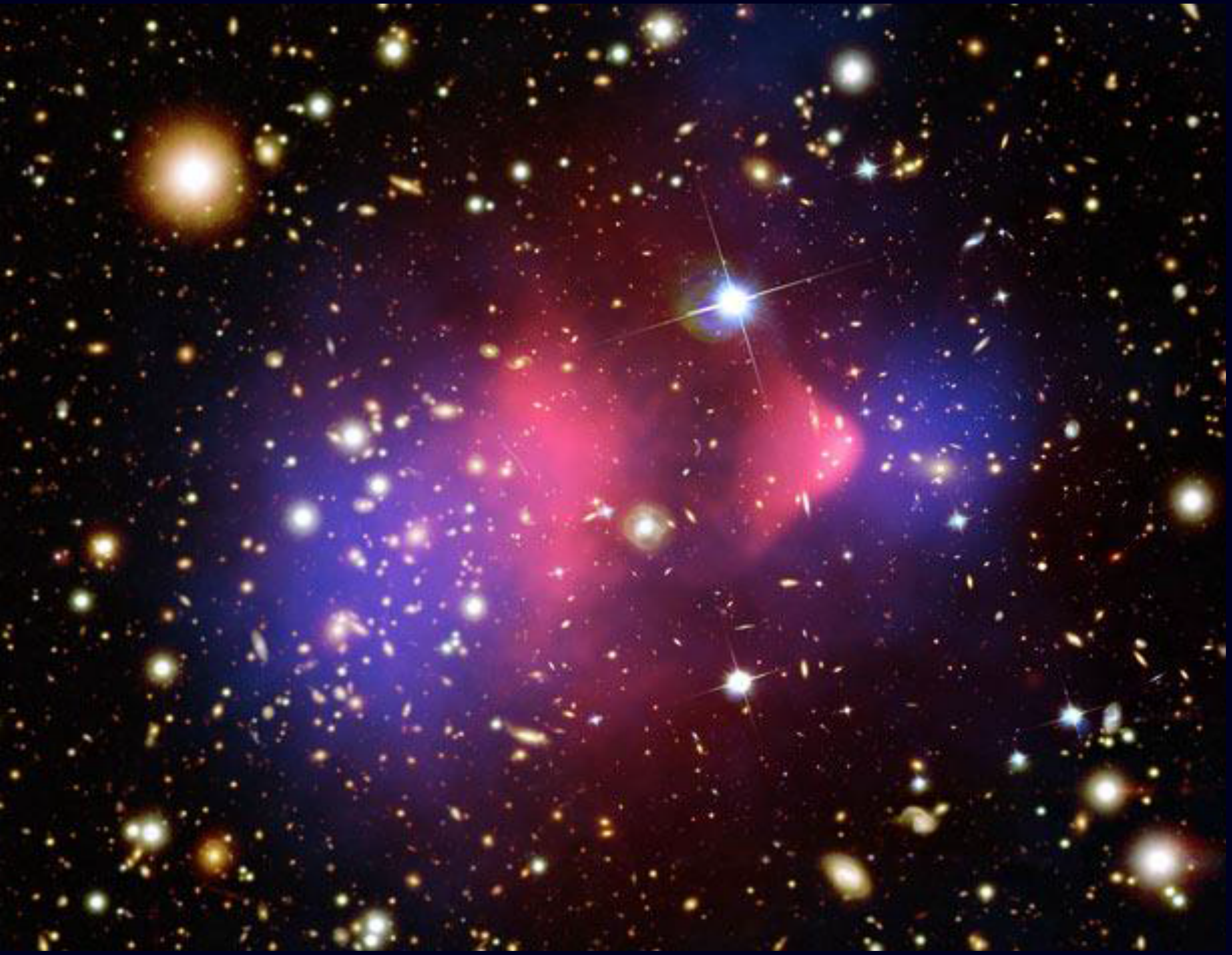


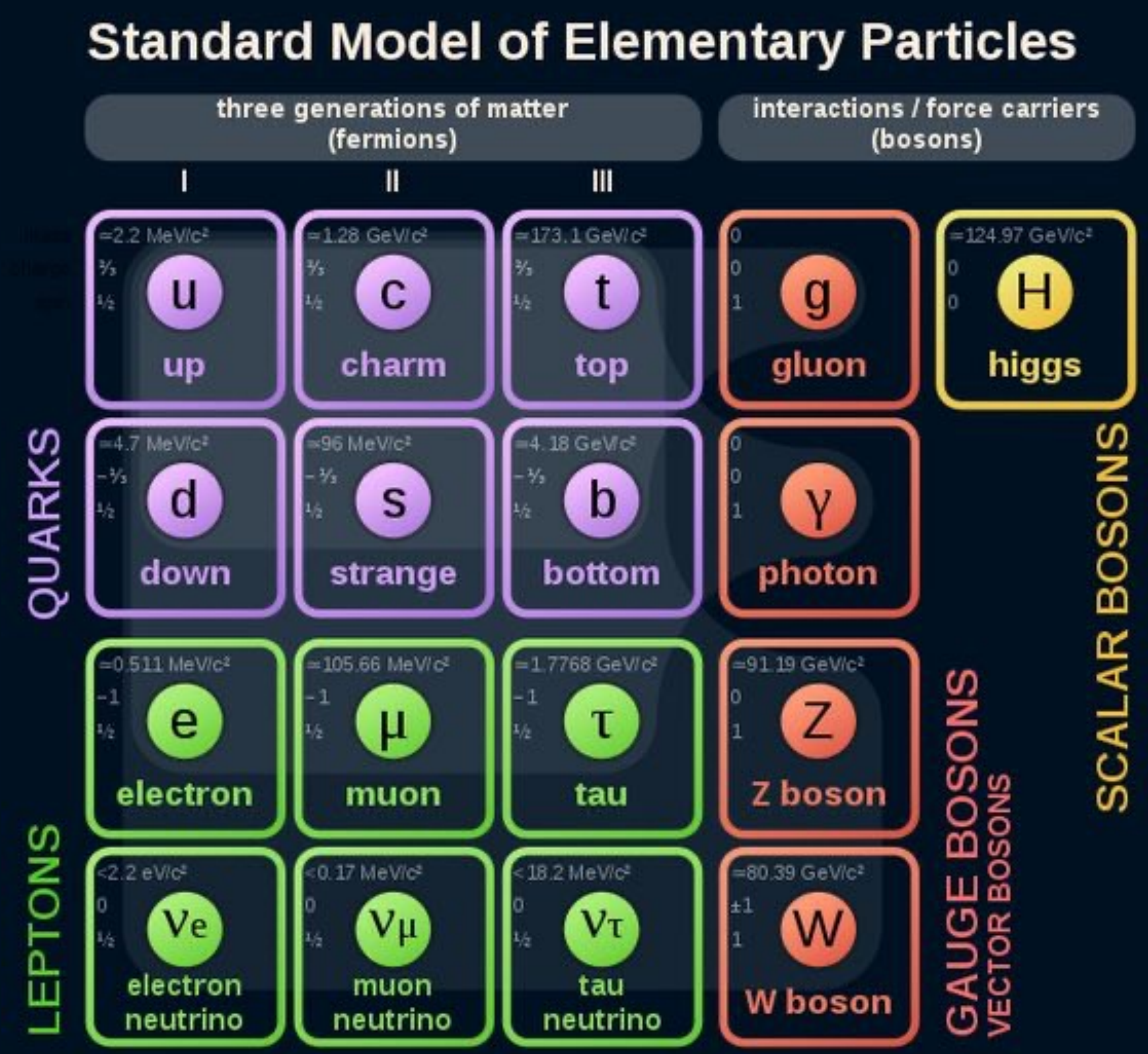
Hunting Dark Matter in a Hidden Valley

Annapaola de Cosa
ETH Zurich

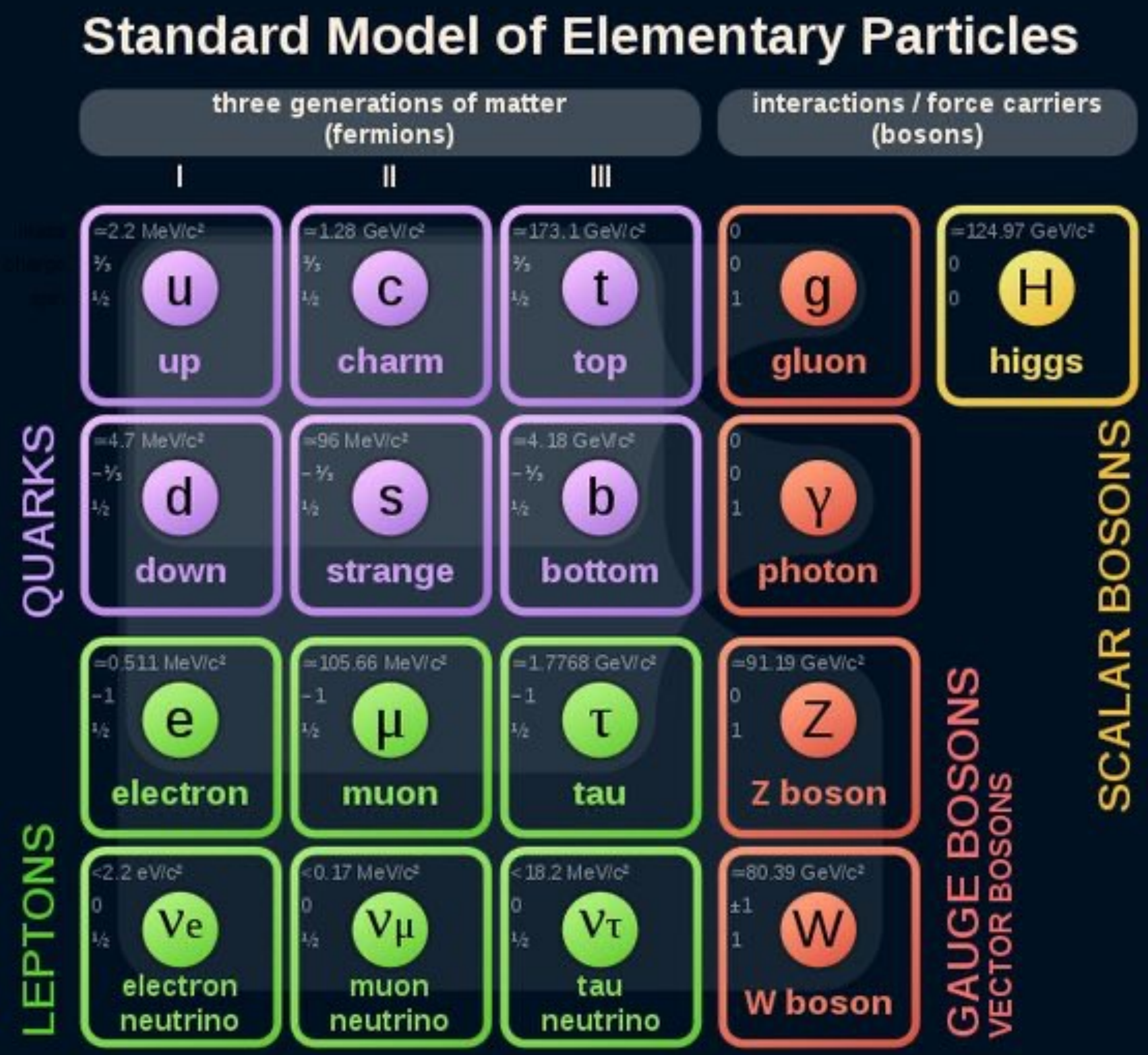
Outlook

- Dark Matter: status of the art of searches at colliders
  - **The quest for strongly-coupled Dark Matter**
- **Novel experimental signatures** to probe
- **Novel techniques** to search for DM and new physics

The Standard Model of Particle Physics



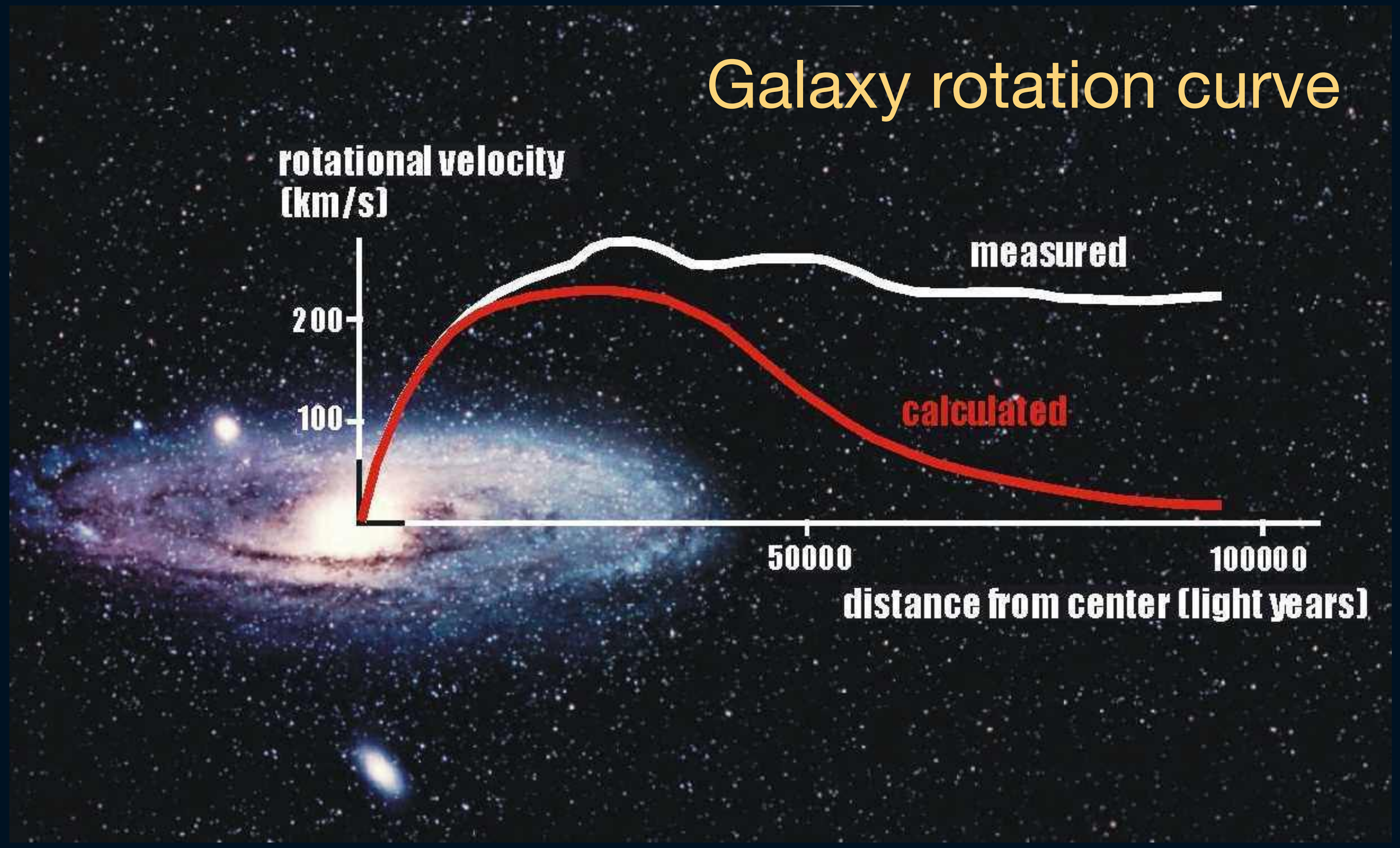
The Standard Model of Particle Physics

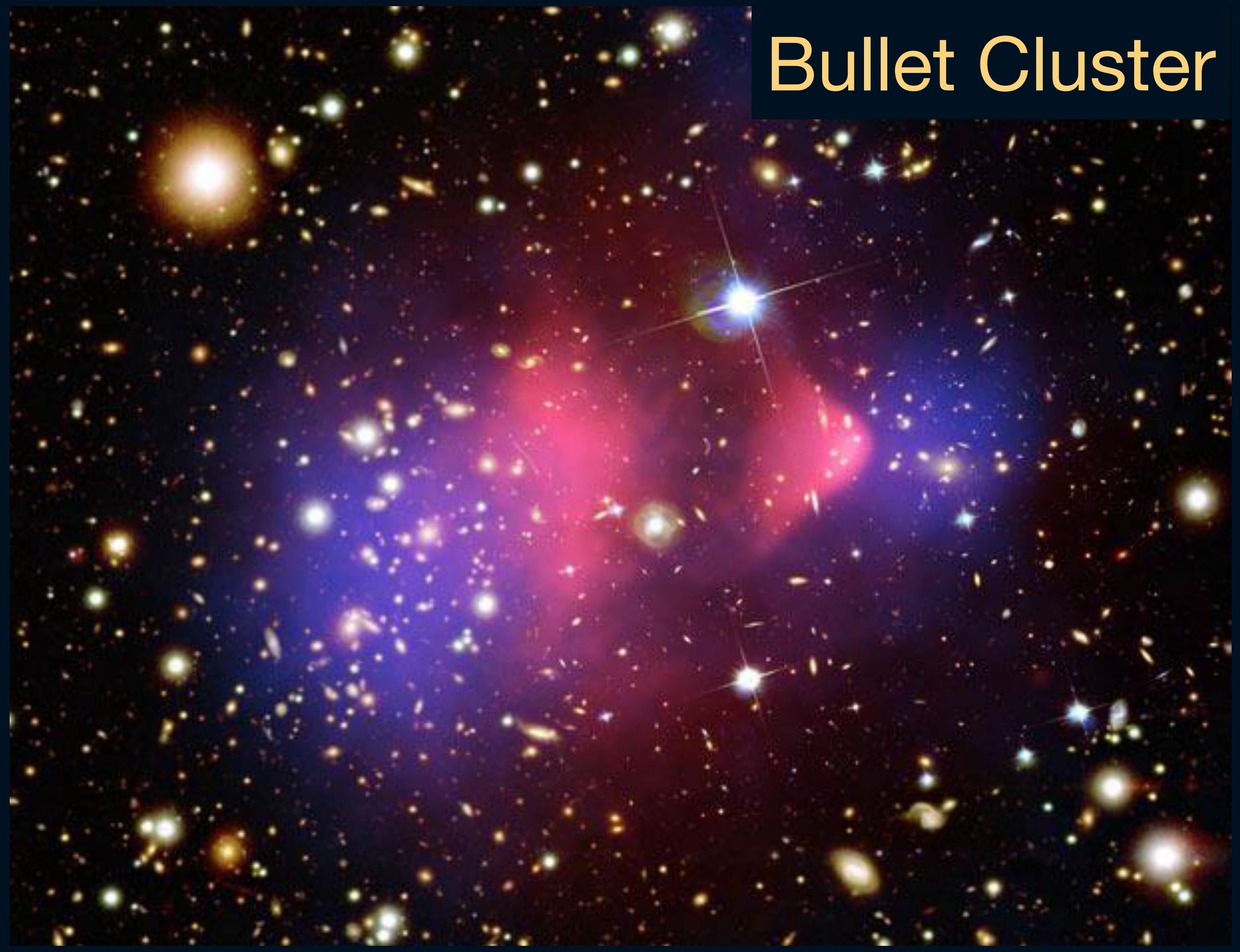


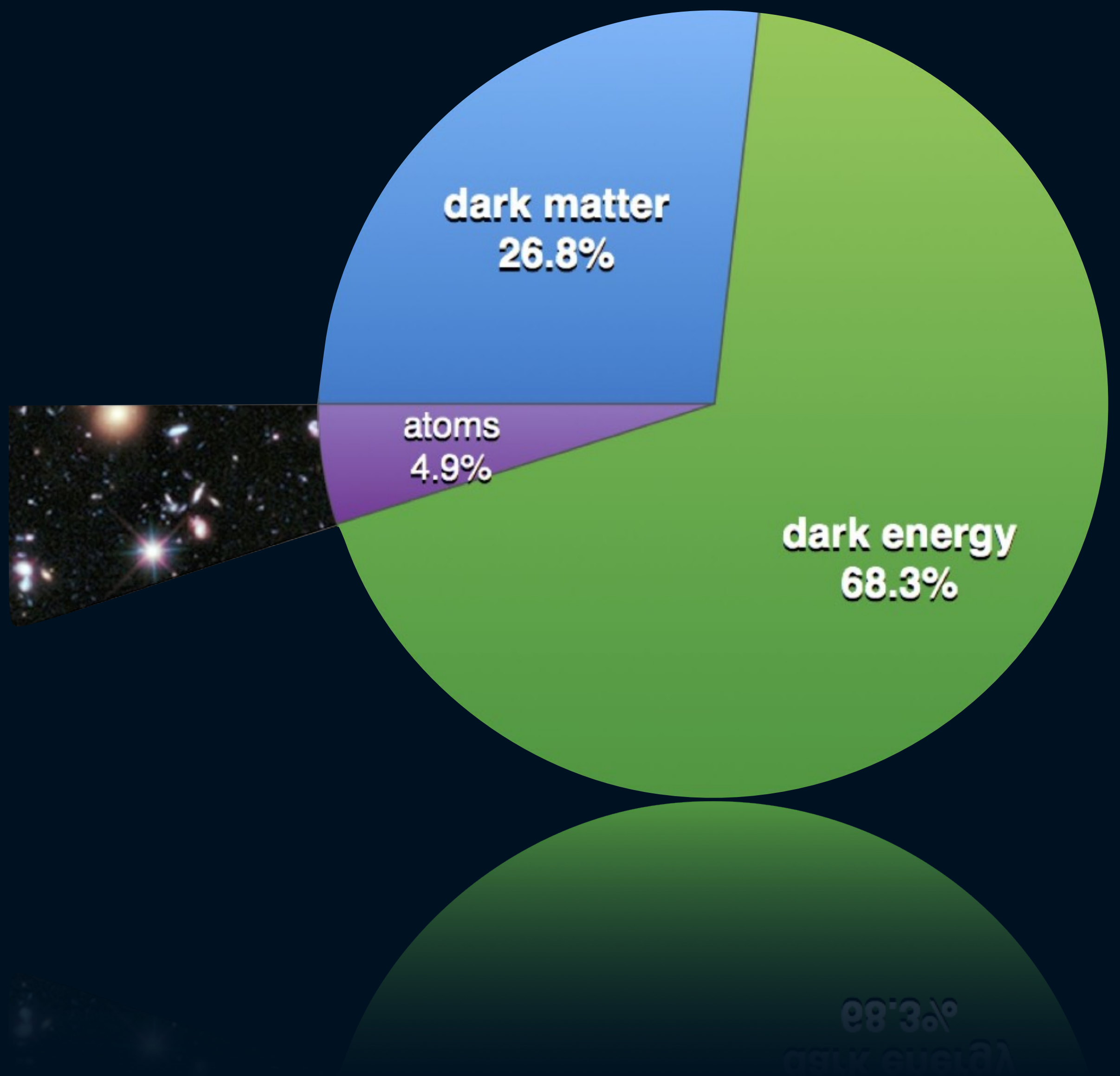
Numerous questions cannot find an explanation in the SM framework

- Can we fit gravity in the SM?
- Why are there three generations?
- What is the origin of particle mass hierarchy?
- Why do matter and antimatter behave differently?
- **What is Dark Matter made of?**

What about Dark Matter?



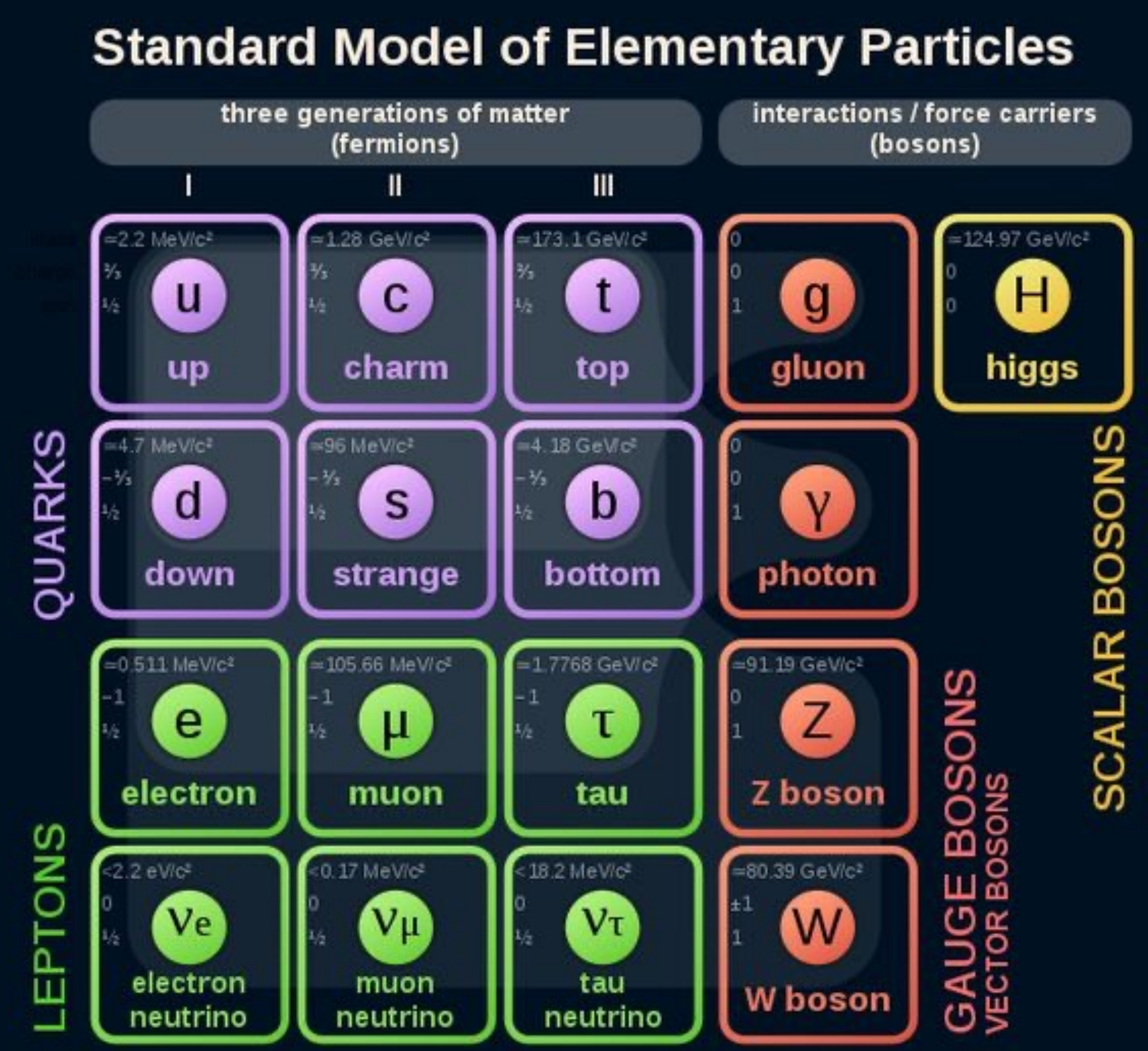




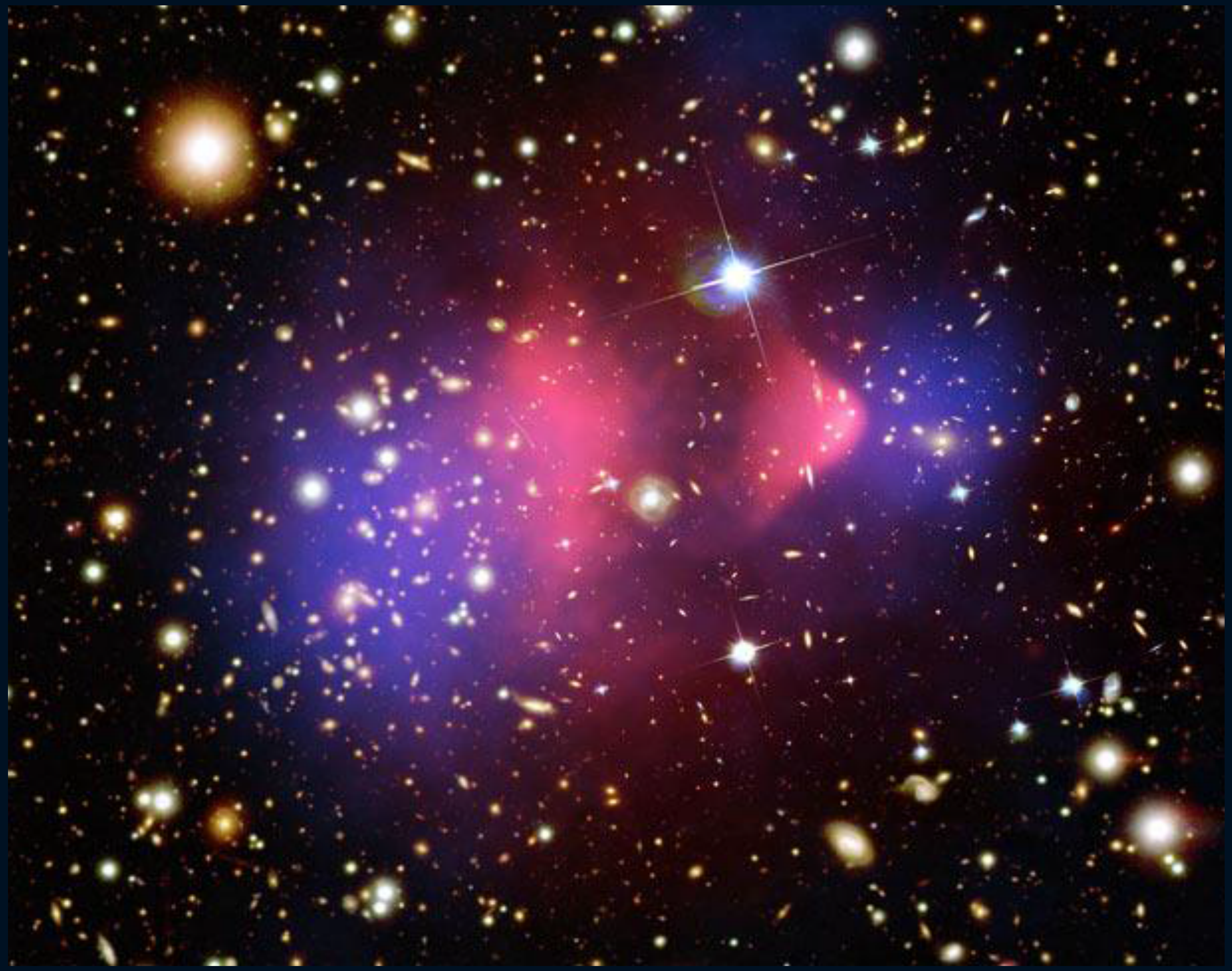
Gravitational lensing

What's Dark Matter nature?

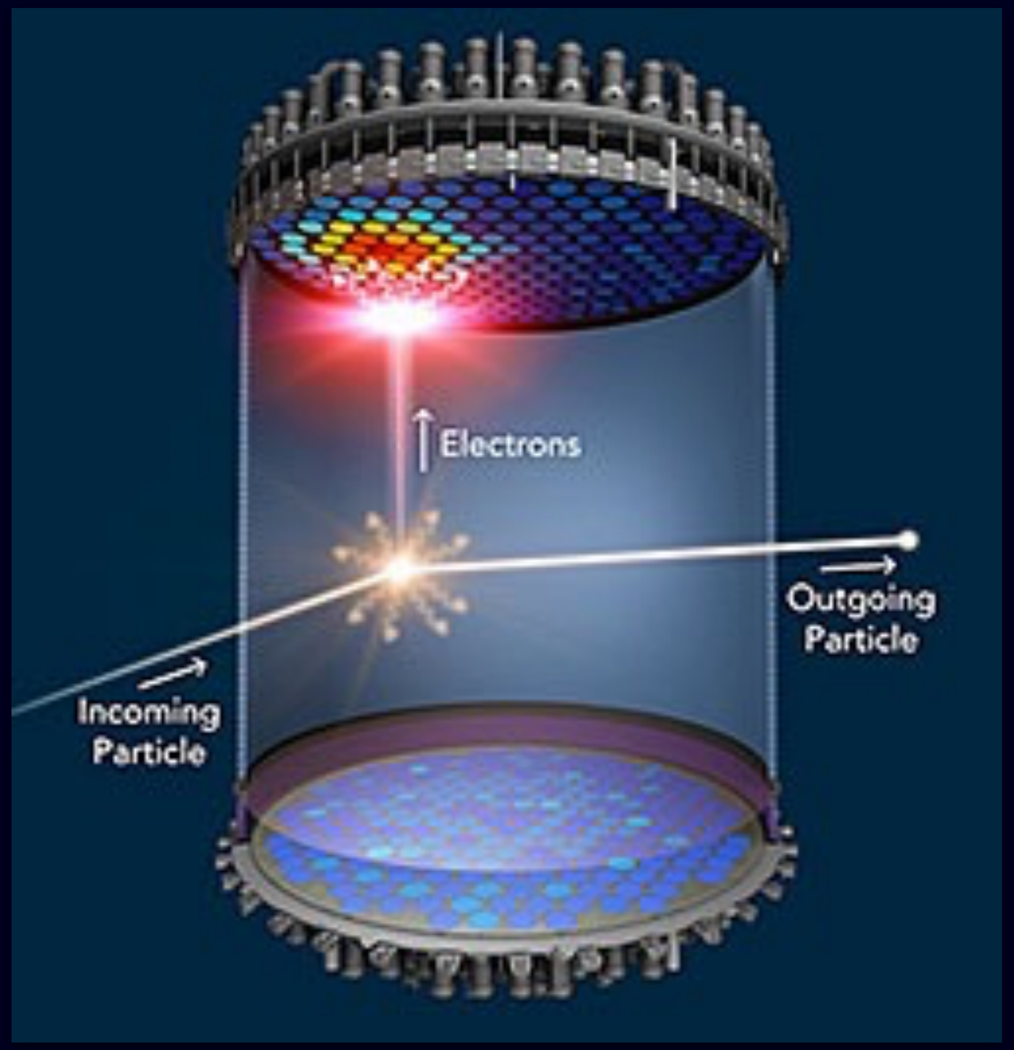
Beyond the Standard Model



Dark Matter



DM quest - Experimental landscape

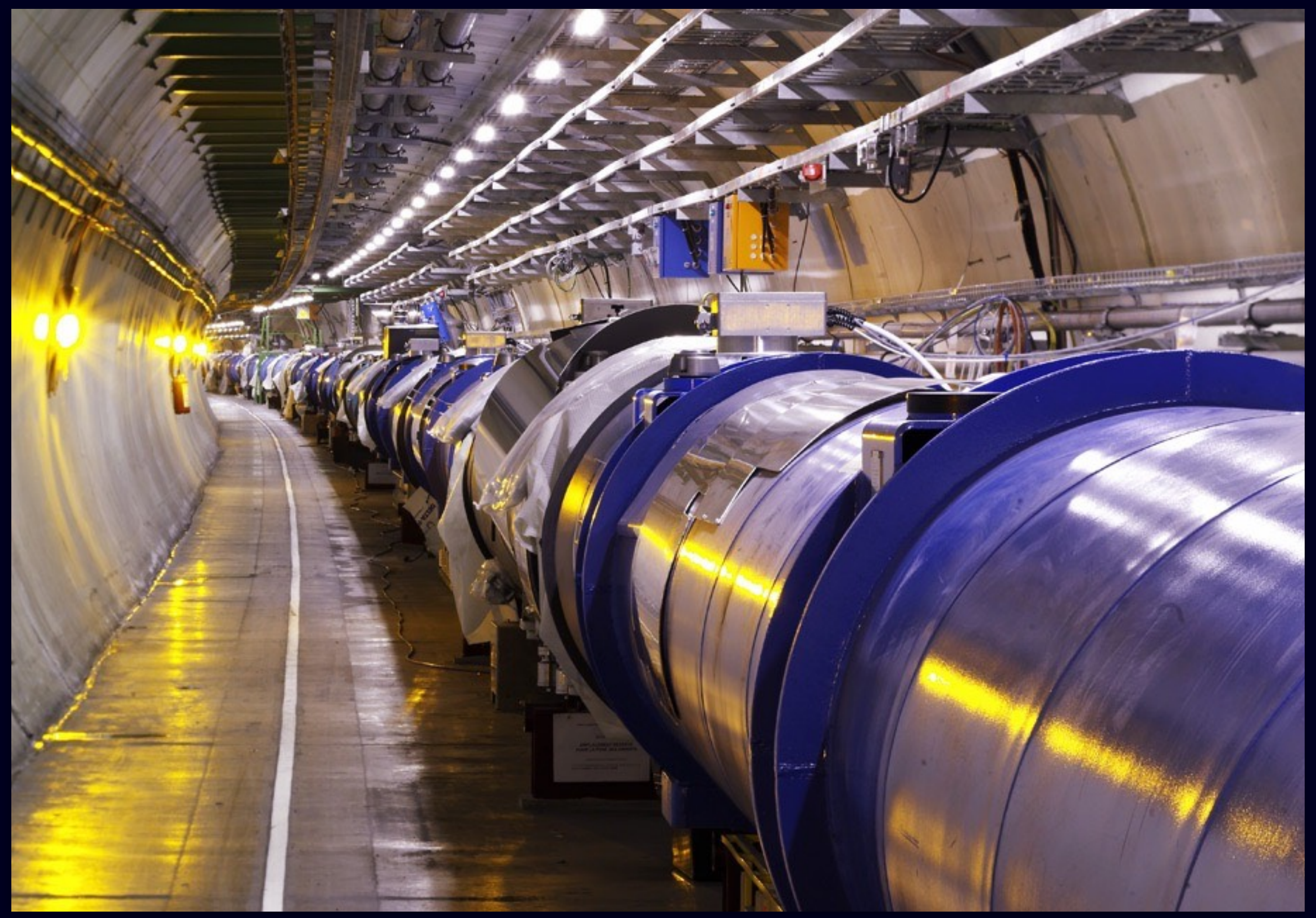


Direct detection

- DM-nucleon scattering

Indirect detection

- DM annihilation products



Particle colliders

- Direct production of DM

DM quest - Theory landscape

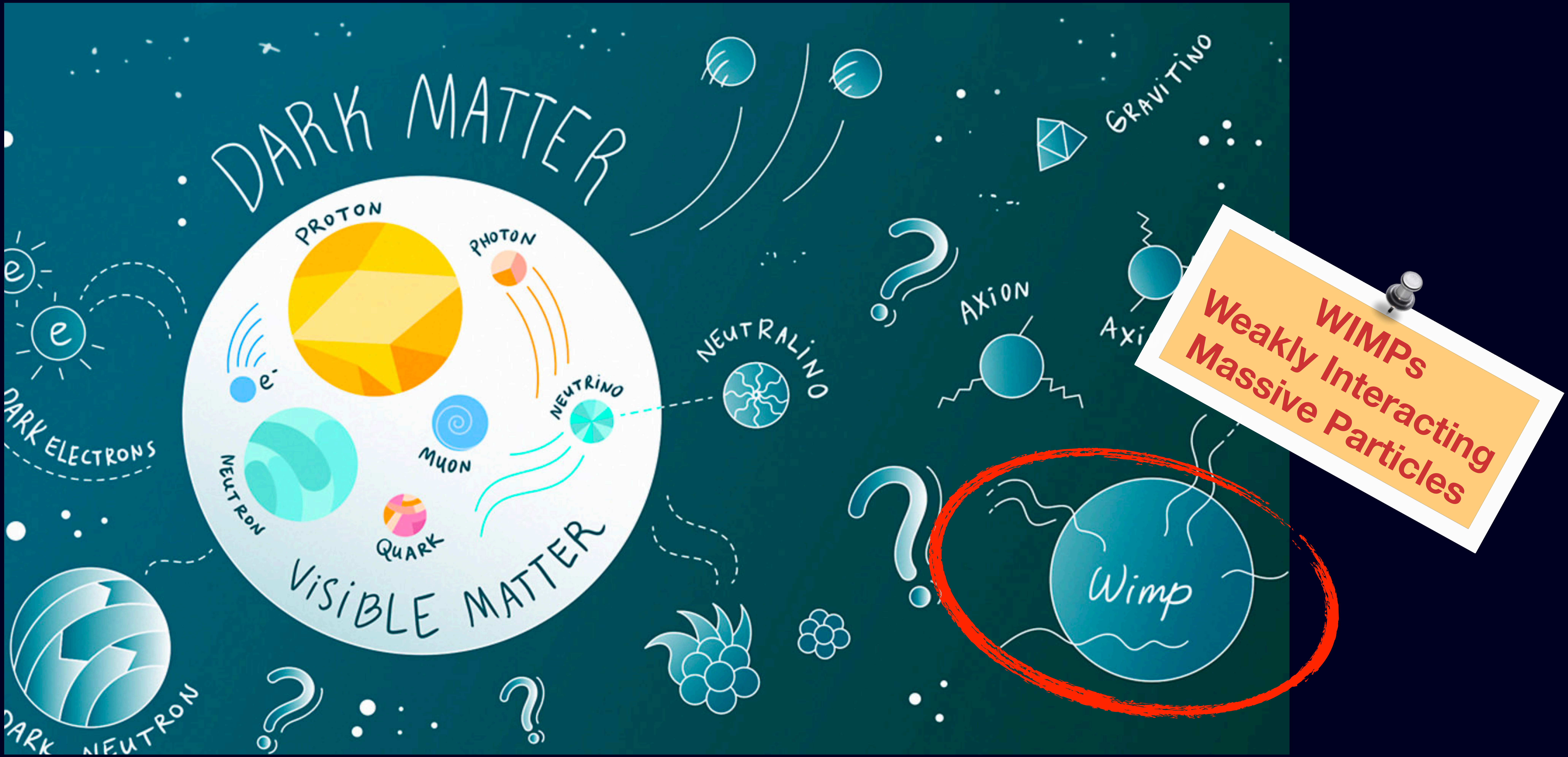
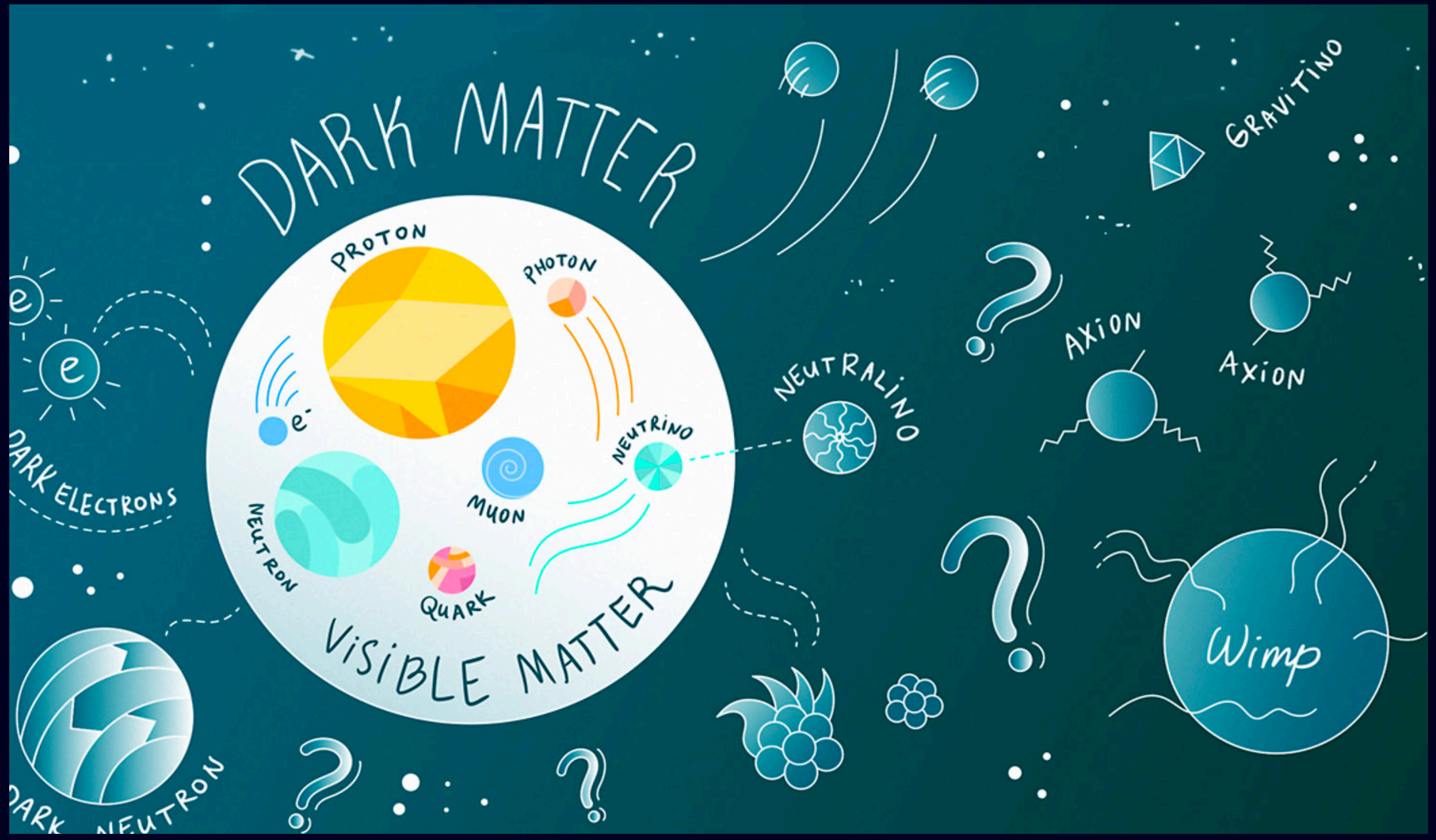
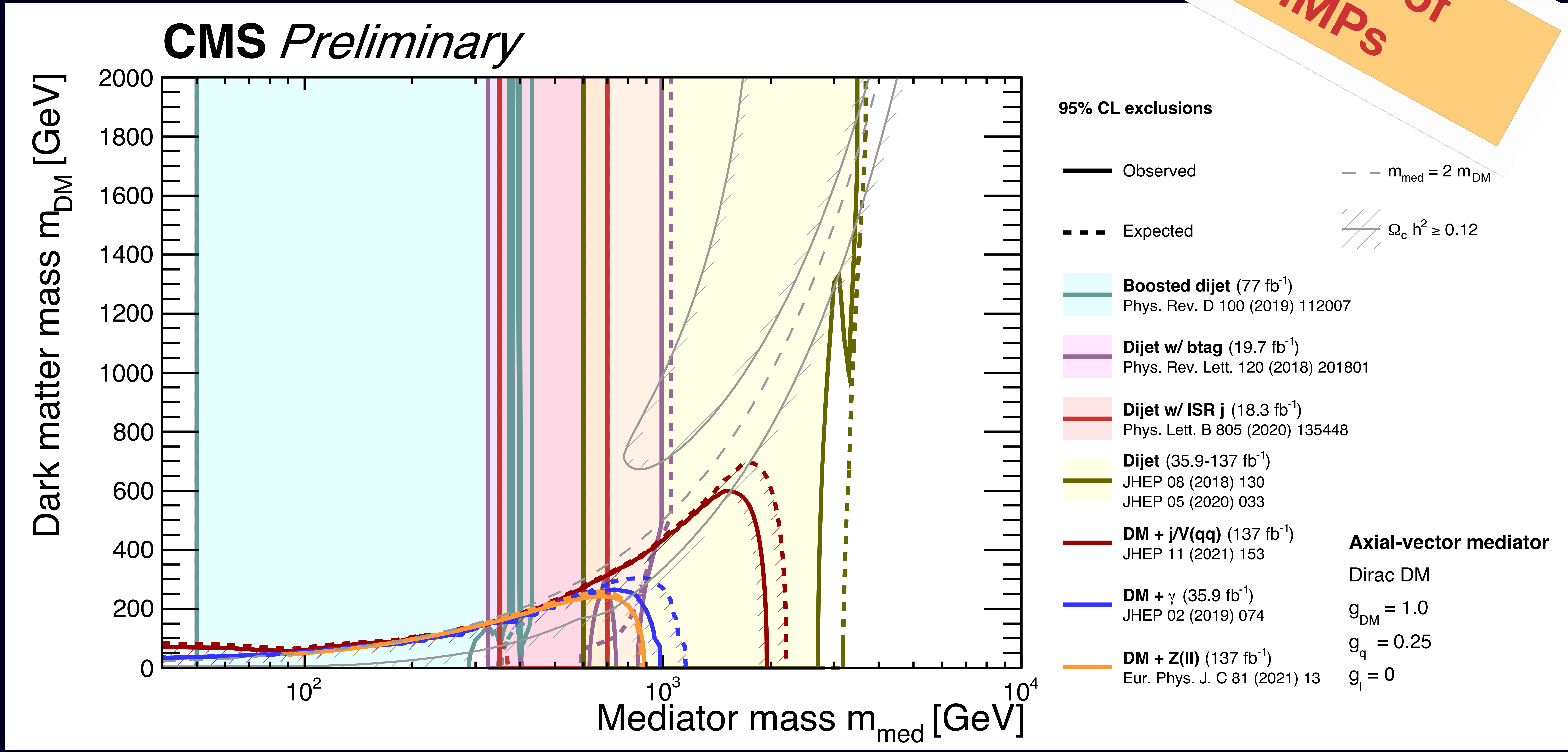


Illustration by Sandbox Studio, Chicago with Ana Kova

DM Quest - Where do we stand



No sign of WIMPs



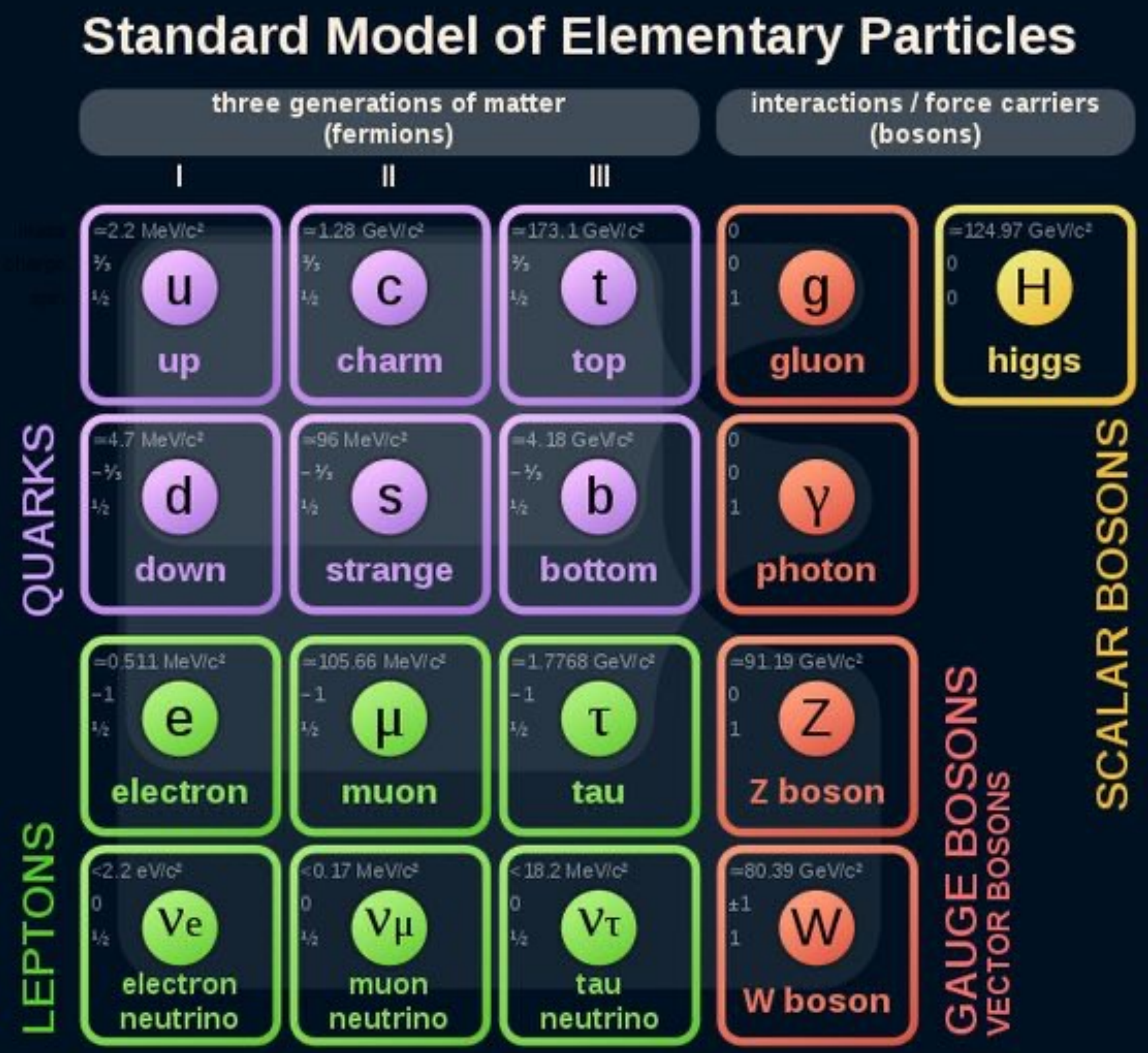
WIMPs widely excluded. Need for:

New theoretical models

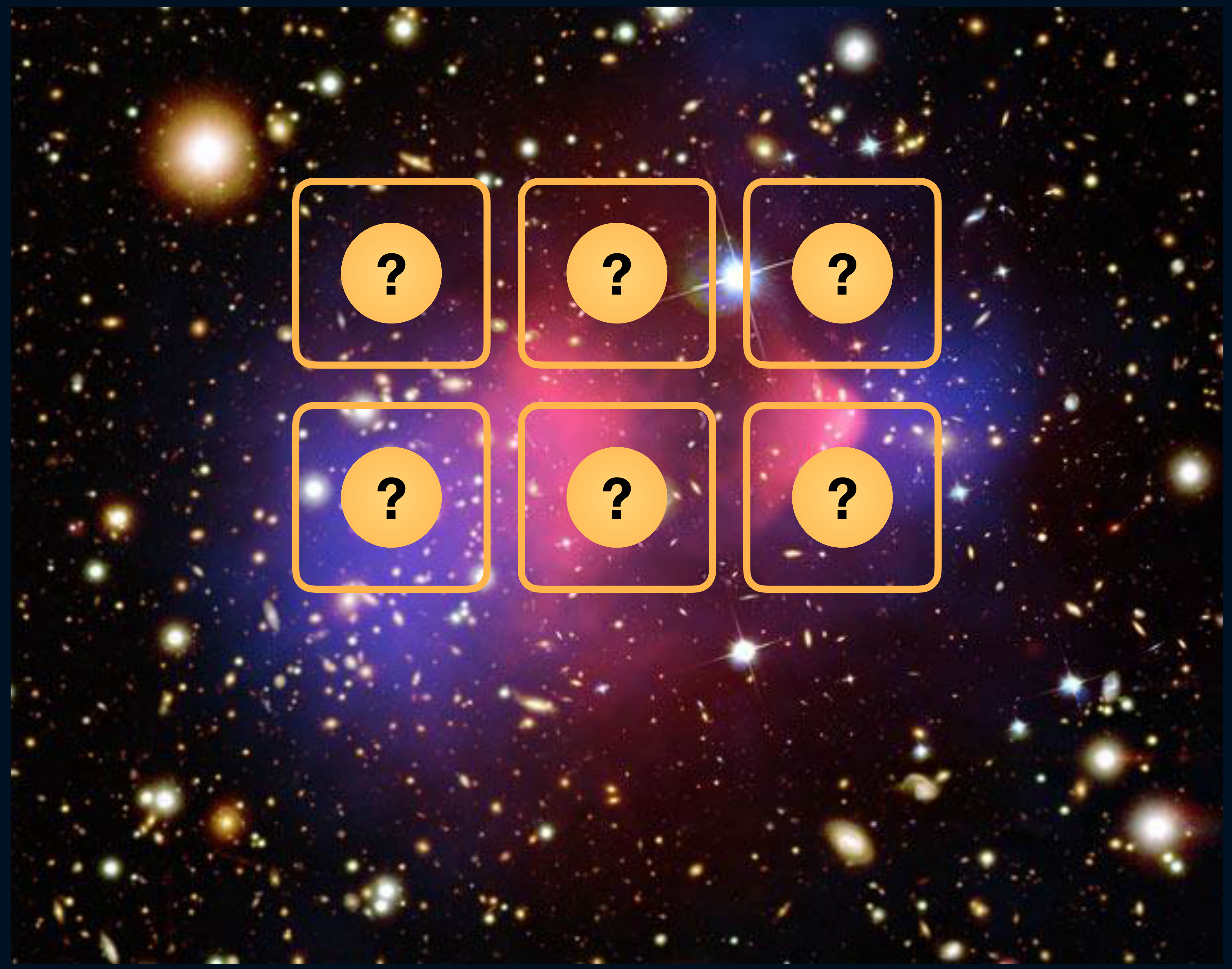
New experimental techniques

Novel theoretical models

What if DM is hiding in a Hidden Valley?

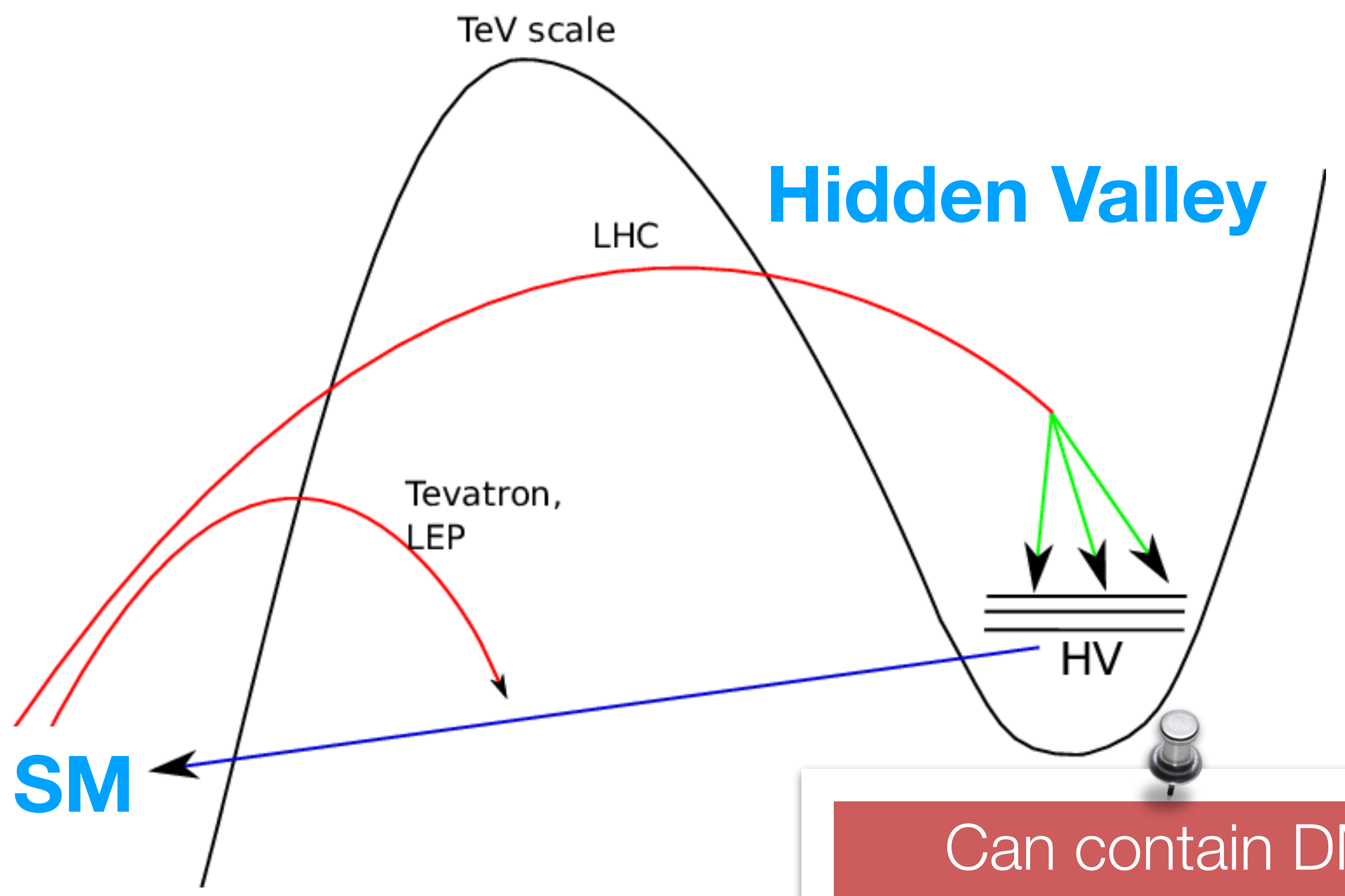


Hidden Valley



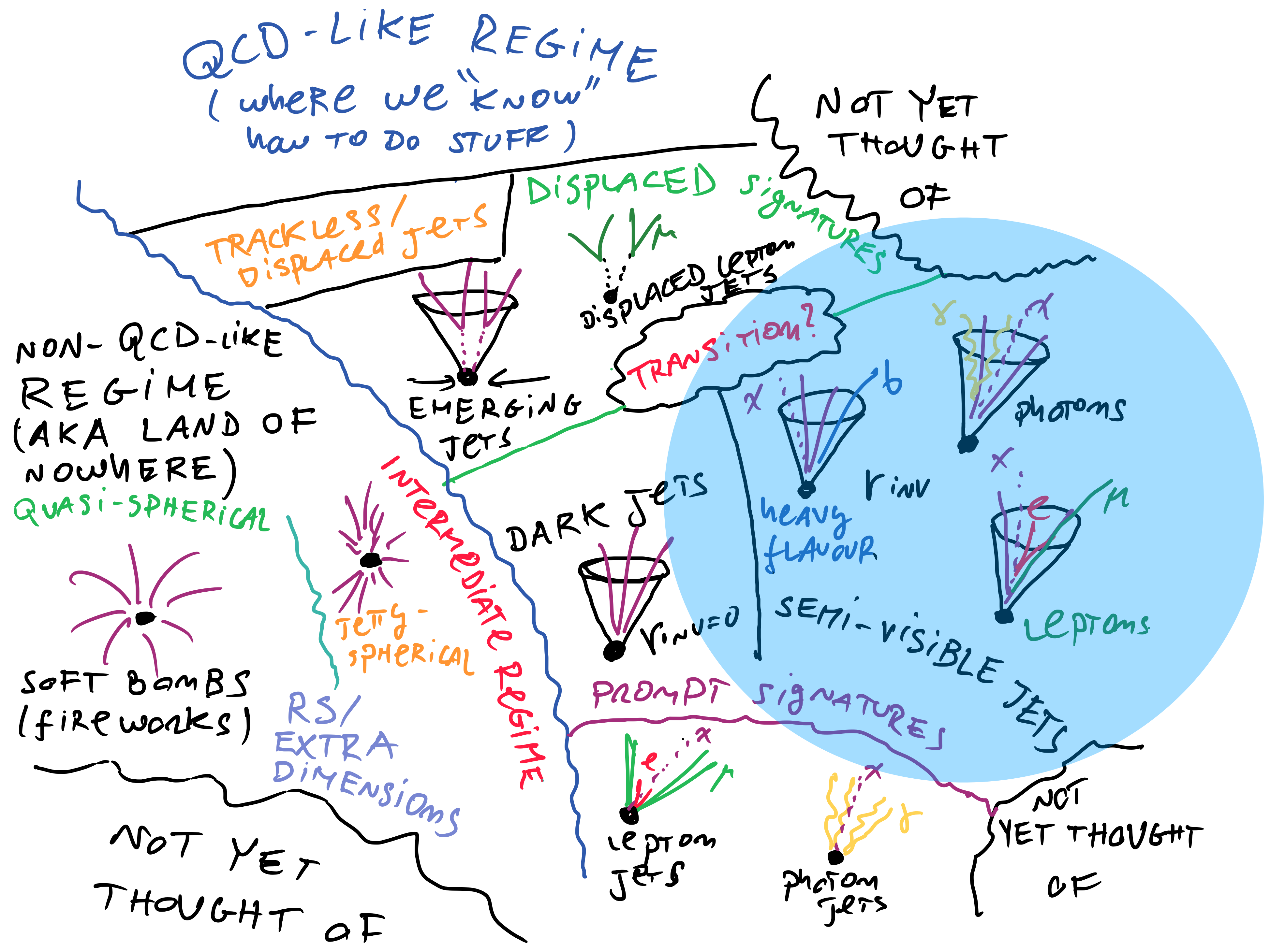
Looking for DM in a Hidden Valley

Hidden Valleys:
Class of theories predicting an Hidden Sector secluded from the SM



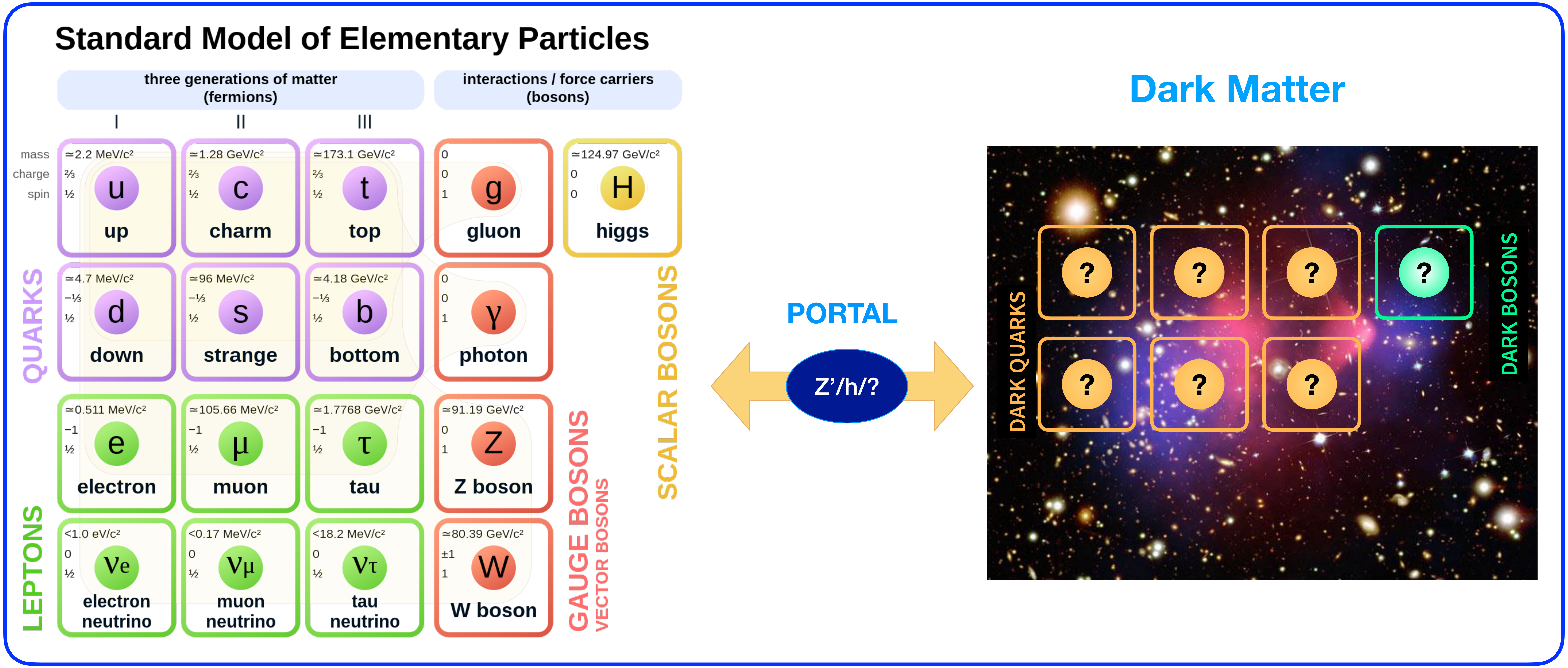
Can contain DM candidates!

- Hidden Valleys very weakly coupled to the SM
- What makes the Hidden Sector?
  - Pletora of possibilities!
  - Here focus on **Strongly coupled Hidden Sectors**



Courtesy of C.Cazzaniga

Strongly coupled Hidden Sectors

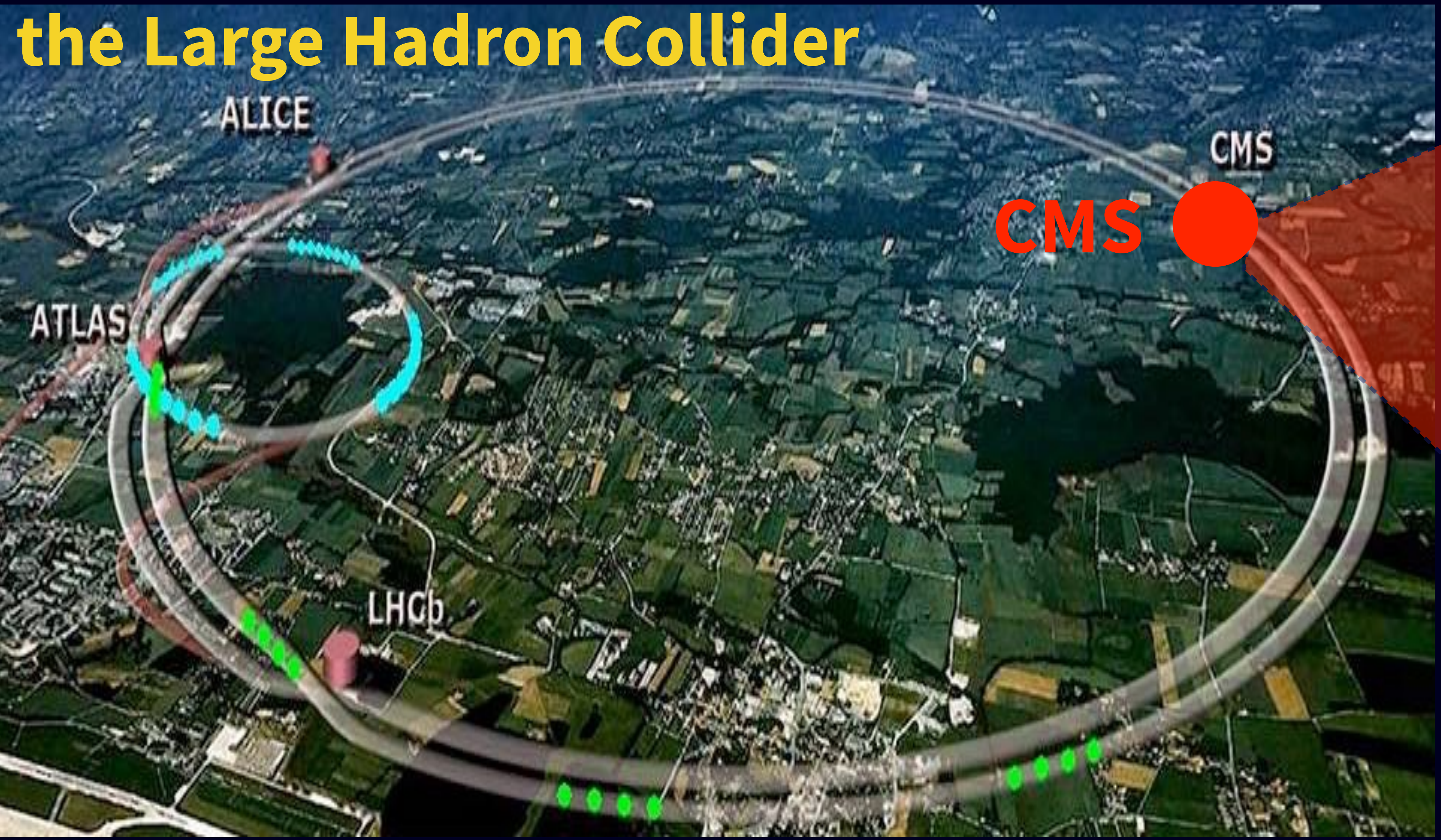


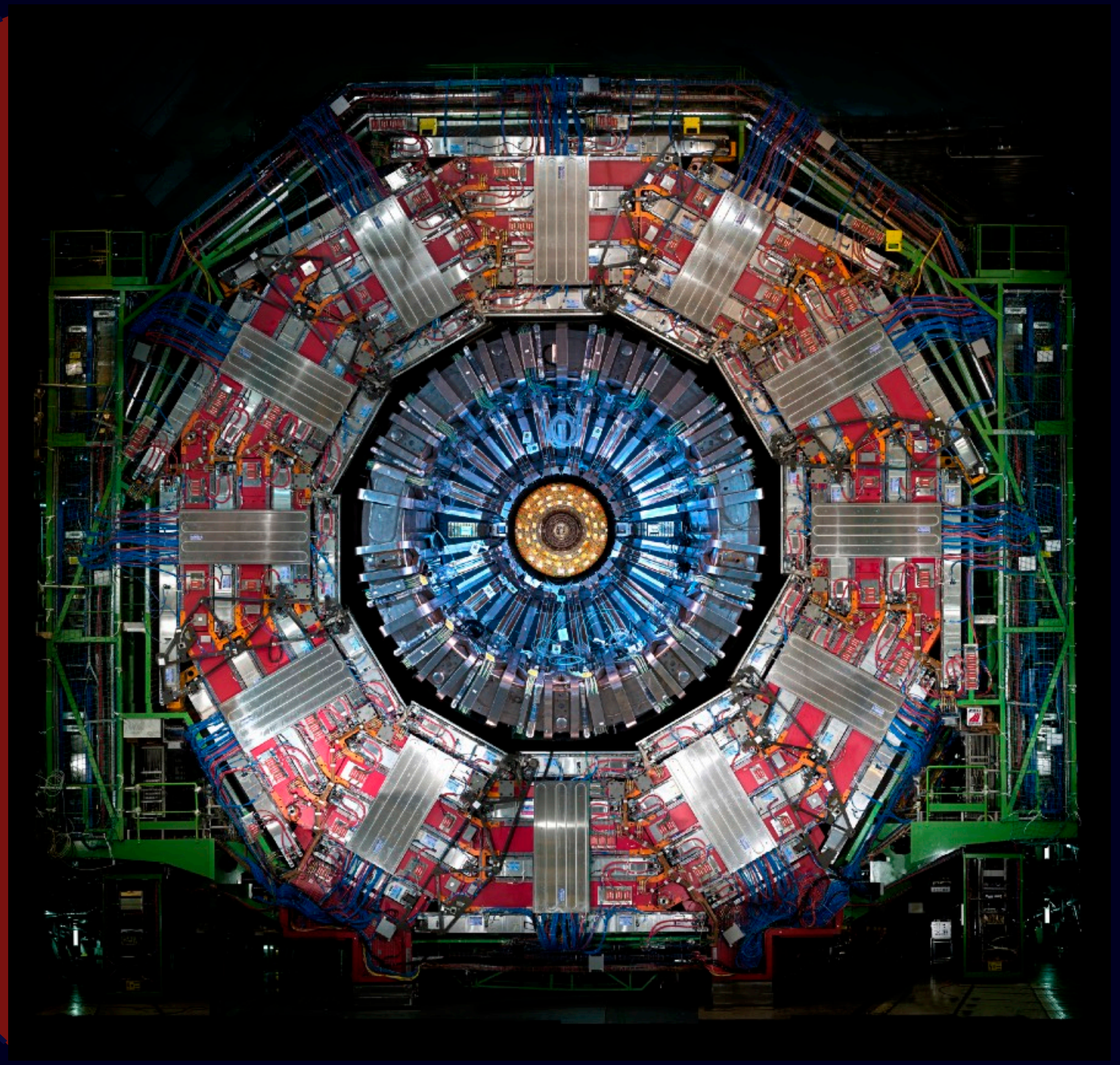
The quest for DM at particle colliders

The Large Hadron Collider

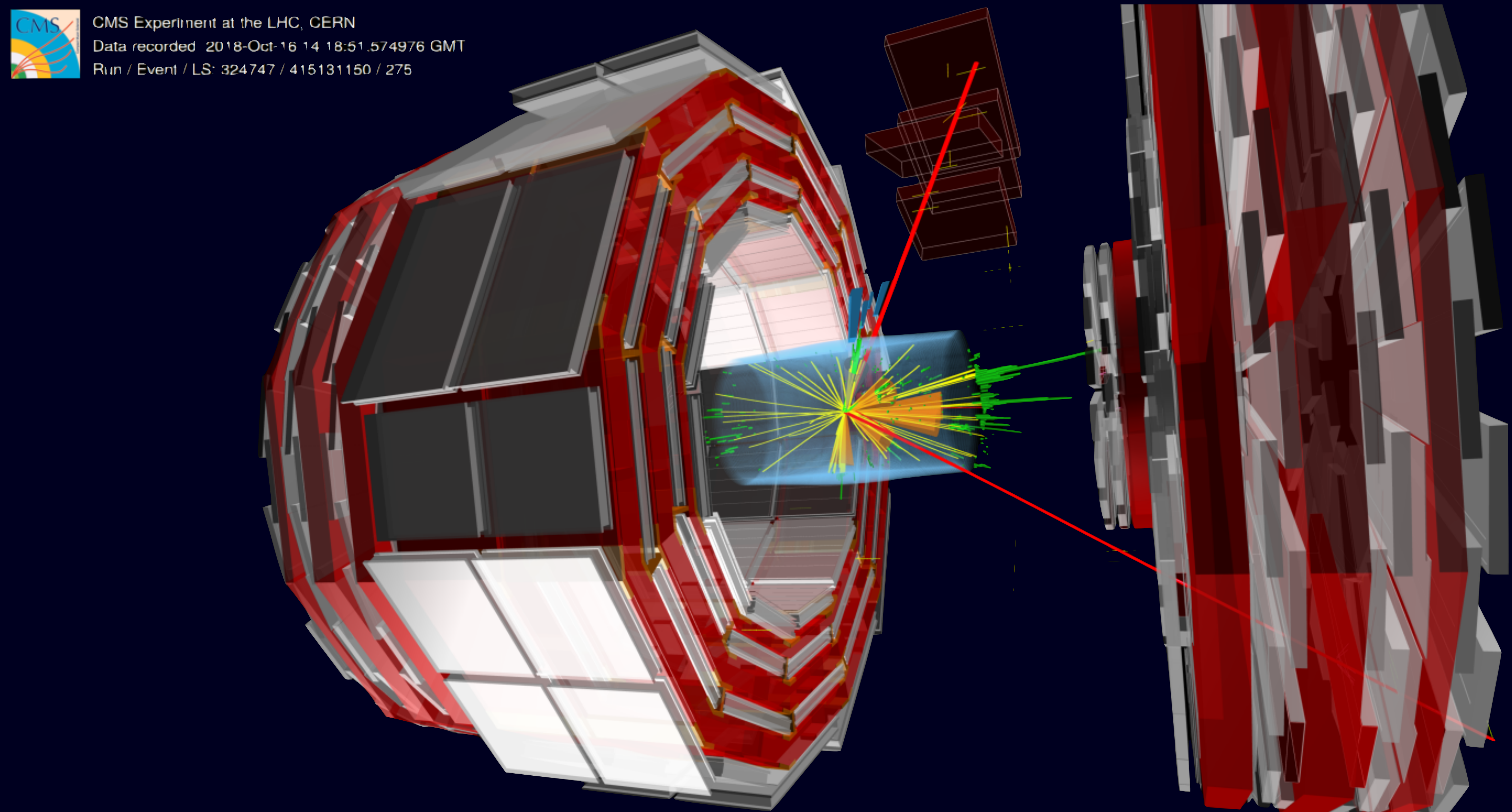
- **27 Km** circumference
- Proton-proton collisions at a centre of mass energy of **14 TeV**

The quest for DM at particle colliders





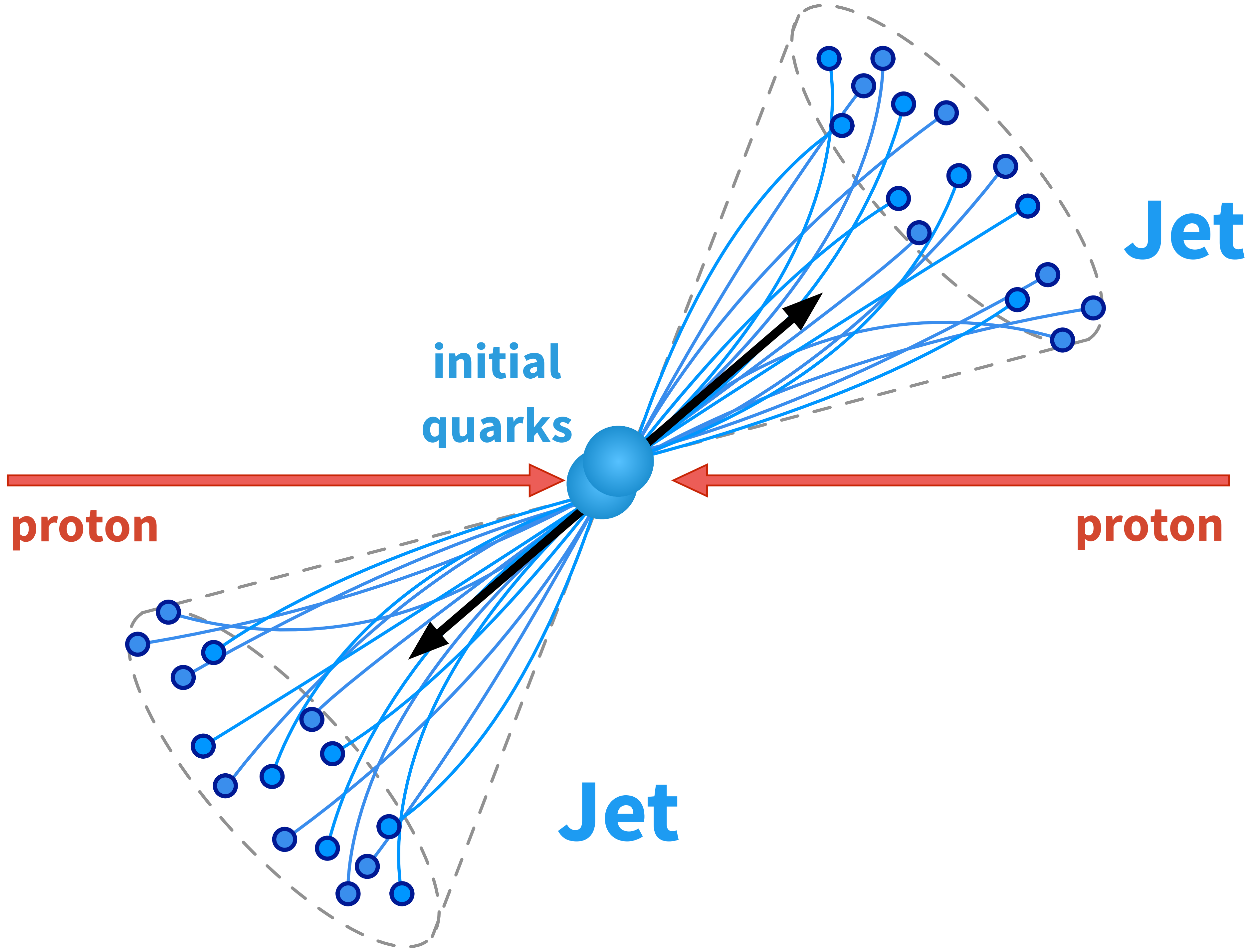
Compact Muon Solenoid (CMS) experiment

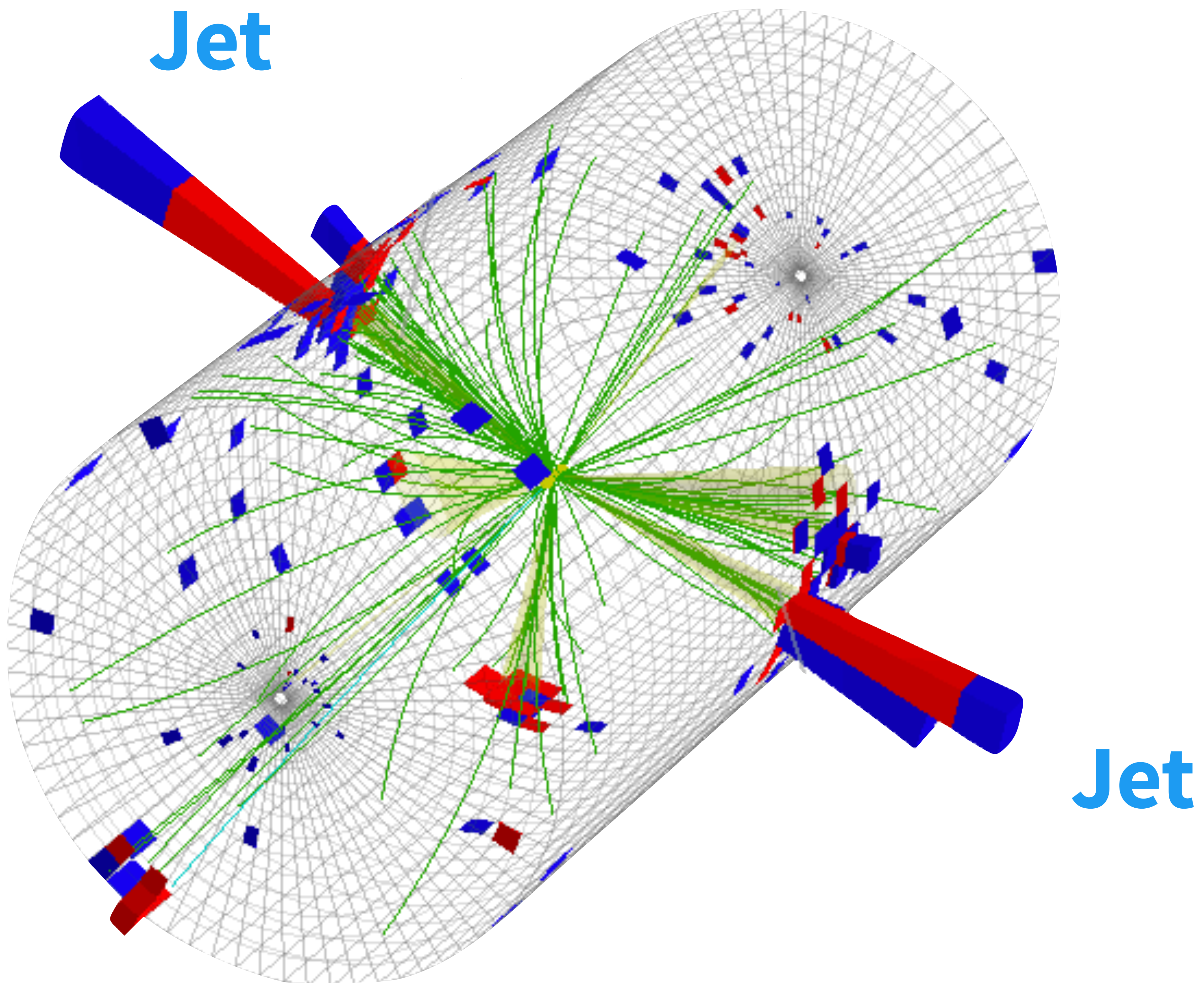


The Large Hadron Collider

- **27 Km** circumference
- Proton-proton collisions at a centre of mass energy of **14 TeV**

How can we see quarks?





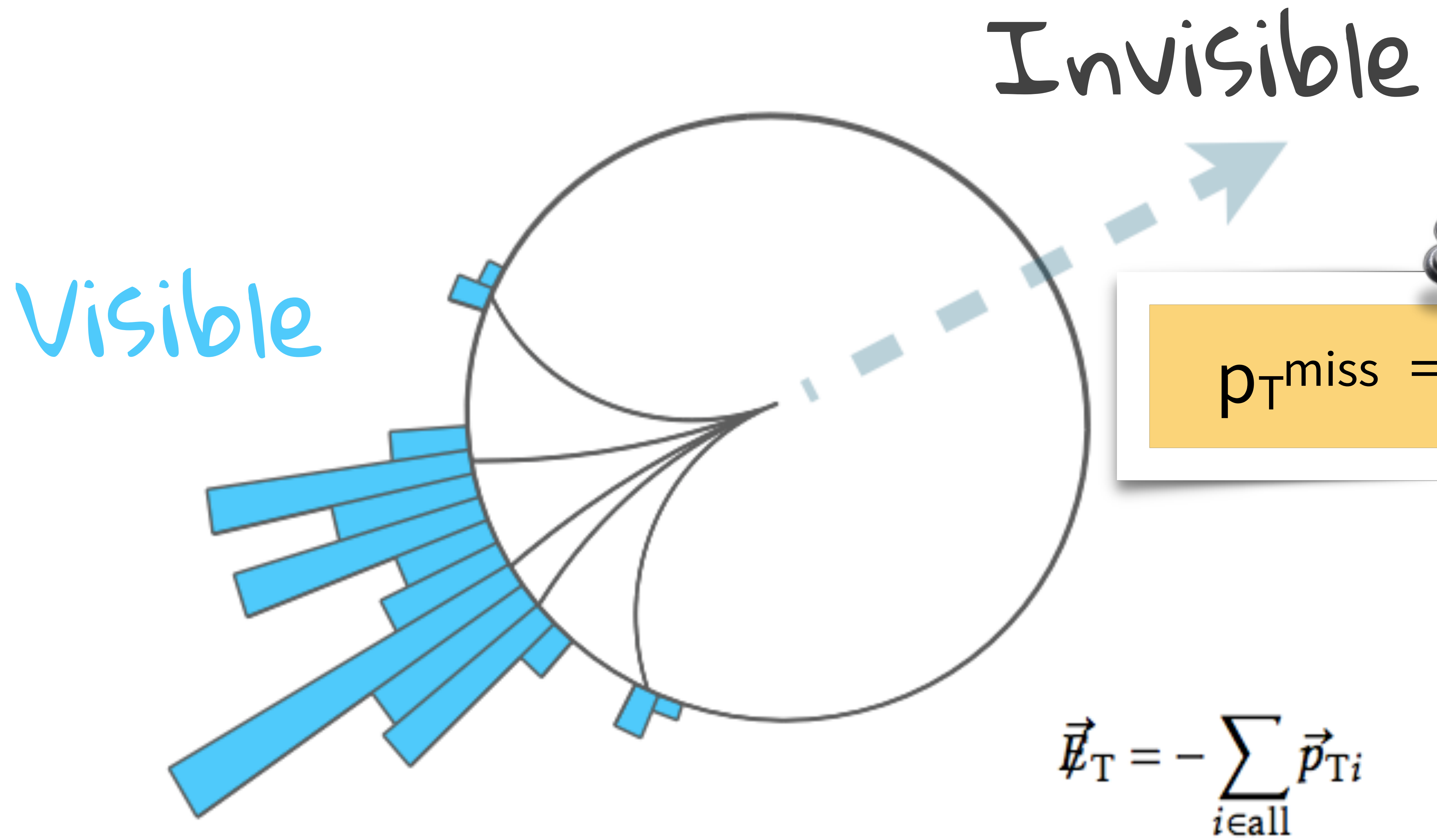
Quarks confined into hadrons by strong interaction:

- cannot be observed directly
- give rise to a ***jet of collimated hadrons***

How can we see invisible particles?

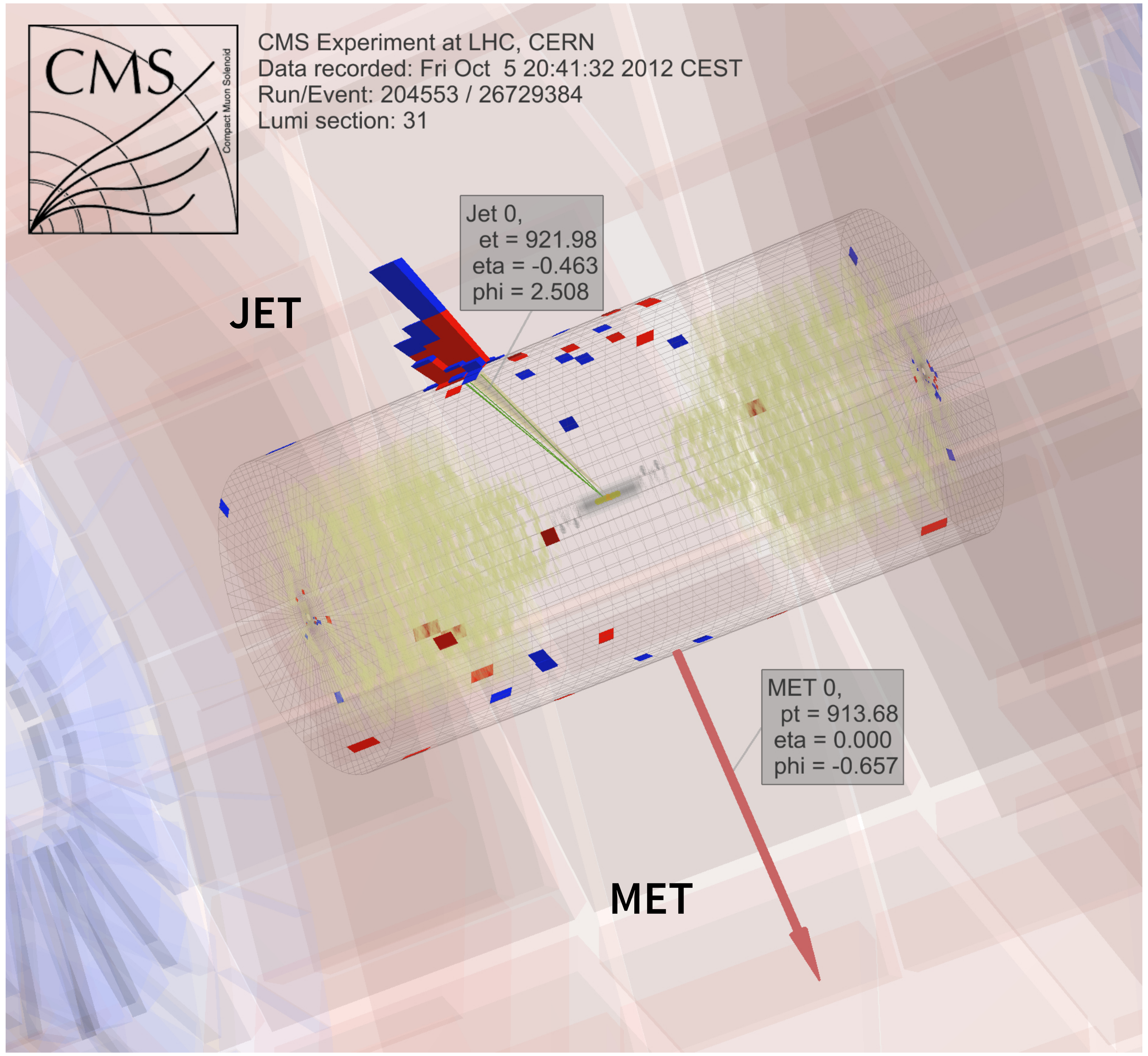
Invisible particles = not reconstructable

Momentum carried out by invisible particles
quantified by an **energy imbalance (MET)** in the event collision products



$$p_T^{miss} = -\sum_i p_T^i$$

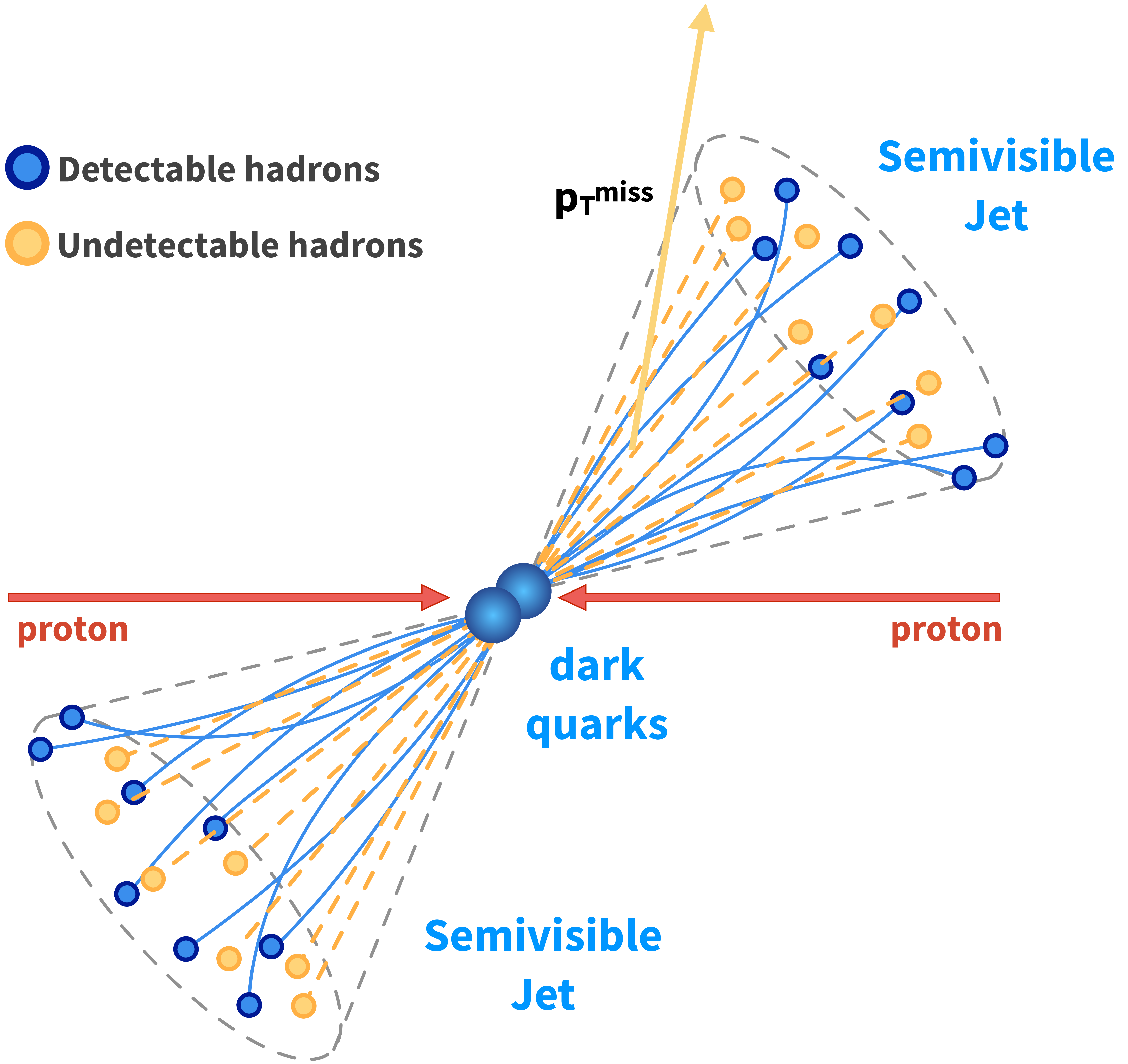
$$\vec{E\!\!\!/}_T = -\sum_{i \in all} \vec{p}_{Ti}$$



Semivisible Jets

Semivisible Jets (SVJs) are an exotic signature arising in a class of strongly coupled Hidden Sectors

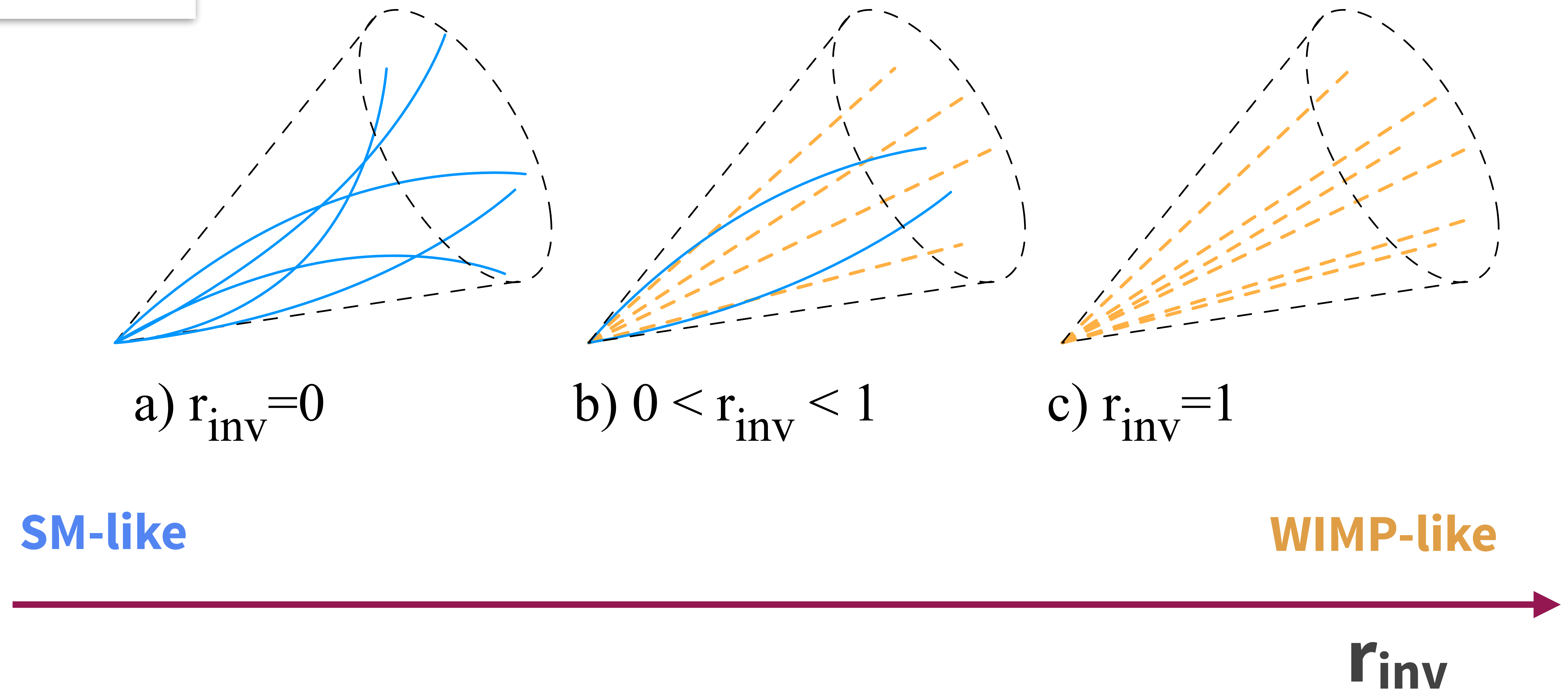
- **Dark quarks hadronise** in the dark sector (**dark QCD**)
  - **Unstable dark mesons** decay to visible particles
  - **Stable dark mesons** remain invisible
- **Challenging experimental signature**:
  - Partially visible jets of particles (***Semivisible jets***)
  - Missing momentum collinear to visible jets



How much invisible?

The jet topology spans from **fully visible to fully invisible** based on the **ratio between stable and unstable dark bound states**

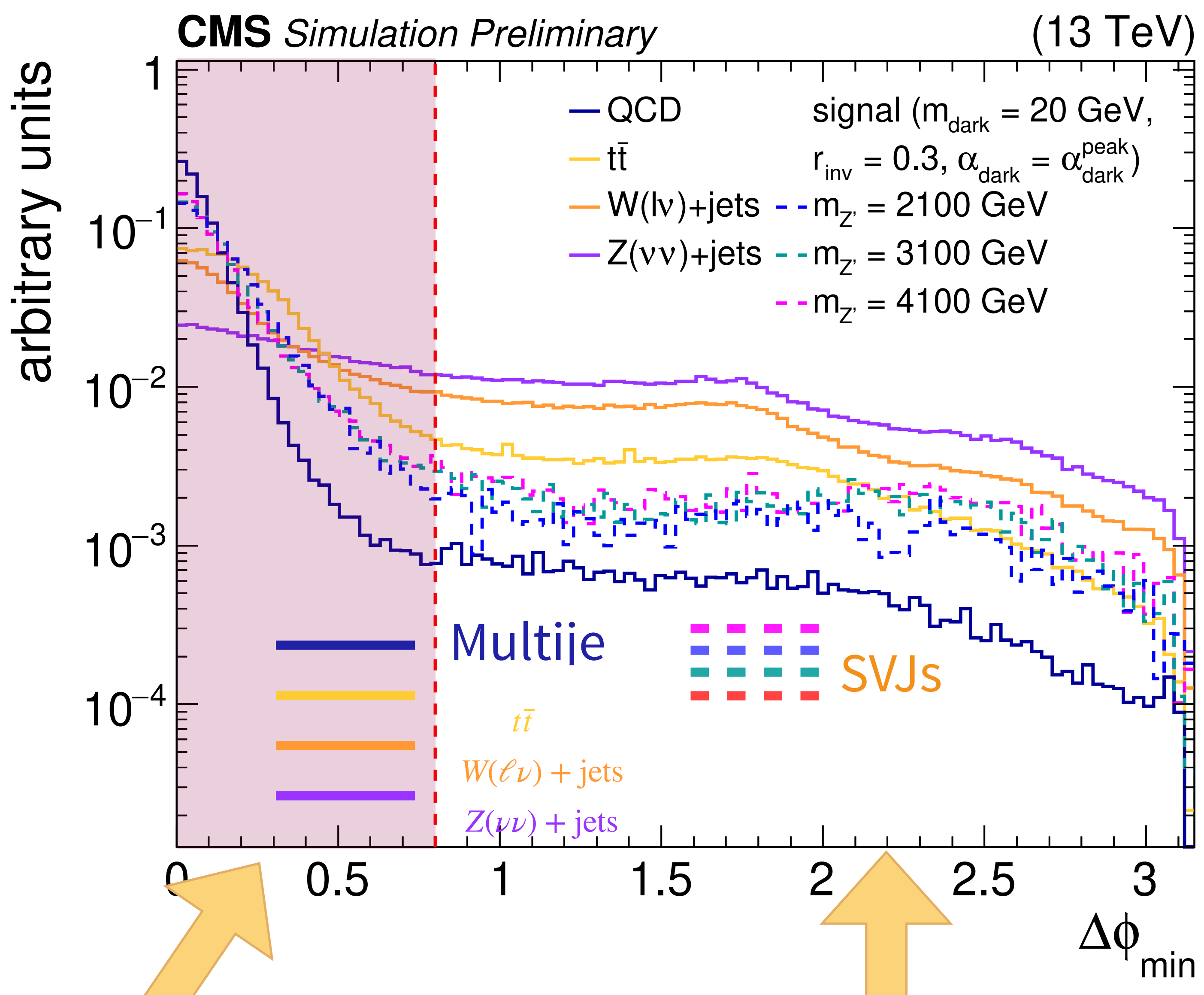
$$r_{inv} = \left\langle \frac{\textit{N stable dark hadrons}}{\textit{N dark hadrons}} \right\rangle$$

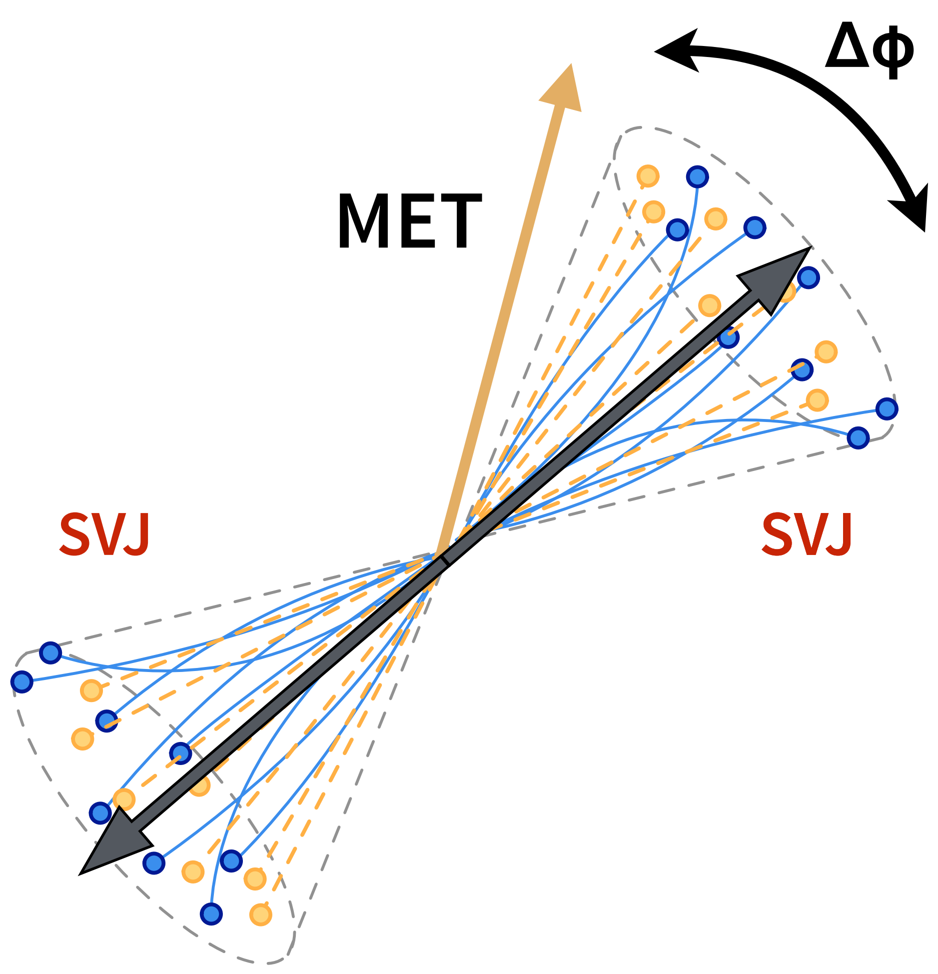


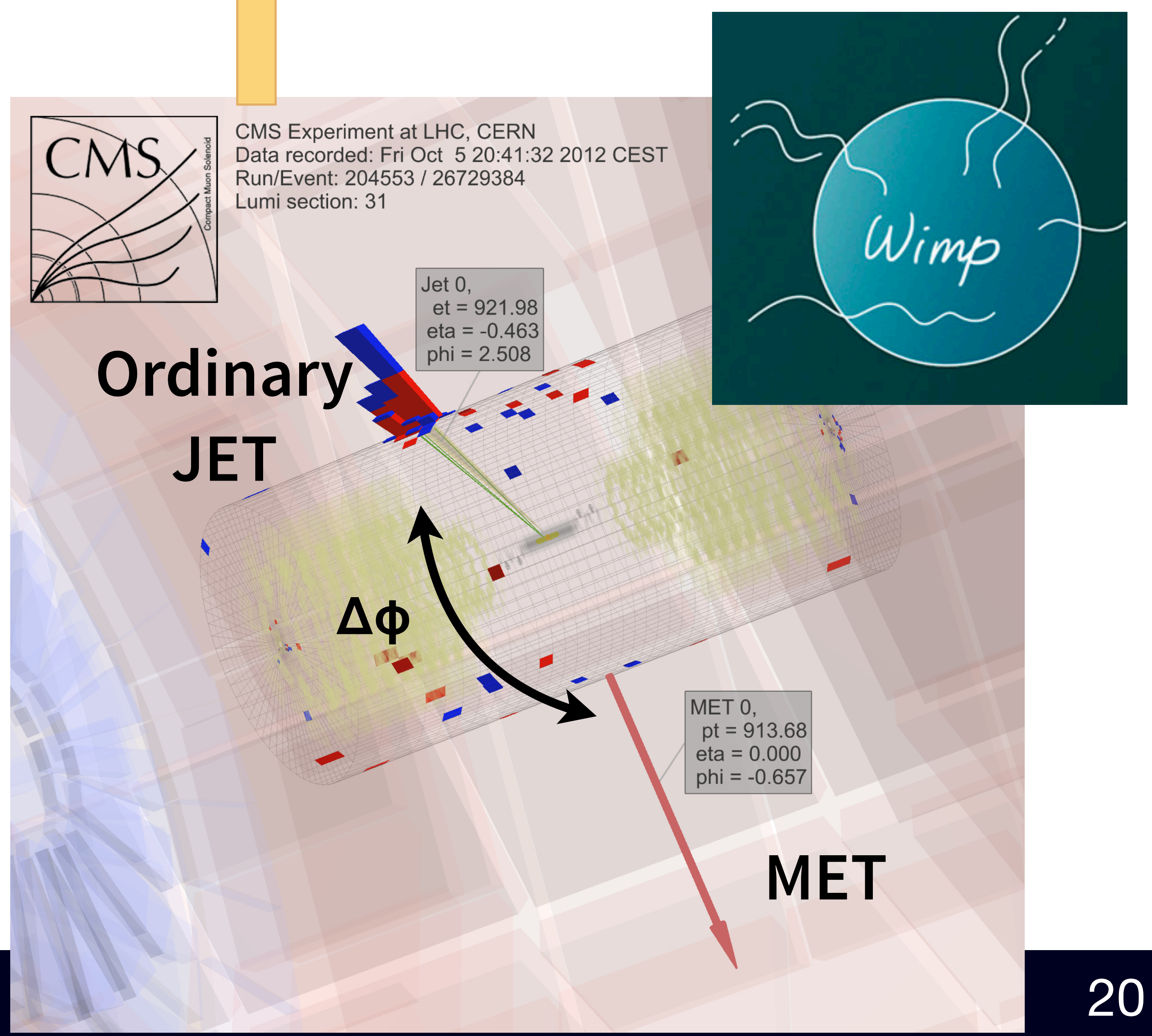
Why it's so interesting?

Unexplored territory:

- SVJs data data lay in a region **discarded by common searches**
- Overwhelming **instrumental background** (Multijet)
- Complex event topology



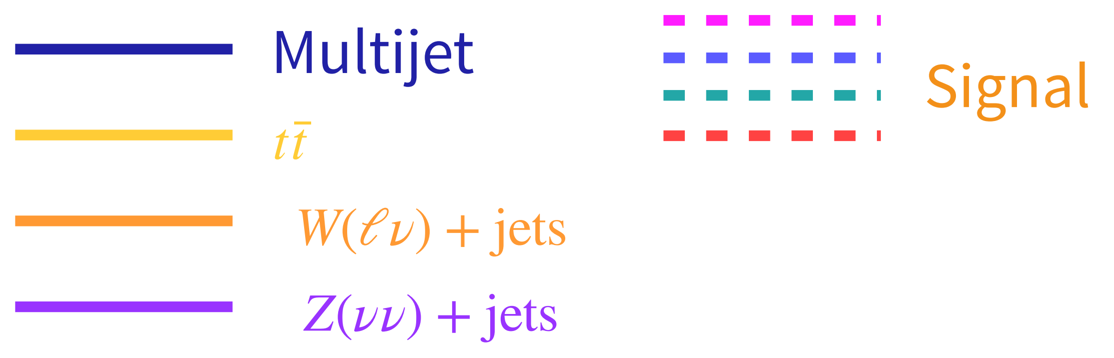


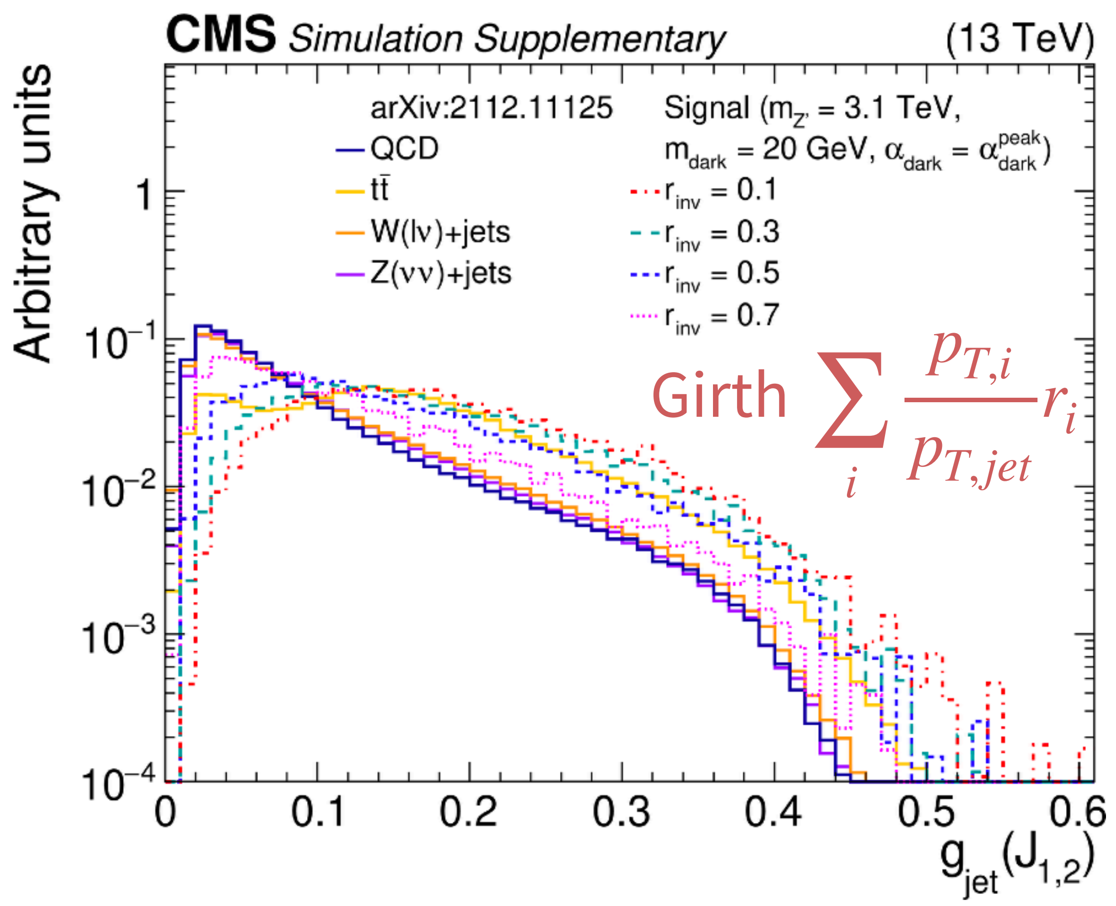


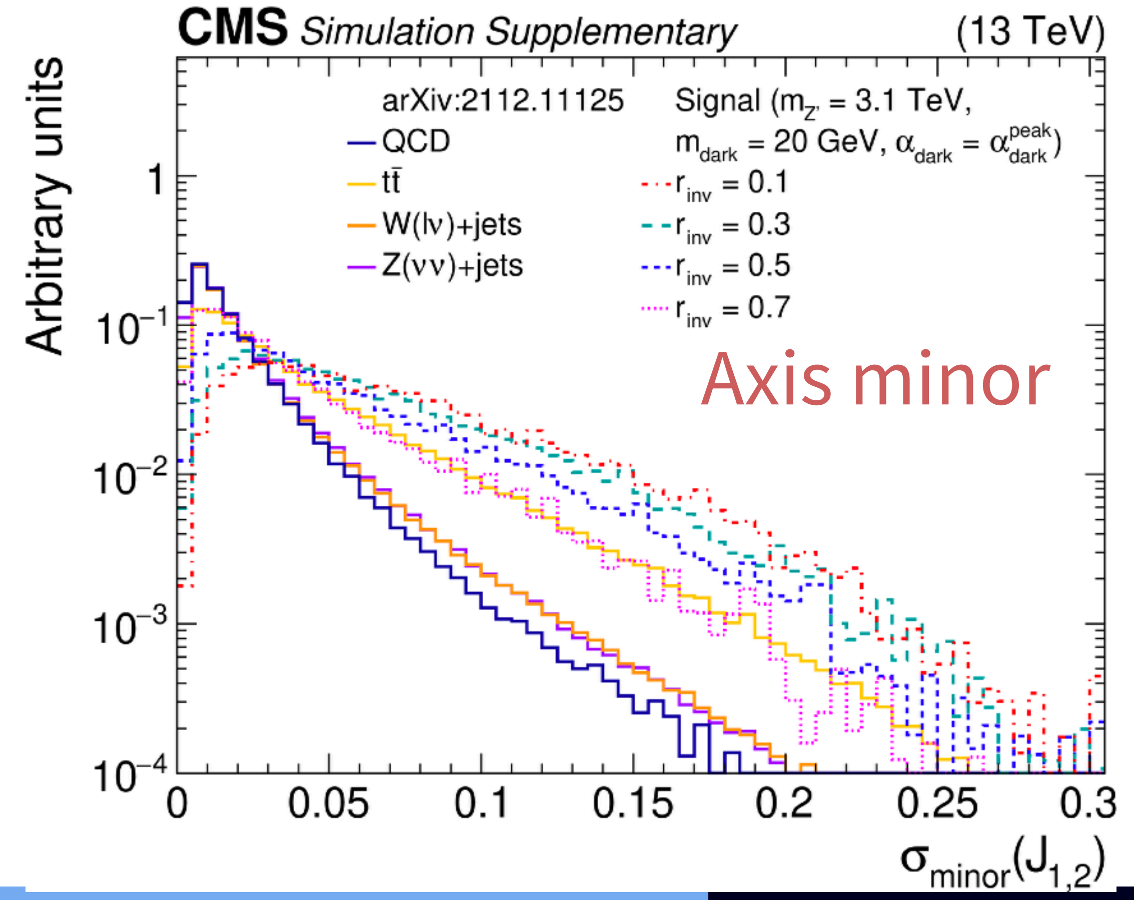
Diving into the jet

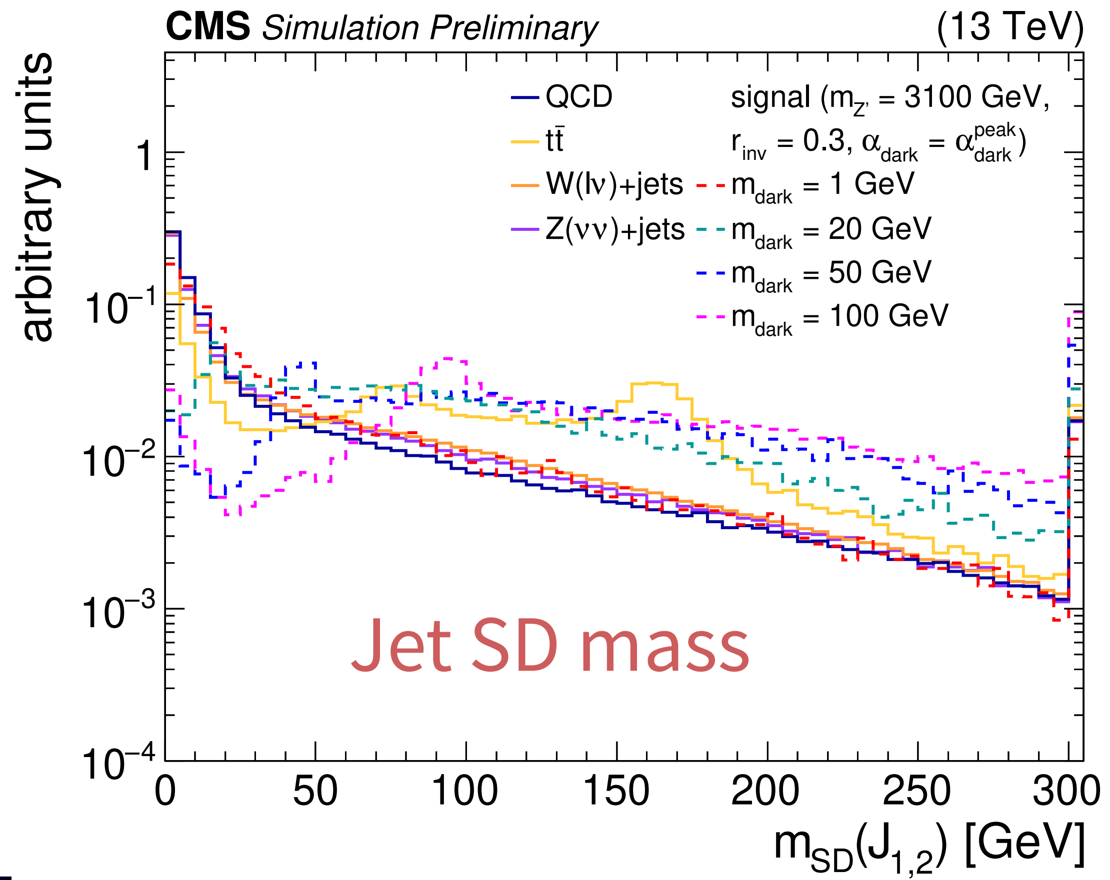
Look inside the jet

- **Underlying interaction** affects the hadronization process
- **Dedicated algorithms** enable investigation of dark jets

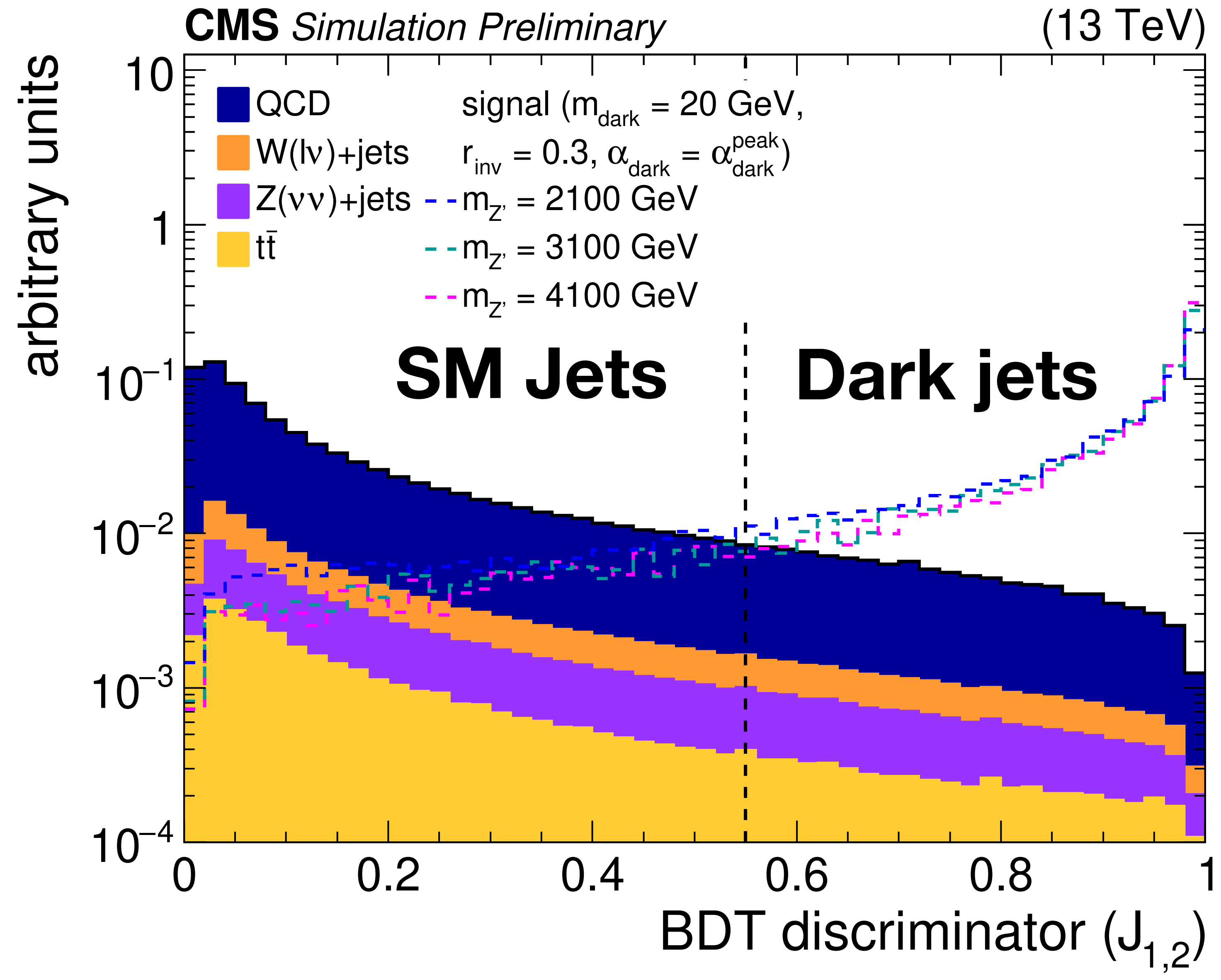








Semivisible jets identification algorithm



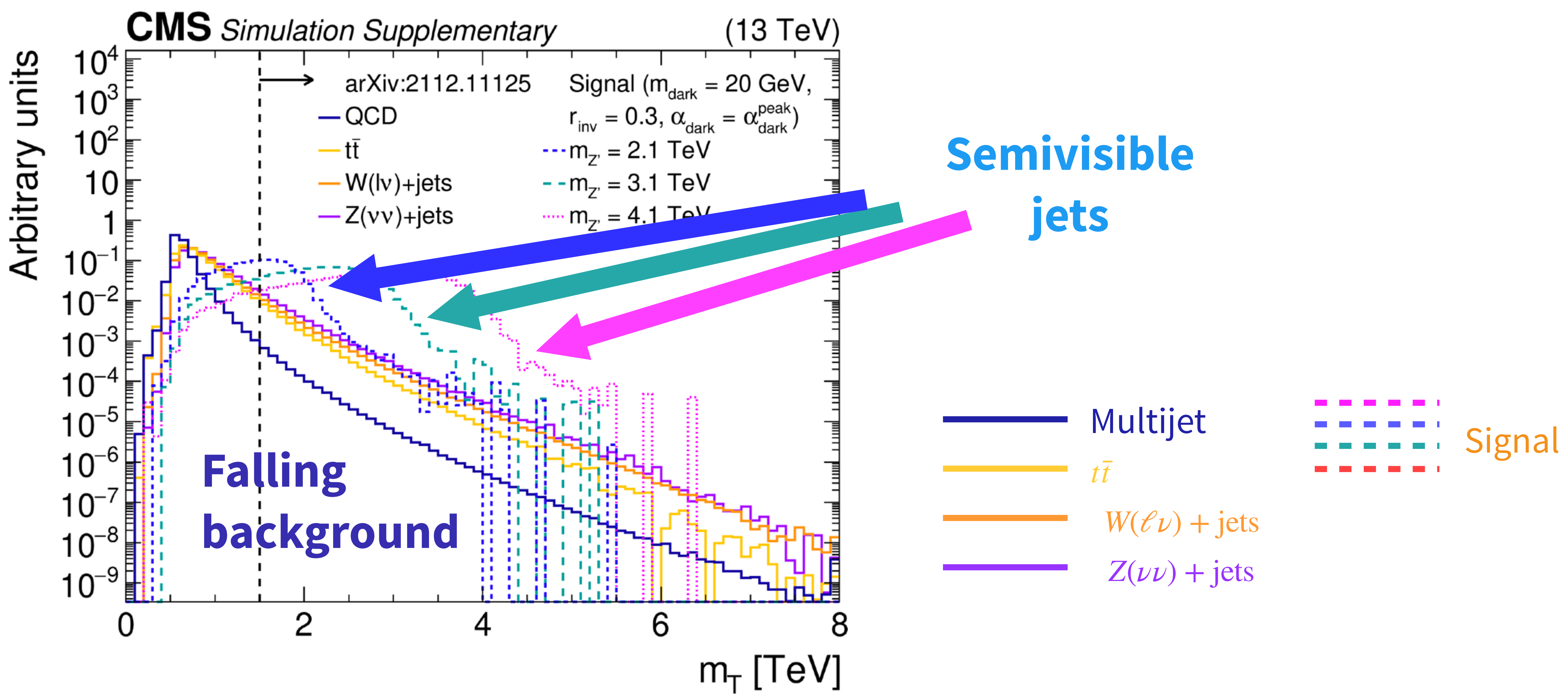
Jet identification algorithm opened the way to SVJs investigation

Reconstruct the mediator

If dark quarks are produced in a resonant model, we can:

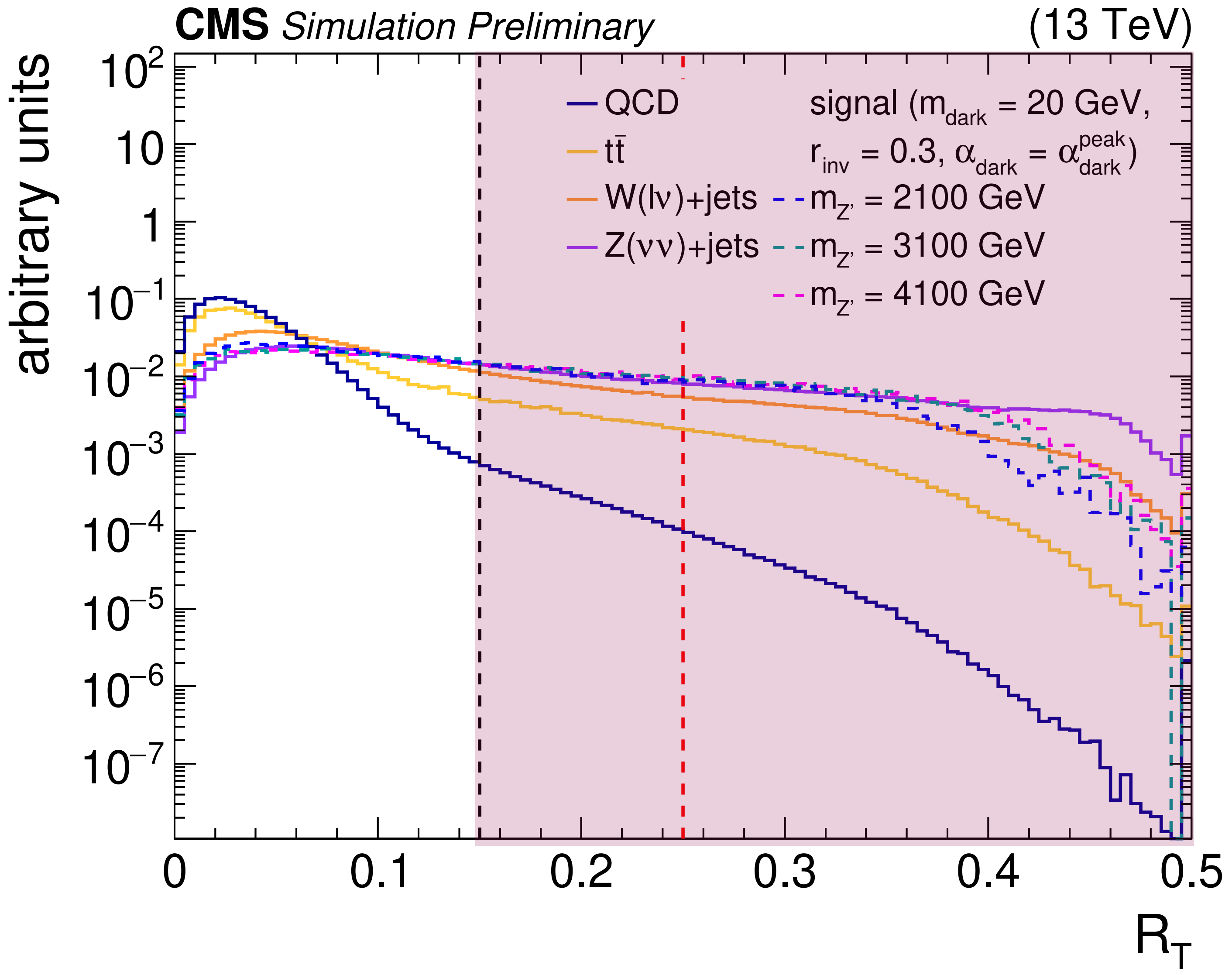
1. (approximately) **reconstruct the mediator mass (M_T)** from visible jets and MET

$$M_T^2 = M_{jj}^2 + 2\left(\sqrt{M_{jj}^2 + p_{T,jj}^2}\, p_T^{miss} - \vec{p}_{T,jj}\cdot\vec{p}_T^{\,miss}\right)$$



2. And exploit **MT vs MET balance**!

R_T = MET/M_T



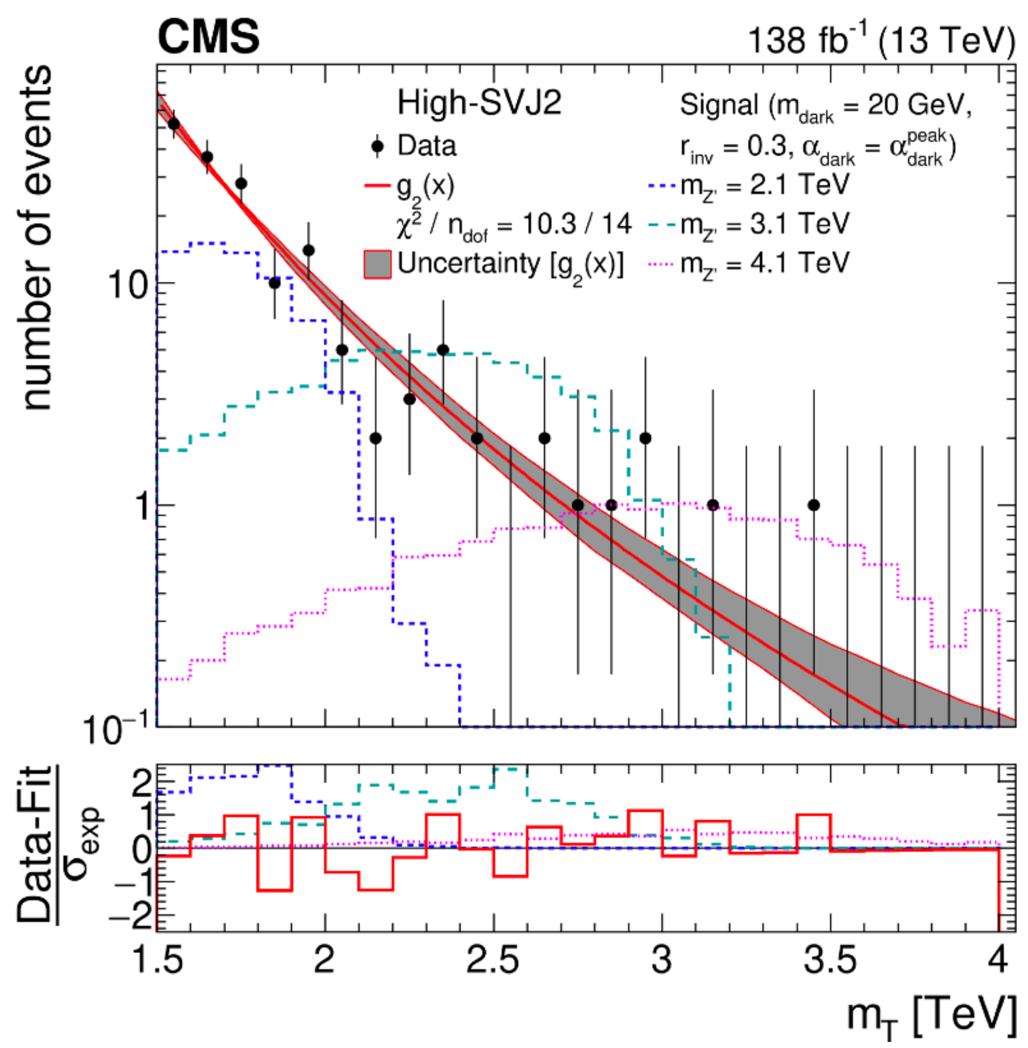
Wide phase space coverage

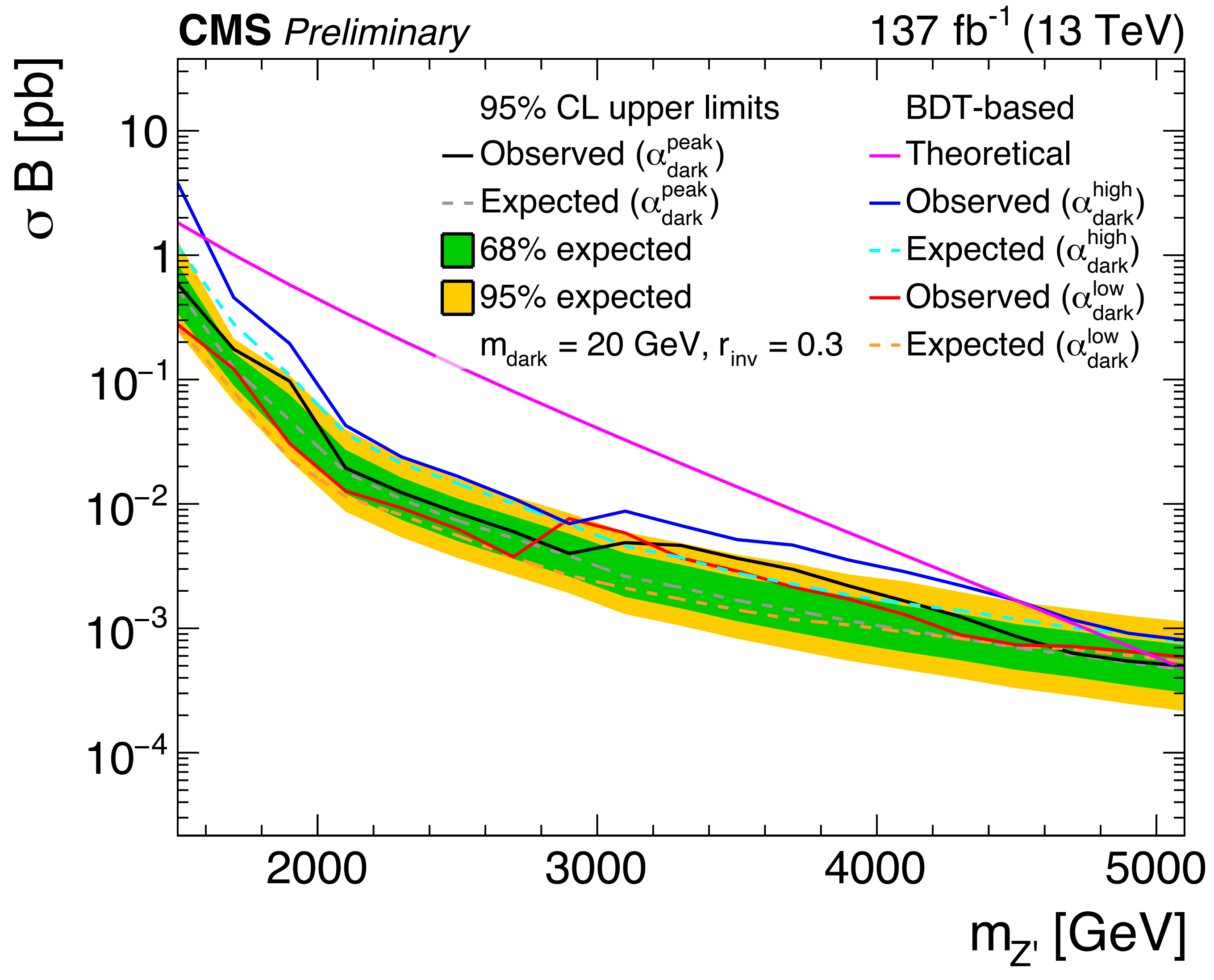
Jet Identifier proved to be key to probe Semivisible Jets

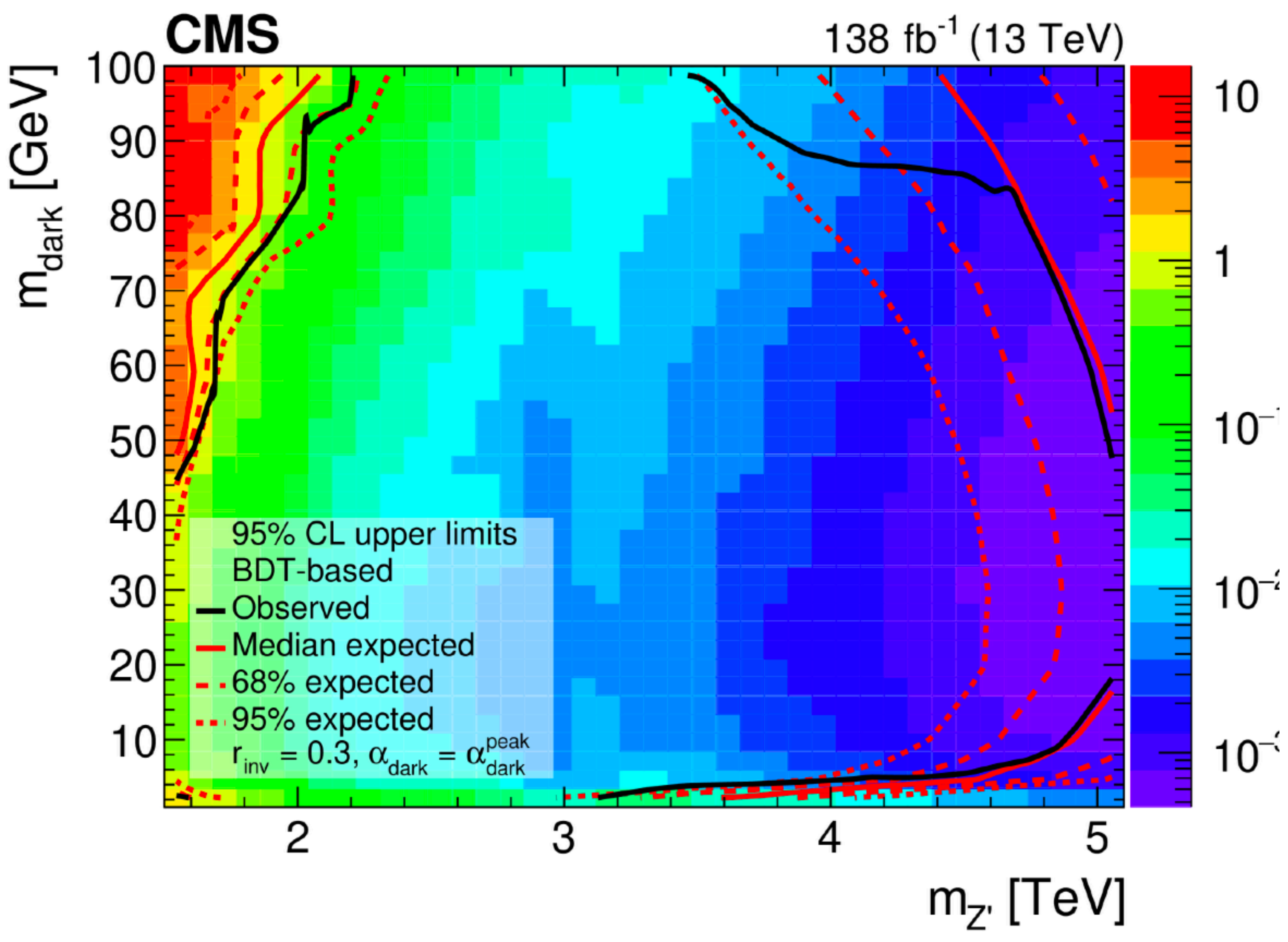
- Improvement in sensitivity by almost a factor 10

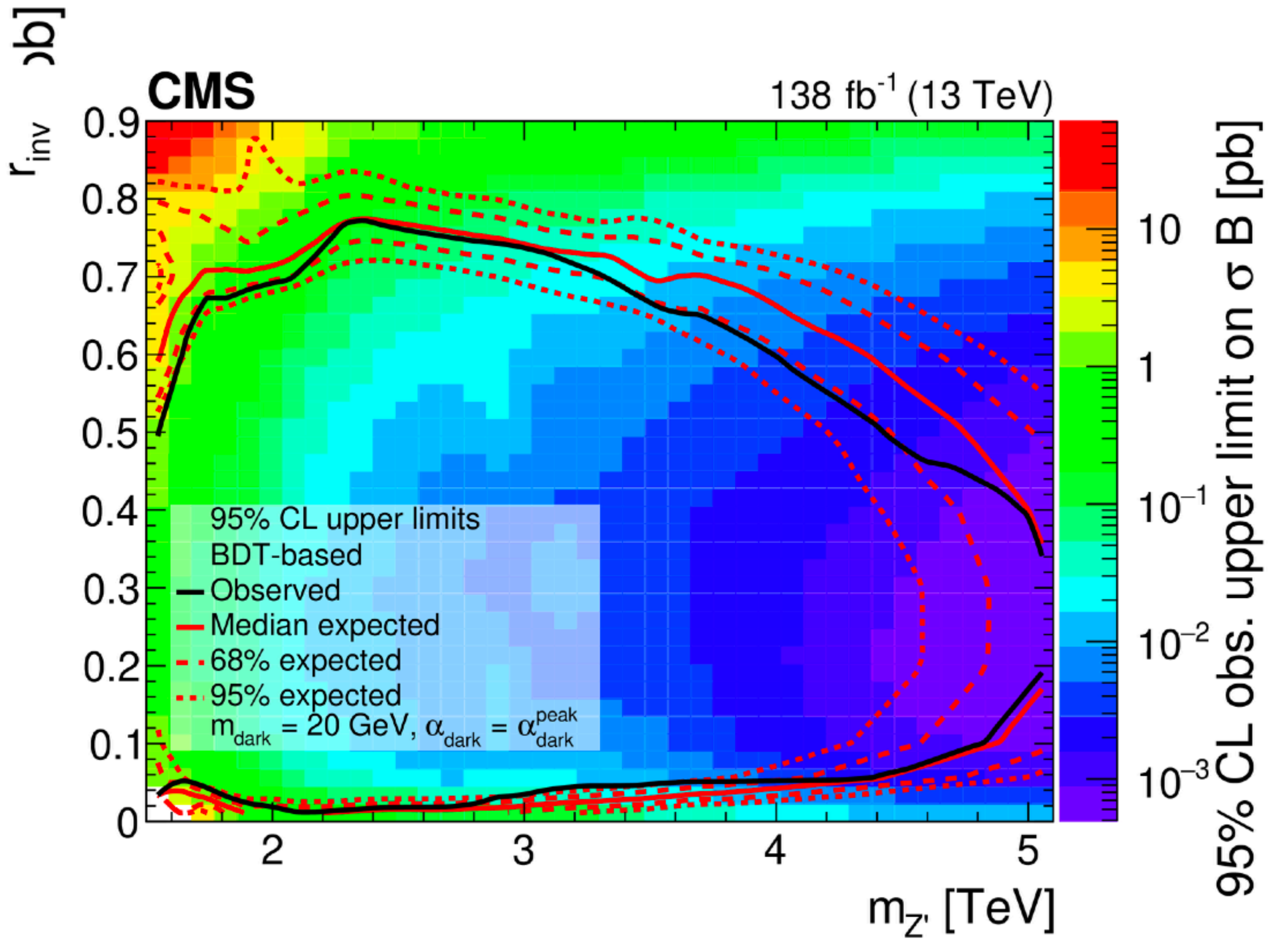
Observed exclusions
(maximum range)

- $1.50 < m_{Z'} < 5.10$ TeV
- $0.01 < r_{inv} < 0.77$



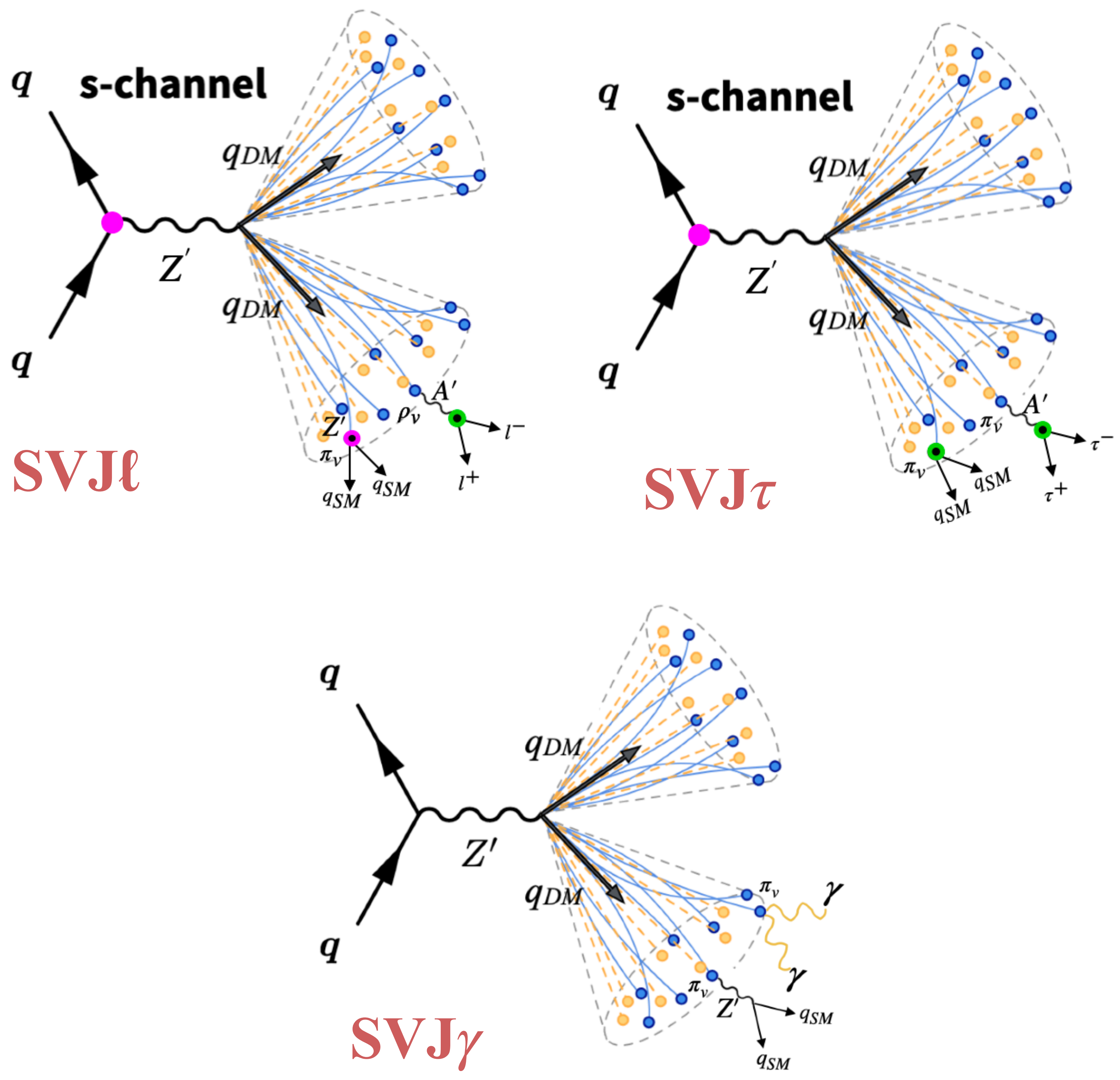


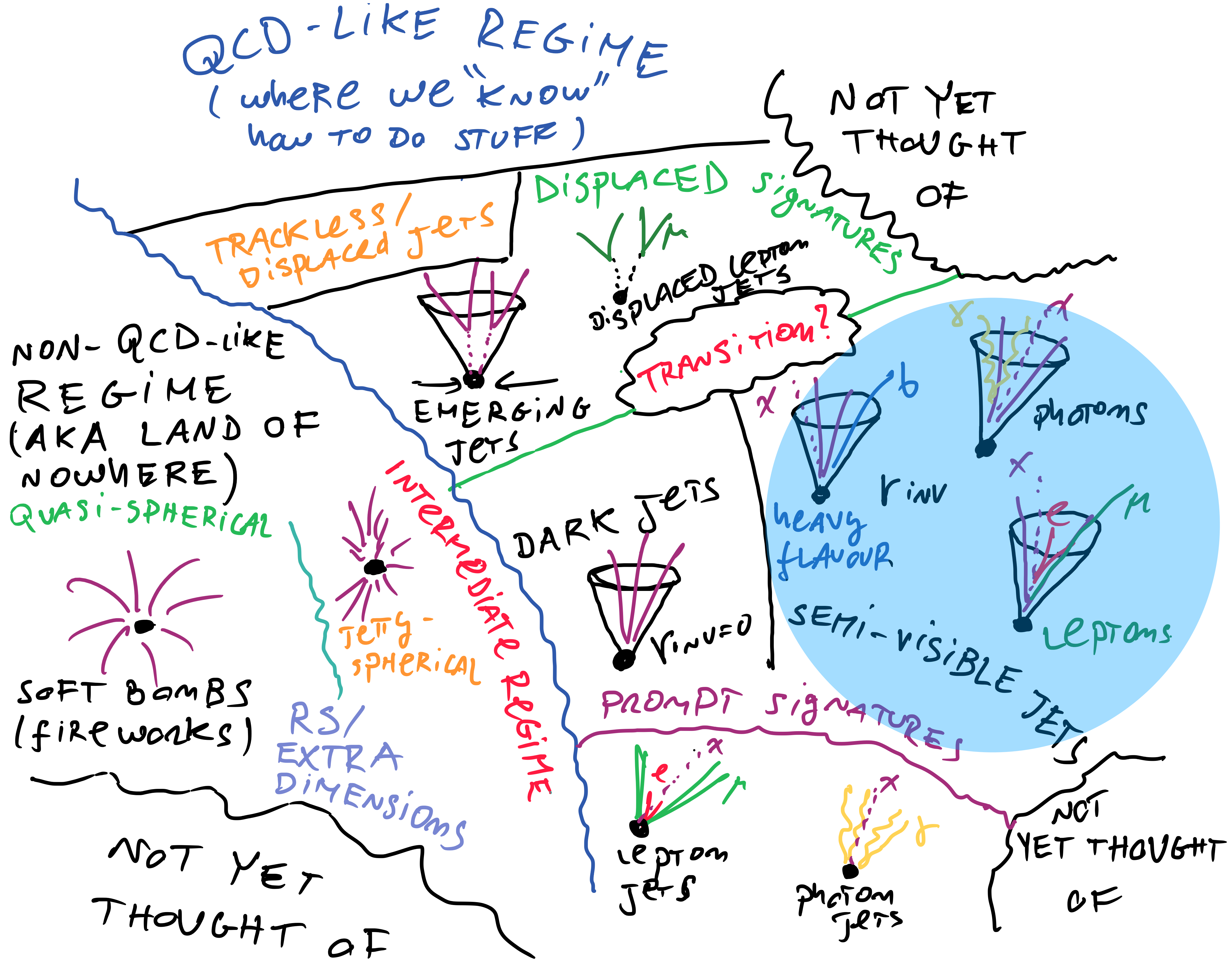




Expanding the territory …

… to novel, uncovered signatures with leptons or photons within the SVJs





Lepton-enriched SVJs

Eur. Phys. J. C 82, 793 (2022) - C.Cazzaniga, A.de Cosa
Eur. Phys. J. C 83, 599 (2023) - H. Beauchesne, C. Cazzaniga, A. de Cosa, C. Doglioni, T. Fitschen, G. Grilli di Cortona, Z. Zhou

Original SVJs models do not allow decays to leptons (leptophbic Z' portal)

New models including decays to leptons:

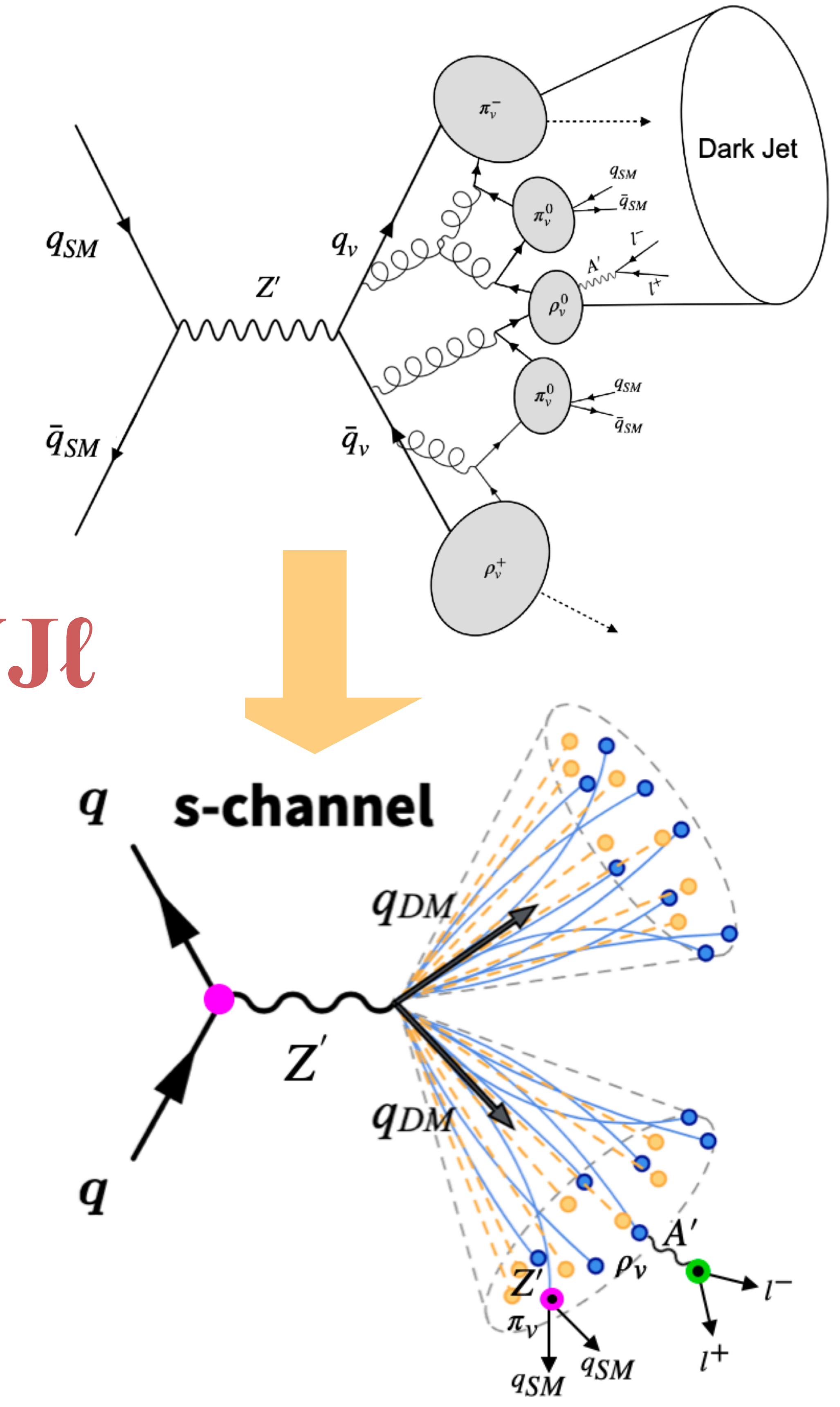
- **SVJℓ**: democratic decay to leptons allows by double messanger model (Z' plus A')
- **SVJτ:** enhanced decay to 3rd lepton generation

- Sensitivity of existing searches very limited to SVJ models with leptons

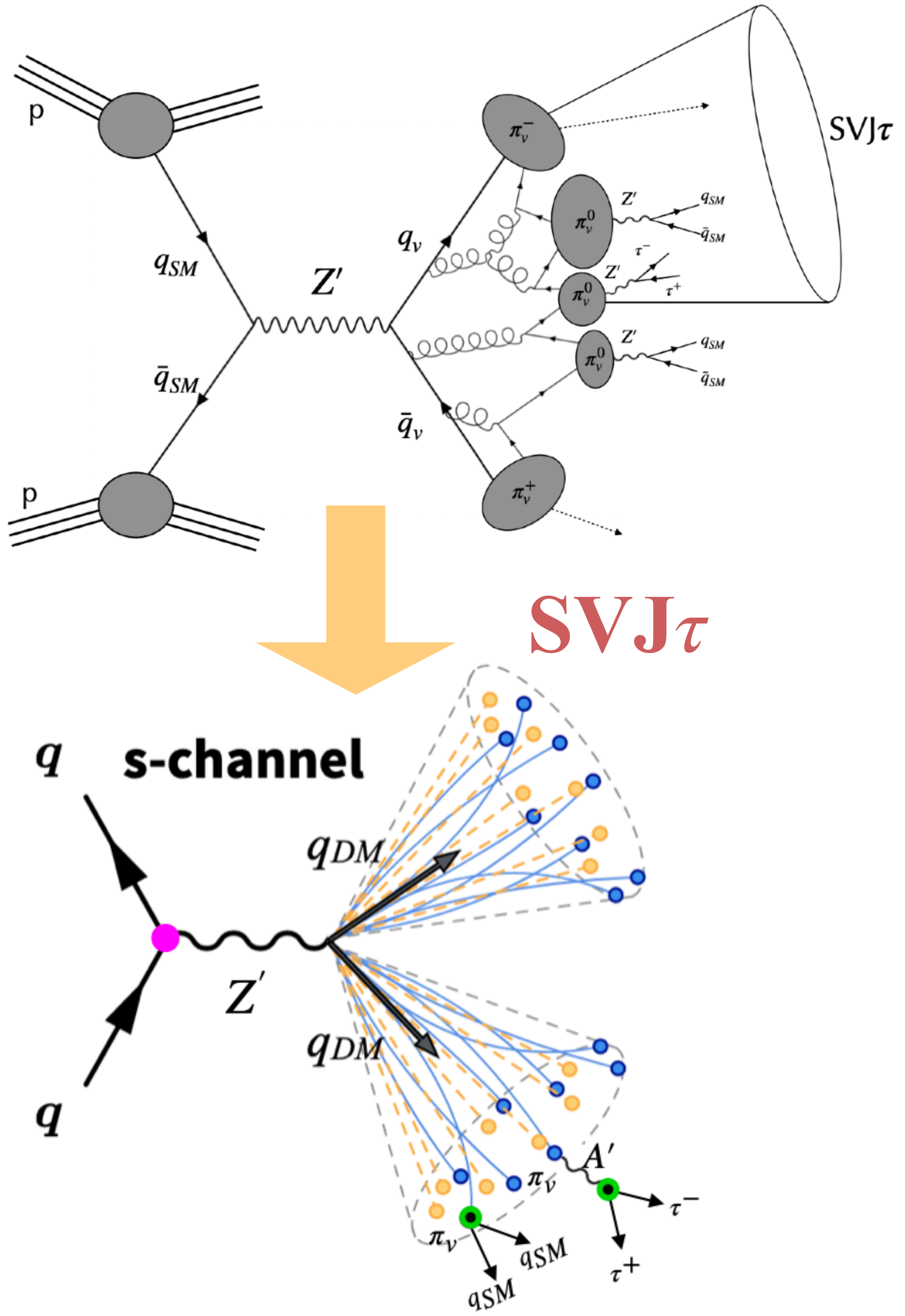
Novel signature: MET aligned to jets with prompt, non-isolated leptons

- Tau final state: additional MET from neutrinos

SVJℓ



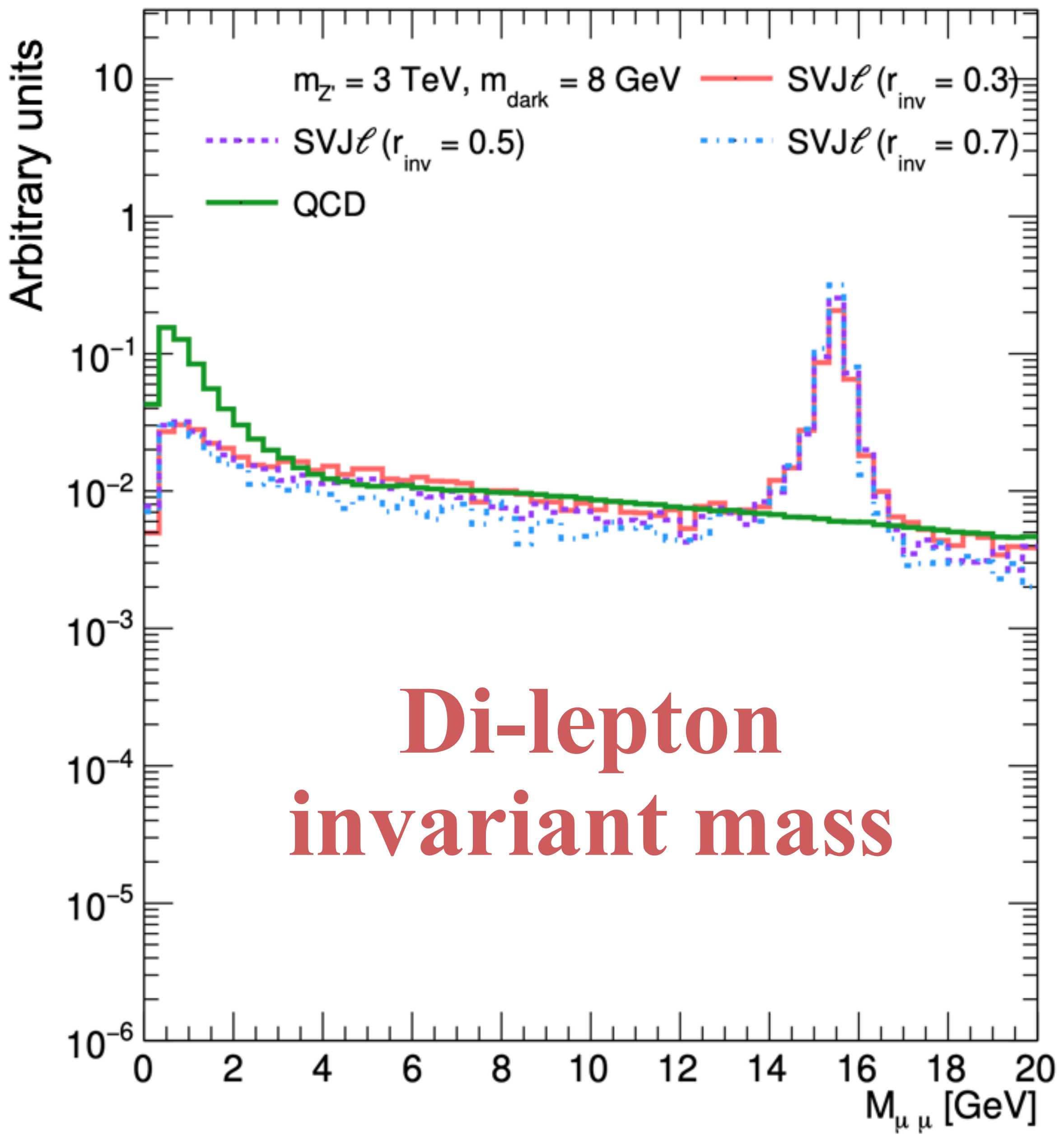
SVJτ



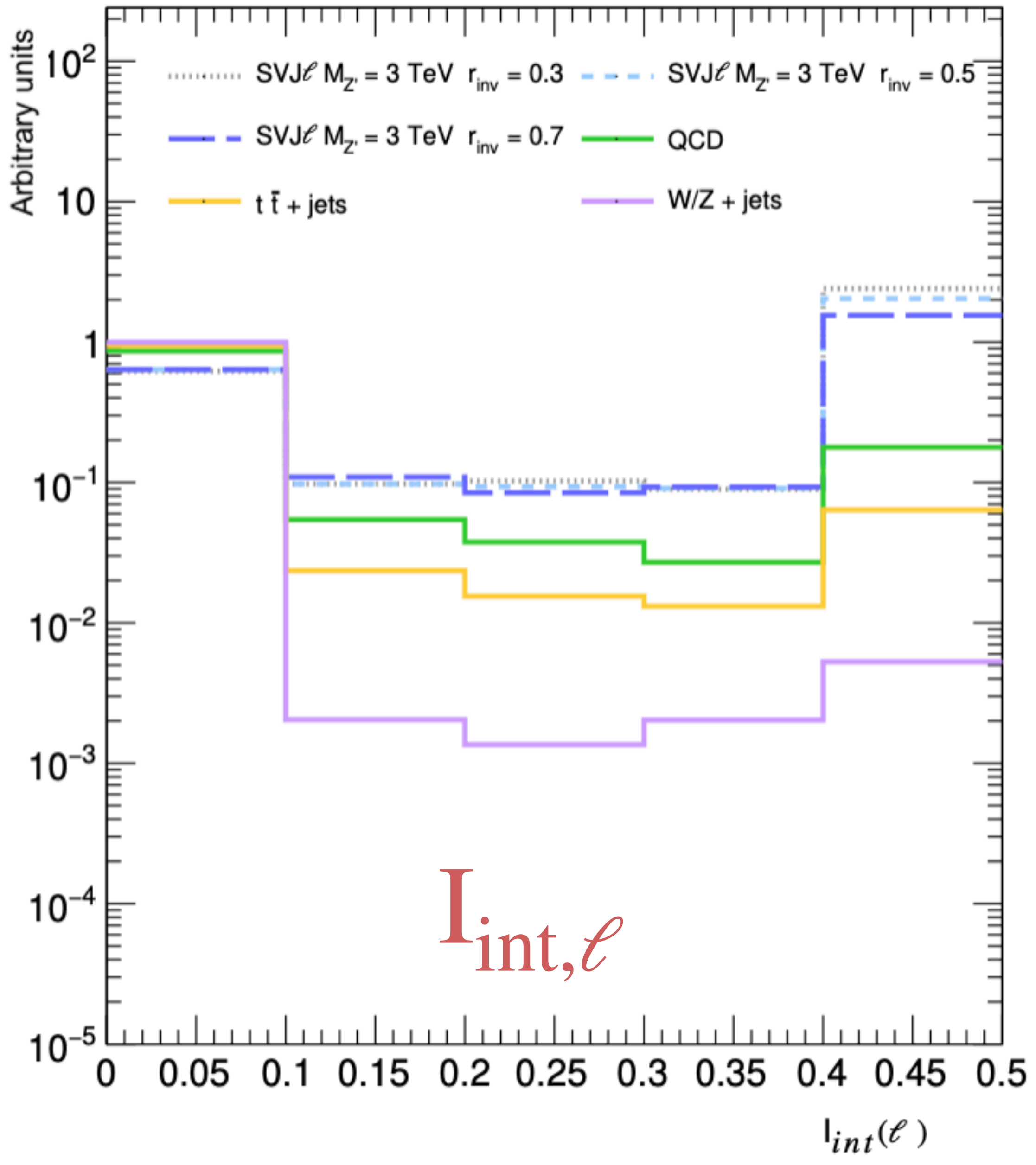
Identifying SVJℓ and SVJτ

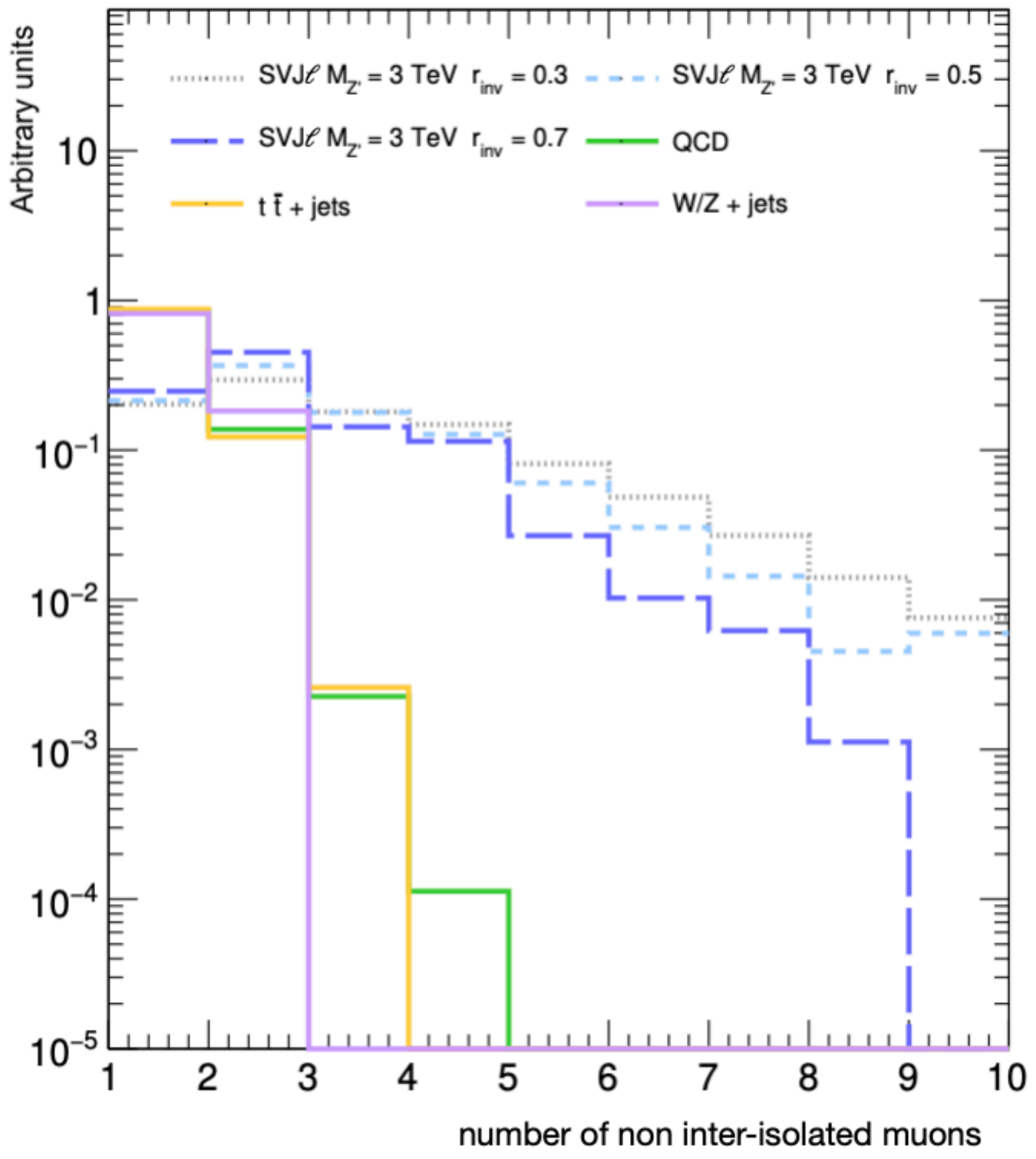
Experimental handles: inter-isolation and di-lepton mass

- **Large number of nearby leptons** in signal jets
  - **Relative inter-isolation:** $I_{int,\ell} = \frac{1}{p_{T,\ell}} \sum^{\Delta R, R^{iso}_{max}} p_{T,i}$
  - **Lepton-based substructure** (*) variables

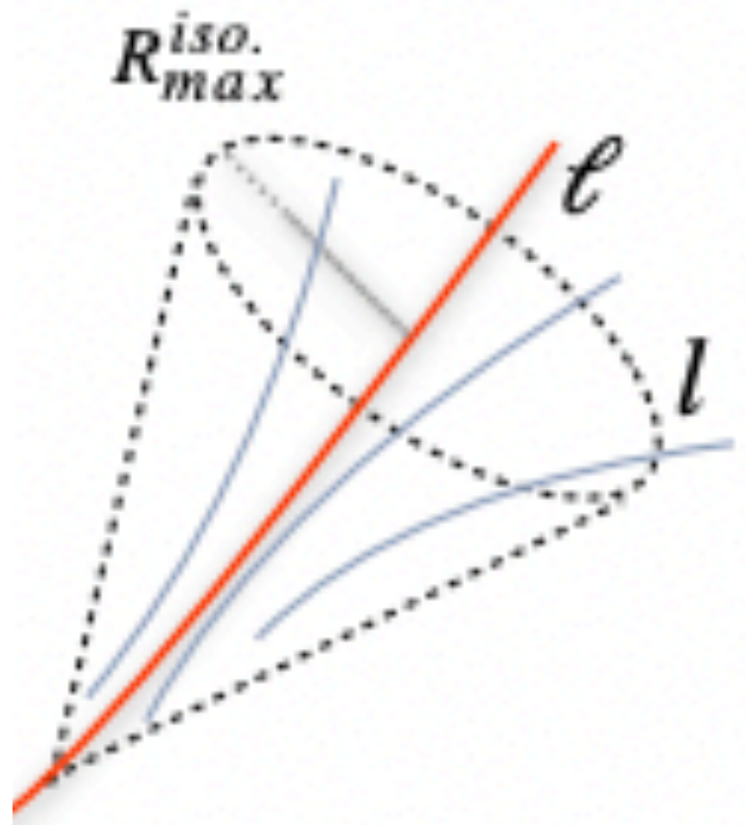


- **Di-lepton mass** inside jets expected to peak around **dark meson**



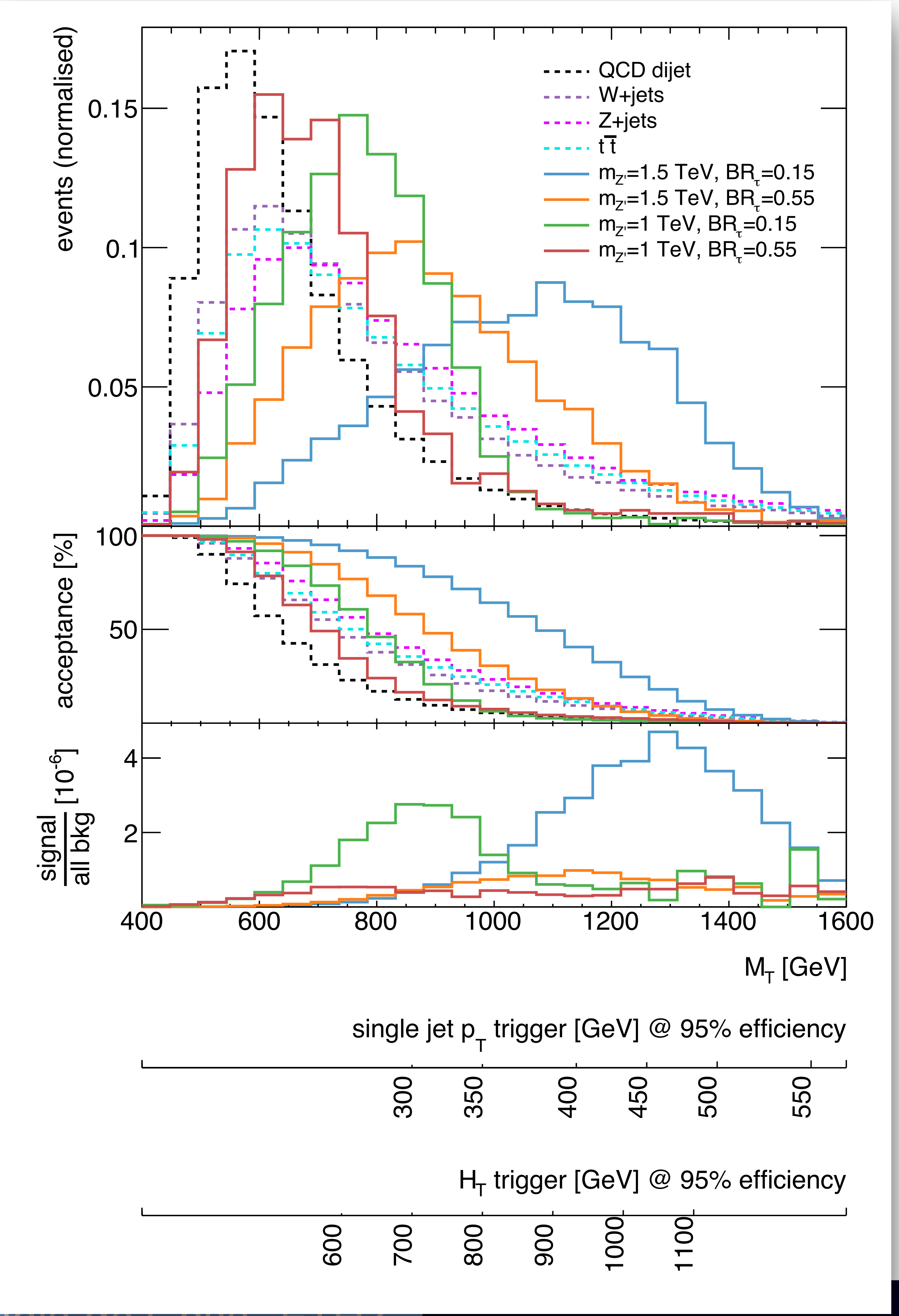


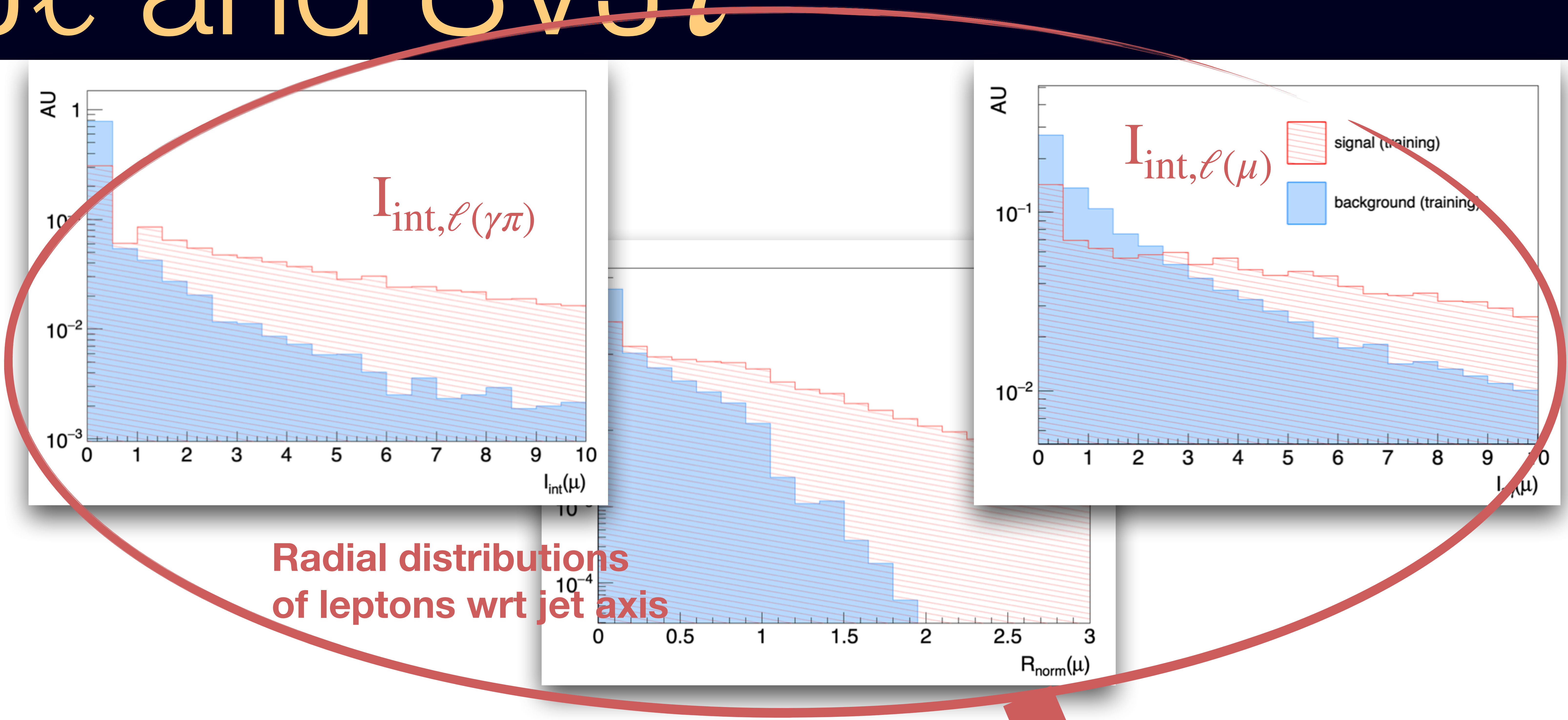
Number Inter-isolated leptons



Identifying SVJℓ and SVJτ

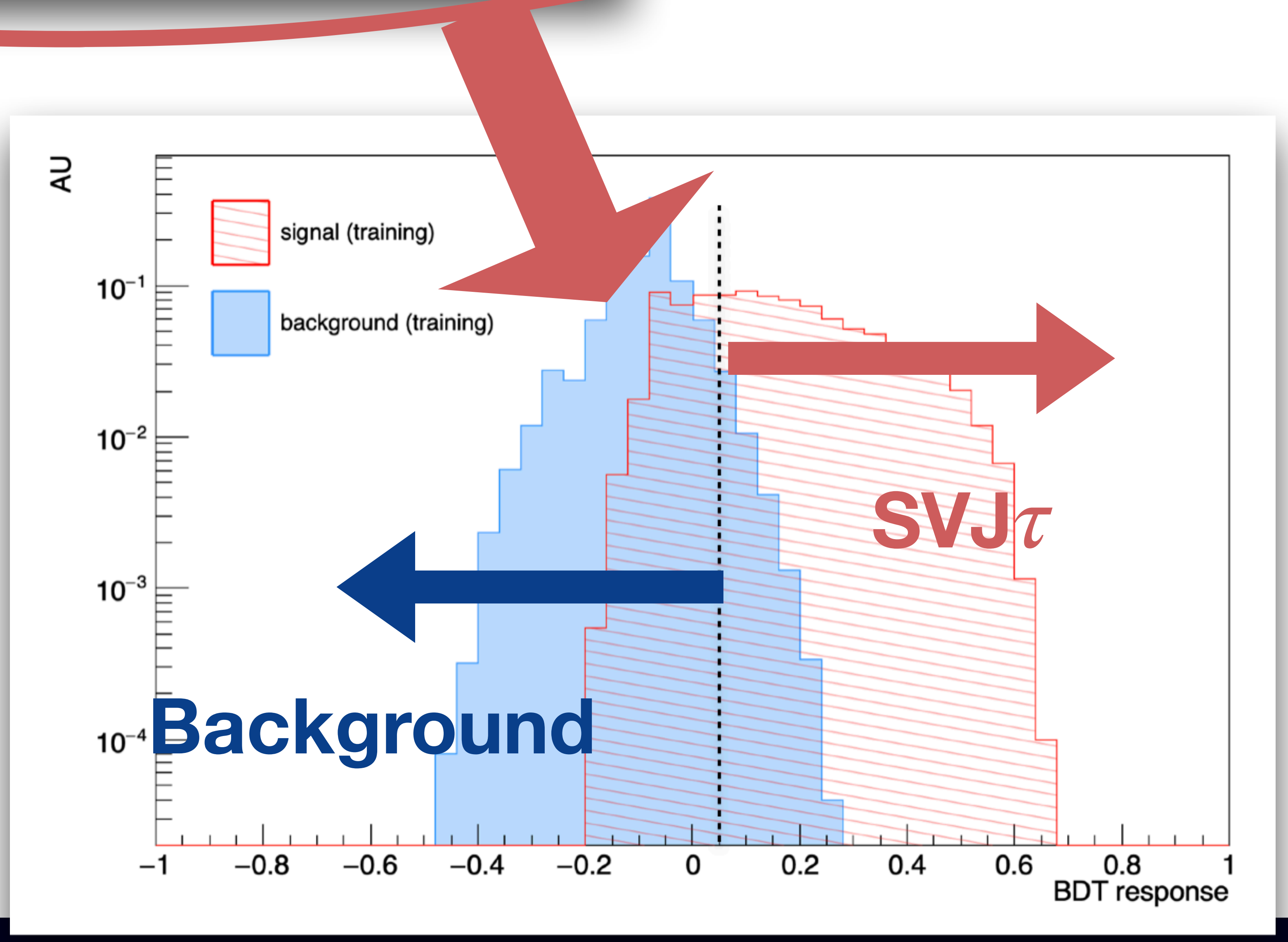
SVJτ are harder to catch due to decay modes variety and neutrino presence

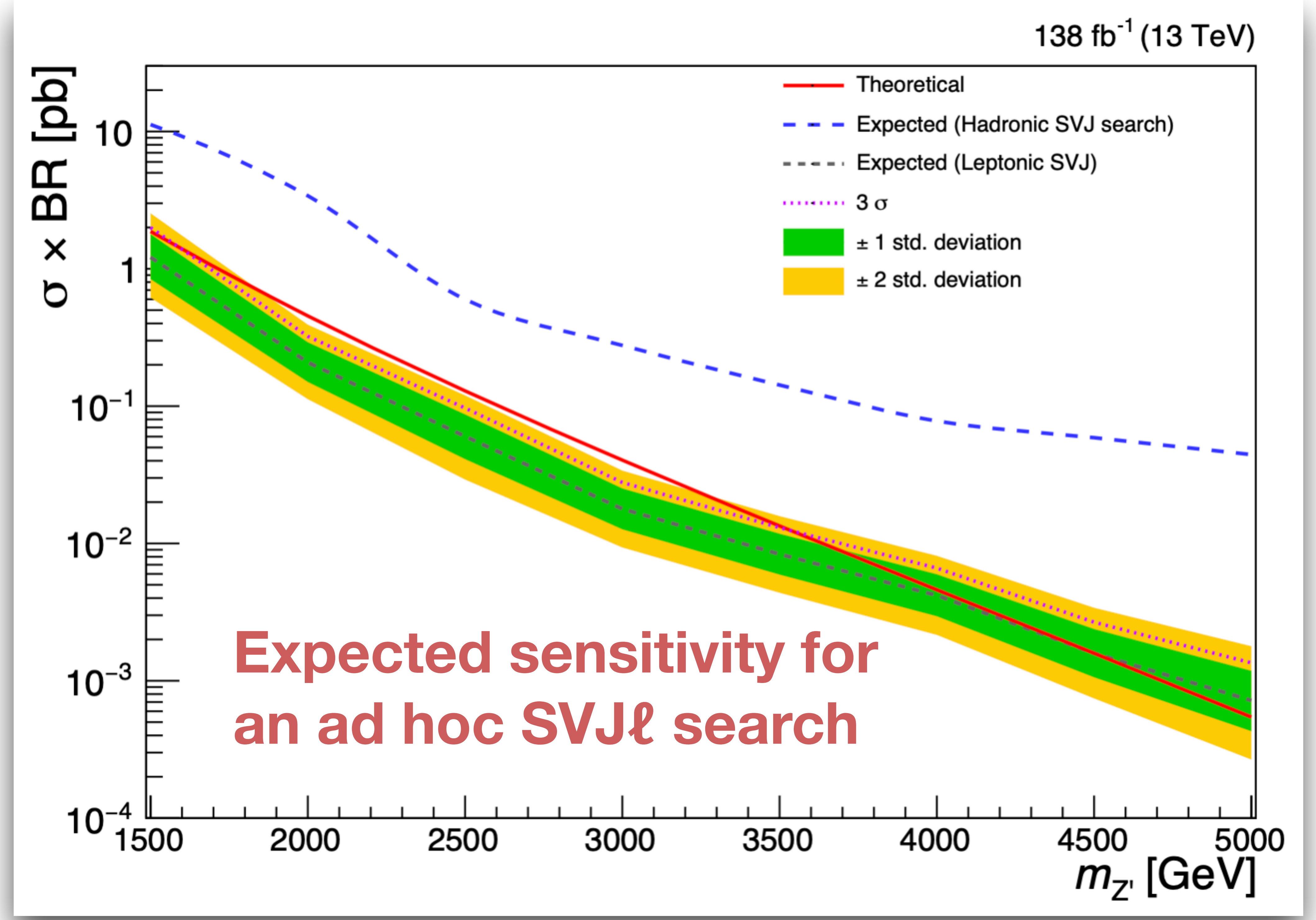




Radial distributions of leptons wrt jet axis

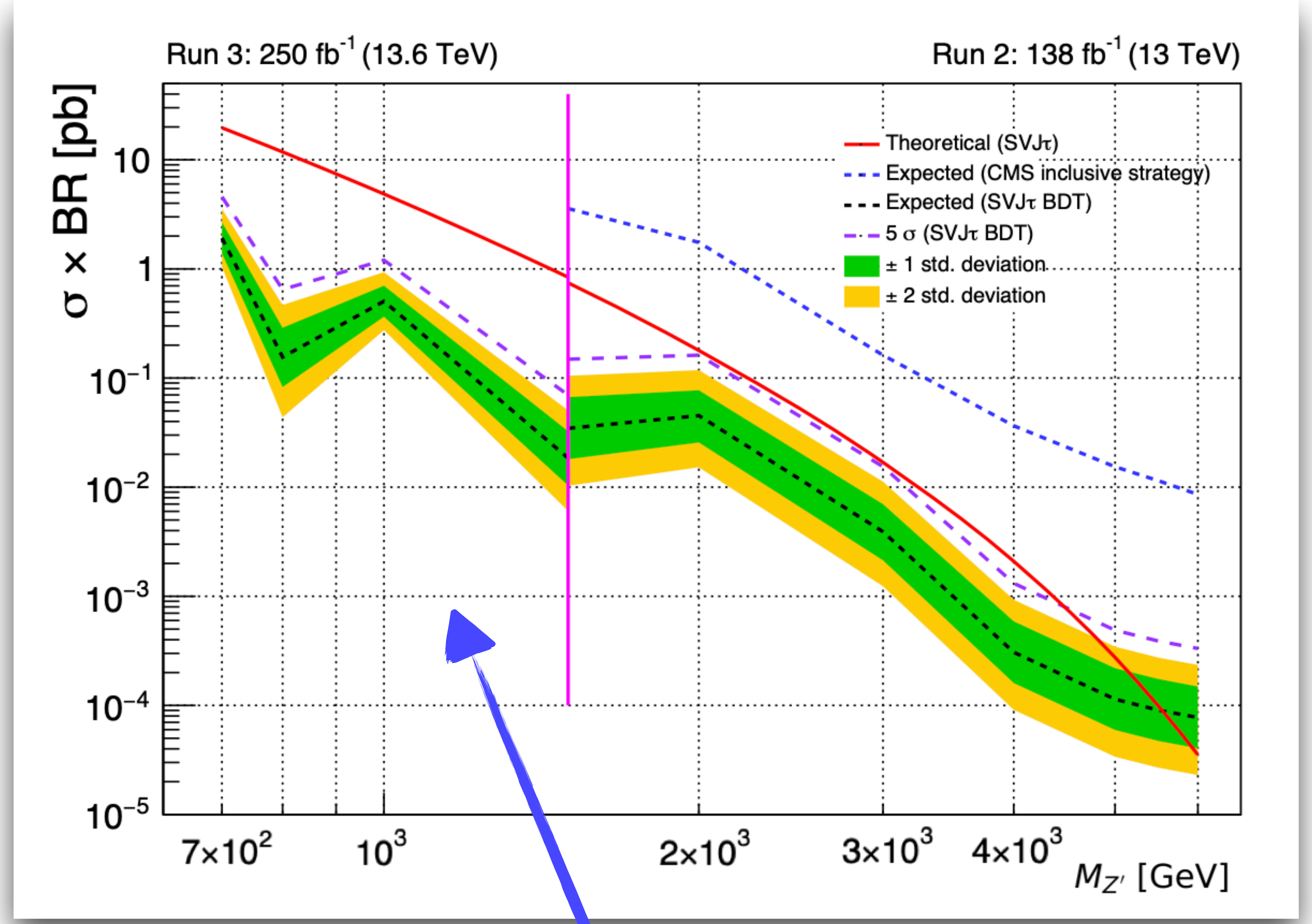
Can still exploit lepton-based substructure





Expected sensitivity for an ad hoc SVJℓ search

Expected sensitivity for an ad hoc SVJτ search



Possible extension using alternative online data selection strategies

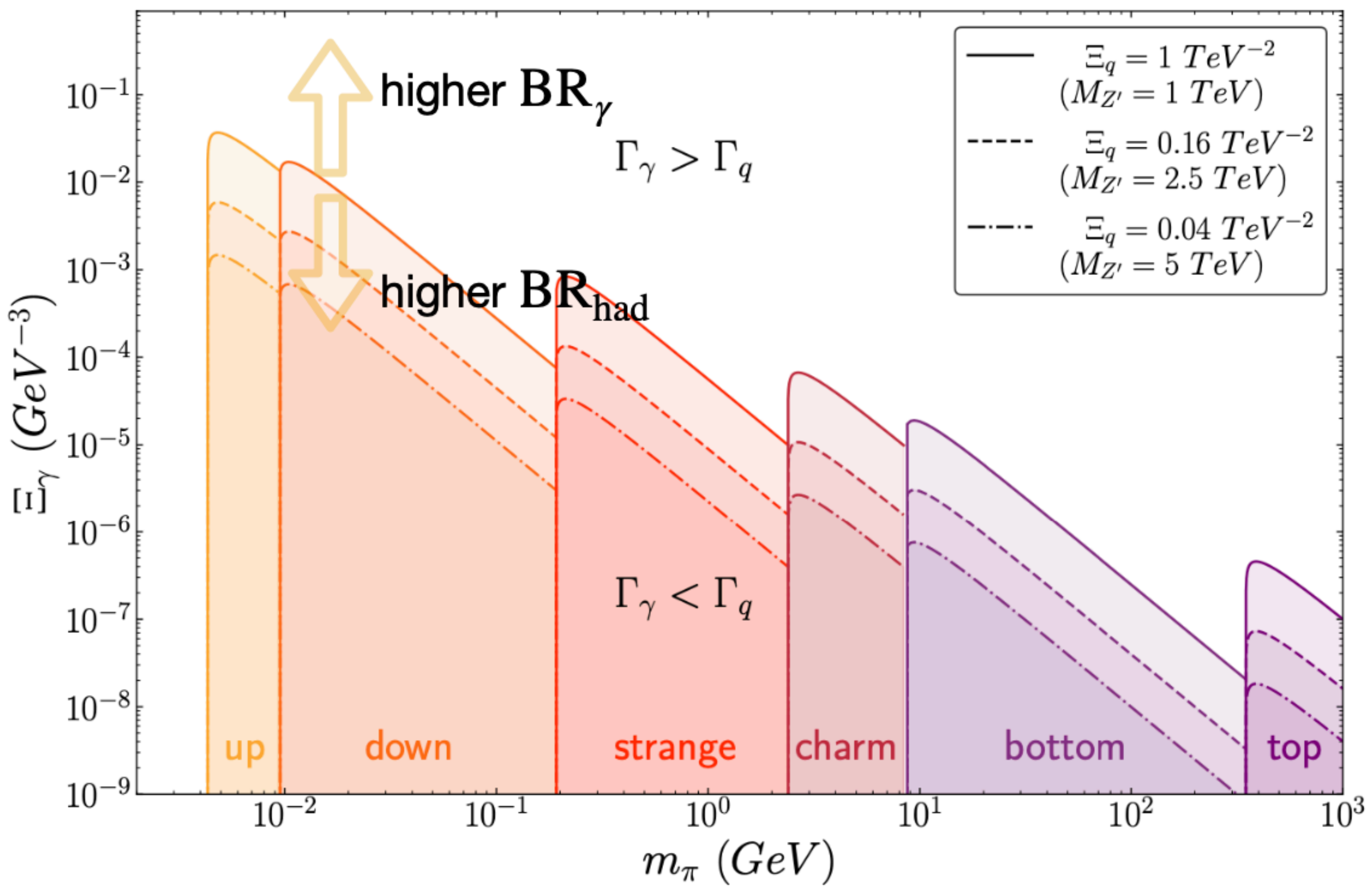
SVJγ

Eur. Phys. J. C 84, 1223 (2024)- C.Cazzaniga, A. Russo[2], E.Sitti[1], A. de Cosa

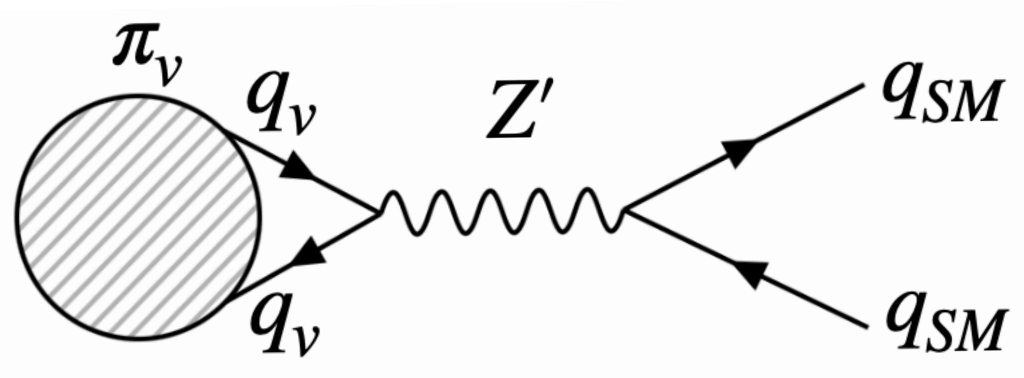
Dark and SM sectors communicate via a Z' boson and a pseudo scalar ALP

- Dark bound states can decay to SM quarks, γ or remain stable
- Signature: hadronic jet with photons and missing momentum aligned

ISOCURVES FOR $BR_\gamma = BR_{had} = 0.5$

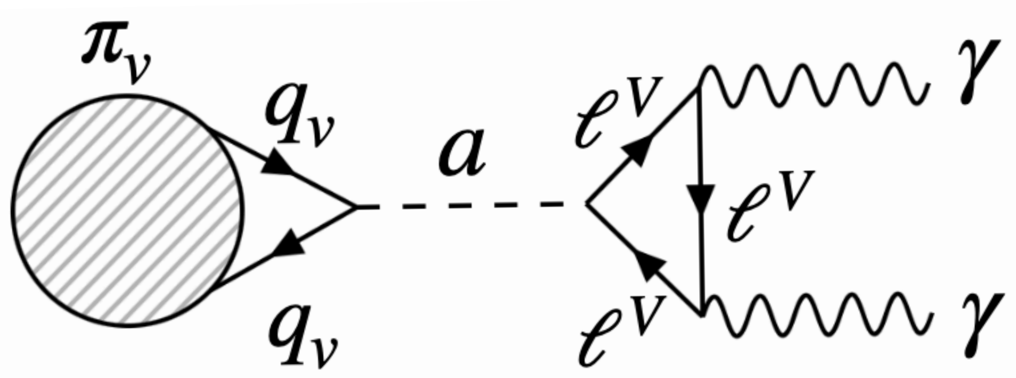


BR_{had}



$$\Gamma(\pi_v \to \bar{q}_i q_i) = \frac{N_c}{32\pi} m_{q_i}^2 m_{\pi_v}^3 \sqrt{1 - \frac{4m_{q_i}^2}{m_{\pi_v}^2}}\, \Xi_q^2$$

BR_γ



$$\Gamma(\pi_v \to \gamma\gamma) = \frac{\alpha_{EM}^2}{64\pi^2} m_{\pi_v}^7 \Xi_\gamma^2$$

Effective couplings

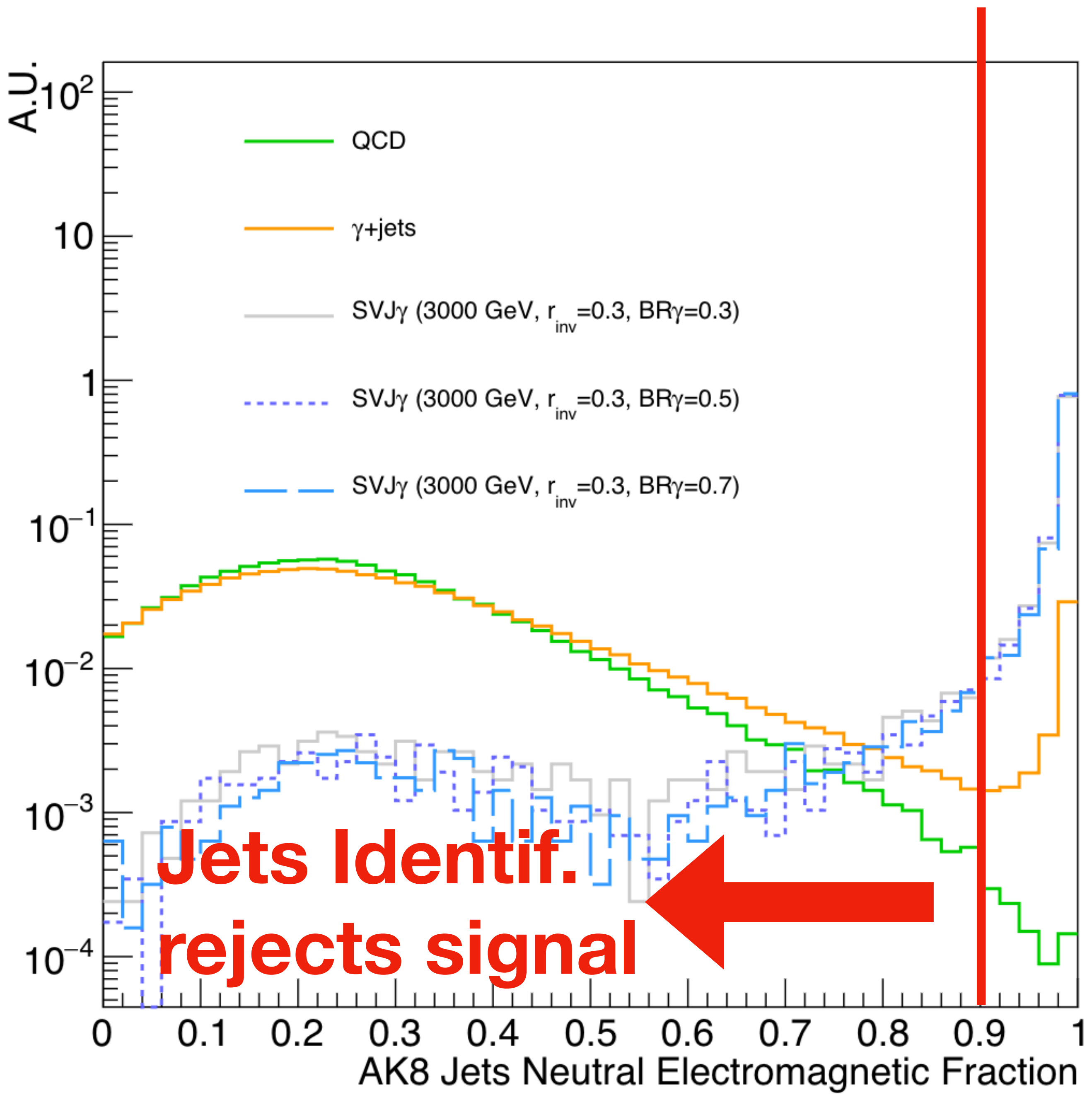
$$\Xi_q := \frac{\Delta_i^q \Delta^\psi}{M_{Z'}^2} \quad \text{and} \quad \Xi_\gamma := \frac{\lambda_L \lambda_\psi}{m_a^2 M_L}$$

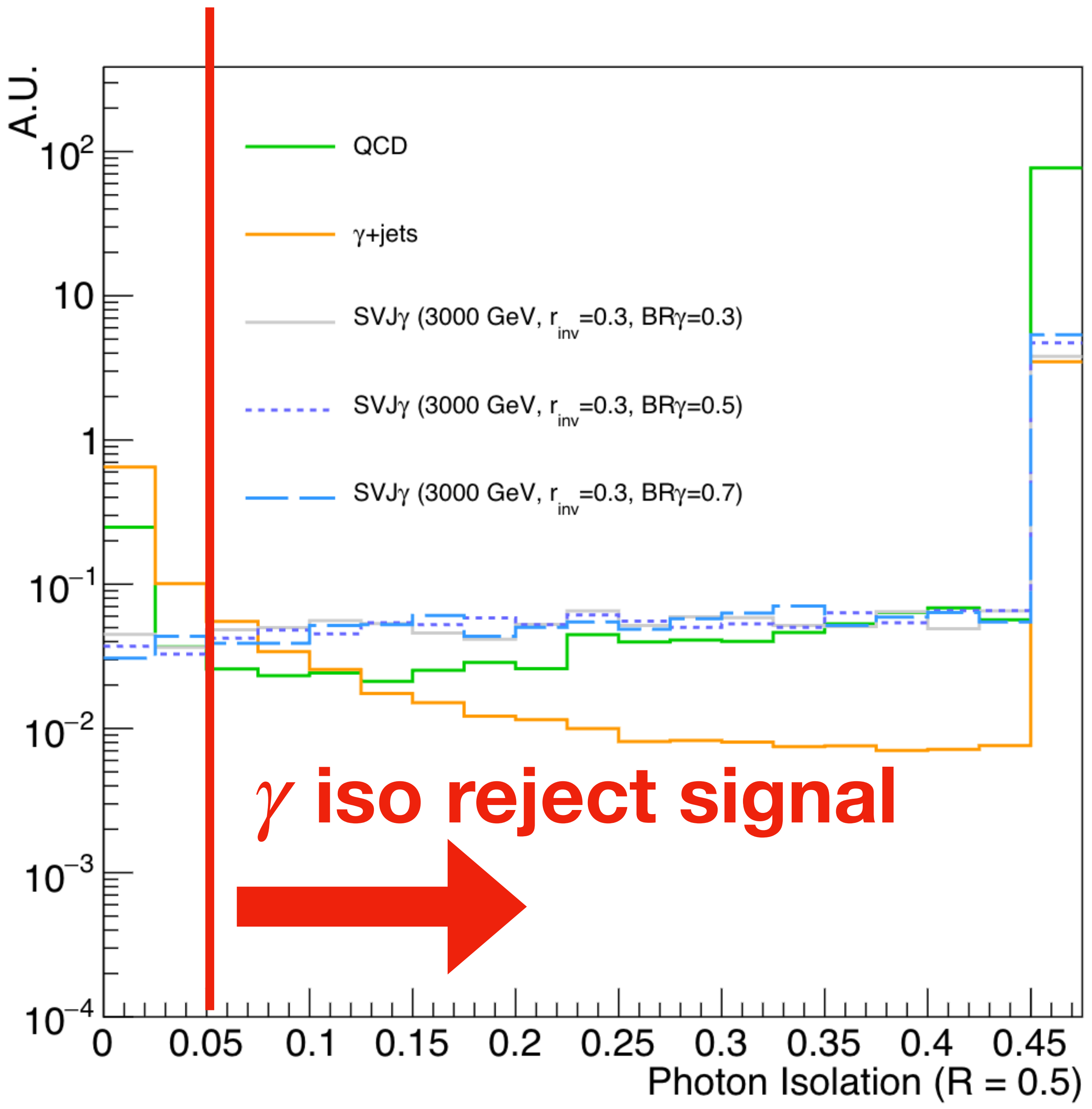
BR_{had} and BR_γ can be tuned

Allowing large coverage of signature space

Inefficiency of standard searches

Searches based on jets or photon final states are insensitive to SVJγ





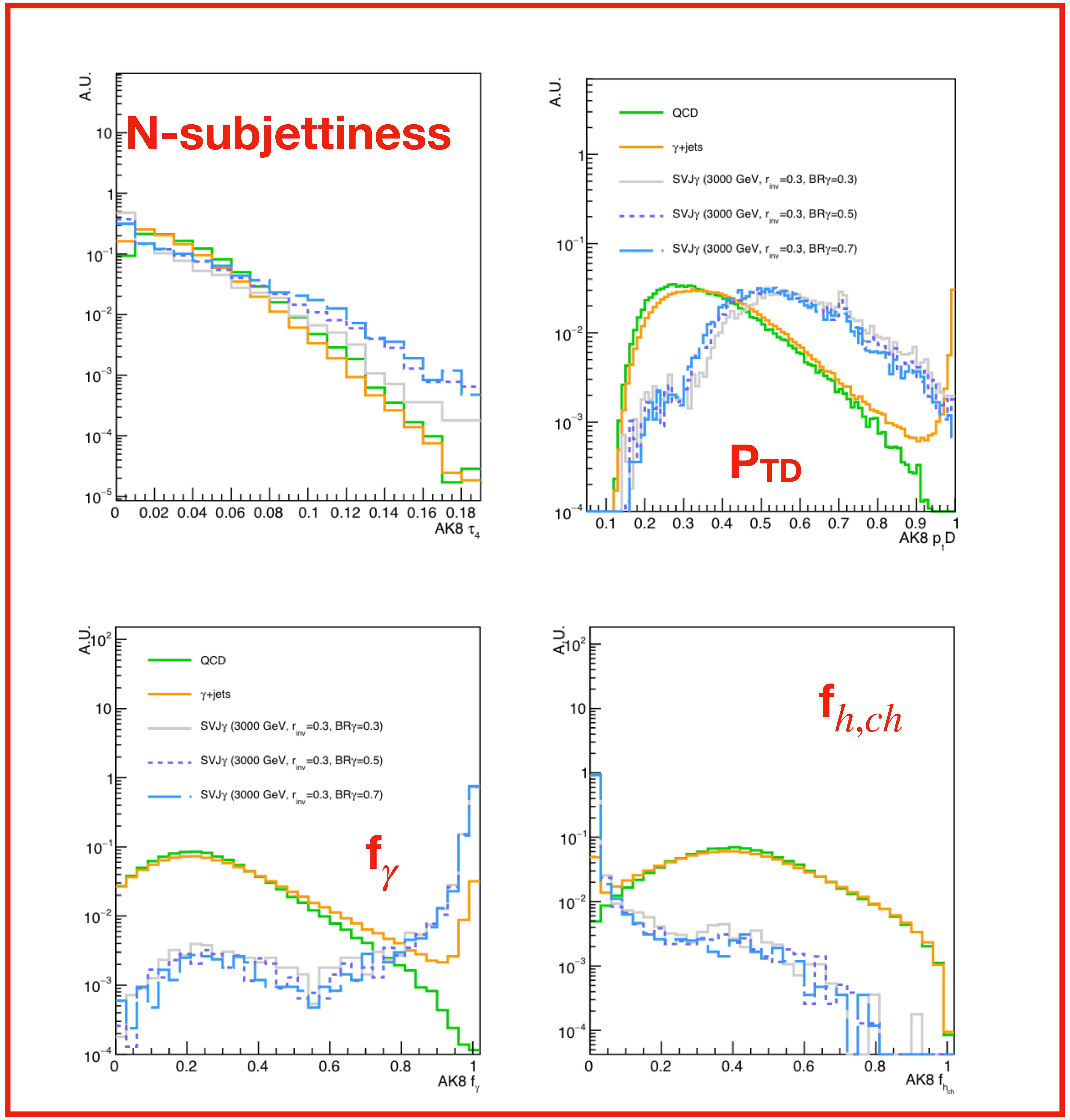
Jets-based searches

Jets identification removes photons emulating jets

γ-based searches

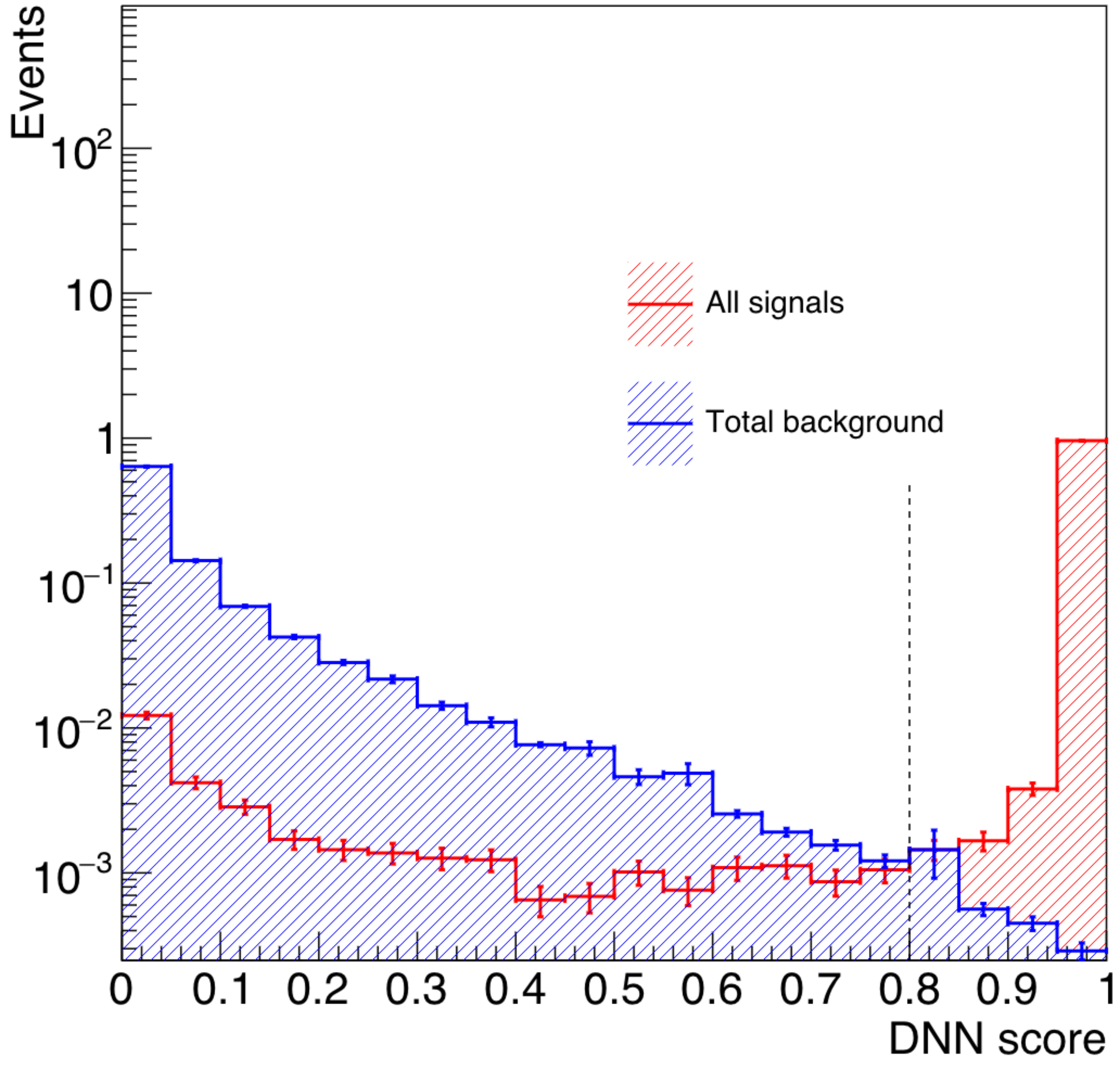
Tight photon isolation criteria rejects misindentified jets

SVJγ tagging algorithm

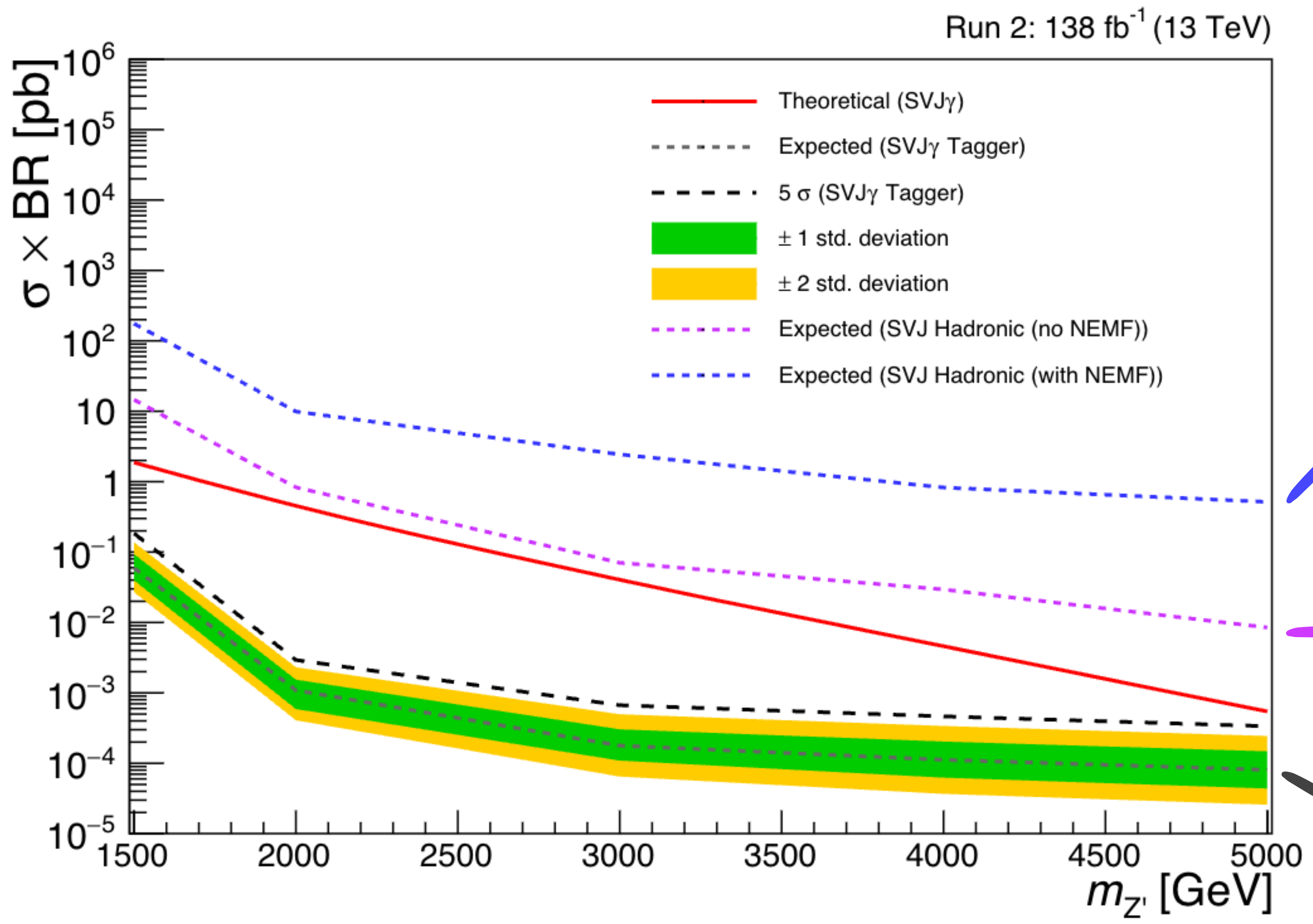


Ad-hoc identification algorithm for SVJγ

- Neural network based on **jet substructure** features helps in identifying SVJγ while efficiently rejecting SM jets and SM photons



Opening the way to discover SVJγ



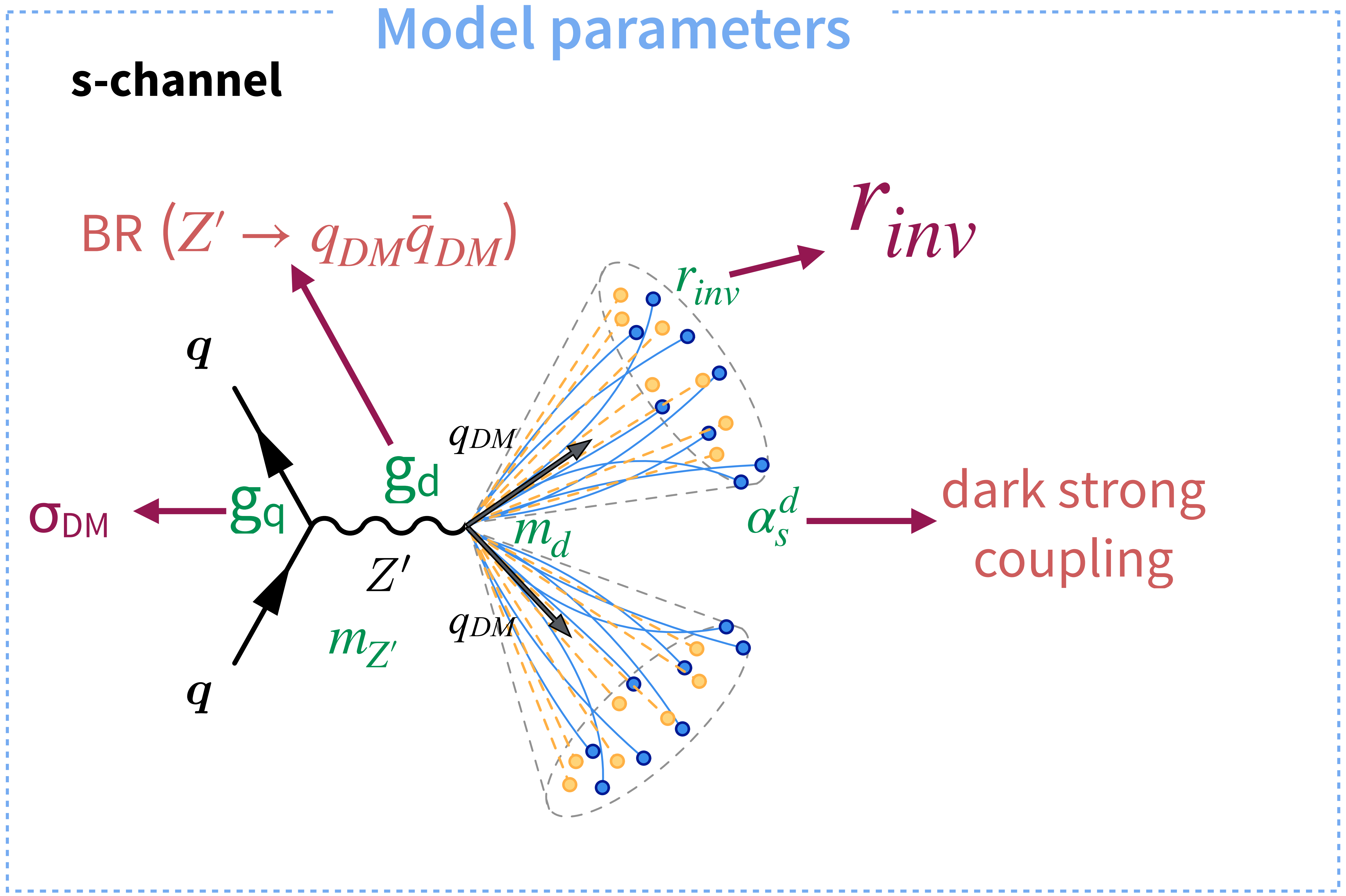
Removing standard identification requirements, and using a **dedicated selection** algorithm **open the way to discover SVJγ** already with Run 2 LHC data

Sensivitiy of CMS hadronic SVJ search

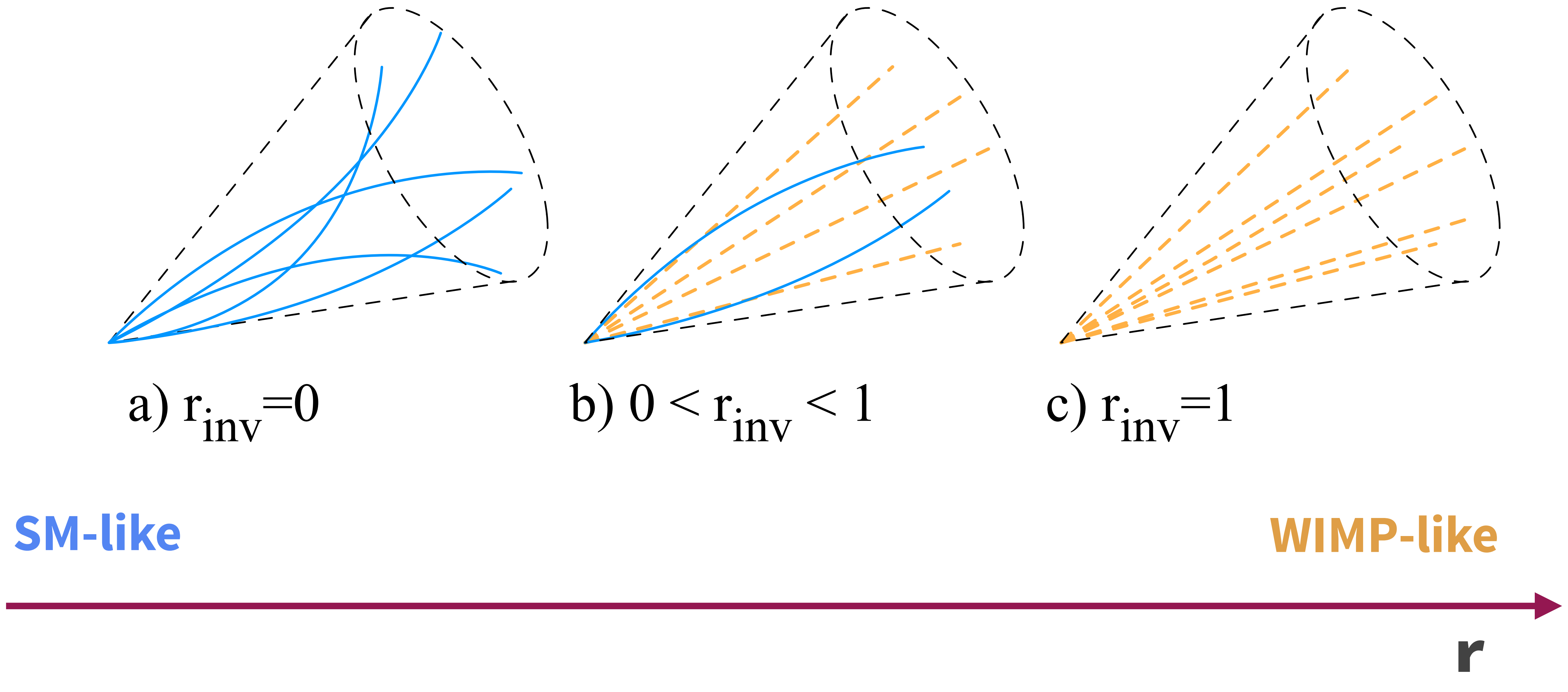
Sensivitiy of CMS search without jet id

Sensivitiy of CMS search without std id. algos and with SVJγ tagger

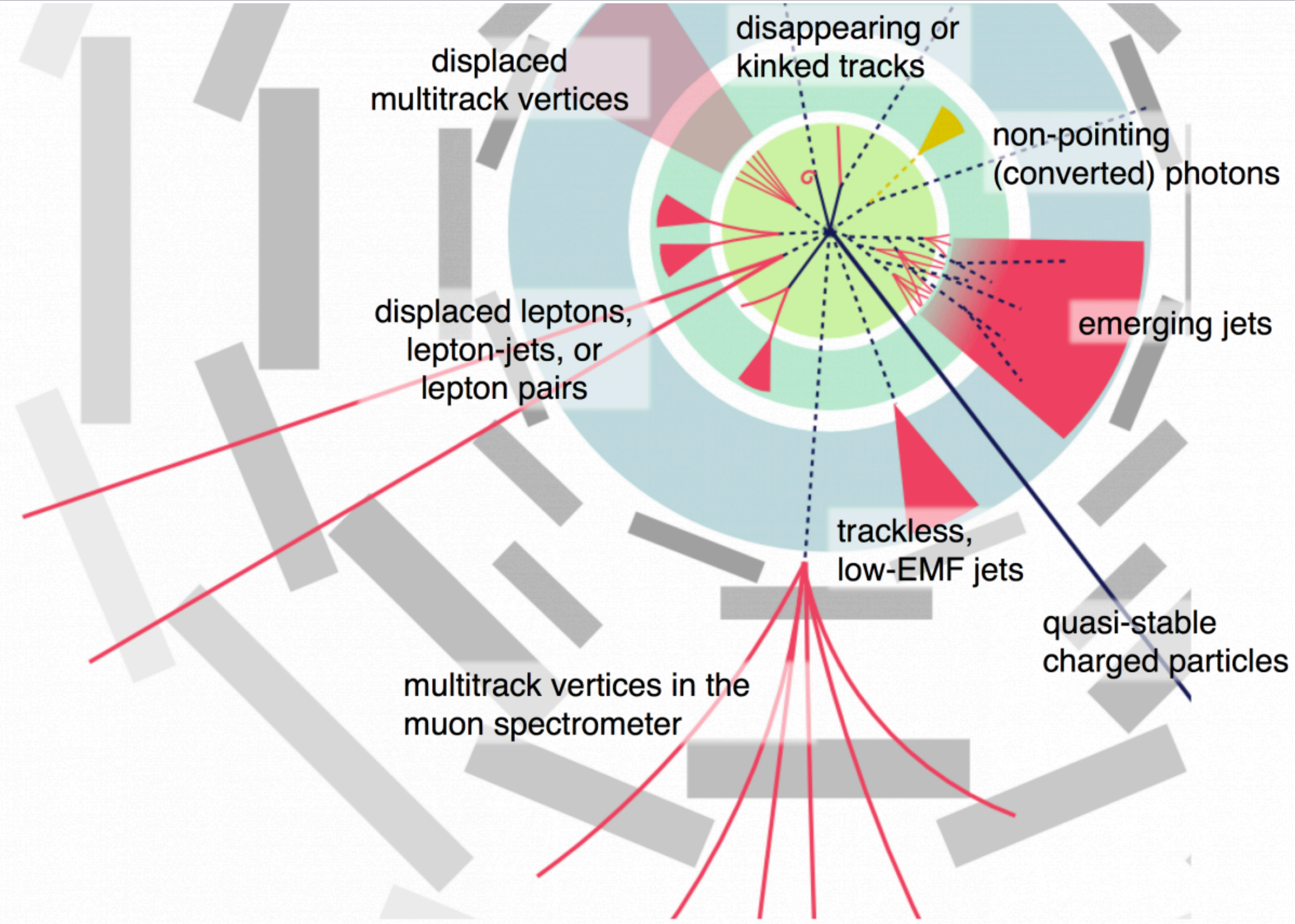
Complexity of Dark QCD



$$r_{inv} = \left\langle \frac{\textit{N stable dark hadrons}}{\textit{N dark hadrons}} \right\rangle$$



A step forward finding new physics



Dark QCD can be as complex as SM QCD

- **Dark showers phenomenology depends on unknown parameters (N_C, N_F, Λ_{dark}, $m_{\pi_{dark}}$, …)**
  - Large parameter space to cover
- **Traditional strategies: Target specific model**
  - Optimal sensitivity to targeted signal
  - Low or null sensitivity to different signatures

Change perspective

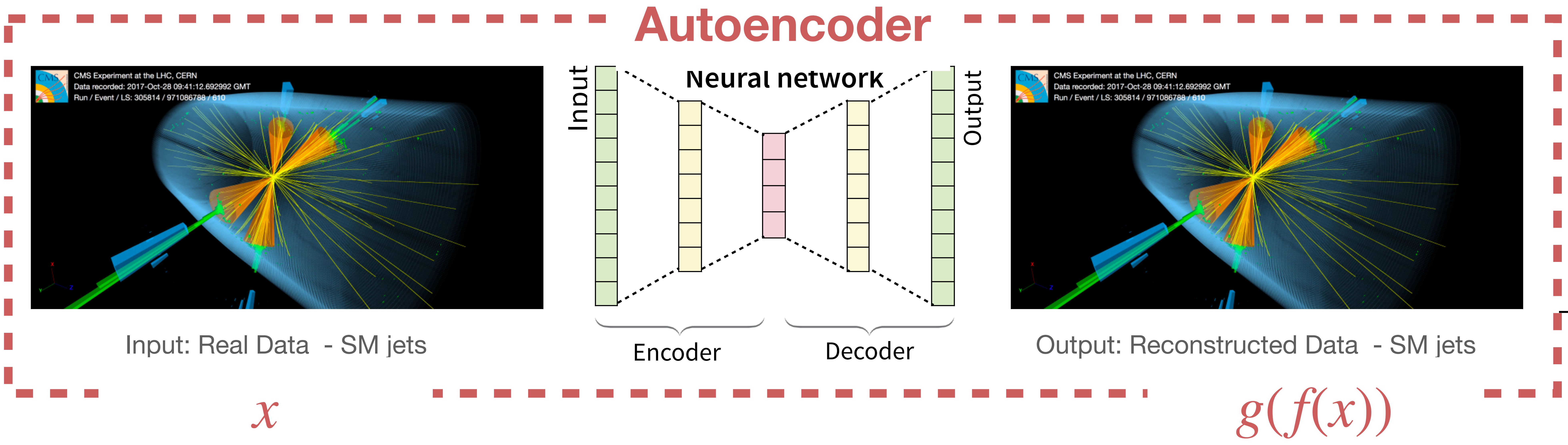
- **Move the focus on known physics!**
- Make no/few assumptions about new physics

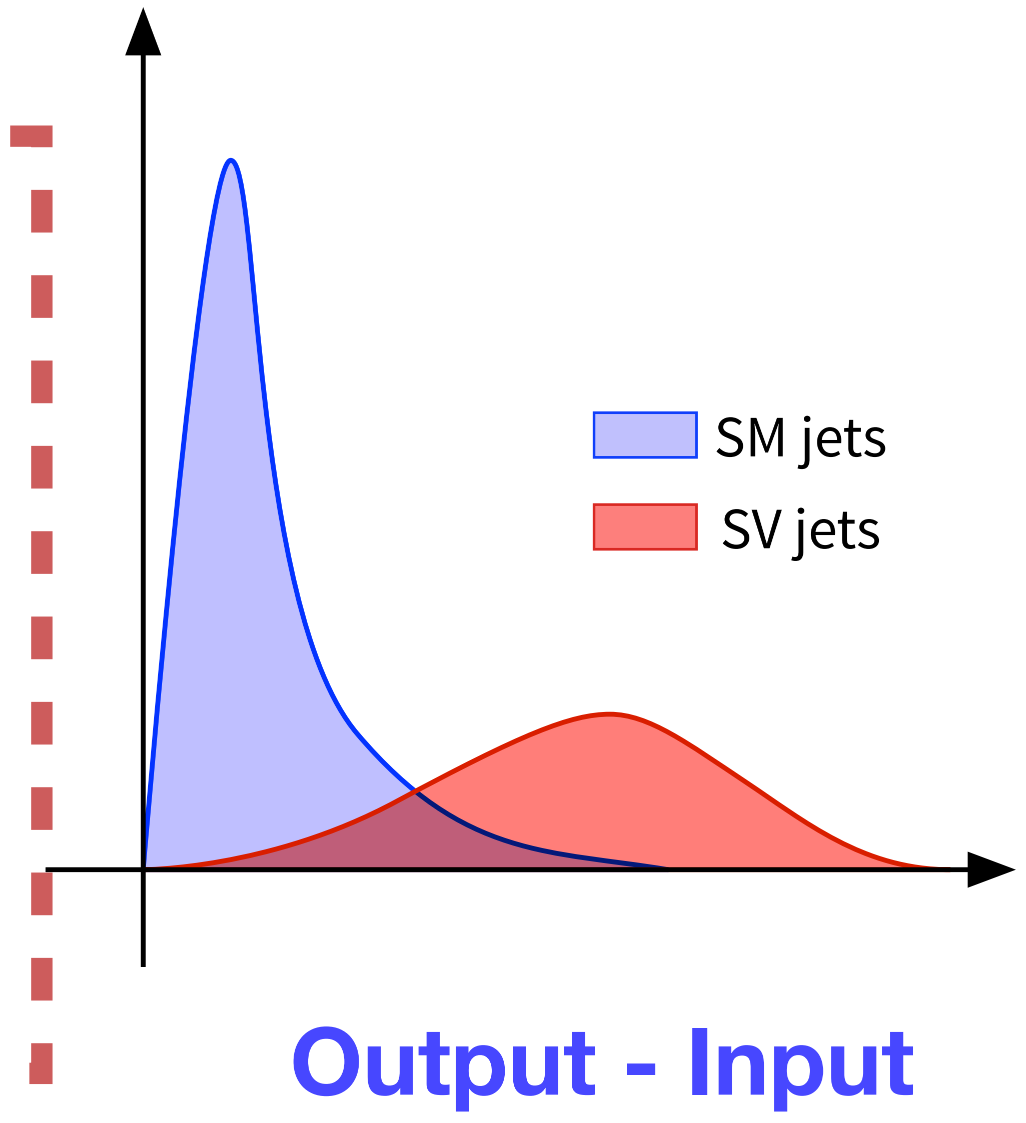
Anomaly detection

- Train a neural network **(Autoencoder)** to learn compressing and decompressing data minimising reconstruction error (e.g. MSE) between input and output
  - $L(x) = ||g(f(x)) - x||$
- **Identify new physics as anomalies**

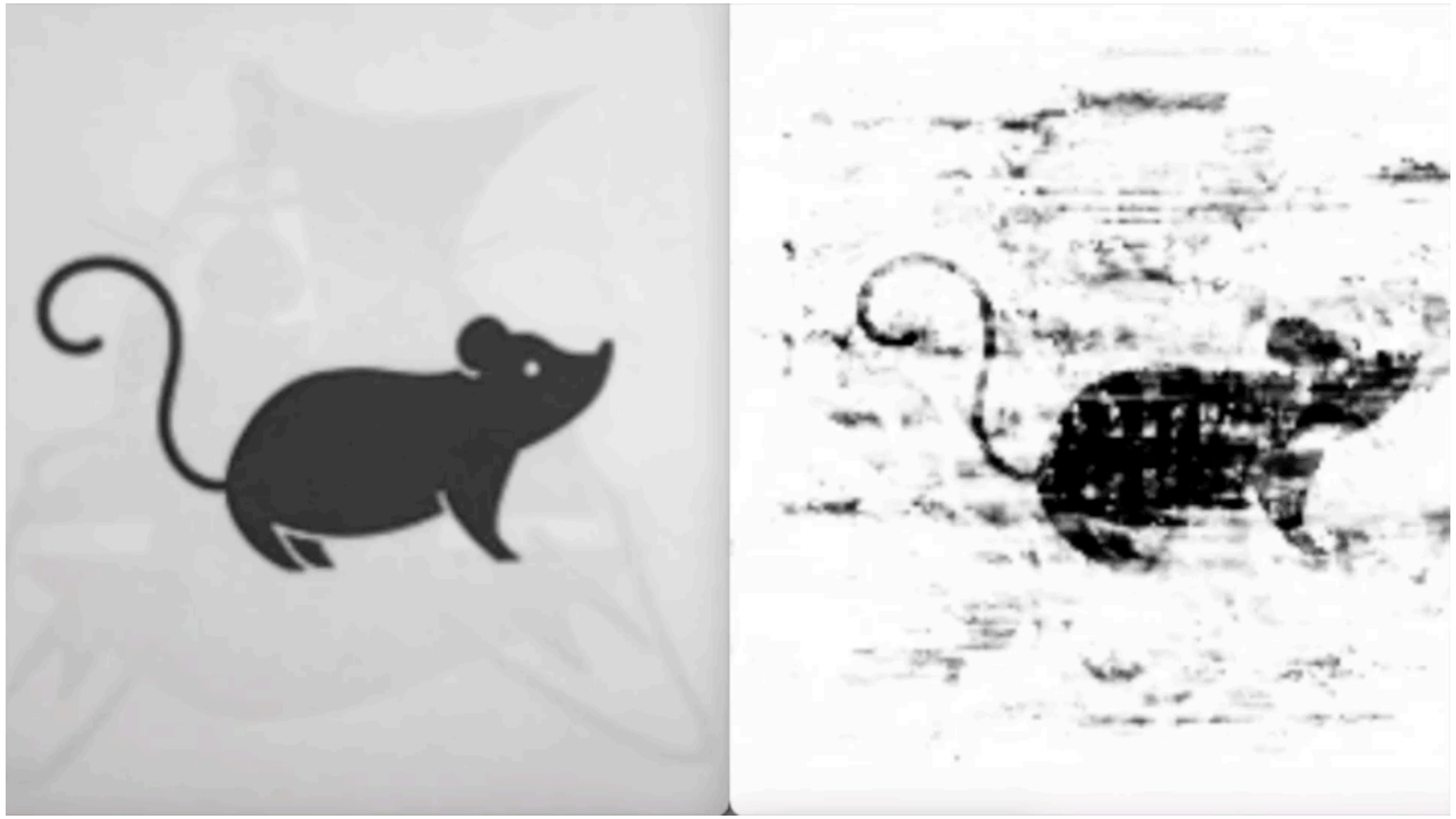
Training on SM data

- Independent on background modelling
- Agnostic about the signal





Autoencoder demo



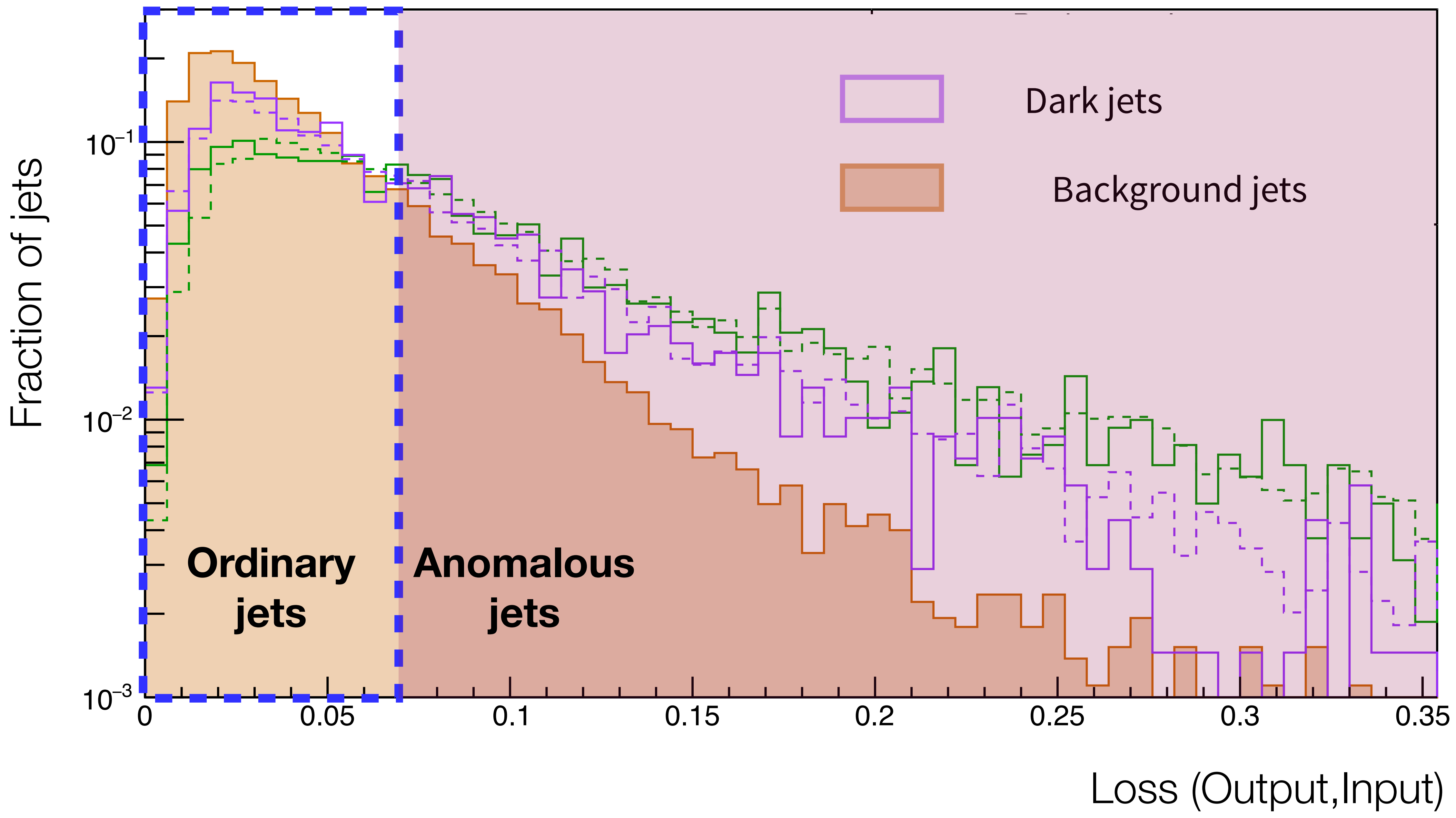
Input: Real Data   Output: Reconstructed Data

Video from Jeremi Niedziela

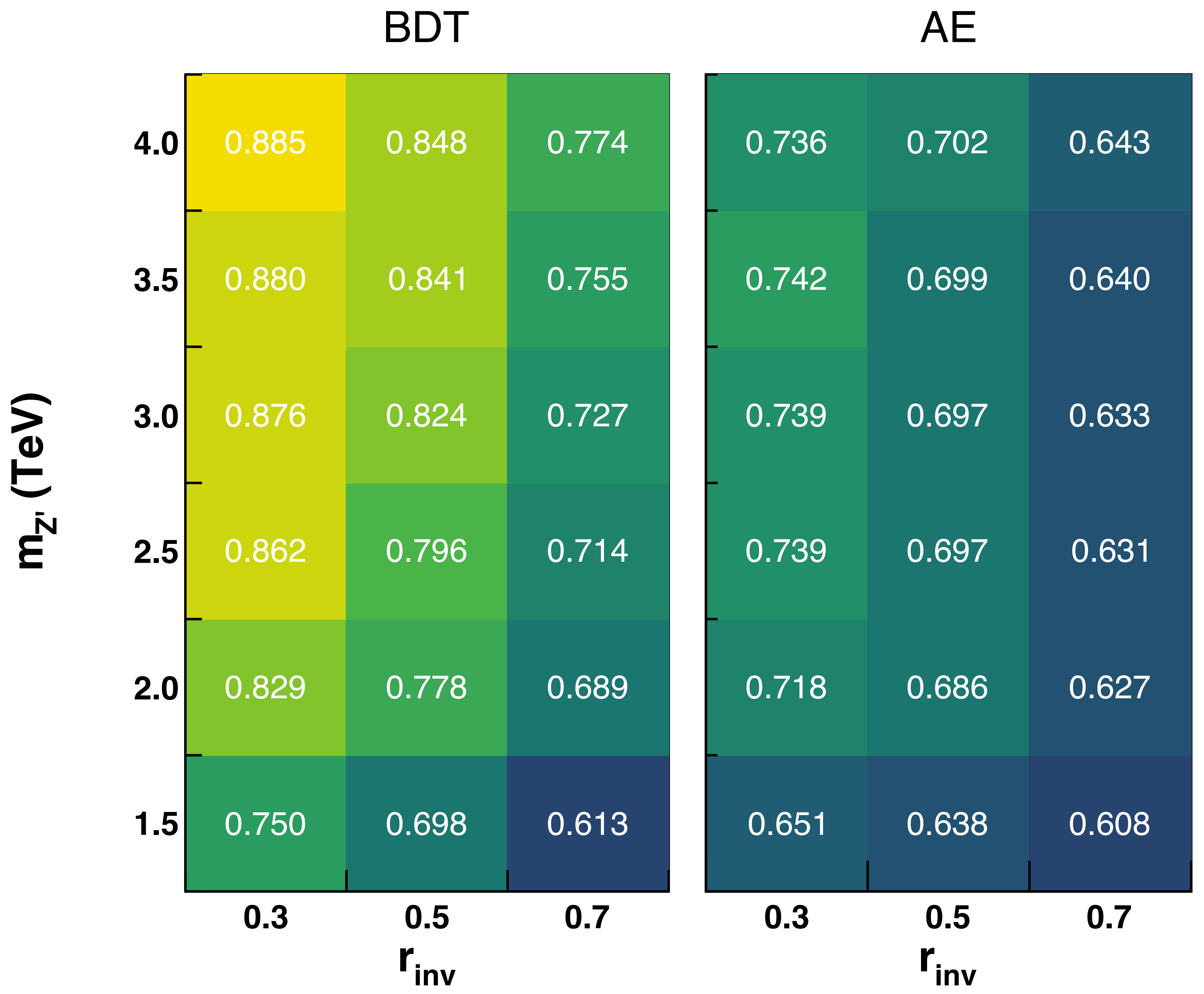
Paving the way for discovery

First investigation of AE application to SVJs gave promising results

- Use high level features from jets internal structure



A. de Cosa et al, "Autoencoders for Semivisible jet detection" JHEP Vol 2022, 74 (2022)

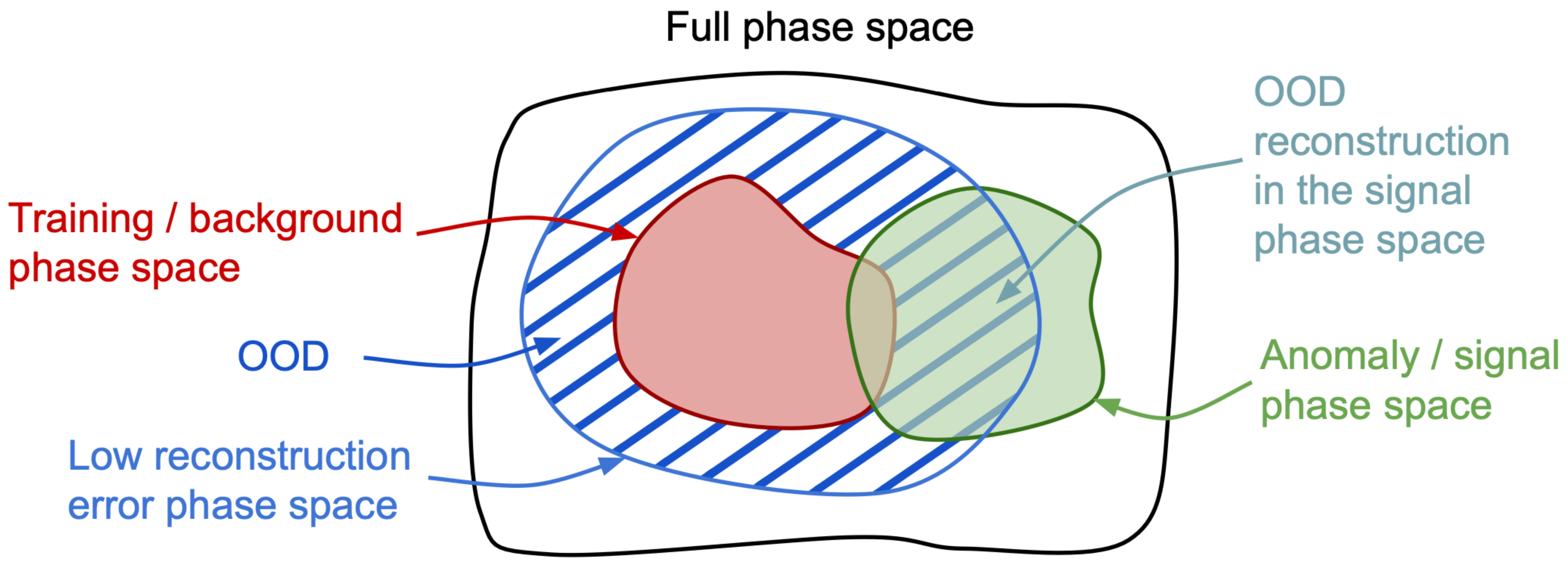


AUC (Area Under the Curve): the higher the better

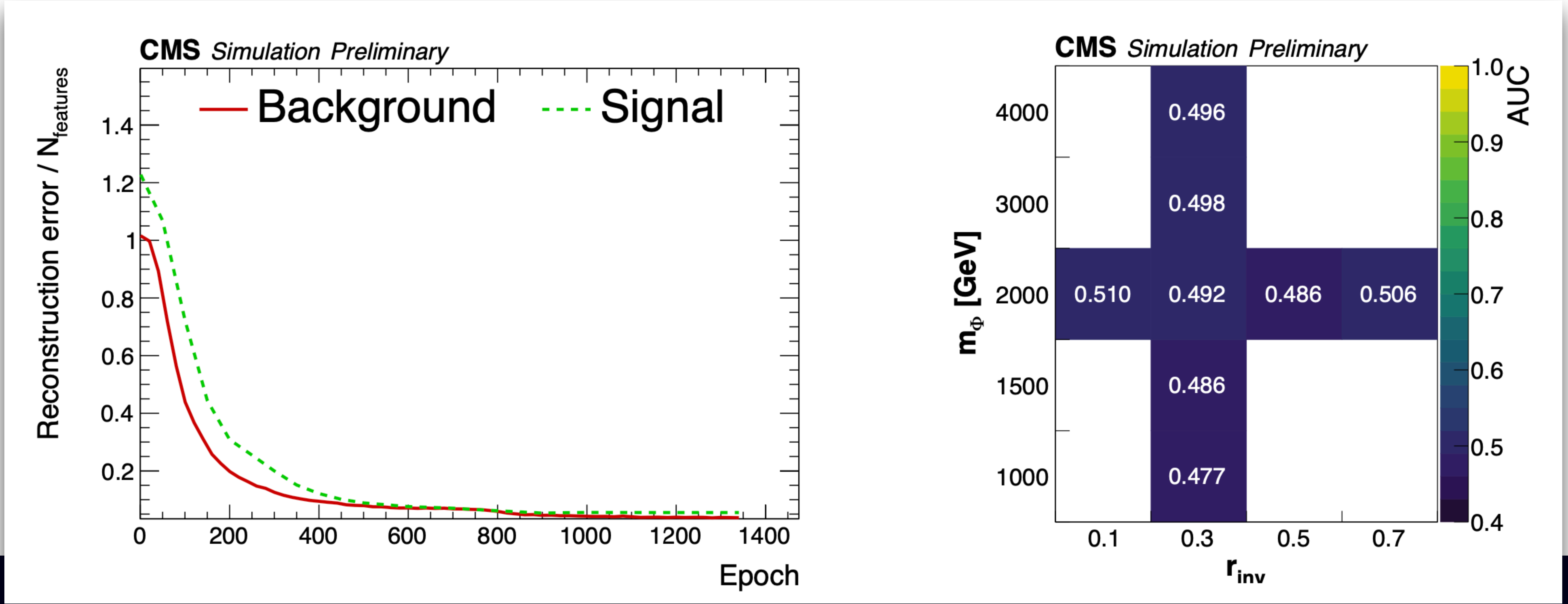
Shortcomings of standard AEs

When trained on complex objects, such as $t\bar{t}$

- **AE tends to generalise**
  - out-of-distribution (OOD) data reconstructed as well as training ones.
- **Critical loss in performance**
- Reconstruction error is not a good metric



$$L(x) = ||g(f(x)) - x||$$



The Normalised AE

We want to ensure matching between training data and low reconstruction error probability distribution

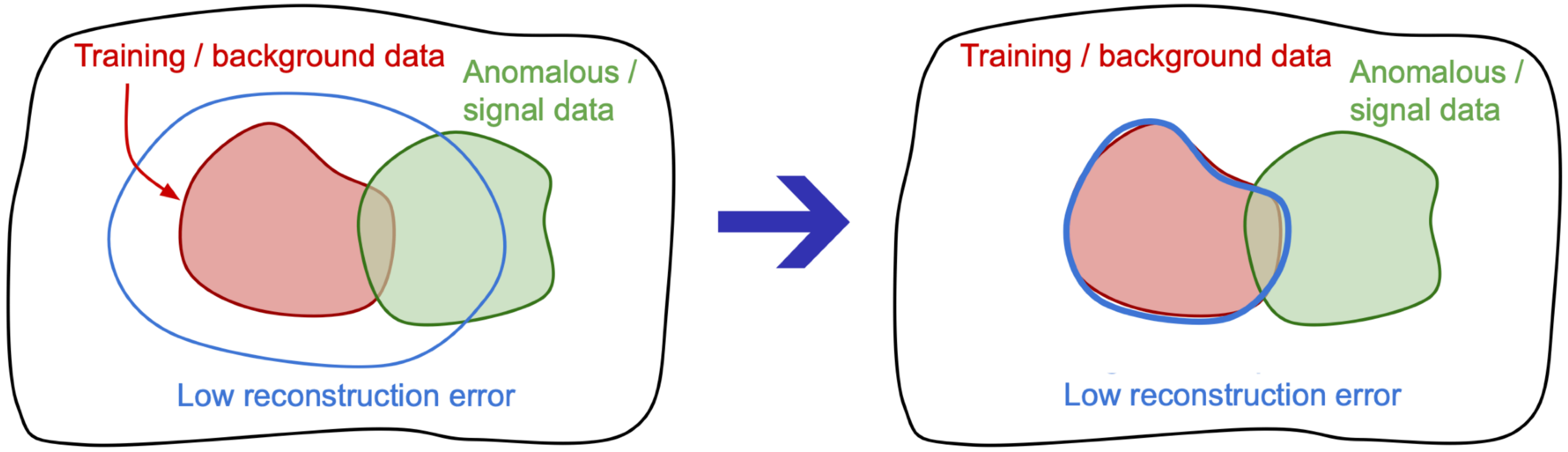
- Define probability distribution p_θ so that regions with low reco error E_θ have high probability
- Define reconstruction error as system Energy (E_θ)
- Loss designed to learn $p_\theta = p_{data}$:

$$p_\theta(x) = \frac{1}{\Omega_\theta} \exp(-E_\theta(x))$$

$$L_\theta = \mathbb{E}_{x \sim p_{\text{data}}}[E_\theta(x)] - \mathbb{E}_{x' \sim p_\theta}[E_\theta(x')]$$

positive energy E_+ negative energy E_-

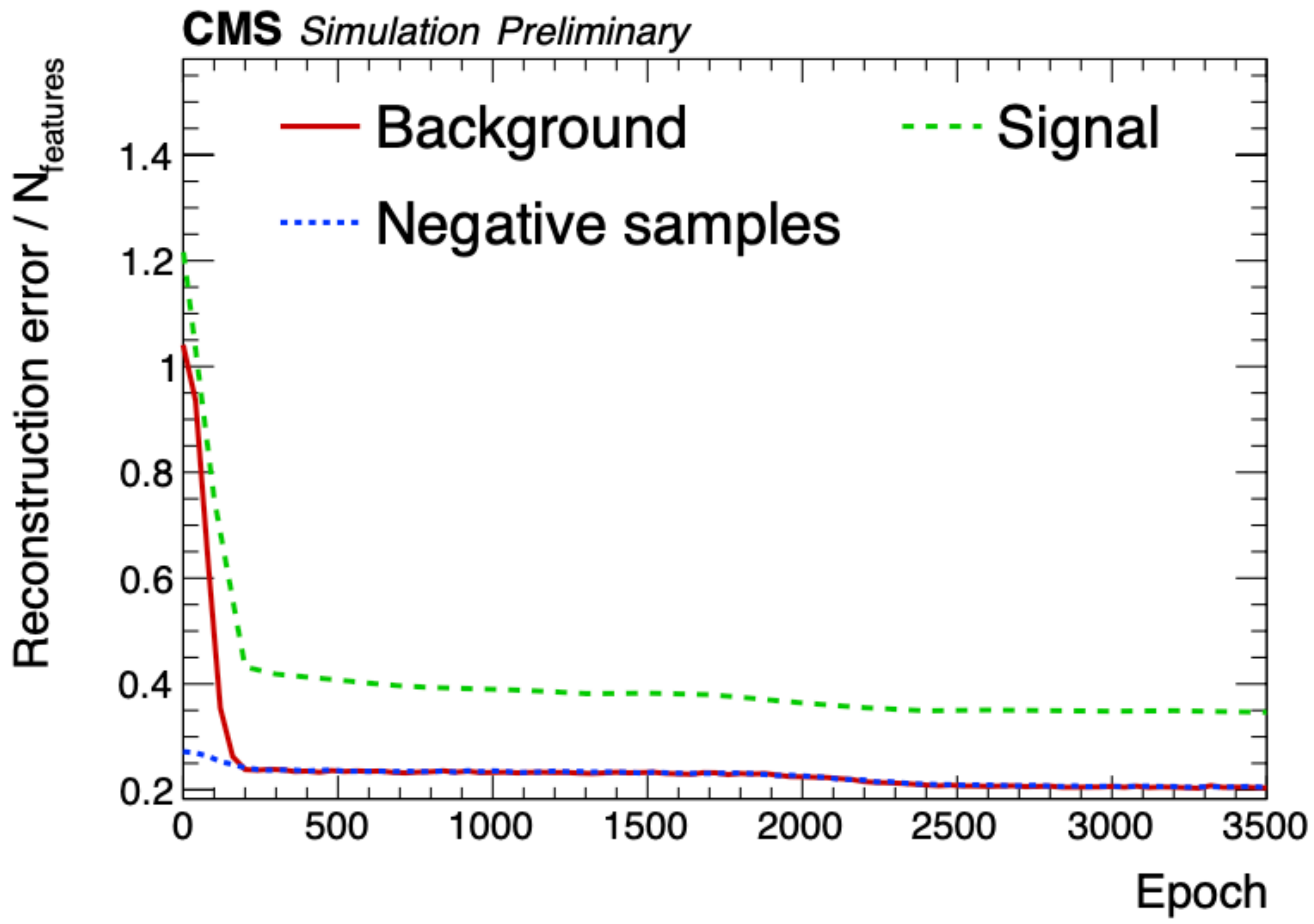
Use MCMC to sample "negative samples" x' from p_θ and compute their reco error E_

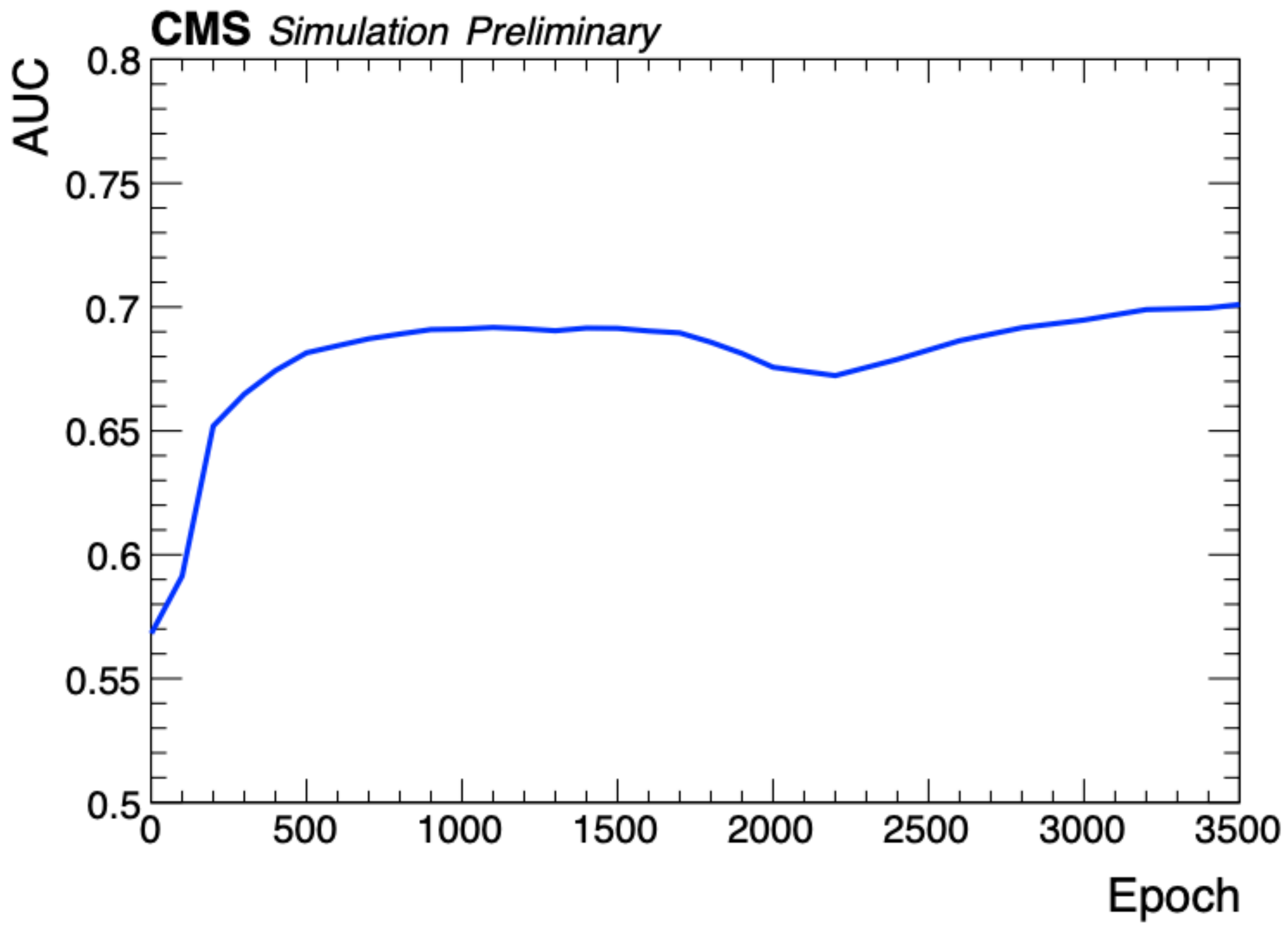


[2] NAE first introduced in arXiv:2105.05735 and used in HEP in arXiv:2206.14225

The Normalised AE

With Loss depending on $E_{-} - E_{+}$, where E_{+} is the reco. error of the training data, Signal SVJ reconstruction is efficiently suppressed





Introducing the Weisserstein distance

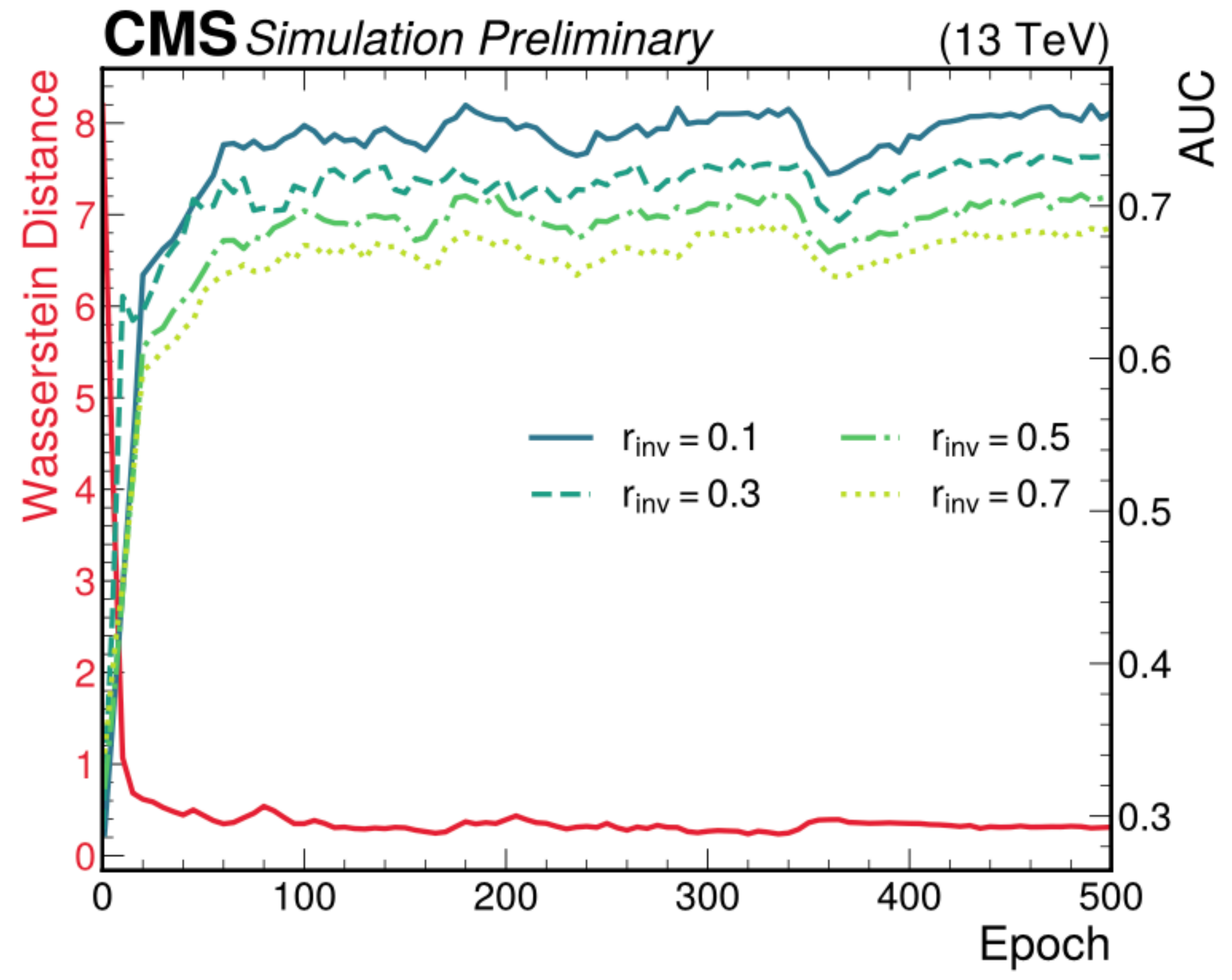
CMS Collaboration,
"Wasserstein Normalised Autoencoder"
CMS-PAS-MLG-24-002

How to determine best number of epochs in a fully agnostic way?

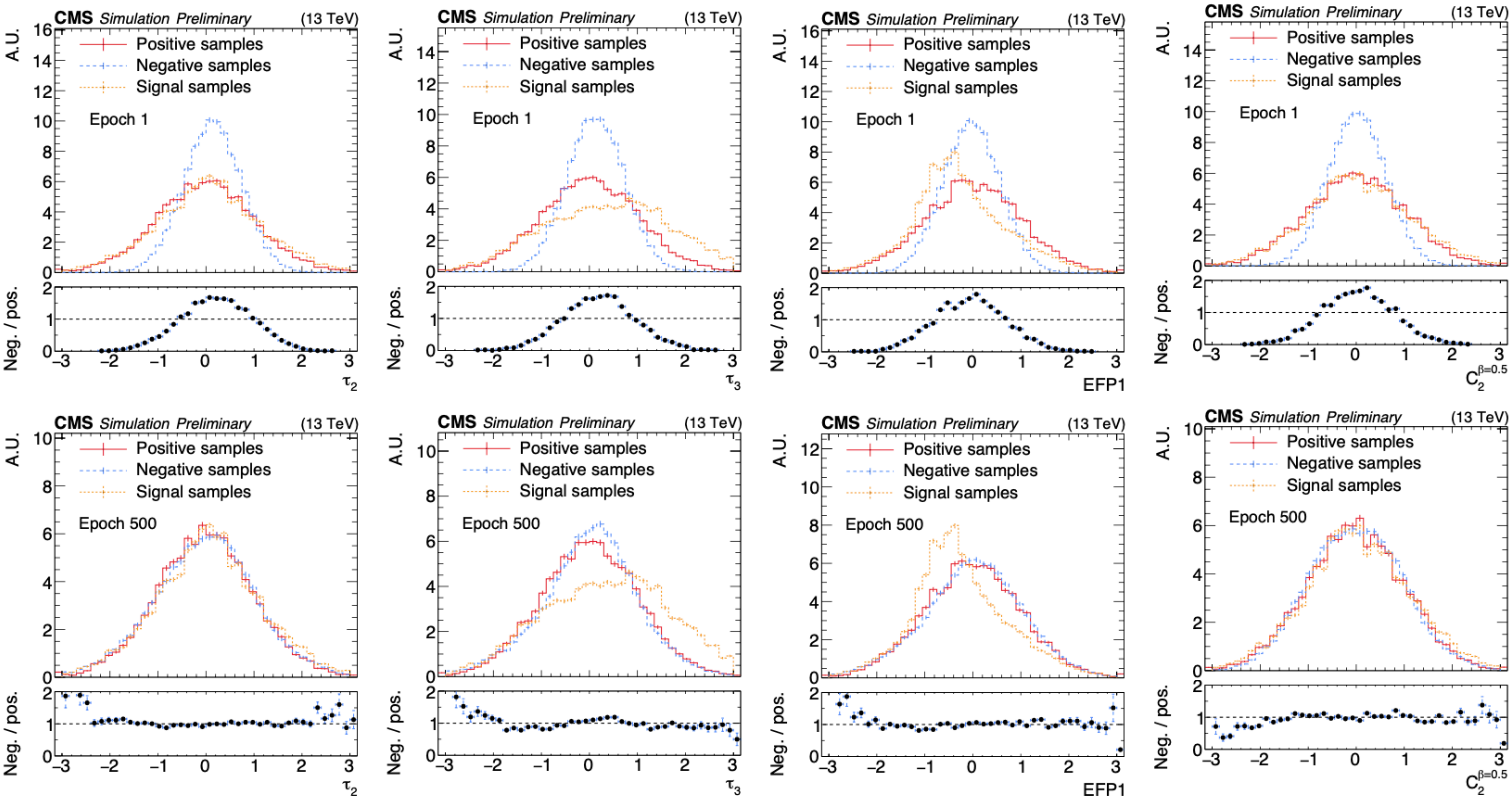
- **Wasserstein Normalised Autoencoder (WNAE)** loss function: use **Wasserstein distance** between $x \sim p_{data}$ and $x' \sim p_\theta$ (**Energy Mover's Distance**) to measure distance between background and AE probability distributions
  - Direct measure of learning $p_\theta = p_{data}$

$$L_\theta^{\mathrm{WNAE}} = \inf_{\gamma \in \Pi(p_{\mathrm{data}}, p_\theta)} \mathbb{E}_{(x,x')\sim\gamma}[\|x - x'_\theta\|]$$

AUC is maximal when W distance is minimal



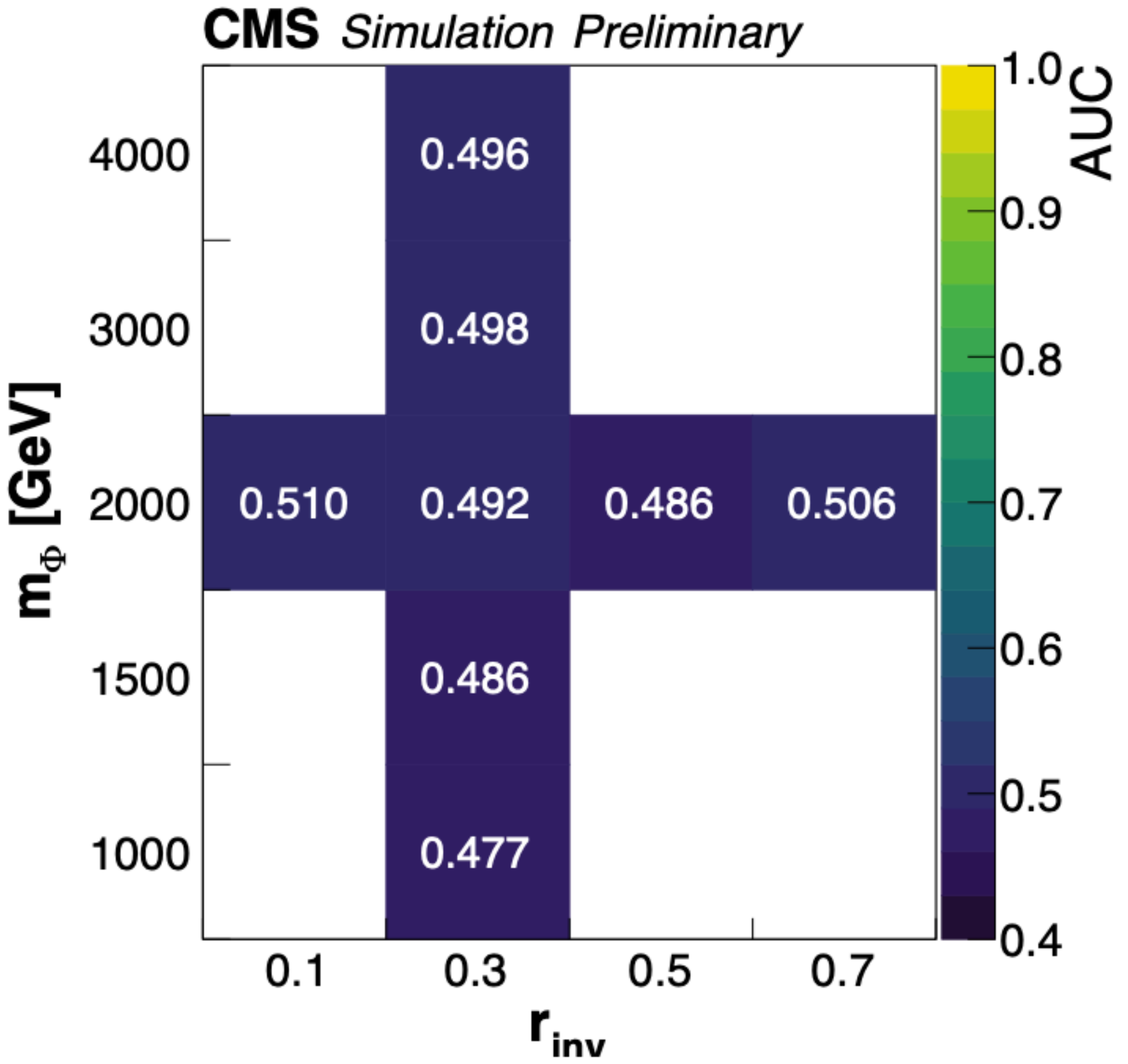
Wasserstein Normalised AE



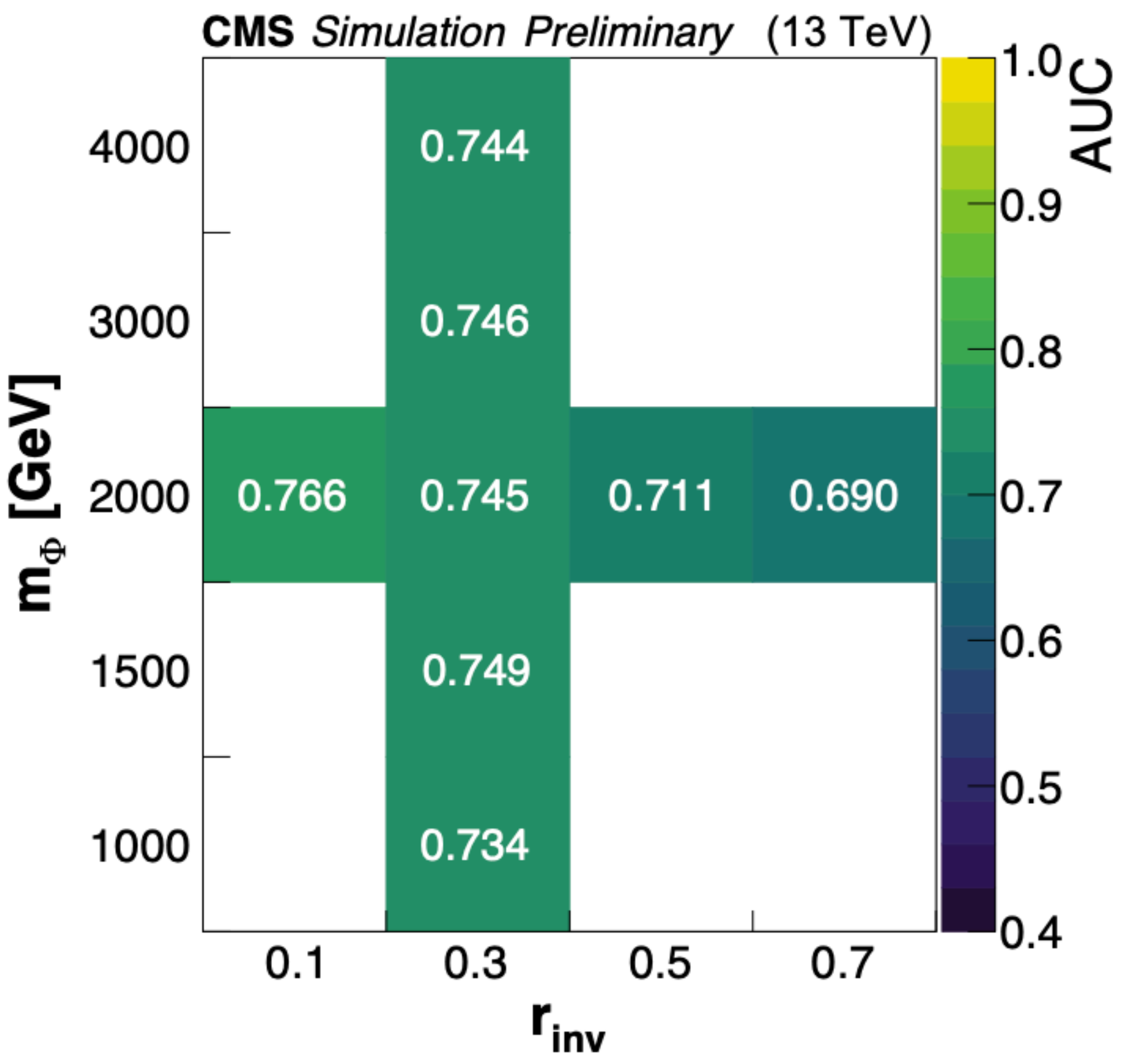
WNAE performance

The WNAE achieves sensible improvement compared to the standard AE

Standard AE



WNAE



Conclusions

- **Dark Matter might be less conventional that we think**
- Huge phenomenology still to investigate
  - **Novel models** (and signatures) to probe
  - **Novel techniques** to use
- Investigation of QCD-like DM still at the dawn
- **Model agnostic techniques might be the keystone for finding new physics**
  - Not limited to QCD-like DM
  - Multitude of applications!

Additional material

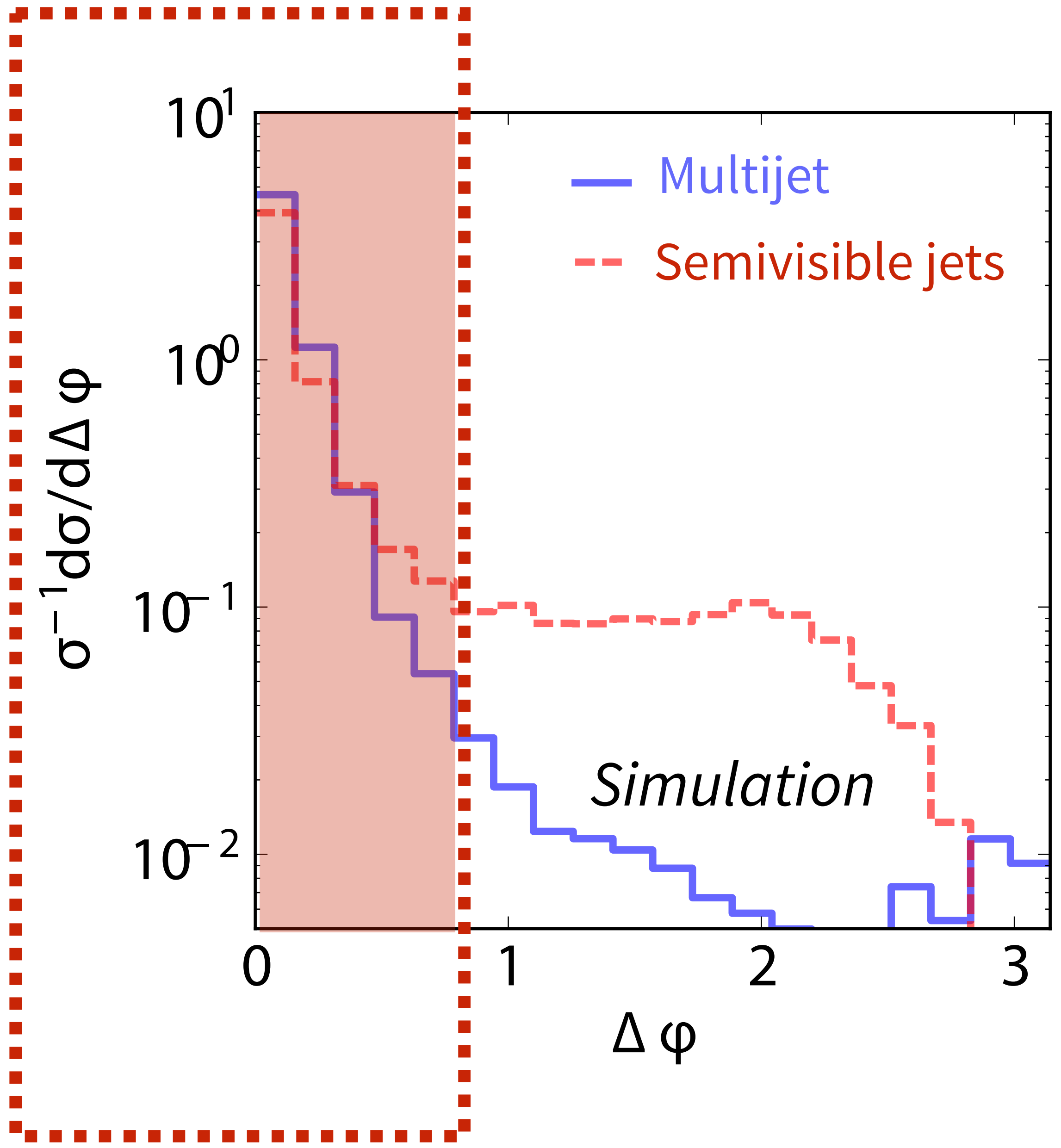
Challenges - Instrumental background

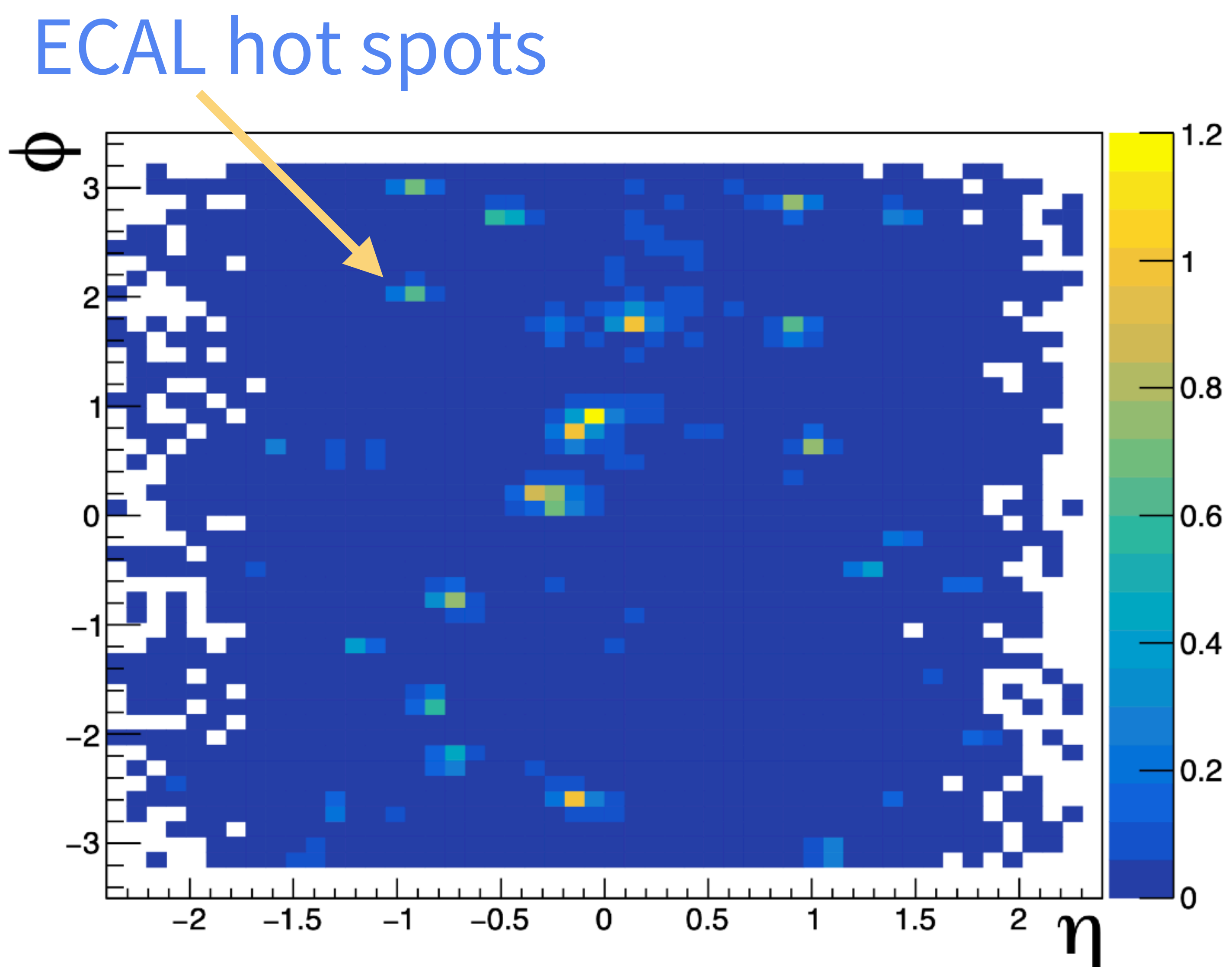
Artificial pTmiss due to instrumental backgrounds

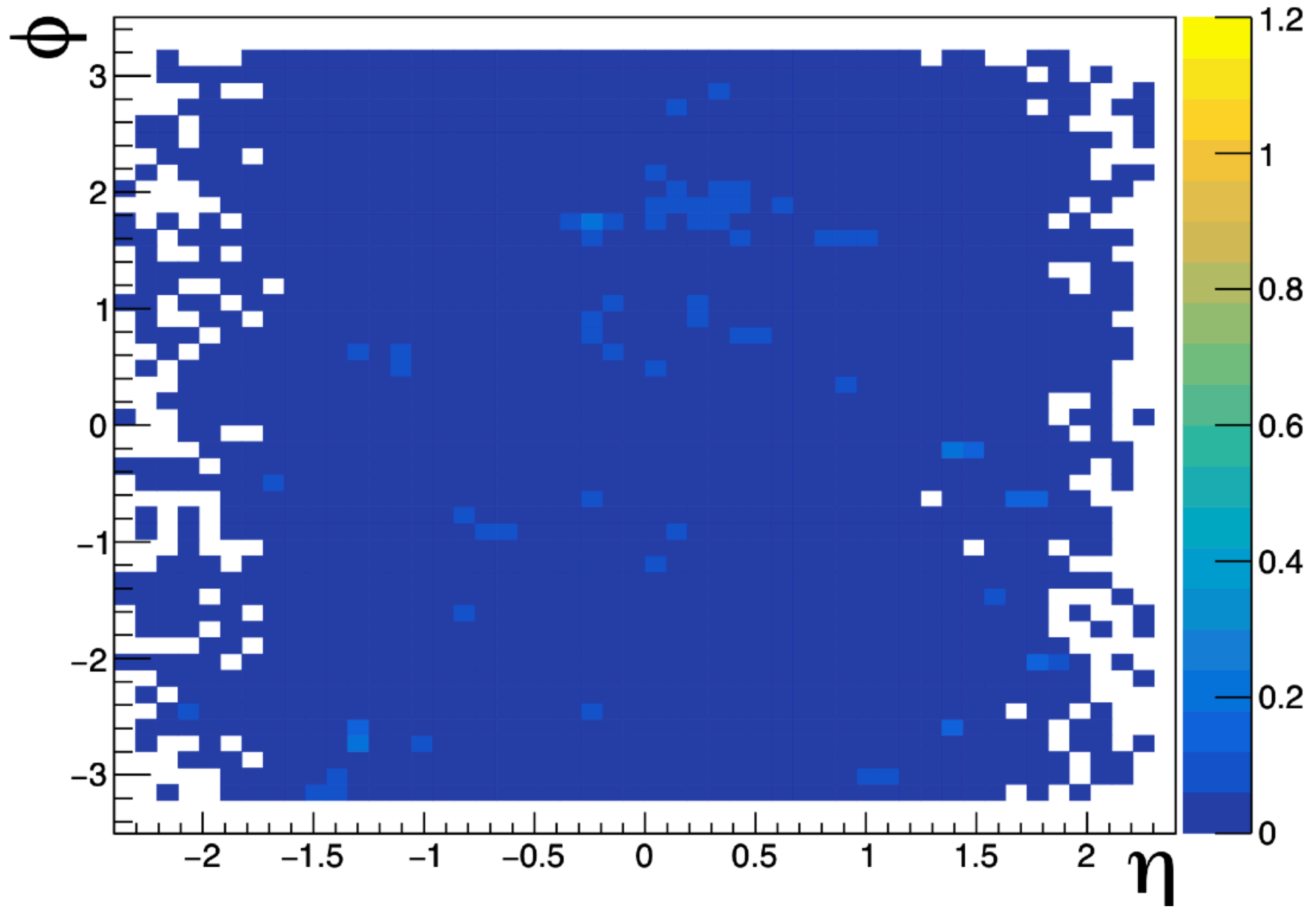
Traditional data quality filters not sufficient for SVJ search

Custom filtering algorithm developed to cope with

- Additional “cold” ECAL cells causing artificial pTmiss
- Remaining signal from direct deposits on APD (spikes) rather than in the crystals in the gap regions (|η|~1.3)







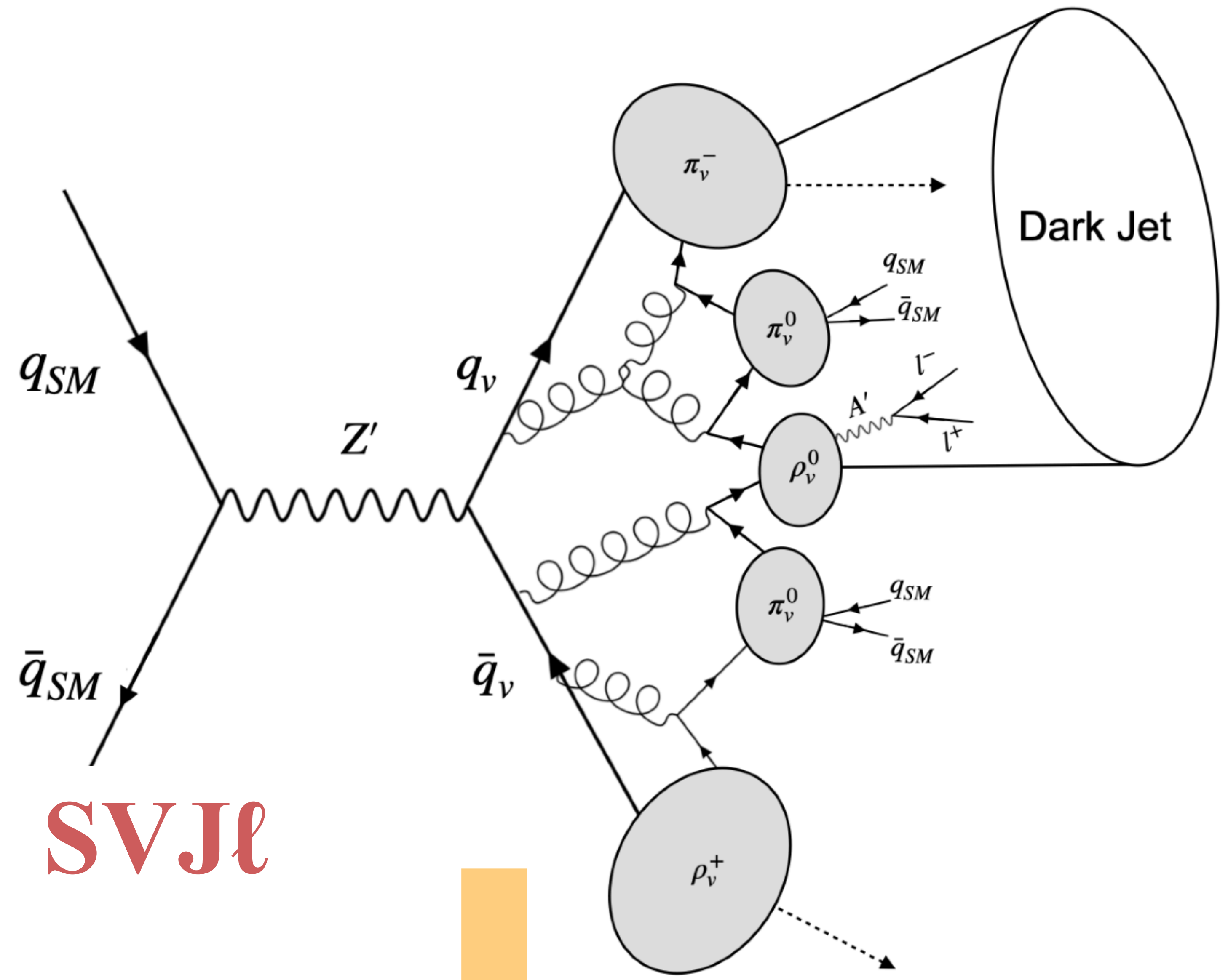
Lepton-enriched SVJs

New models including decays to leptons:

- **SVJℓ**: democratic decay to leptons allows by double messanger model (Z' plus A')

PARAMETER	DESCRIPTION	BENCHMARK
$M_{Z'}$	Z' pole mass	1.5-5 TeV
$\epsilon_{\text{eff,v}}$	effective mixing	0.03
r_{inv}	invisible fraction	0.3, 0.5, 0.7
Λ_v	dark confinement scale	5 GeV
m_{π_v}/Λ_v	pseudo-scalar mass ratio	1.6

Table 1 Signal model parameterization.



SVJℓ