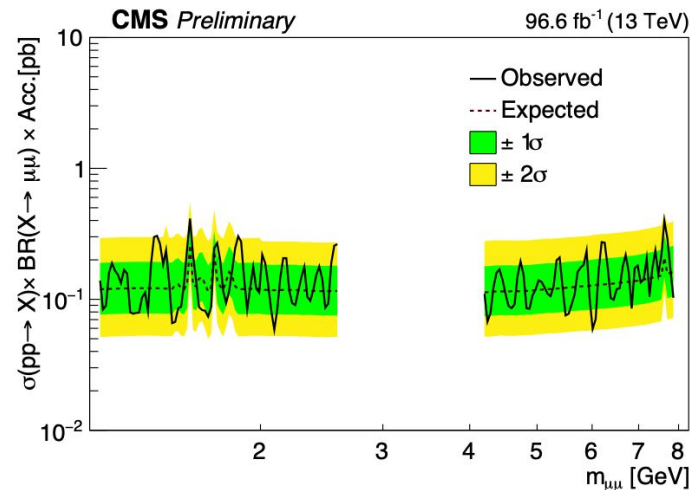
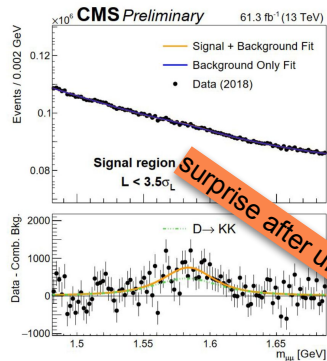
Fully exploiting Scouting data by selecting events with two muons $p_T > 4$ GeV (very soft!)

Novel MVA muon identification increases sensitivity by 30%

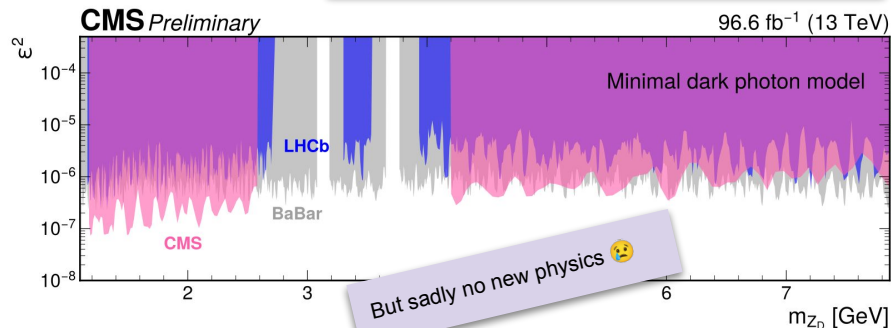
- Track quality
- Isolation
- Vertex information

Fit to dimuon mass distribution:

- signal: double Crystal Ball + Gaussian,
- background: empirical functions (e.g. polynomial times exponential)



More sensitive than LHCb or BaBar with “2nd-grade” data 😊



But sadly no new physics 😞

Lifetime frontier

Unlocking displaced signatures

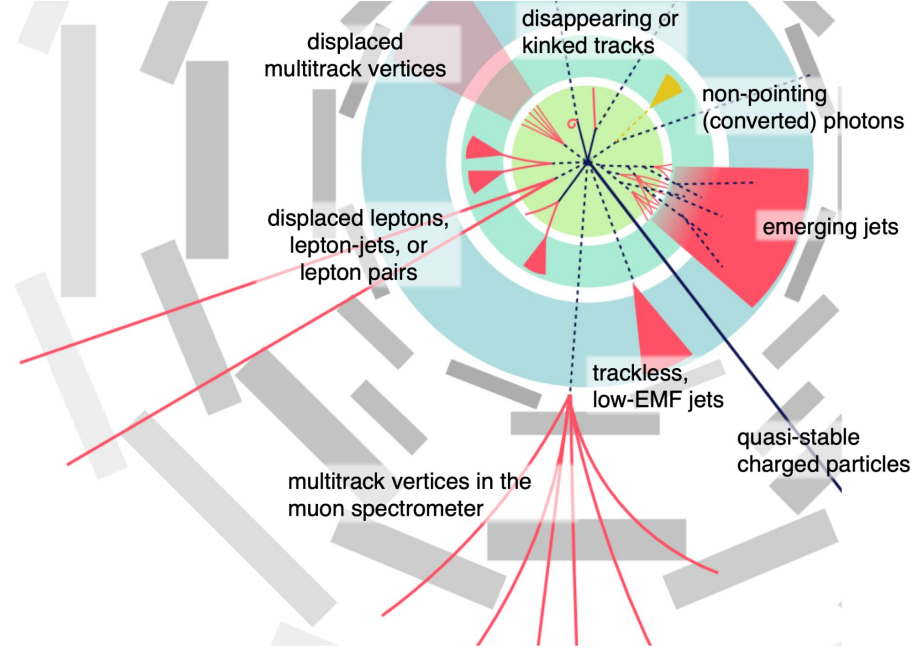
$$\Gamma \propto \varepsilon^2 \left(\frac{m}{\Lambda} \right)^{2n} \Phi$$

My time has come.

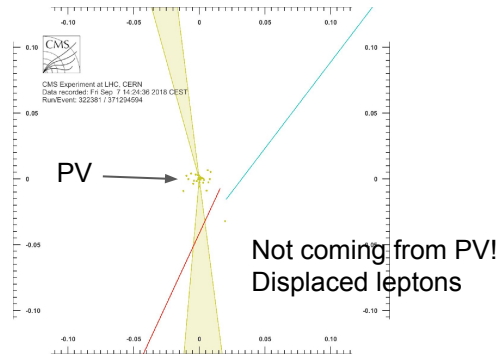
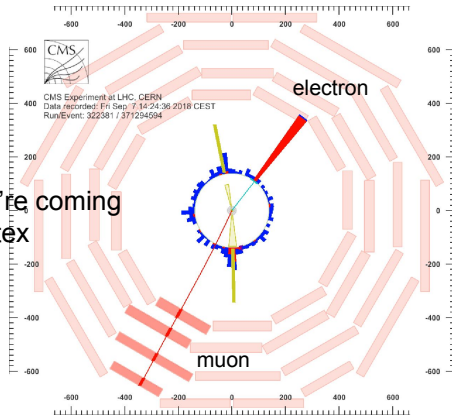
Why long-lived particles are tricky

CMS was **not designed** to look for **displaced** new physics

Reconstruction algorithms, cylindrical geometry, trigger, all designed assuming particles emerge from the collision point



Looking like they're coming from primary vertex



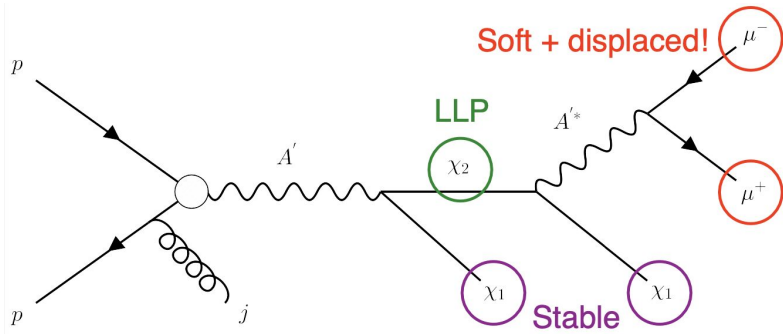
Not coming from PV!
Displaced leptons

Inelastic Dark Matter with displaced muons

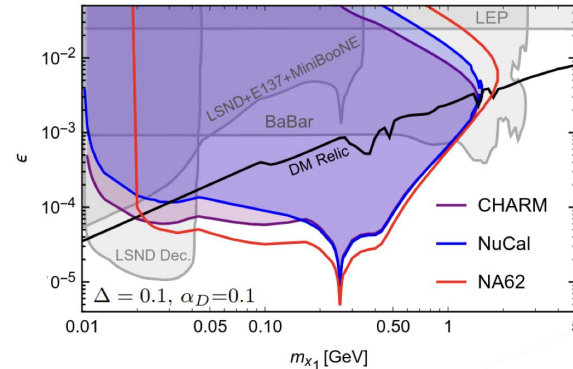
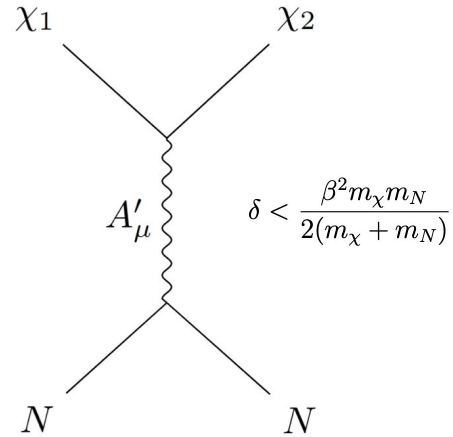
Two dark matter states with **small mass gap** coupled via dark photon. Depending on Δm_{DM} : no direct detection (“inelastic”)

Heavier DM (χ_2) has **long lifetime** due to **small mass splitting**
 → Pair of soft displaced muons overlapping with MET (from χ_1)

$$c\tau \propto \frac{(m_{A'})^4}{(\Delta m_{\text{DM}})^5} ; \text{ target } 1\text{mm} - 1000\text{mm range}$$



Off-diagonal matrix element dominating



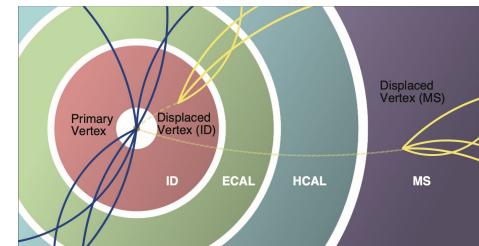
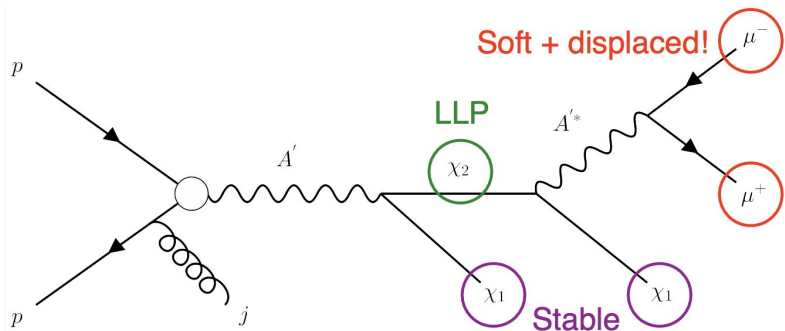
Highly constrained by beam-dump experiments for < 1 GeV

Inelastic Dark Matter with displaced muons

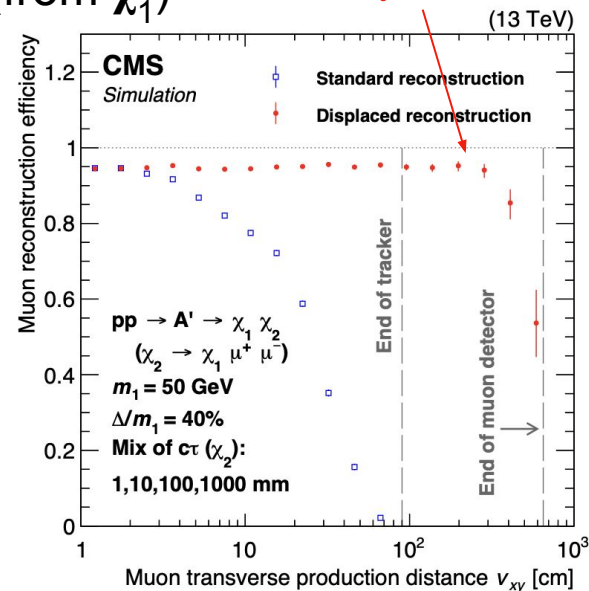
Two dark matter states with **small mass gap** coupled via dark photon. Depending on Δm_{DM} : no direct detection (“inelastic”)

Heavier DM (χ_2) has **long lifetime** due to **small mass splitting**
 → Pair of soft displaced muons overlapping with MET (from χ_1)

$$c\tau \propto \frac{(m_{A'})^4}{(\Delta m_{\text{DM}})^5} ; \text{target } 1\text{mm} - 1000\text{mm range}$$



Dedicated muon reconstruction using only muon system



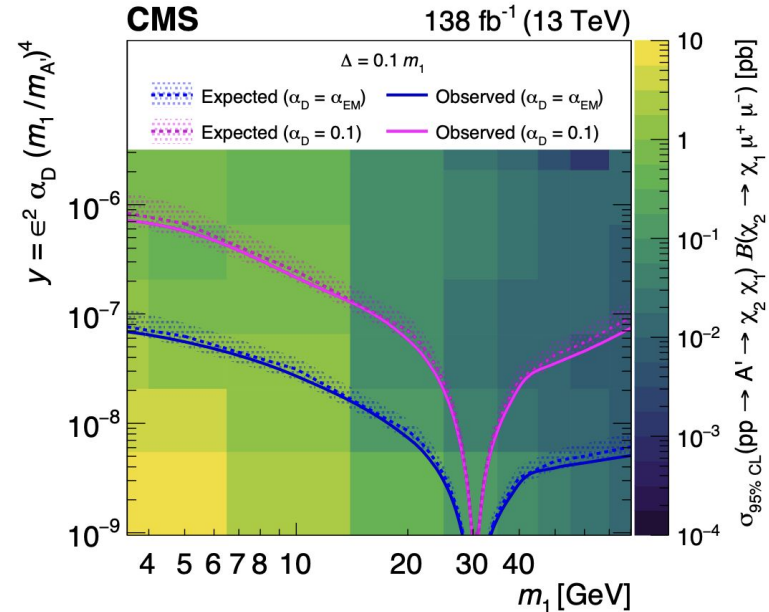
Inelastic Dark Matter with displaced muons

$$m_{A'} = 3 m_1$$


Categorize by number of displaced muons that match in direction to a standard muon

- Fewer matches \rightarrow larger displacement
- Use kinematics, isolation, displacement parameter to both suppress and estimate background (mostly QCD)

	Events per SR category		
	0-match	1-match	2-match
Pred.	1.2 ± 0.6	0.5 ± 0.3	0.5 ± 0.3
Obs.	2	0	0



First such search at a hadron collider!



Preparing for the High-Luminosity LHC
Novel Hardware

Phase-II upgrade: **replacement of CMS endcap** instrumentation (→ very congested region) with a novel detector

Let's talk about **our needs**:

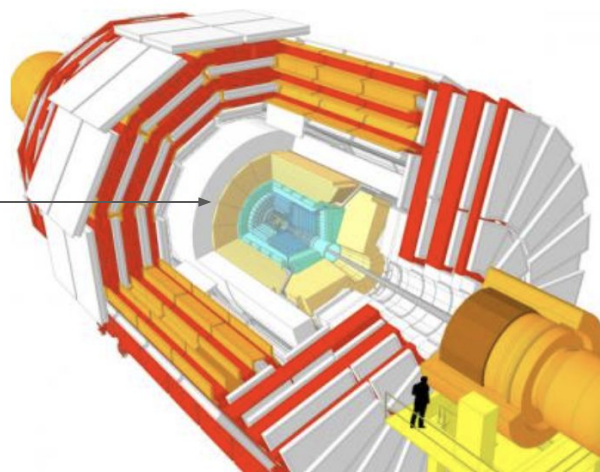
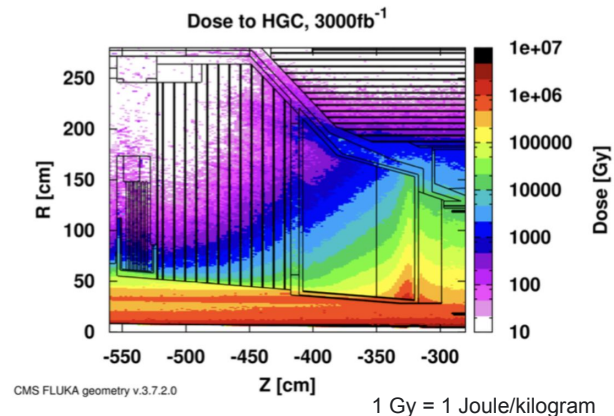
expensive endcaps

- Withstand extreme exposure
- Good pileup rejection capability
- Superior shower reconstruction

cheap endcaps



Current endcaps only designed for $< 500 \text{ fb}^{-1}$



Phase-II upgrade: replacement of CMS endcap instrumentation (→ very congested region) with a Highly Granular Calorimeter (“HGCal”)

What **silicon** has to offer:

Radiation hardness
Fast response

↔

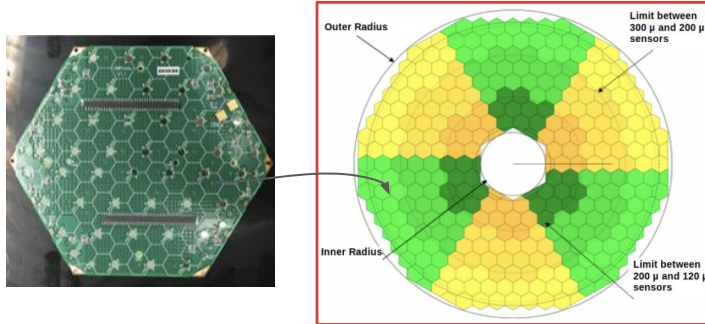
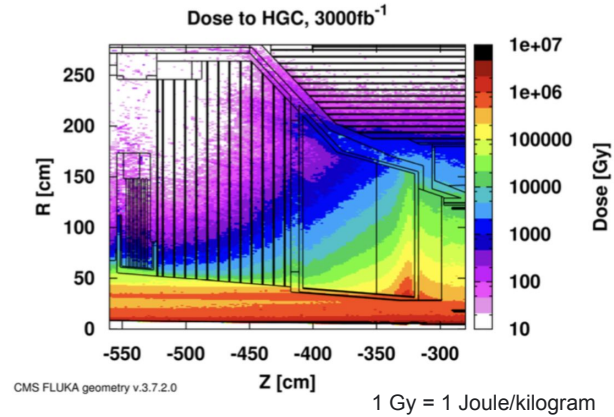
Withstand extreme exposure
Good pileup rejection capability

↔

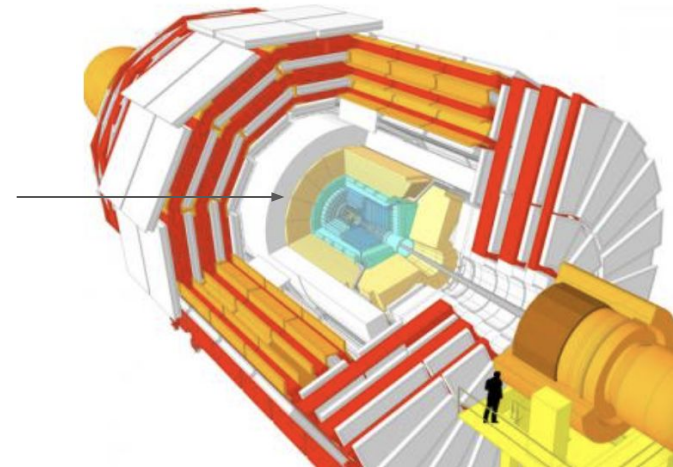
Can be finely segmented

↔

Superior shower reconstruction



Current endcaps only designed for $< 500 \text{ fb}^{-1}$

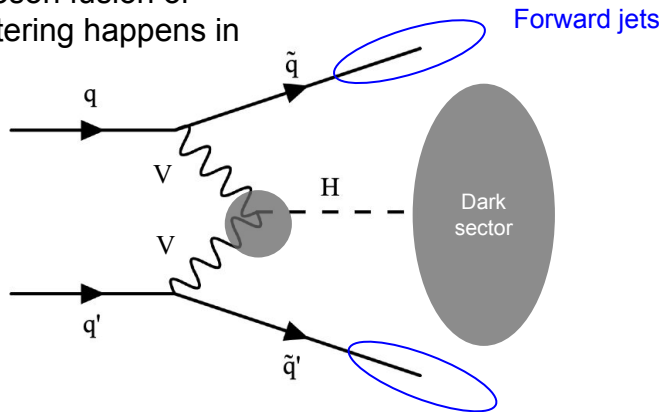


Highly Granular Calorimeter (“HGCal”) Silicon wafer

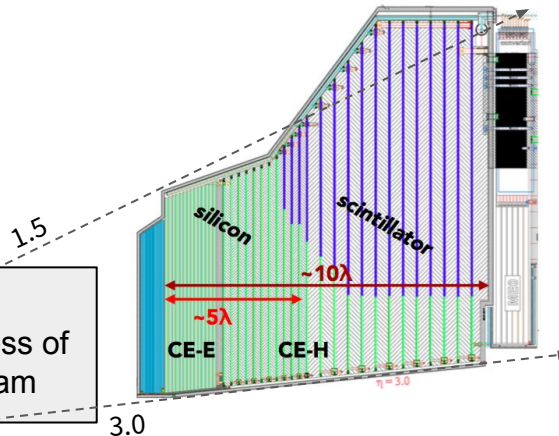
How dark sectors will gain

Anything vector boson fusion or vector boson scattering happens in that region!

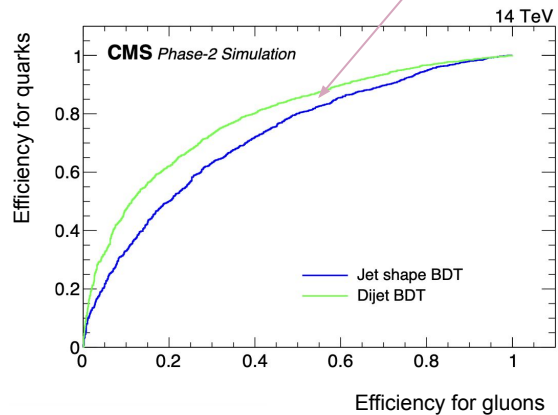
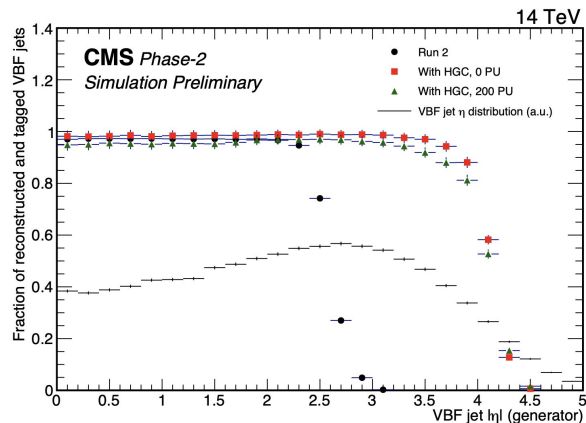
VBF H
VBF HH
VBS ...




$1.5 < \eta < 3.0$
→ crucial for success of LHC physics program



Highly segmented detector allows for good quark-gluon discrimination





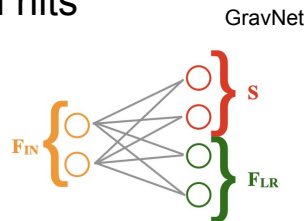
Preparing for the High-Luminosity LHC
Novel Software

Reconstruction with AI

6M readout channels → traditional algorithms fail

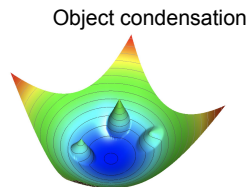
End-to-end neural networks to perform ML-driven particle reconstruction based on hits

Graph neural networks with efficient kNN aggregation in latent space

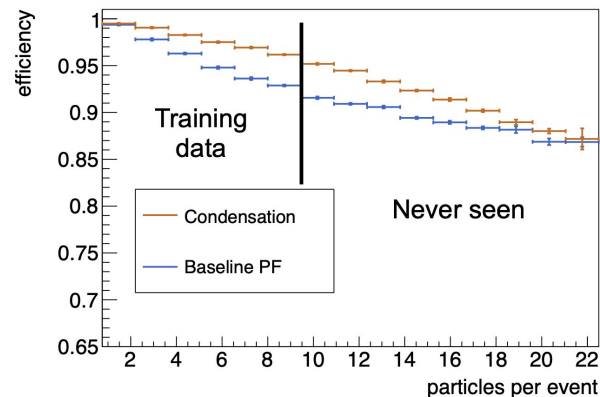
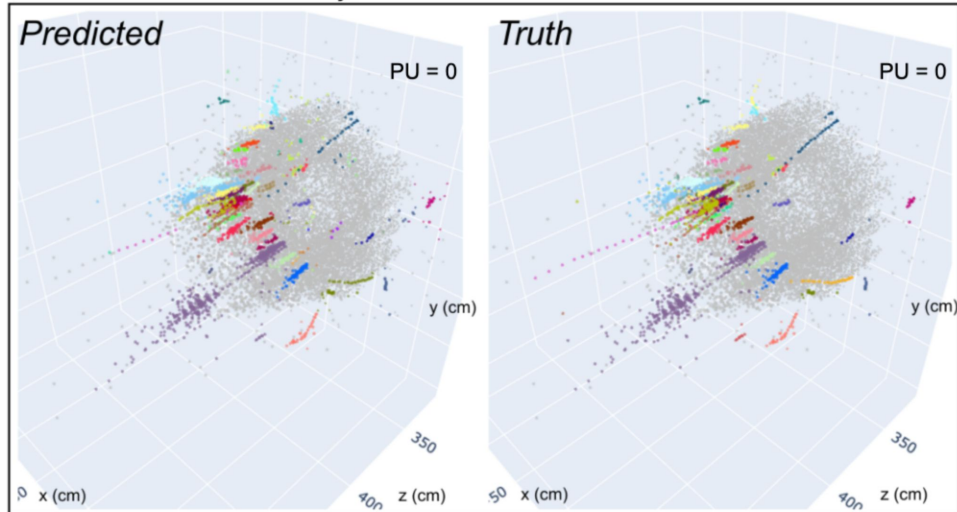



Physics-inspired loss functions to promote single hits around which other hits cluster and predict E/type of incident particles

- Outperforms classical ParticleFlow
- Excellent extrapolation (adaptation) properties



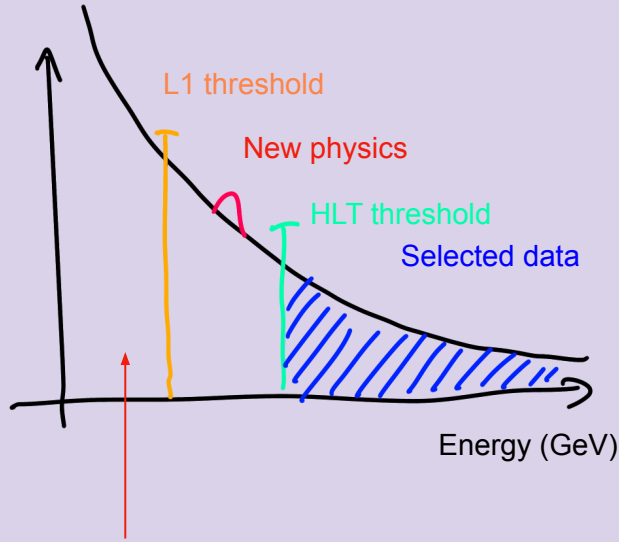
CMS Simulation Preliminary



The background is a dark blue field filled with a complex, dense network of thin, light blue lines that radiate from a central point, creating a starburst or web-like pattern. A horizontal line of small, light blue dots is positioned across the middle of the image, slightly above the text.

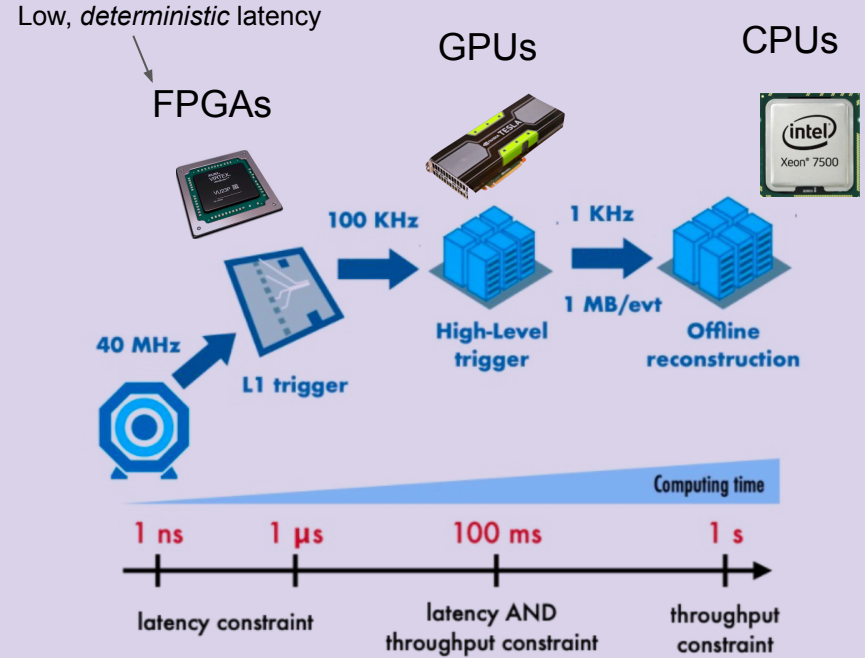
Preparing for the High-Luminosity LHC
Novel Firmware

Coming back to this

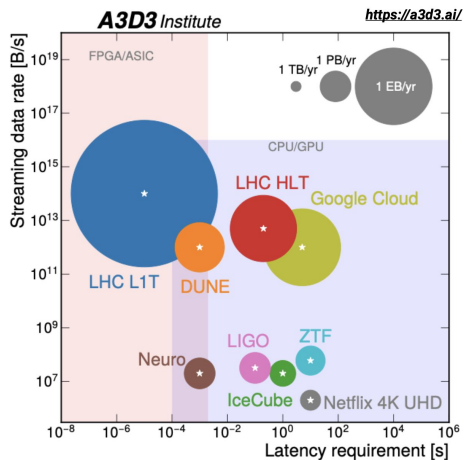


What if new physics resides below even L1 threshold?

Can we make the **hardware** detect interesting events through **AI**?



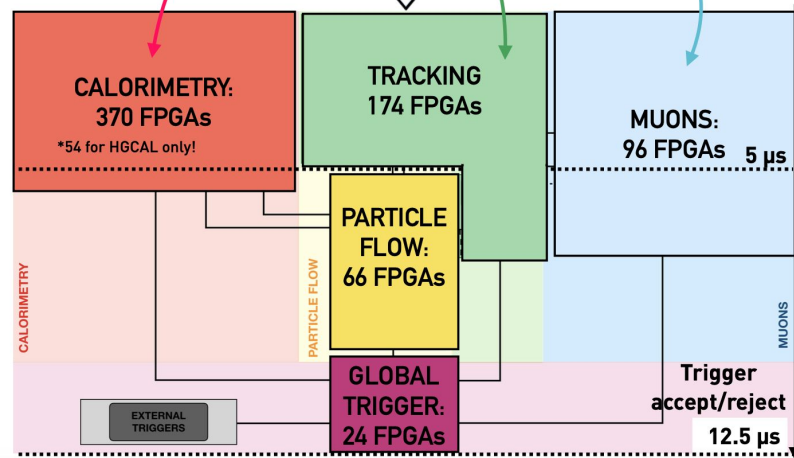
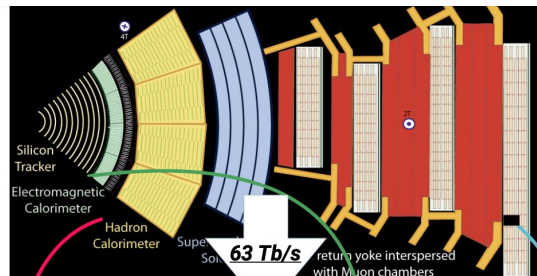
Hardware and AI melt together



The problem clearly is the low time (microseconds) available at L1 to look at the events

Yet, data intensity (\geq Google) asks for AI-based solutions

FPGAs not very easy to work with ...
hls4ml to the rescue

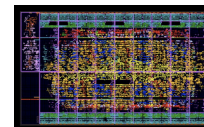


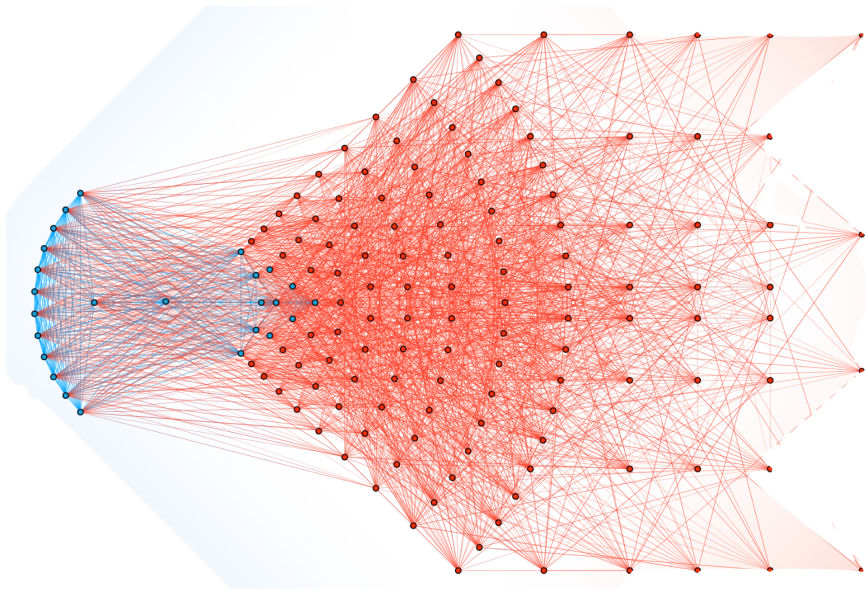
```

hls_model = hls4ml.converters.convert_from_keras_model(
    model12,
    hls_config=hls_config,
    output_dir=output_dir,
    backend=backend,
    io_type=io_type,
)

X = np.random.rand(1, 28, 28, 1)

hls_model.compile()
    
```





How can we fit this on that?



Model quantization

Article | [Published: 21 June 2021](#)

Automatic heterogeneous quantization of deep neural networks for low-latency inference on the edge for particle detectors

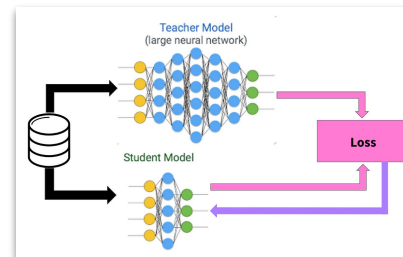
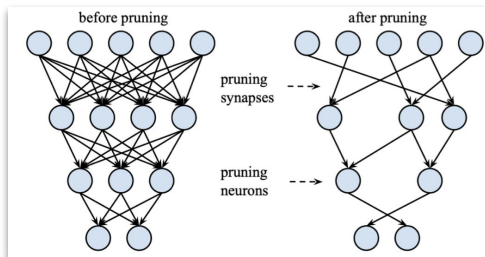
[Claudionor N. Coelho Jr](#), [Aki Kuusela](#), [Shan Li](#), [Hao Zhuang](#), [Jennifer Ngadiuba](#), [Thea Klæboe Aarrestad](#)

✉ [Vladimir Loncar](#), [Maurizio Pierini](#), [Adrian Alan Pol](#) & [Sioni Summers](#)

Nature Machine Intelligence **3**, 675–686 (2021) | [Cite this article](#)

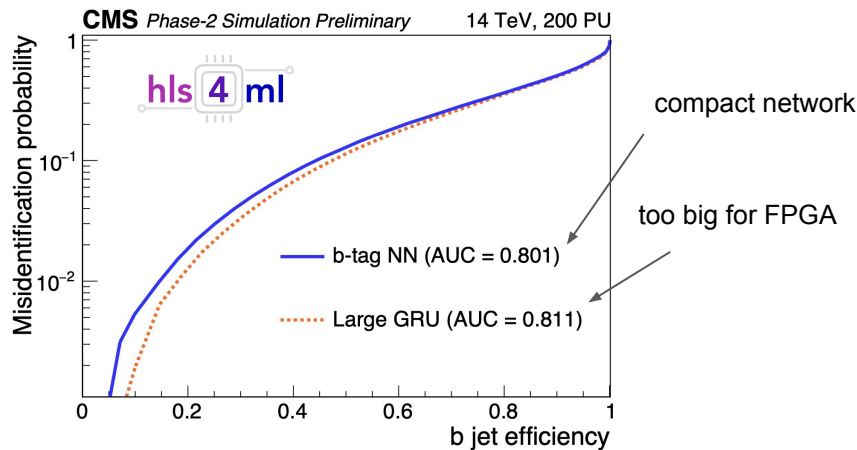
32bit → 8bit
1/2 memory
20x less power

Model pruning



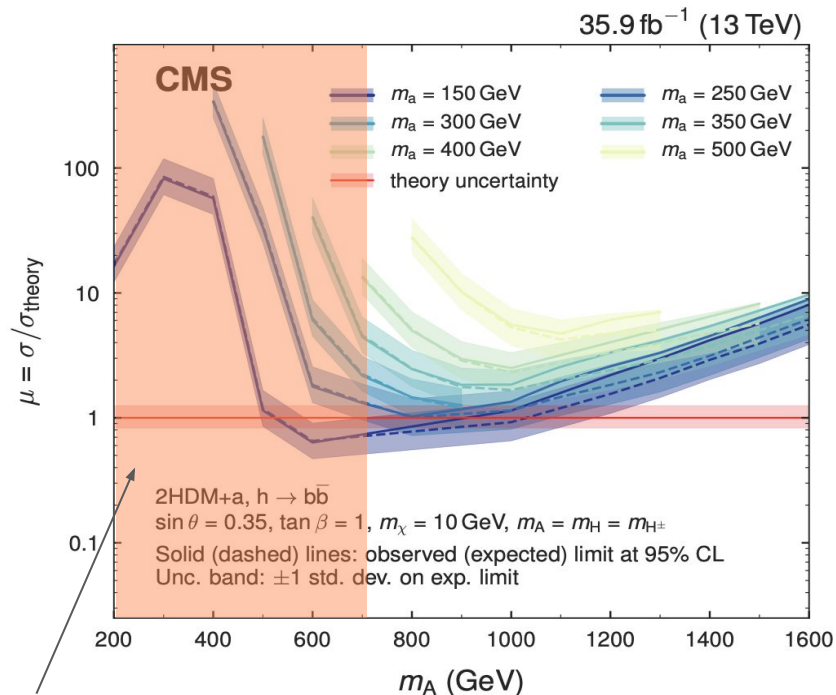
Model distillation

How dark sectors will gain



AI on FPGAs will result in (among others):

- higher yield in true Higgs boson (bb) events
- better object reconstruction
→ better MET resolution



Heavily limited by poor MET resolution at L1/HLT

Triggers fully efficient only for True MET > 250 GeV

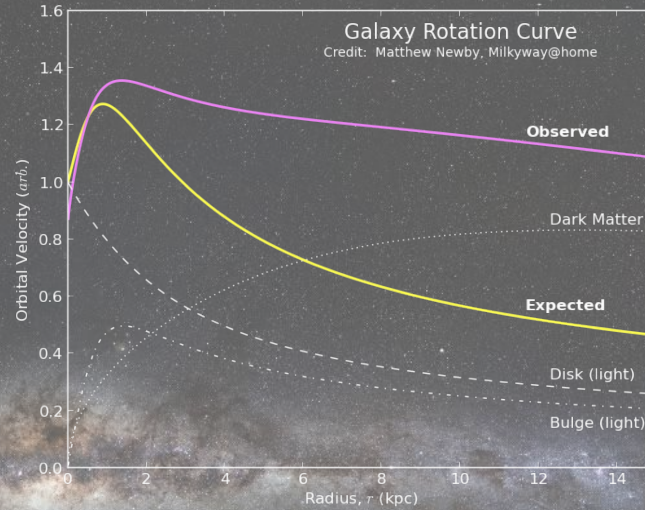
Conclusion

CMS operates & develops at the energy, intensity, and lifetime frontiers

Pushing both software and hardware (and the merger of both) to the limits

This allows for a rich search program for dark physics

The path towards the HL-LHC is pretty clear



~~A Shot in the Dark~~ *We have a plan*
Searching for a Dark Sector with the CMS
Experiment

Benedikt Maier, KIT
SLAC, Sept 5, 2023