Searches for Dark Matter with CMS

ICNFP2023: 12th International Conference on New Frontiers in Physics

Work supported by DOE Award: DE-SC0010103

Pedro J. Fernández Manteca

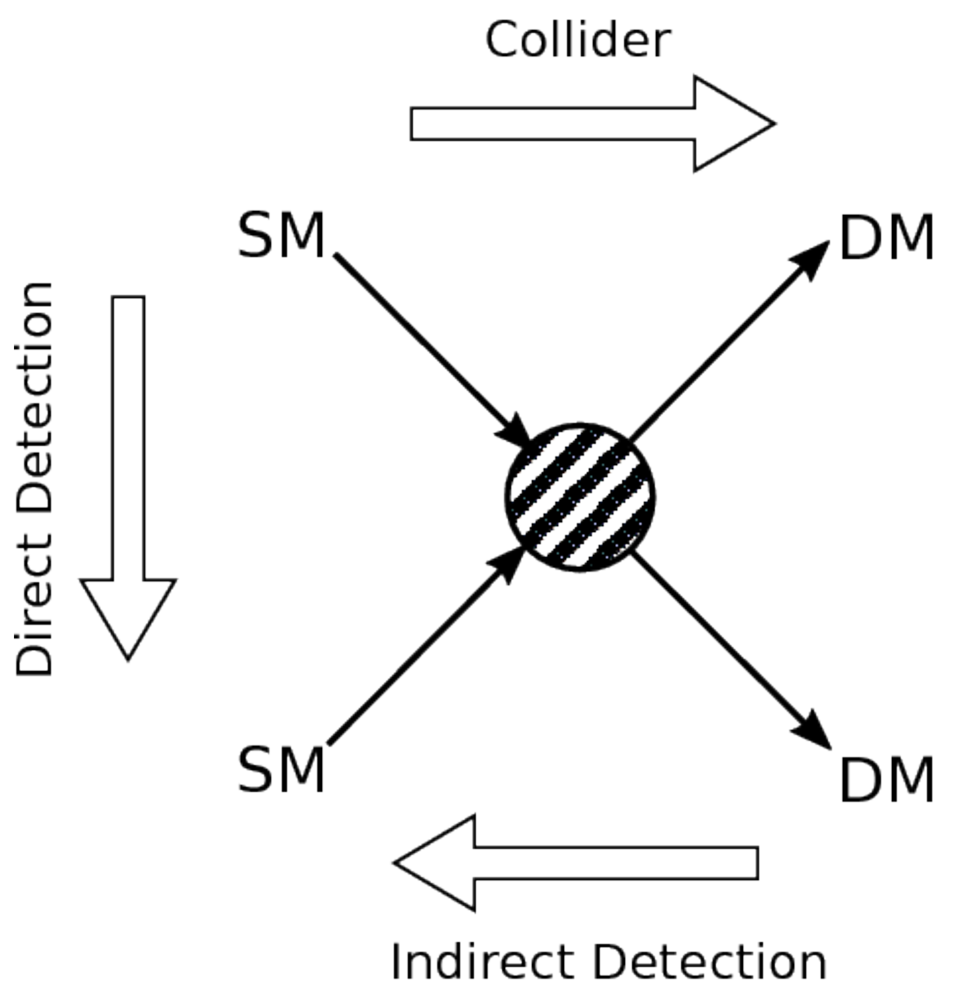
Rice University

10th - 23rd July, 2023

Probing Dark Matter

- **Evidence of Dark Matter** (DM) on several astronomical observations
- DM is expected to compose about 25% of our universe
- **Underlying assumption:** DM interacts with the Standard Model particles (SM) at least by the weak interaction

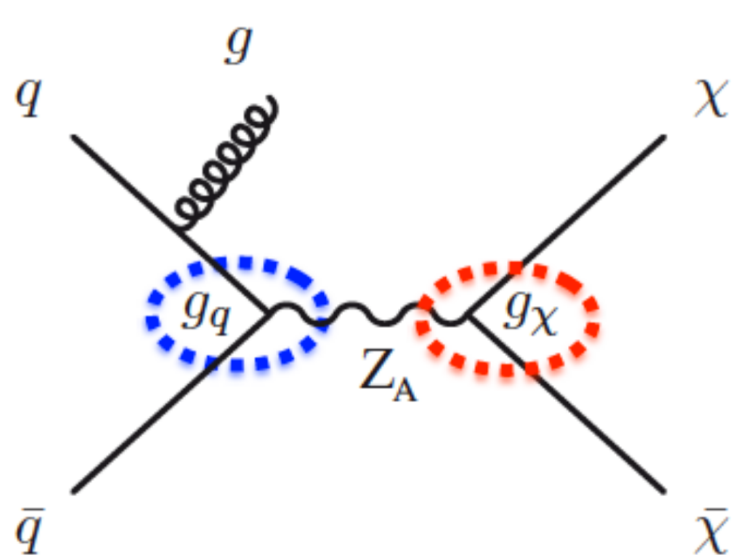
- **Direct detection**: scattering of DM particles on nuclei
- **Indirect detection**: annihilation products out of DM
- **Collider search**: produce DM through collision of SM particles



Hunting Dark Matter at the LHC

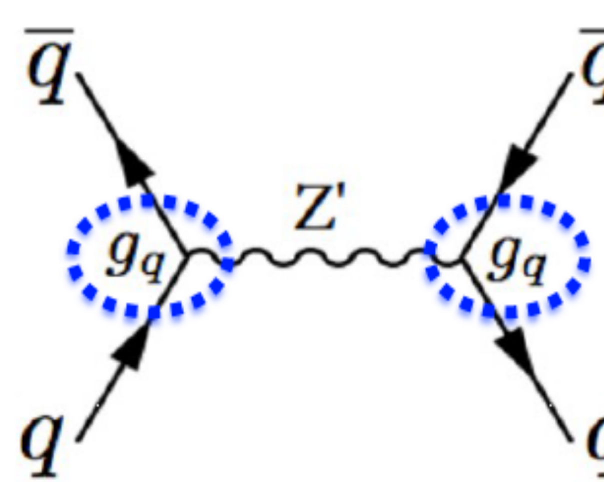
- The **LHC** is currently **the best machine** to produce high energy physics both for its energy and for its luminosity → Possibility of study particles up to high masses with low production cross sections
- **Classical searches:** Weakly Interacting Massive Particle (**WIMP**). It is not expected to interact in the detectors → Missing Energy
- Simplified DM models: Explicit definition of the mediator and couplings
  - Free parameters: masses, spins, coupling structure and strength

mono-X searches



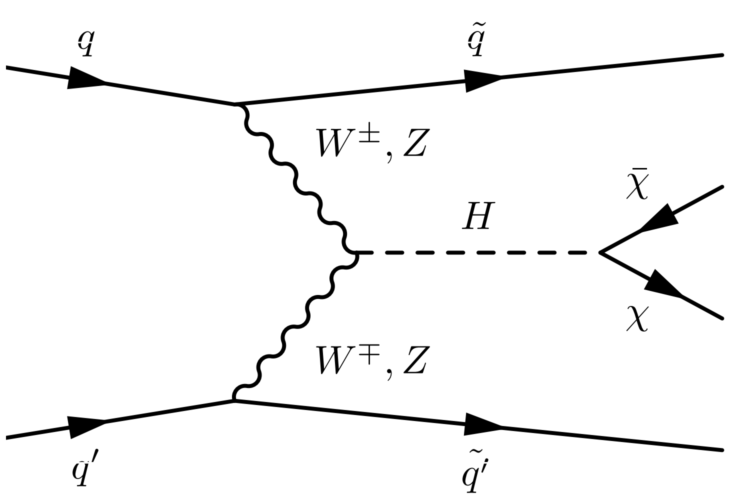
- DM particles are produced together with Standard Model particles
- Look for an energetic SM particle recoiling against the invisible DM system (p_T^{miss} + X signature)

Mediator searches



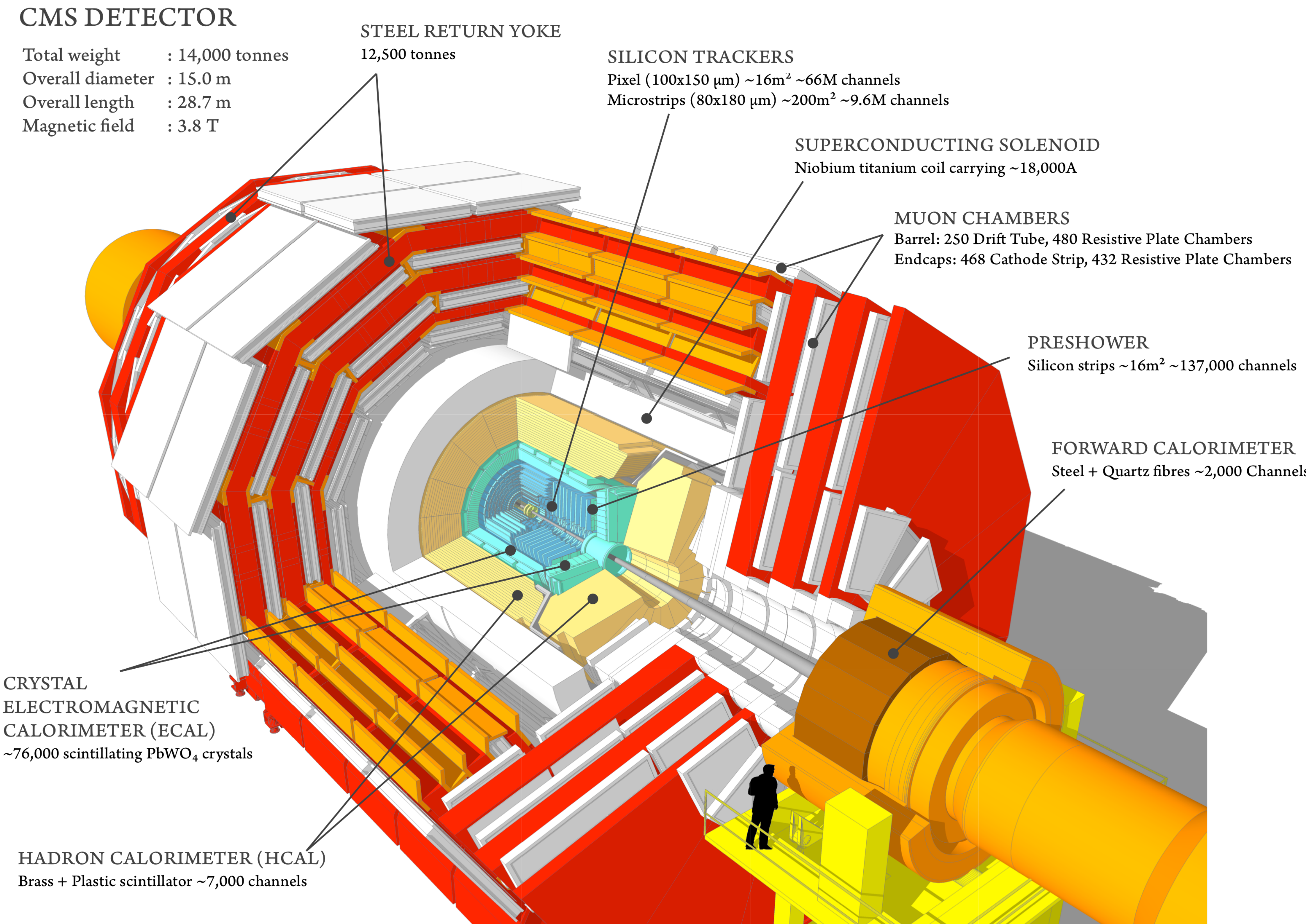
- DM mediators are produced and decay to pair of SM particles, typically quarks
- Search for bumps in the m_{jj} spectrum

Higgs portal



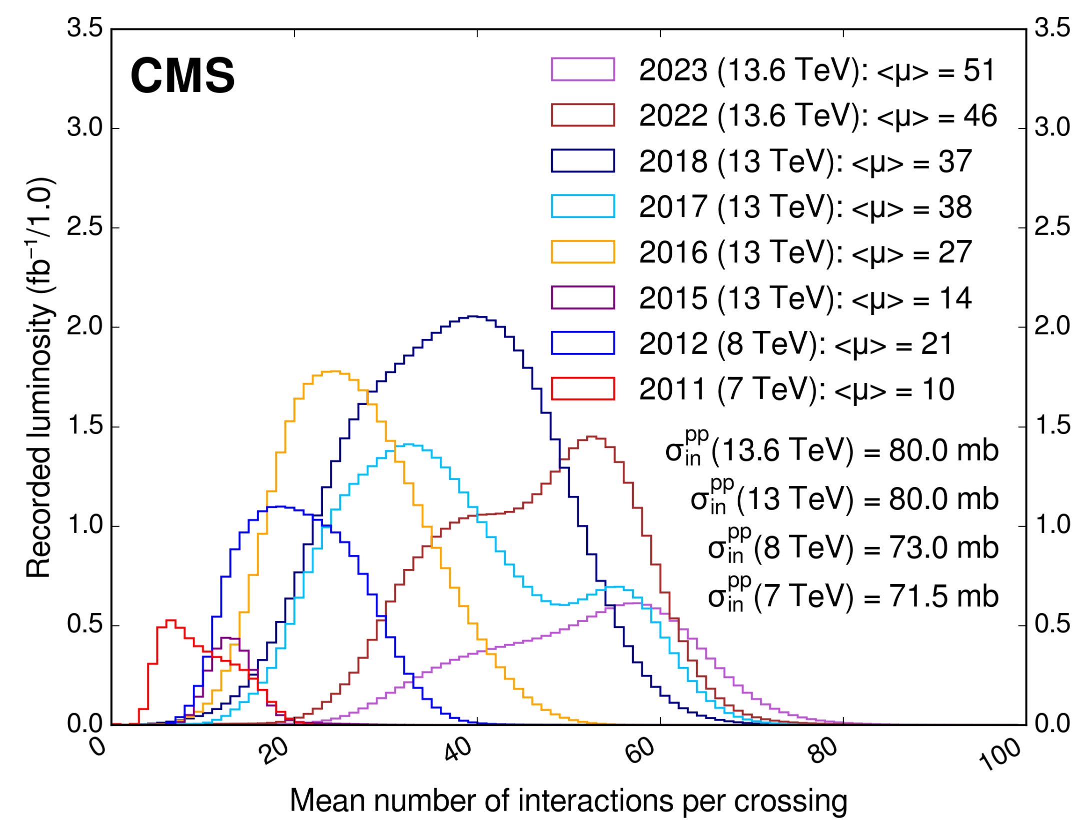
- DM production through the Higgs portal
- Higgs Boson can decay into DM particles

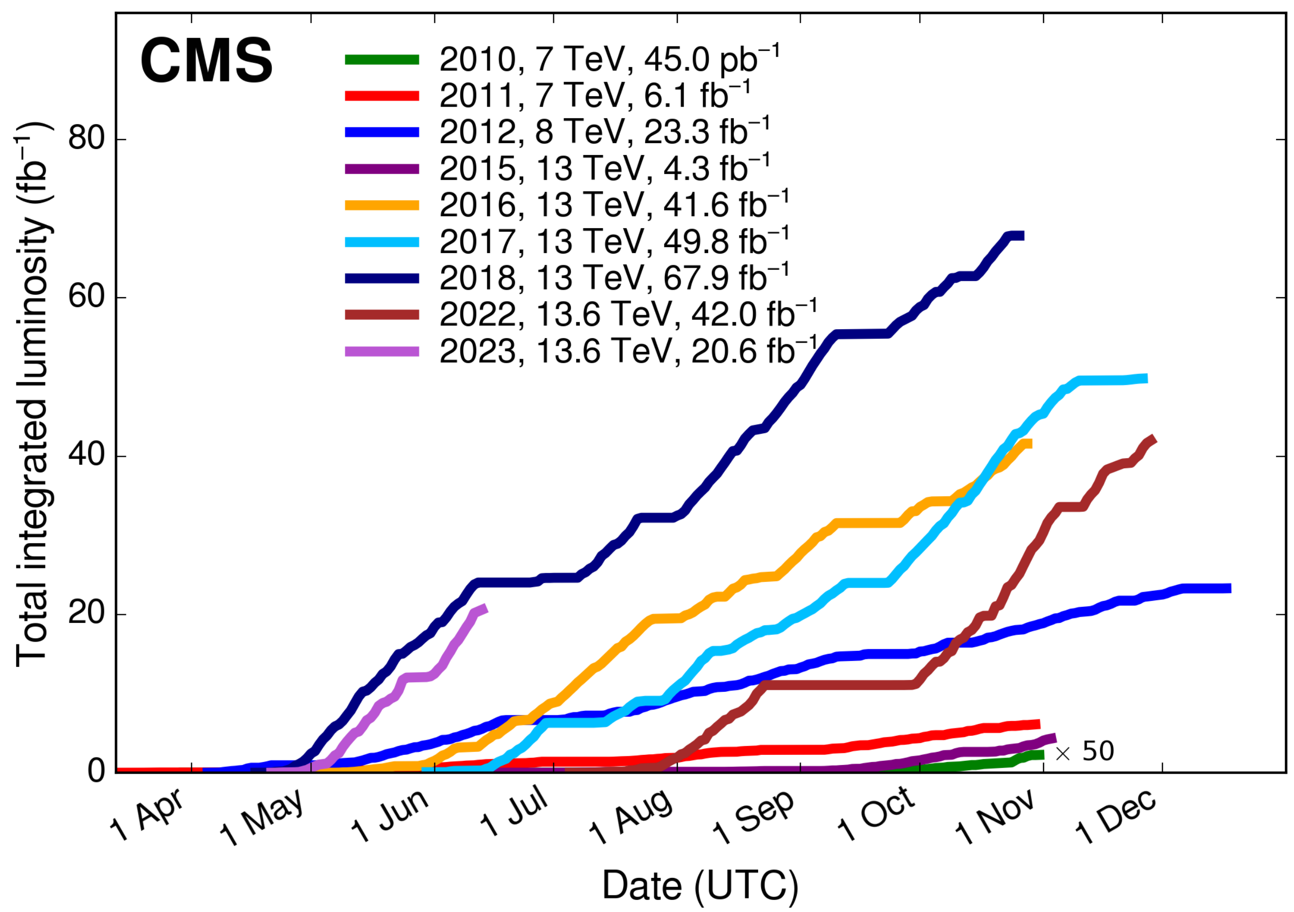
The CMS experiment



CMS data & Triggers

- **CMS collected 138 fb-1 of good data at 13 TeV (Run 2)**
- Already collected about 60 fb-1 at 13.6 TeV (Run 3)
    - Larger datasets, but with larger pileup. p_T^{miss} and triggers very sensitive to this
- How do we collect the interesting events? **Triggers**
    - Most DM searches rely on p_T^{miss} triggers

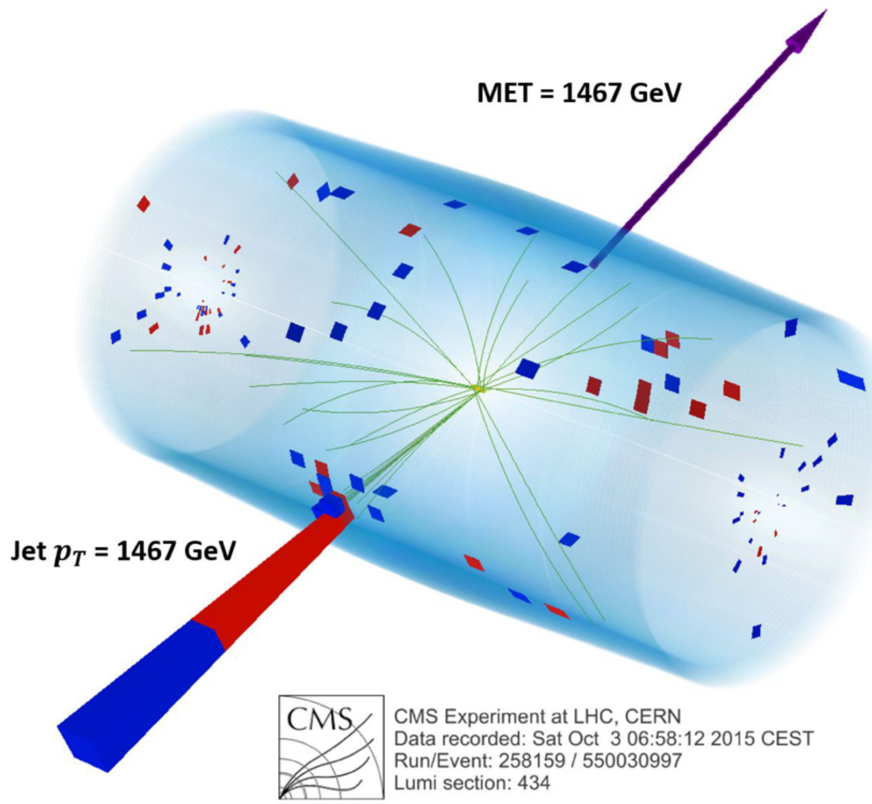




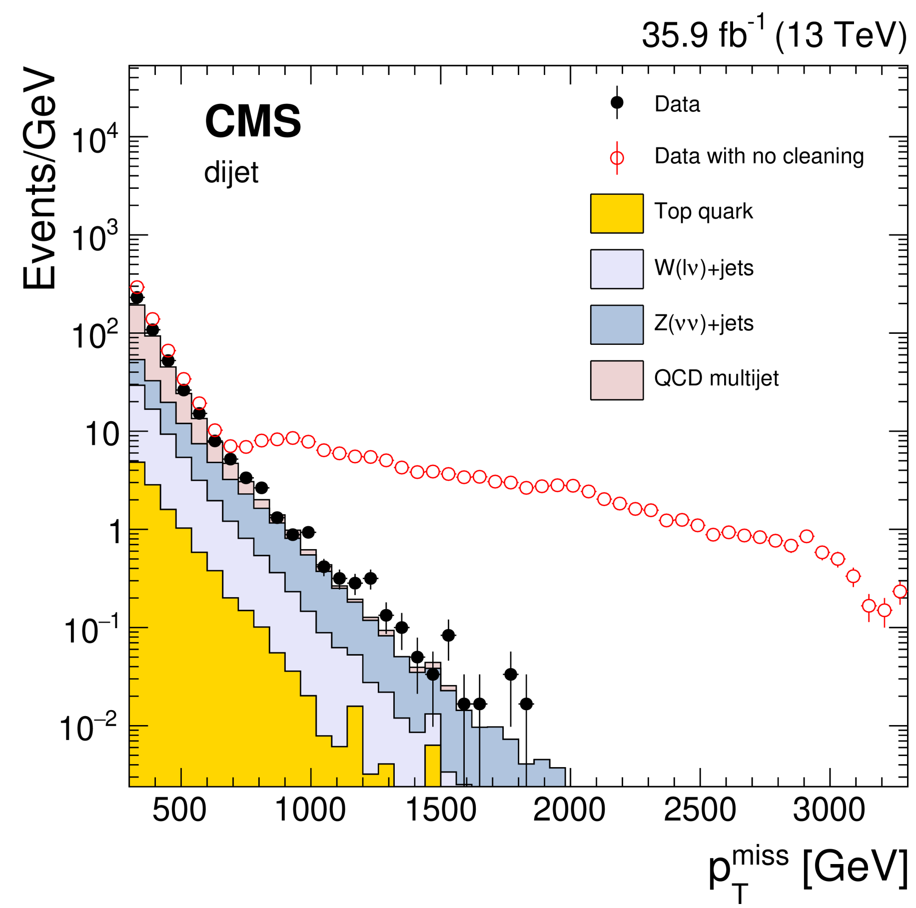
Most searches have exploited the Run 2 (2016 + 2017 + 2018) dataset for now

The Key variable: p_T^{miss} (MET)

- The sum of all final-state particles transverse momentum is expected to be null
- Since DM particles are not expected to interact with the detector → $p_T^{miss} \neq 0$
- Most analyses rely on large p_T^{miss} balancing against "visible" (X) objects (jets, boosted-jets, b-jets, photons, charged leptons…)



$$\Delta\phi\left(p_T^{miss}, X\right) \sim \pi$$
$$p_T^{miss} \sim p_T^X$$



- Events with non-null p_T^{miss} can arise from limited detector resolution, presence of neutrinos in final state, non-collisional background events…
    - **Understanding p_T^{miss} is a critical task**

Main Dark Matter analysis in CMS

The goal is to search for DM in all possible / feasible final states

- **Bunch of classical searches** during Run 2 presented in several conferences since 2017. No significant discrepancies with SM predictions observed → results interpreted as limits on DM mediators, simplified model parameters, and on DM-nuclei interaction cross sections

p_T^{miss} + X

- mono-X:
  - Mono-jet & Mono-V (CMS-EXO-20-004)
  - Mono-photon (CMS-EXO-16-053)
  - Mono-Z(→ℓℓ) (CMS-EXO-19-003)
  - Mono-top (CMS-EXO-16-051)
  - Mono-Higgs (CMS-EXO-18-011)
- DM + t / tt (CMS-EXO-18-010)

Mediator searches

- Di-jet X → jj bump hunting (CMS-EXO-16-056)
- Di-jet light resonances (CMS-EXO-17-001)
- Angular Di-jet (CMS-EXO-16-046)

Higgs Portal

- Higgs → invisible (CMS-HIG-20-003)

Dark Photon

- Dark Photons in VBF Higgs (CMS-EXO-20-005)
- in Z(ℓℓ) + HY + MET (CMS-EXO-19-017)
- Resonance decaying to two muons (CMS-EXO-19-018)

- **Covered in this talk:** latest CMS results on Run 2 data (became public within the last year or so) involving more exotic signatures and / or novel interpretations:
  - Search for inelastic Dark Matter with long-lived particles (CMS-EXO-20-010)
  - Search for semi-visible jets (CMS-EXO-19-020)
  - Search for dark Higgs (CMS-EXO-21-012)

Search for inelastic Dark Matter

arXiv:2305.11649

- **Two dark matter states χ_1 and χ_2 with near masses**
- Vector current couples χ_1 and χ_2 as the Majorana mass term $\delta/m << 1$ (Smith & Weiner, PRD 64 (2001) 043502)

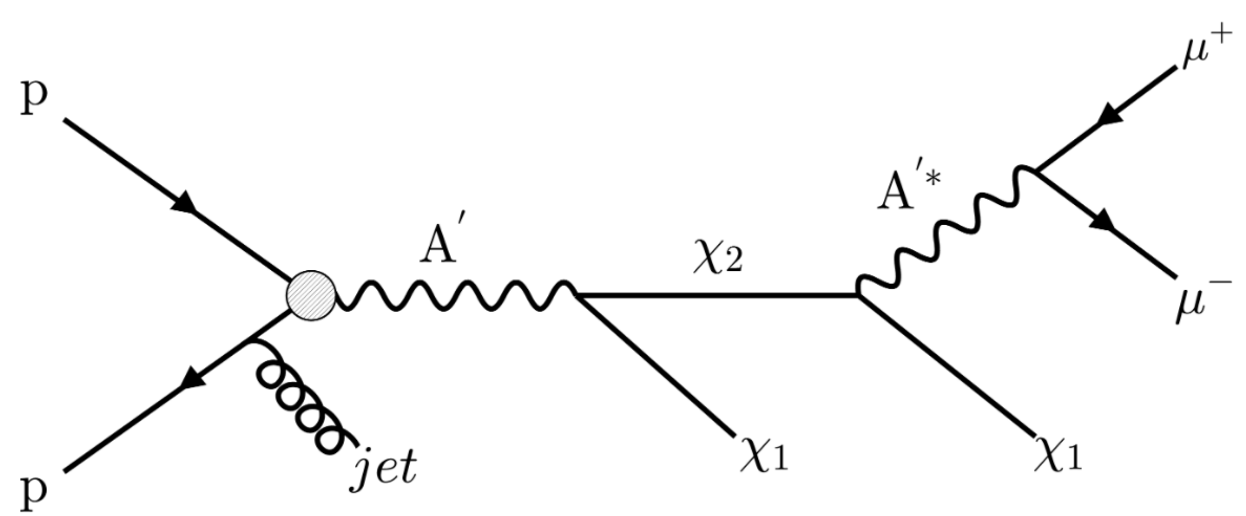
inelastic — elastic

$$\overline{\psi}\gamma_\mu\psi \simeq i(\overline{\chi}_1\overline{\sigma}_\mu\chi_2 - \overline{\chi}_2\overline{\sigma}_\mu\chi_1) + \frac{\delta}{2m}(\overline{\chi}_2\overline{\sigma}_\mu\chi_2 - \overline{\chi}_1\overline{\sigma}_\mu\chi_1)$$

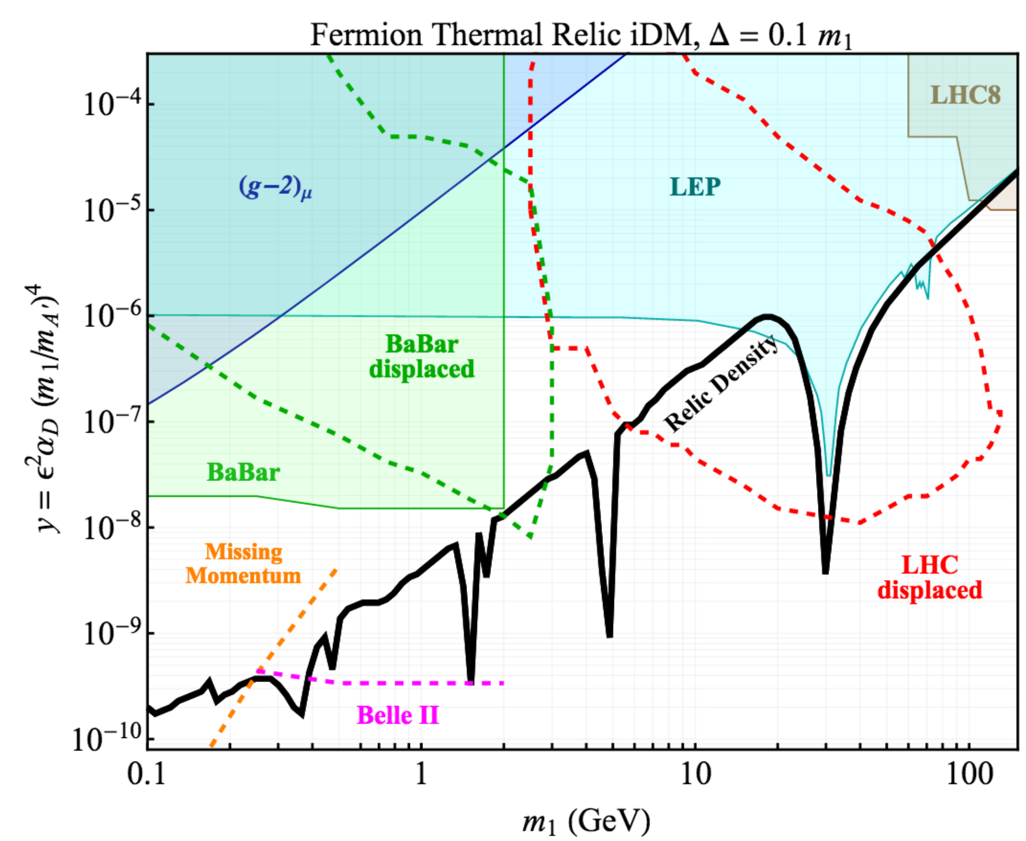
Model parameters:

m1 = [3-80] GeV
m2 - m1 = δ = {0.1, 0.4} * m1
lifetime χ2= [1-1000] mm
mA' = 3 * m1
coupling strength of dark QCD α_D = {0.1, α_{EM}}
fixed kinematic mixing ε between ϒ, Z and A'

- Dominant ineslastic coupling → χ_1 and χ_2 can be produced simultaneously at the LHC via dark photon

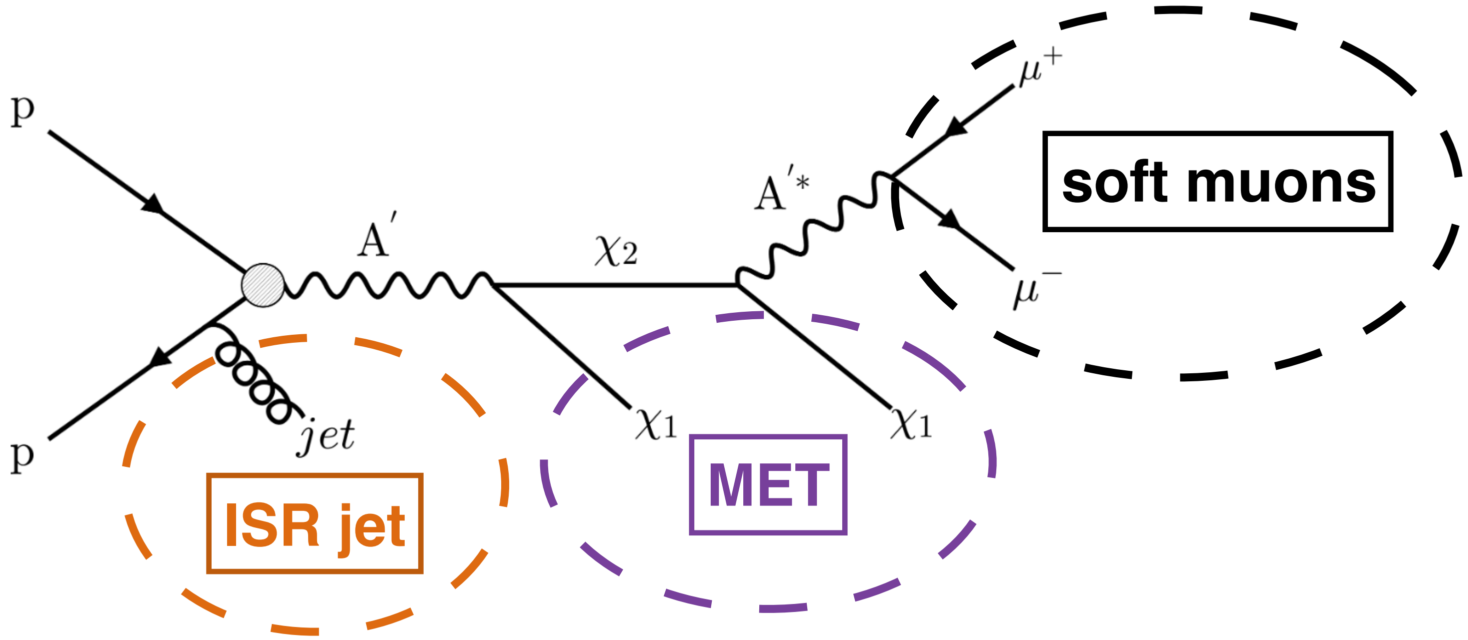


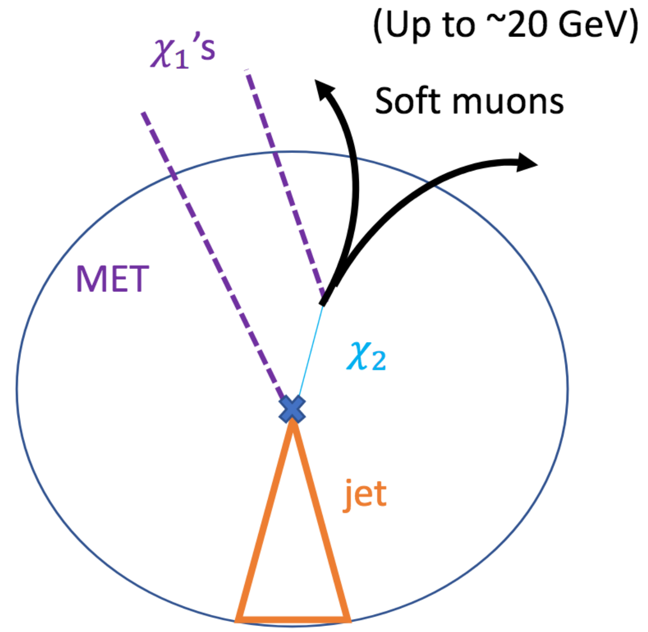
Proj studies: Izaguirre, Krnjaic & Shuve, PRD 93 (2015) 063523



- **Compatible with observed relic DM abundance from cosmological measurements**

Events topology

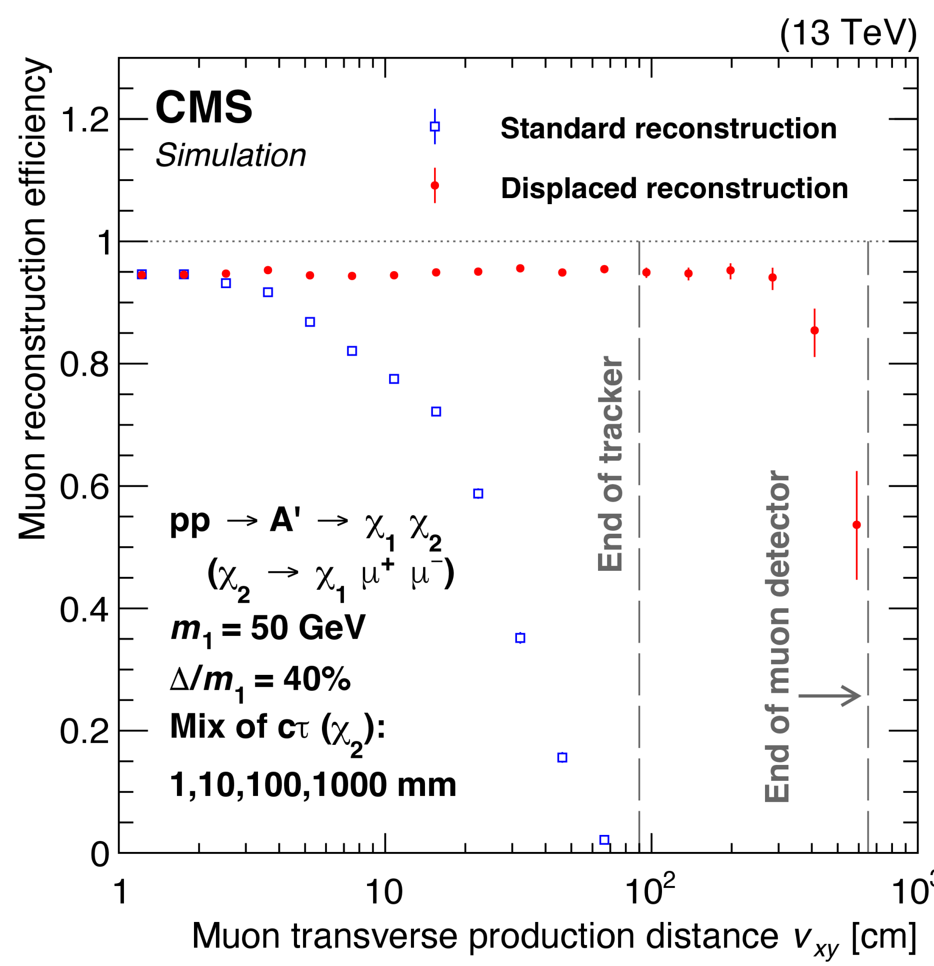




- Soft decay muons due to small mass difference δ between χ_1 and χ_2
- Decay width of χ_2 proportional to δ^5 → **long-lived muons**

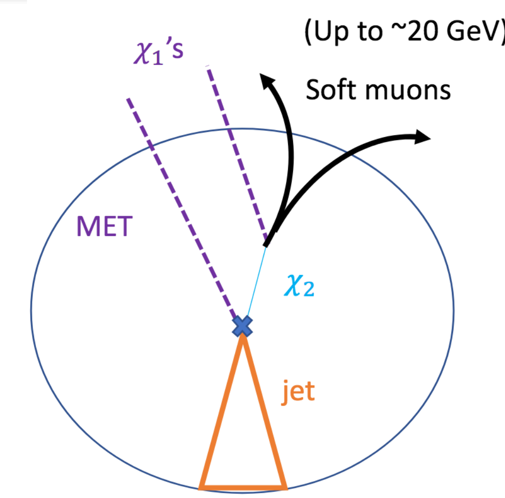
- **Displaced muon reconstruction:**
  - Used hits from muon detectors only
  - No beam constraints
  - Worse momentum & position resolution but much better reco efficiency for large displacements

CMS transverse cross-section

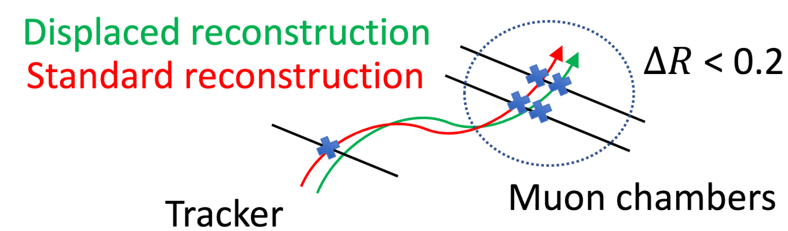


Analysis strategy

- **Event selection:**
  - Two displaced muons, MET > 200 GeV, 1 or 2 reconstructed jets, Δϕ(MET, leading jet) > 1.5, Δϕ(MET, sub-leading jet) > 1.5, no b-tagged jets, ΔR(μ+, μ-) < 0.9, Δϕ(MET, μ+μ-) < 0.5



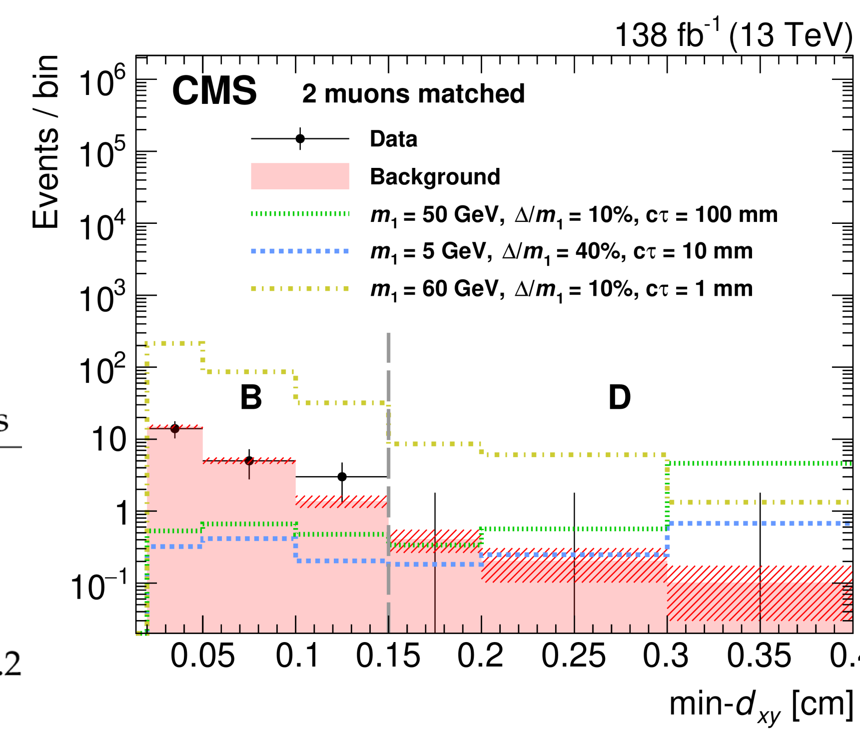
- Analysis categorization based on geometrical matching with standard muons (0-match, 1-match, 2-math)
  - **Maximize the resolution & signal sensitivity on 0-match**



$$N_{bkg}^{A} / N_{bkg}^{B} = N_{bkg}^{C} / N_{bkg}^{D}$$

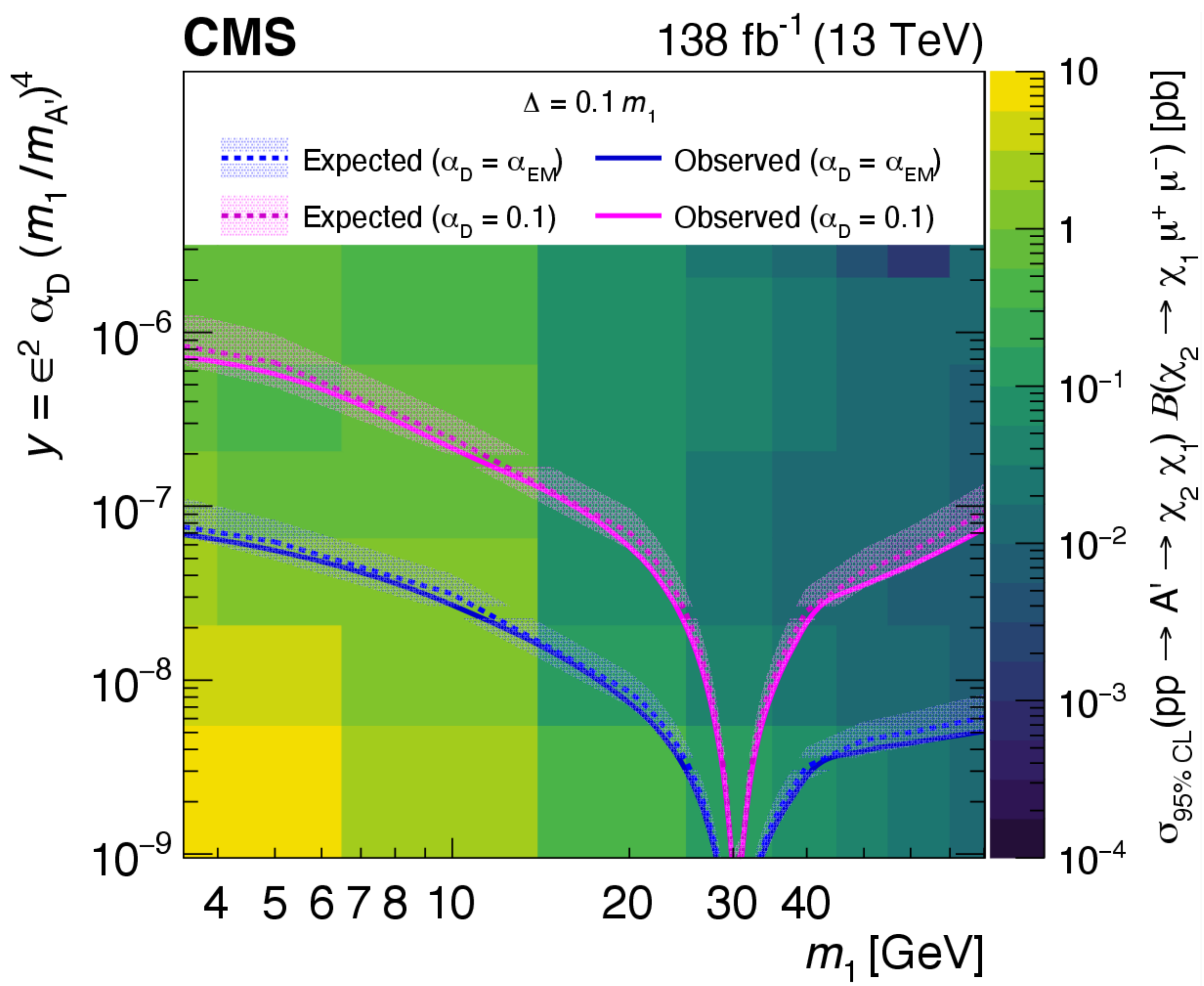
- **Background estimation:** ABCD data-driven approach
  - For 1-match & 2-match: v1 = muon displacement, v2 = muon isolation
  - For 0-match: v1 = muon displacement, v2 = Δϕ(MET, μ+μ-)

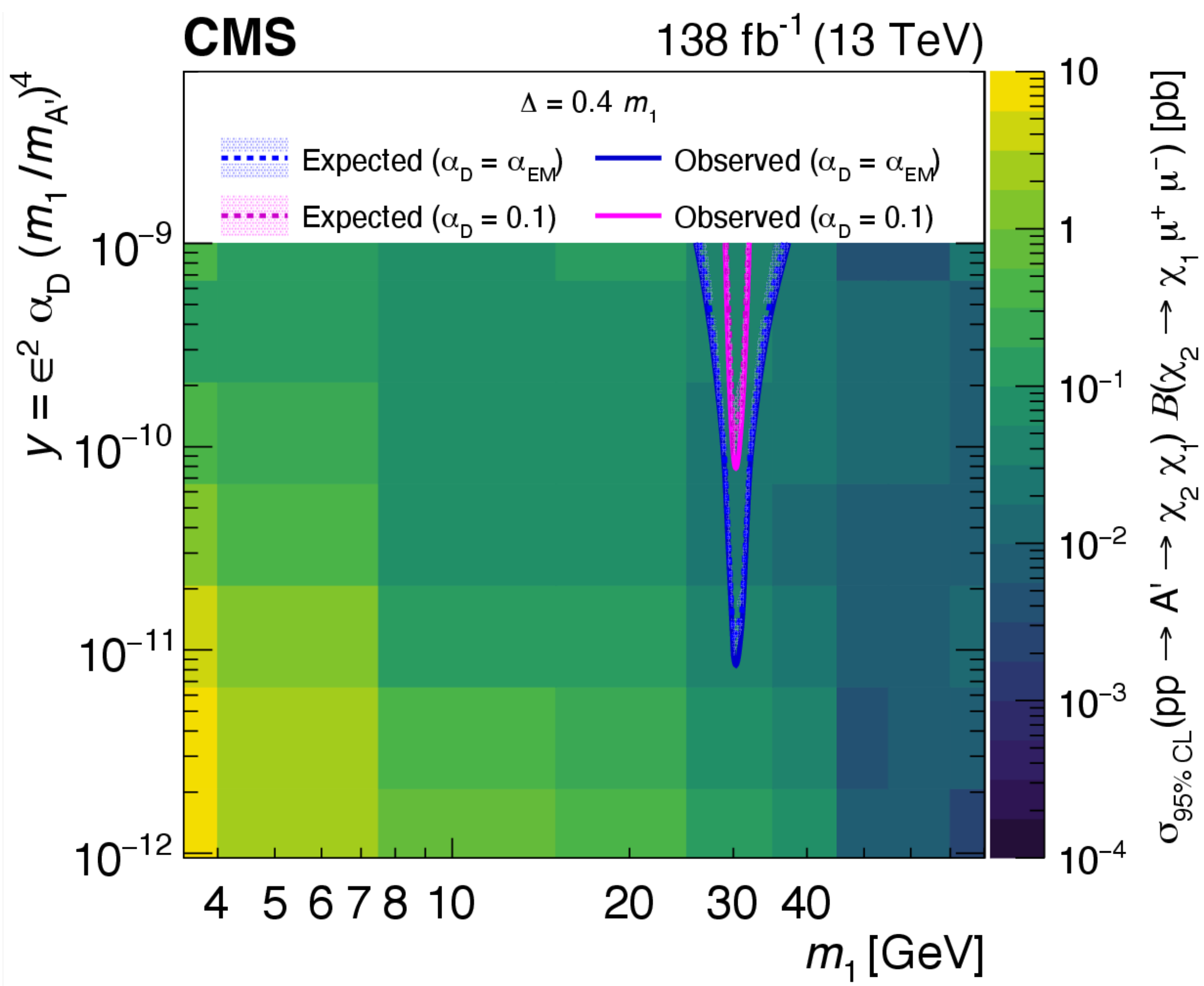
Bin	0-match			1-match			2-match		
	$\Delta\phi^{miss}_{\mu\mu}$ [rad]	min-d_{xy} [cm]	Events	I^{rel}_{PF}	min-d_{xy} [cm]	Events	I^{rel}_{PF}	min-d_{xy} [cm]	Events
Obs. A	0–0.25	3–15	68	>0.25	0.02–0.75	716	>0.25	0.02–0.15	424
Obs. B	0.25–0.50	3–15	9	<0.25	0.02–0.75	33	<0.25	0.02–0.15	22
Obs. C	0–0.25	>15	9	>0.25	>0.75	12	>0.25	>0.15	10
Obs. D	0.25–0.50	>15	2	<0.25	>0.75	0	<0.25	>0.15	0
Pred. D			1.2 ± 0.6			0.5 ± 0.2			0.5 ± 0.2



Results

- **First attempt to search for inelastic DM at the LHC**
- Upper limits on σ(pp → A' → $\chi_2\chi_1$) B($\chi_2 \rightarrow \chi_1\mu^+\mu^-$)
  - Tighter constraint around 30 GeV as mA' = 3*mχ_1 ~ m_Z (larger xsec from mixing with Z boson)
  - $\alpha_D \propto 1/\varepsilon^2$ → smaller xsec for α_D = 0.1

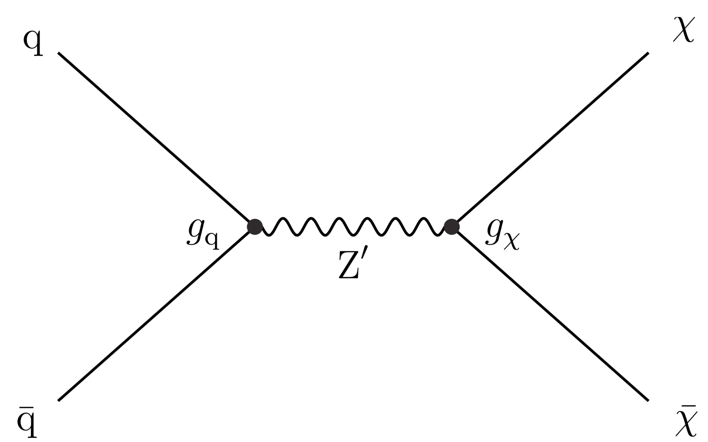




Search for Semi-visible jets

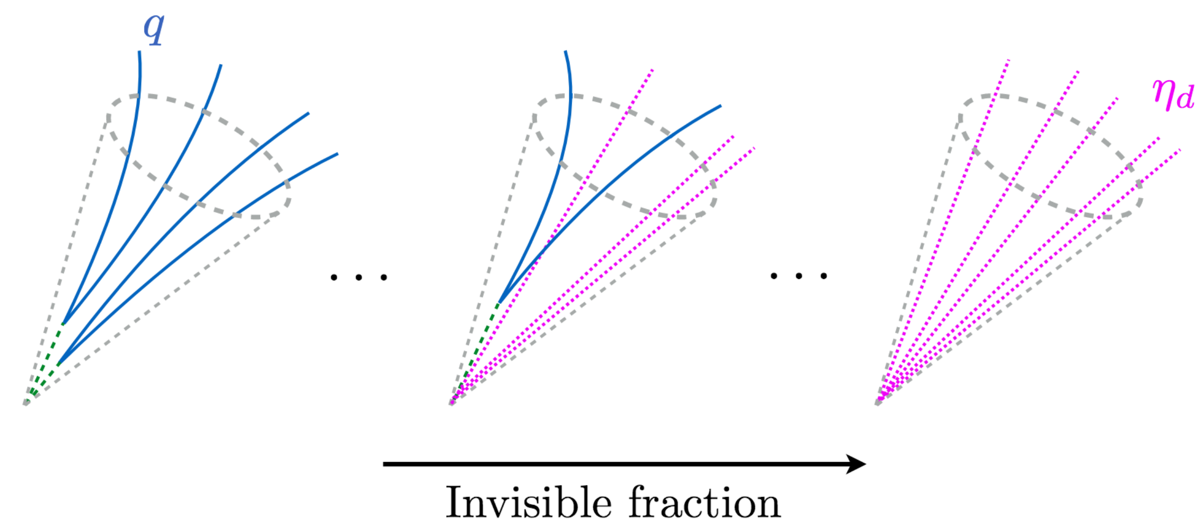
JHEP06(2022)156

- Z' boson mediator with couplings to the SM quarks (g_q) and dark quarks (g_χ)
- Several flavors of quarks (χ_1, χ_2...) contained by dark sector which can yield to stable or unestable dark hadrons:
    - Unstable hadrons → decay to SM particles
    - Stable hadrons (DM candidates) → no interaction with the detector
- **Collimated mixture of visible and invisible particles** so called "semi-visible jets". Wider and more energetic than SM jets
- **Jets aligned with MET** (commonly rejected by other DM searches). Overall MET will cancel as Z' is produced at rest
    - Back-to-back semi-visible jets



Model parameters:

mZ' = [1.5-5.1] TeV
md = [1-100] GeV
rinv = [0-1]
mX = md/2
gq = 0.25, gχ = 0.5
α_D = fixed by md

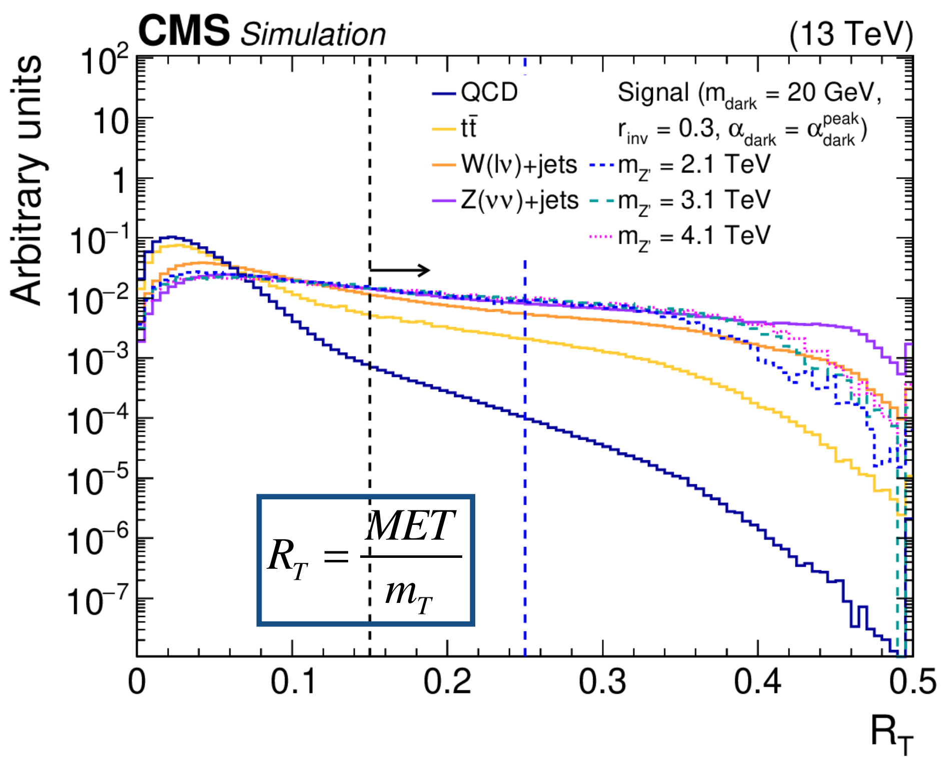


$$r_{\text{Inv}} = \frac{N_{\text{stable}}}{N_{\text{stable}} + N_{\text{unstable}}}$$

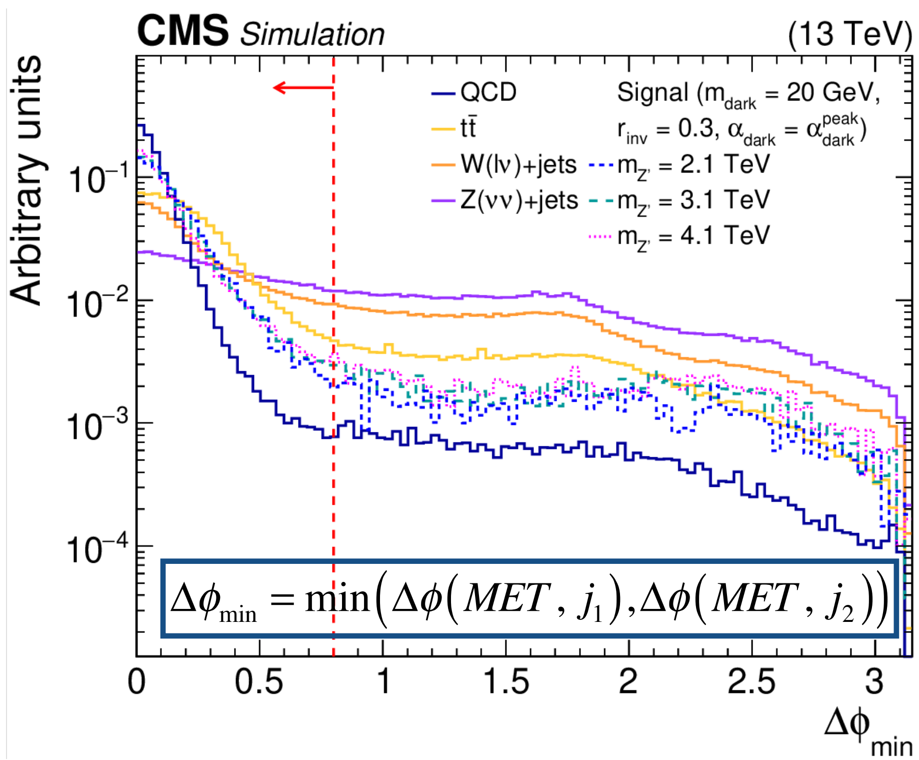
- r_{Inv} ~ 1 → covered by classical searches (high MET along with ISR radiation)
- r_{Inv} ~ 0 → covered by di-jet searches (two jets and low MET)

Analysis strategy

- **Event selection:**
  - Two central wide jets with pT > 200 GeV (signal topology), dijet mT > 1.5 TeV, Δη(j1, j2) < 1.5 (trigger)
  - Lepton veto (reject t$\bar{t}$, W+jets)
  - $R_T > 0.15$ (reject QCD), $\Delta\phi_{min} < 0.8$ (reject t$\bar{t}$, W+jets, Z(νν)+jets)



Decorrelate MET and mT



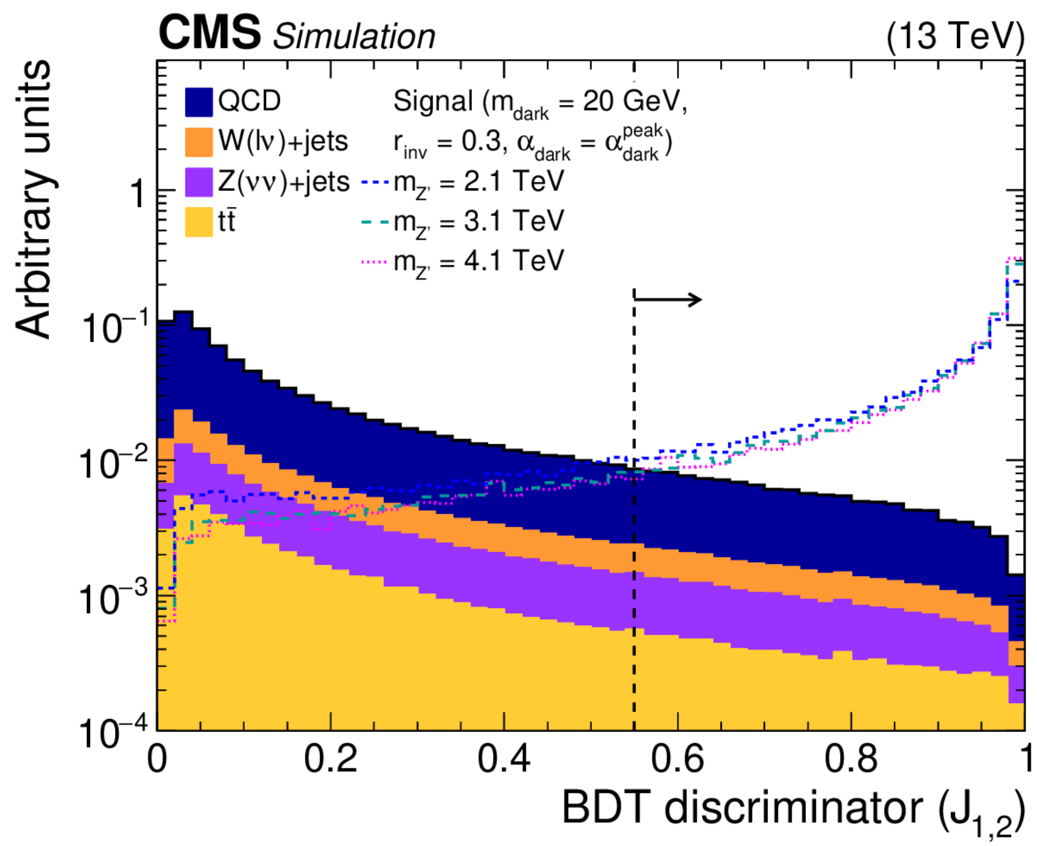
Select for events with MET aligned to a jet

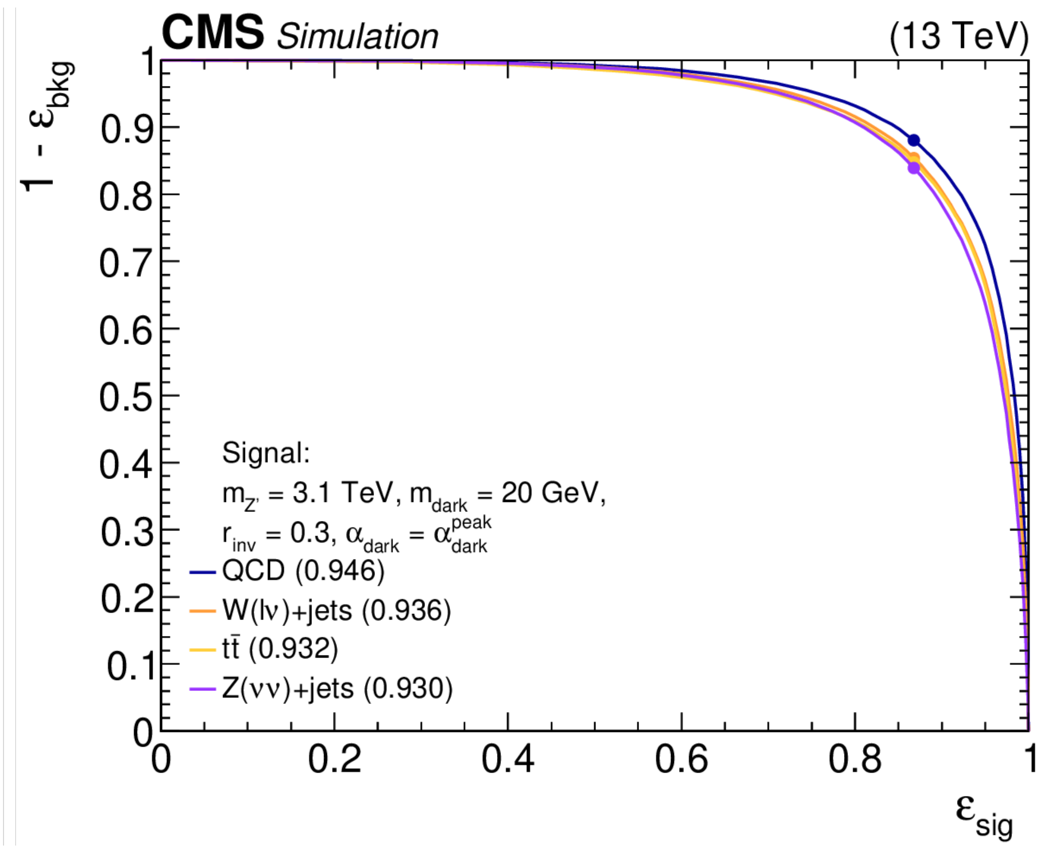
- **Two signal regions** to maximize the signal sensitivity:
  - Low-R_T: $0.15 < R_T < 0.25$
  - High-R_T: $R_T > 0.25$

- **Two analysis approaches:**
  - Inclusive / Generic analysis (sensitive to any model)
  - BDT based (semi-visible jets tagger. Dark sector model dependent)

BDT training & performance

- **Jet tagger trained to discriminate semi-visible jets from lighter SM jets** (different underlaying interactions between SM and dark quarks ➝ make use of jet internal structure features)
  - Trained on a mixture of different signal samples and simulated multijet and $t\bar{t}$ events
  - 15 input variables related to jet flavor, substructure, q/g discrimination… etc
  - Reweighted background jet p_T spectrum to match signal ➝ flat efficiency vs mT





- Selected jets with **BDT score > 0.55** ➝ rejects ~84-88% background jets, keeps ~87% of signal semi-visible jets
- Same signal region categorization as in the inclusive analysis: Low-R_T ($0.15 < R_T < 0.25$); High-R_T: ($R_T > 0.25$) for events with 2 jets with BDT score > 0.55.

Results

- **Signal extraction: fit to mT shape**

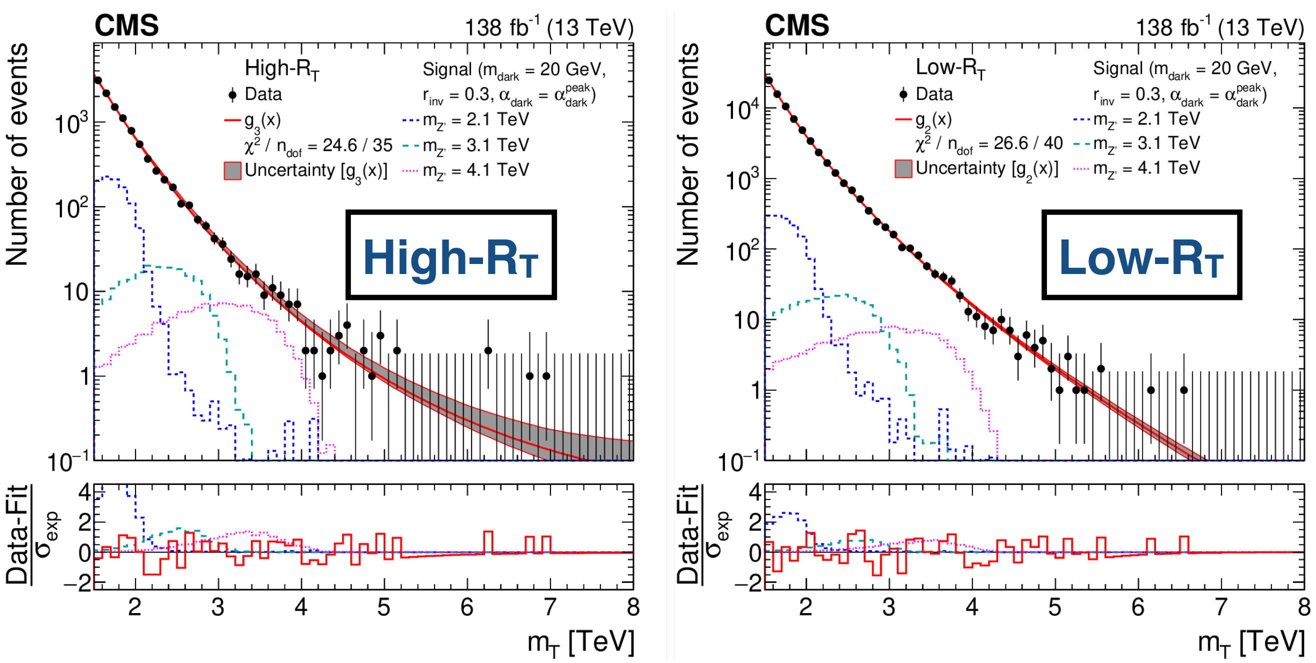
$$m_T^2 = m_{JJ}^2 + 2\ MET \left[\sqrt{m_{JJ}^2 + p_{T,JJ}^2} - p_{T,JJ} \cos\left(\varphi_{JJ,MET}\right) \right]$$

  - Background estimated with analytic smoothly falling function

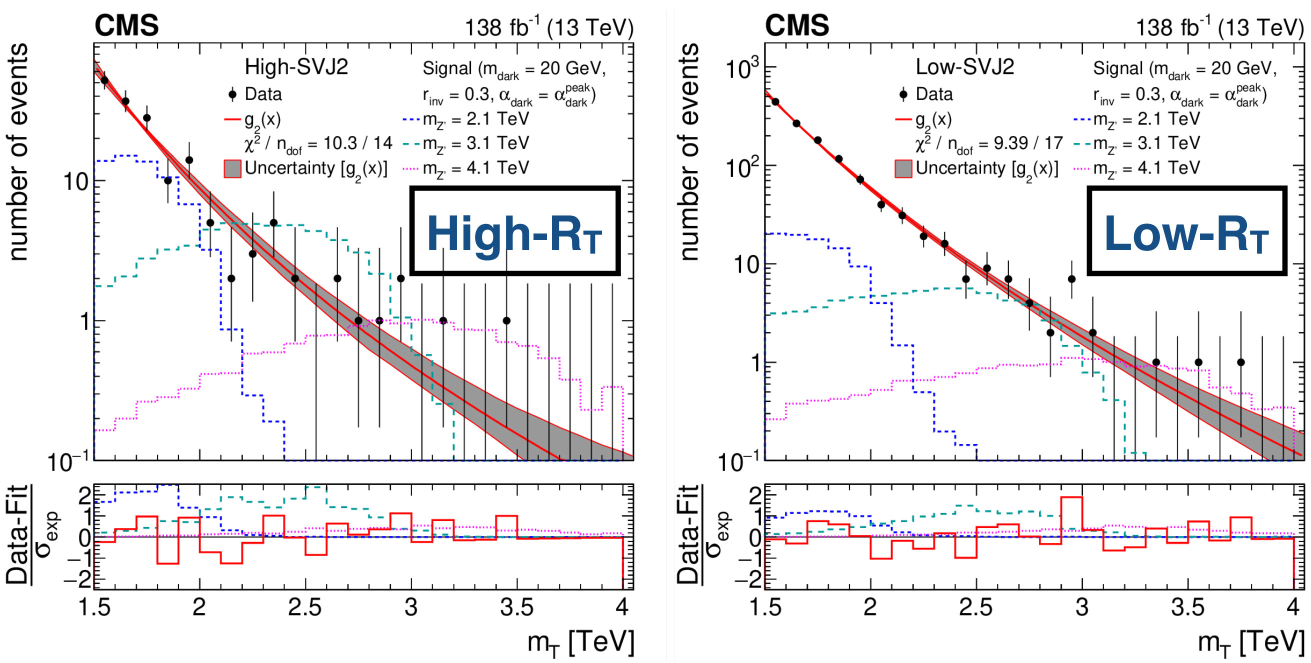
$$g(x) = \exp(p_1 x) x^{p_2[1+p_3 \ln(x)]} \qquad x = m_T/\sqrt{s}$$

  - Background normalization can freely float to the data

Inclusive analysis



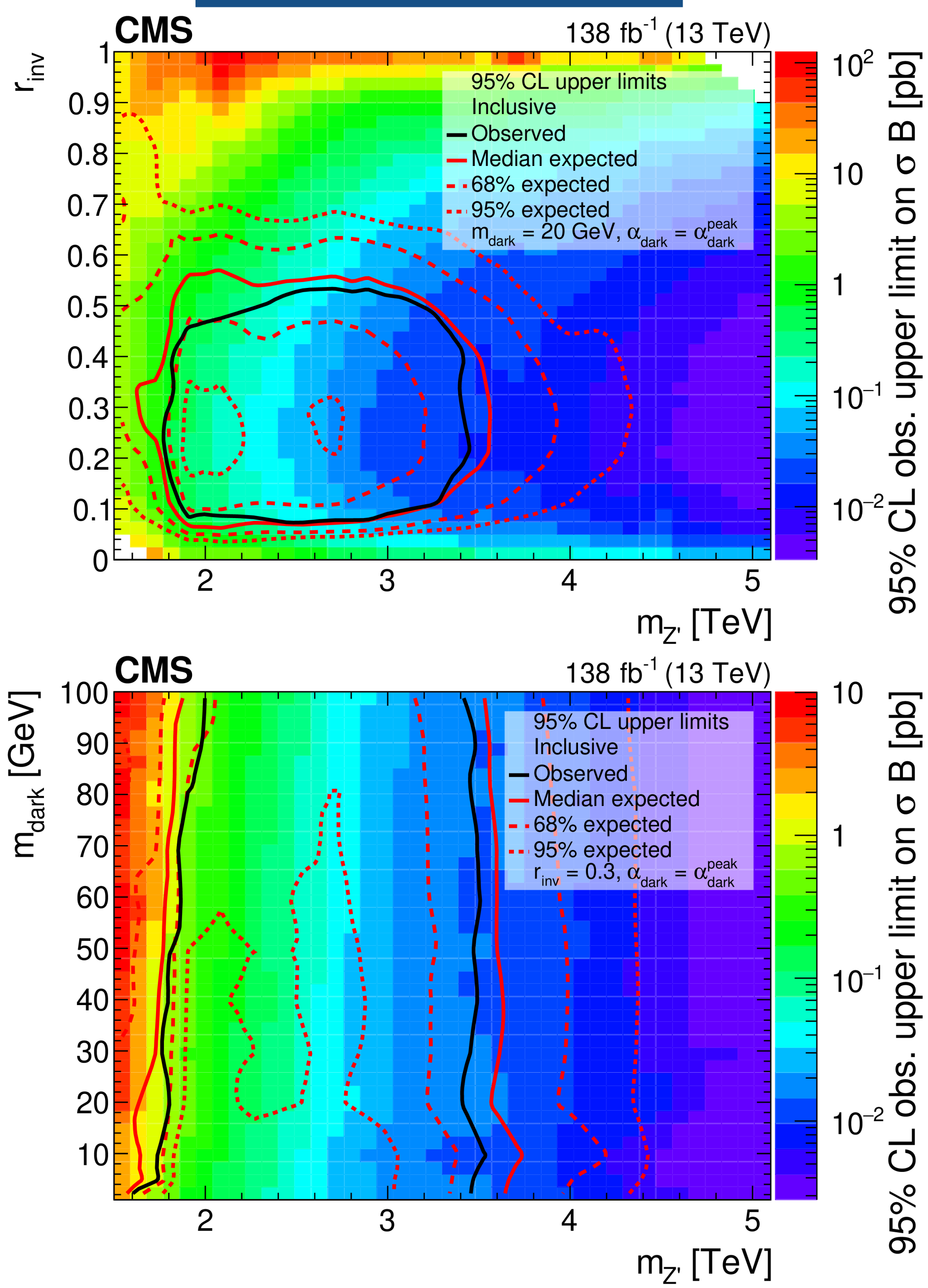
BDT-based analysis



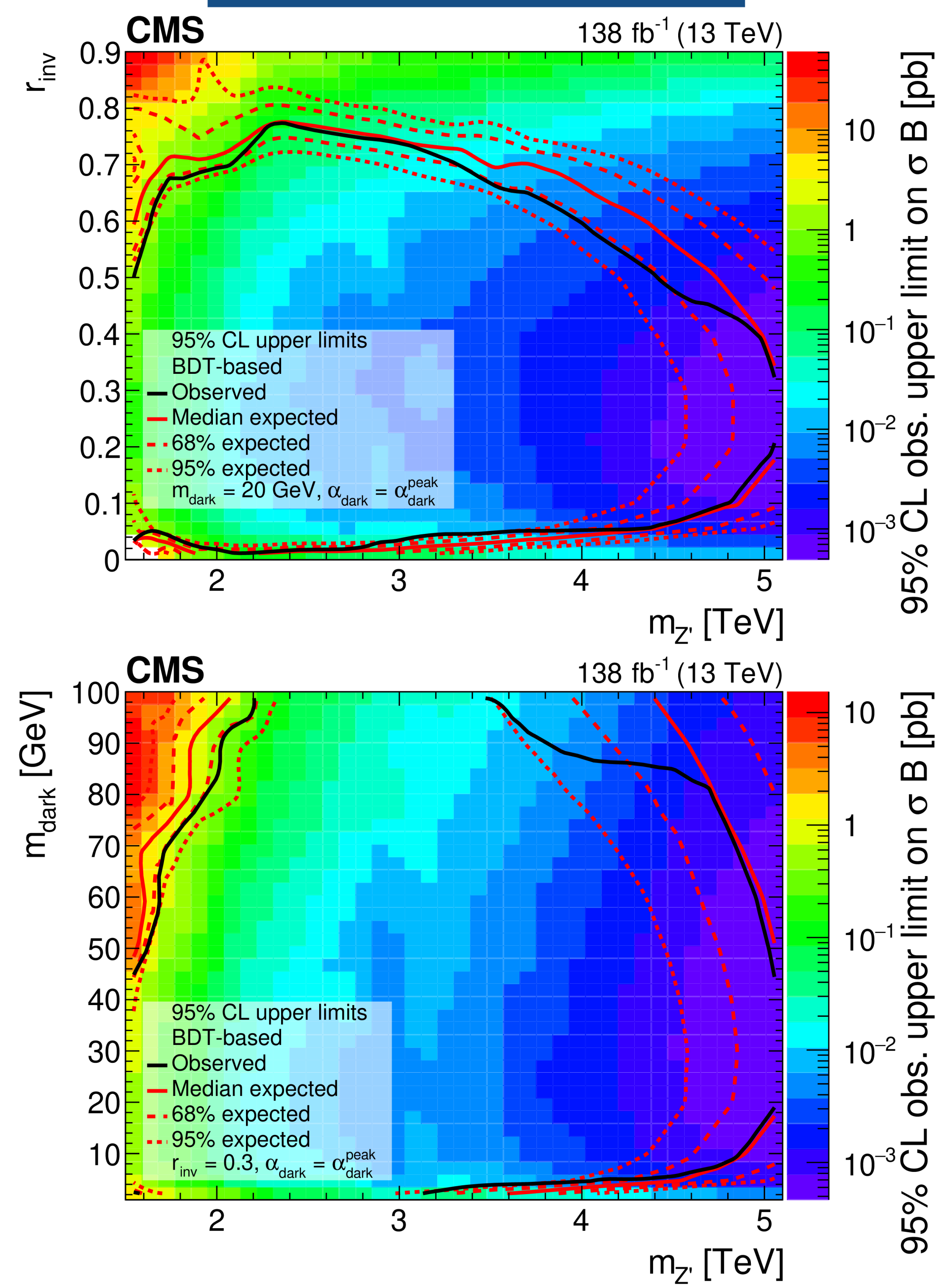
Results

- Excluded models with $1.5 < M_{Z'} < 4.0$ TeV and $0.07 < r_{inv} < 0.53$ (inclusive), and $1.5 < M_{Z'} < 5.1$ TeV and $0.01 < r_{inv} < 0.77$ (BDT-based) for $m_{dark} = 20$ GeV
- BDT efficiency depends on m_{dark} → the inclusive search provides stronger limits at very high and very low m_{dark} values. **Inclusive also provides reinterpretability** for other signal models with MET aligned with jets

Inclusive analysis



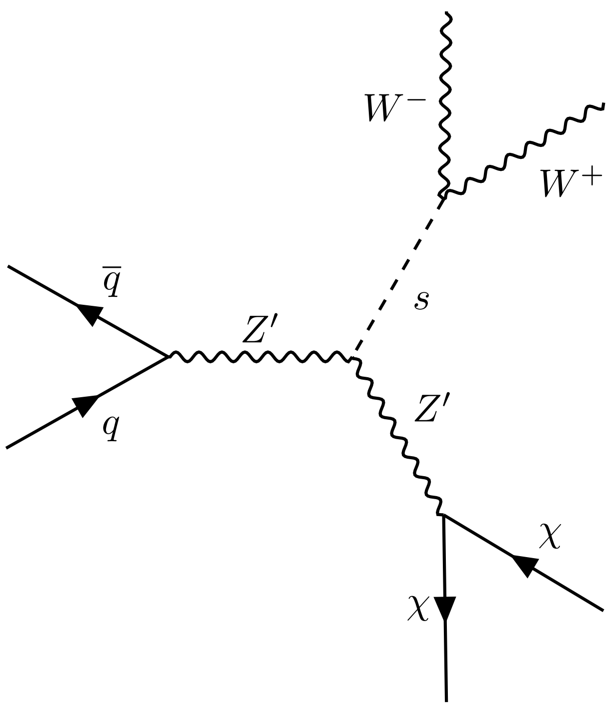
BDT-based analysis



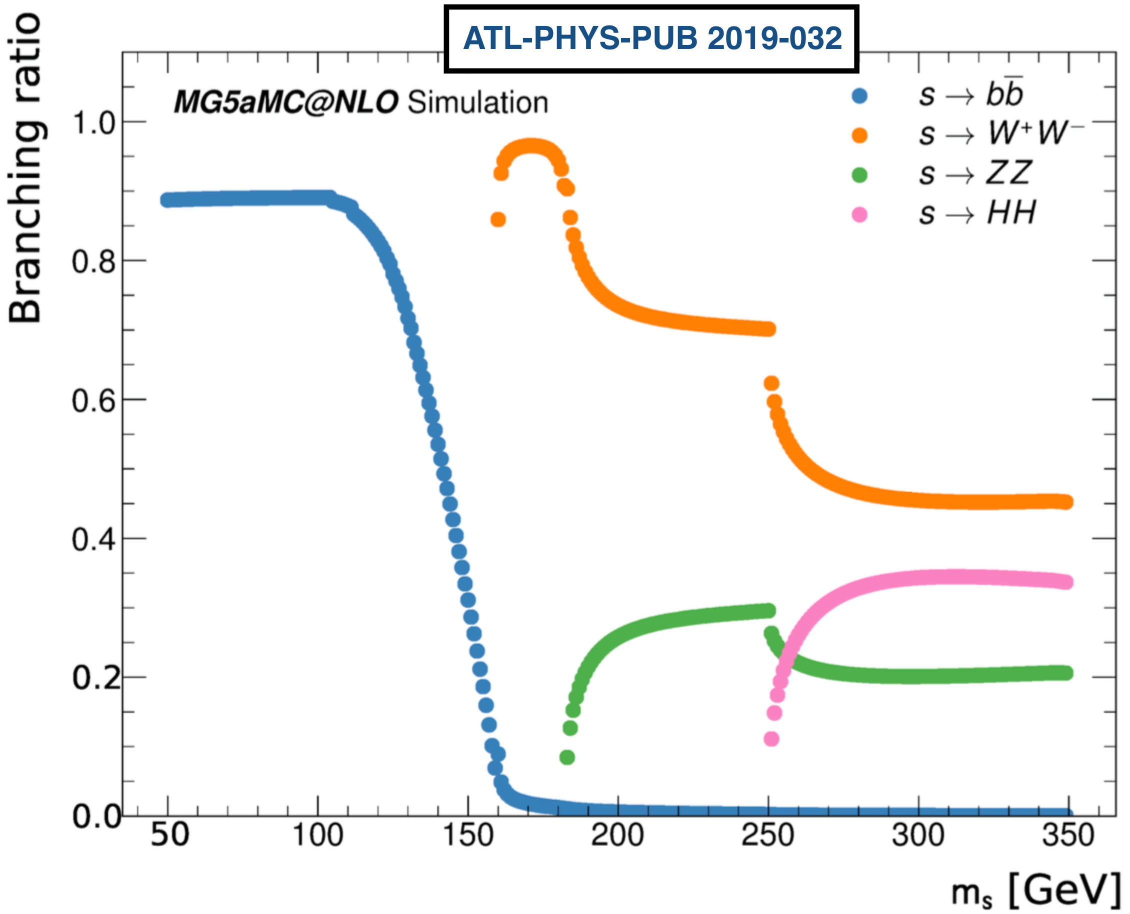
Search for Dark Higgs

CMS-PAS-EXO-21-012

- Search for DM using the dark Higgs simplified model as benchmark
- Emission of a dark Higgs boson, *S*, that mixes with the SM boson and **would provide mass to the DM (WIMPs) particles**
  - *S* can decay into SM states → can be searched at the LHC.
  - Mass scan [m*S*, mZ', m*χ*]. Fixed couplings
- Since *S* can be lighter than DM particles, it **could relax the DM relic abundance** constraints by introducing a new annihilation channel *χχ* → *SS*
- WW decay mode dominates the dominates the picture for m*S* ≳ 160 GeV (resonance). Ongoing CMS bb analysis is exploring the low mass region m*S* < 160 GeV
- Semi- and di-leptonic W^+W^- final states have been studied



ATL-PHYS-PUB 2019-032



Event selection: dileptonic

- The analysis targets the dileptonic + semileptonic decay of the W^+W^- boson pair
- Same event **selection** for the three data periods:

Target signature: two opposite charged isolated leptons, and large transverse missing energy from the neutrinos + χ_s

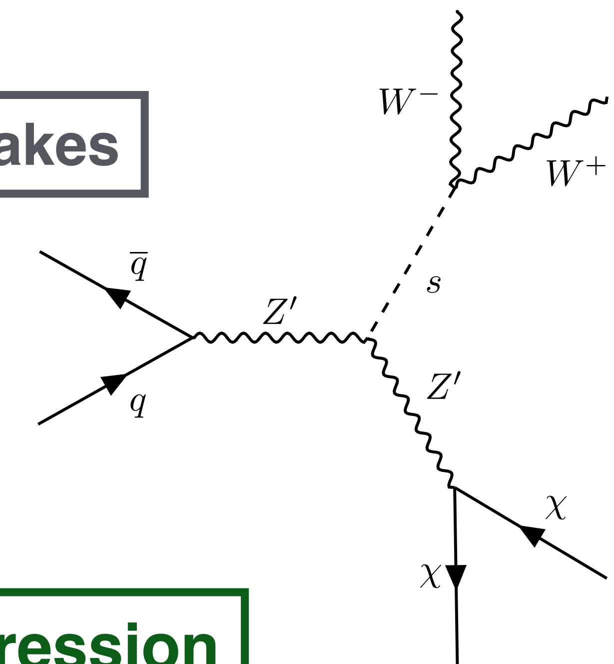
Dilep Selection
2 leptons, Different flavour, opposite signed
$p_{\ell_1 T}$ / $p_{\ell_2 T}$ > 25 / 20 GeV
Vetoed additional leptons with $p_{\ell_3 T}$ > 10 GeV
$p_T^{\ell\ell}$ > 30 GeV
$m_{\ell\ell}$ > 12 GeV
p_T^{miss} > 20 GeV
$m_T(\ell\ell + p_T^{miss})$ > 50 GeV
$\Delta R(\ell\ell)$ < 2.5
Number of b-tagged jets = 0

Reduce Fakes ($p_{\ell_1 T}$ / $p_{\ell_2 T}$; $p_T^{\ell\ell}$)

DY suppression ($p_T^{\ell\ell}$; $m_{\ell\ell}$; p_T^{miss}; $m_T(\ell\ell + p_T^{miss})$)

suppress non-resonant W^+W^- ($\Delta R(\ell\ell)$)

suppress ttbar (Number of b-tagged jets)



Event selection: semileptonic

- The analysis targets the dileptonic + semileptonic decay of the W^+W^- boson pair
- Same event **selection** for the three data periods:

Target signature: one isolated leptons, two jets, and large transverse missing energy from the neutrinos + χ_s

Semilep Selection
1 isolated lepton, ≥2 jets
$p_{\ell_1 T}$ > trigger thresholds
Vetoed additional leptons with $p_{\ell_2 T} > 10$ GeV
Number of b-tagged jets = 0
$65 < m_{jj} < 105$ GeV
$p_T^{miss} > 60$ GeV
$p_{T\ell jj} > 60$ GeV
$\Delta\phi(jj,\ell) < 1.8$, $\Delta R(jj,\ell) < 3$
$\Delta\phi(\ell jj, p_T^{miss}) > 2$

Number of b-tagged jets = 0 → **suppress ttbar**

$65 < m_{jj} < 105$ GeV → **suppress W+jets**

p_T^{miss}, $p_{T\ell jj}$, $\Delta\phi(jj,\ell)$/$\Delta R(jj,\ell)$, $\Delta\phi(\ell jj, p_T^{miss})$ cuts → **target dark Higgs topology**

Background estimation overview

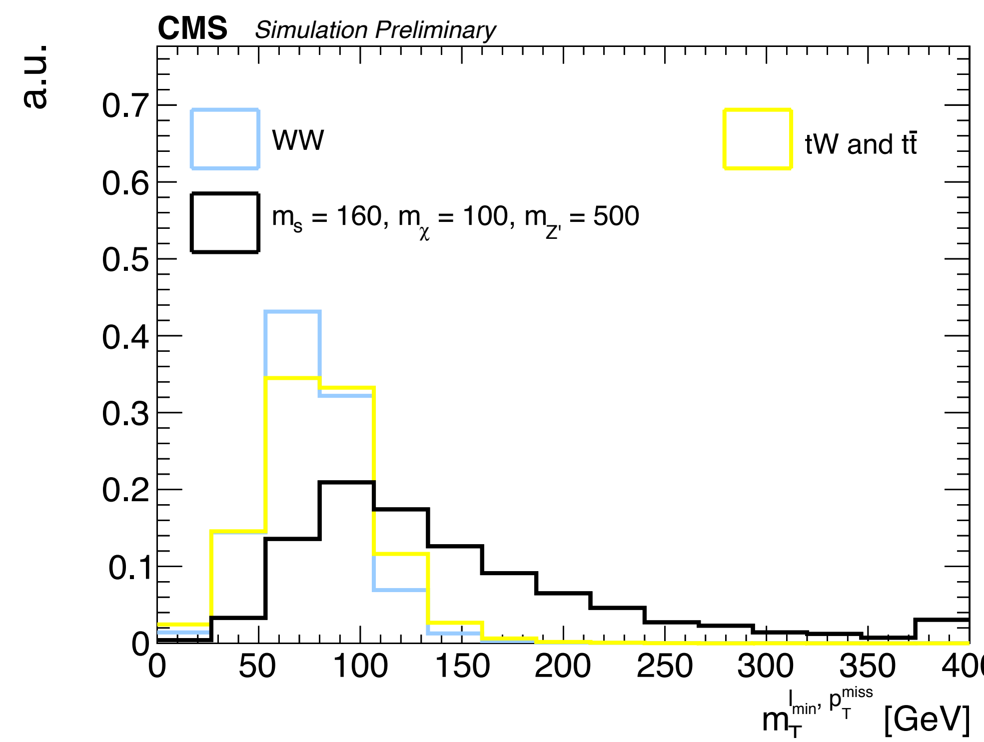
- **Dileptonic:**
  - **Non-prompt** leptons: estimated with a fully data-driven method, and validated in same-signed validation region
  - **Top-quark:** the normalization is measured using top-tagged events in data control region
  - **Non-resonant WW**: the normalization is measured using events with large angular distance between the two leptons in data control region
  - $\mathbf{Z/\gamma^* \rightarrow \tau^+\tau^-}$: the normalization is measured using low mT(ℓℓ+MET) events in data control region
- **Semileptonic:**
  - **Non-prompt:** same strategy as for dileptonic
  - **Top-quark:** same strategy as for dileptonic
  - **W+Jets:** the normalization is measured using events with mjj side band in data control region
- All other (small) processes are estimated directly from simulation: HWW, VΥ/VΥ*, VZ, VVV

Orthogonal selection:
Same lepton charges
Number of b-tagged jets > 0
$\Delta R(\ell\ell) > 2.5$
$m_T(\ell\ell+MET) < 50$ GeV
mT(l + MET) < 30 GeV && MET < 30 GeV
Number of b-tagged jets > 0
$m_{jj} < 65$ or $m_{jj} > 105$ GeV

(keeping the other preselection requirements in each case)

Analysis strategy: dileptonic

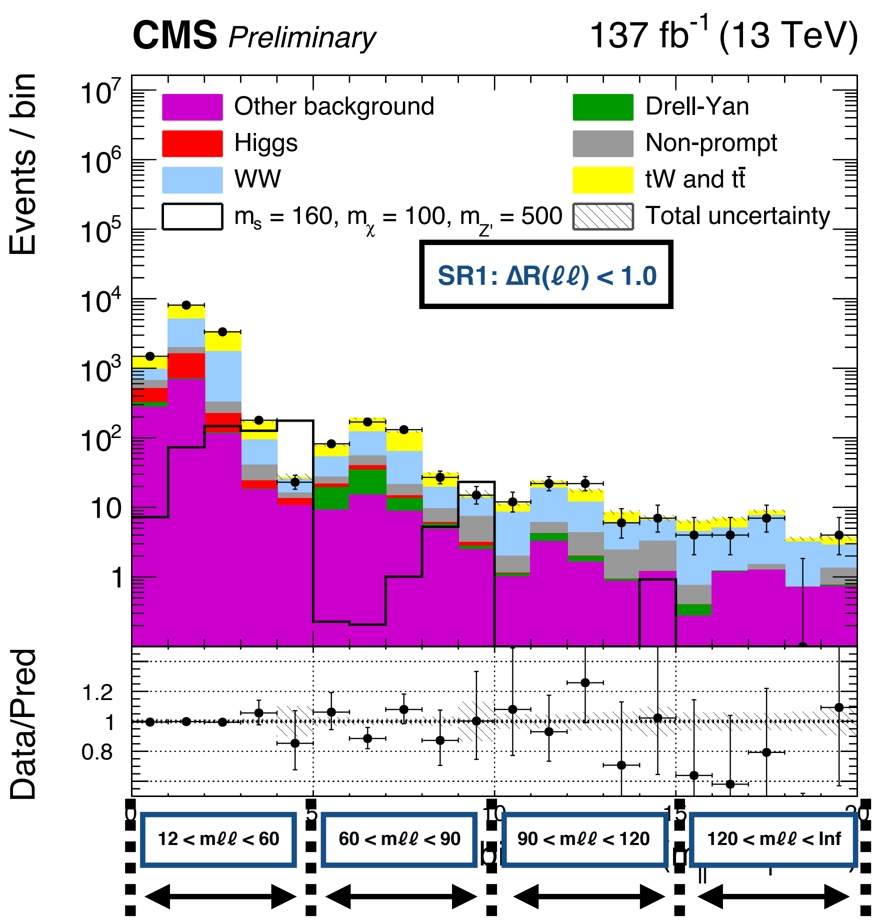
- **Dileptonic signal extraction:** 3D ML fit to $\Delta R(\ell\ell)$ - $m_{\ell\ell}$ - mT(ℓmin + p_T^{mis})
- More sensitive to the dark Higgs signal prediction than other quantities based on lepton kinematics and/or p_T^{miss}
- **Optimized procedure:** strong kinematic dependence on m_S
  - Three signal regions are defined in $\Delta R(\ell\ell)$, based on the S/sqrt(S+B) curves vs $\Delta R(\ell\ell)$ for each dark Higgs mass
    - **SR1:** $\Delta R(\ell\ell) < 1.0$, **SR2:** $1.0 < \Delta R(\ell\ell) < 1.5$, **SR3:** $1.5 < \Delta R(\ell\ell) < 2.5$

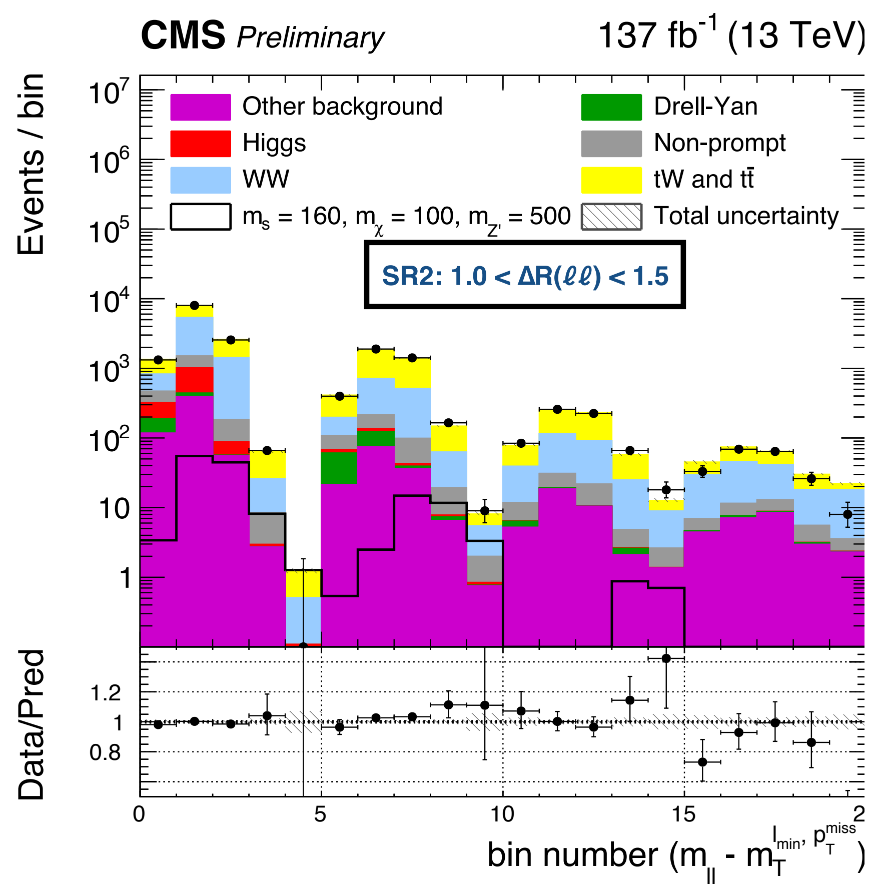


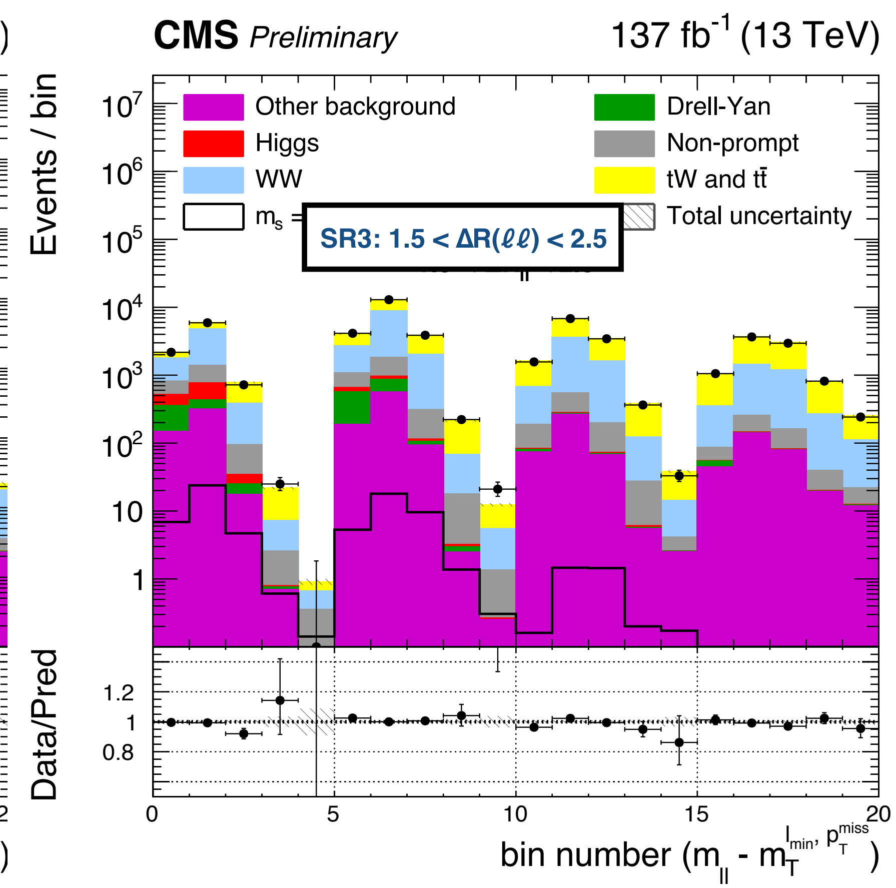
- For each SR, a 2D template of **$m_{\ell\ell}$ - mT(ℓmin + p_T^{mis})** is defined:
  - The $m_{\ell\ell}$ binning is set from significance S/sqrt(S+B) curves vs $m_{\ell\ell}$ for each dark Higgs mass
  - The mT(ℓmin + pTmiss) binning is set by squeezing the sensitivity for each data period
- Allow the different signal mass points to **freely populate the 3D phase space** while using the same background modeling procedure

Results: dileptonic

- The signal strength is extracted by fitting the predicted yields to the observed events
- **ML fit:** 3 Signal Regions, 1 Top Control Region, 1 DY Control Region, 1 WW Control Region for each data period
  - Signal regions information entering in the fit: 2D histograms of $m_{\ell\ell}$ - $m_T(\ell min + p_T^{mis})$ from SR1, SR2 and SR3
  - Control regions information entering in the fit: 1-bin distributions. Top, WW, and DY normalization freely float within the global fit







- No significant excess over the SM prediction

Analysis strategy: semileptonic

- **Semileptonic signal extraction:** fit to the shape of BDT output score
- BDT trained with set of variables that showed most separation power between signal and background, based on final state objects (lepton, 2 jets, MET)

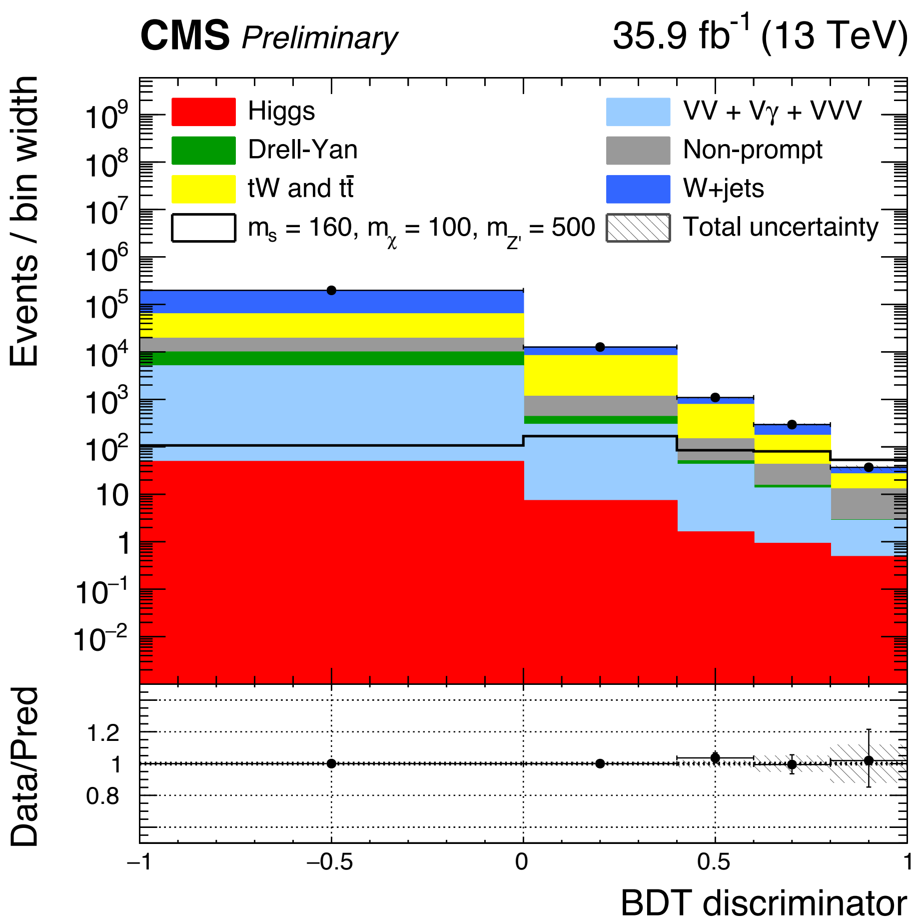
Variable	Definition
p_T^{jj}	p_T of the vectorial sum of the W candidate jets
$p_T^{\ell jj}$	p_T of the vectorial sum of the visible particles
p_T^{miss}	Size of the missing transverse momentum vector
$\Delta\eta_{\ell,jj}$ and $\Delta\phi_{\ell,jj}$	$\Delta\eta$ and $\Delta\phi$ between the lepton and the di-jet system
$\Delta\eta_{j,j}$ and $\Delta\phi_{j,j}$	$\Delta\eta$ and $\Delta\phi$ between the W candidate jets
$\Delta\eta_{\ell,p_T^{miss}}$ and $\Delta\phi_{\ell,p_T^{miss}}$	$\Delta\eta$ and $\Delta\phi$ between the lepton and $\vec{p}_T^{miss}$
$\Delta\phi_{\ell jj,p_T^{miss}}$	$\Delta\phi$ between the vectorial sum of the visible particles and $\vec{p}_T^{miss}$
$min(p_T^{\ell}, p_T^{j_2})/p_T^{miss}$	Minimum of the lepton p_T and the trailing jet p_T, divided by p_T^{miss}
$max(p_T^{\ell}, p_T^{j_2})/p_T^{miss}$	Maximum of the lepton p_T and the leading jet p_T, divided by p_T^{miss}
$max(p_T^{\ell}, p_T^{j_1})/m_{\ell jj p_T^{miss}}$	Maximum of the lepton p_T and the leading jet p_T, divided by the invariant mass of the vectorial sum of the visible particles and the p_T^{miss} where the missing energy is considered to be massless

- Binning optimized based on S/sqrt(S+B) curves for the three data periods

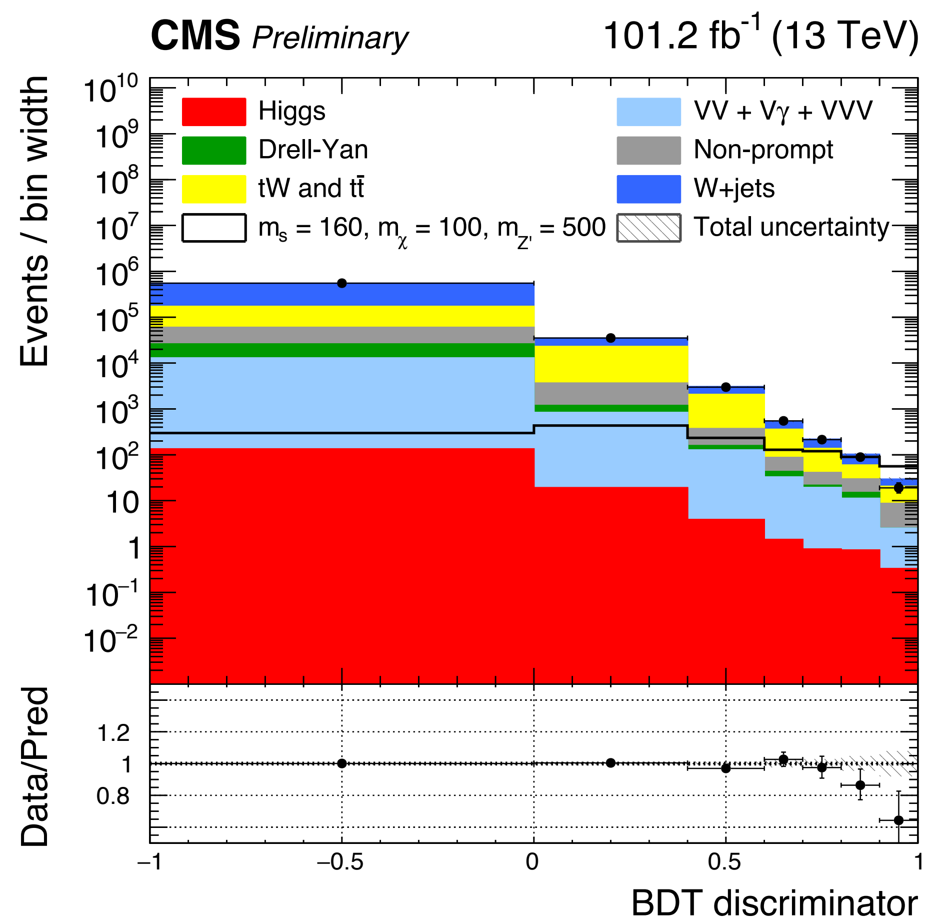
Results: semileptonic

- **ML fit:** 1 Signal Regions, 1 Top Control Regions, 1 DY Control Region, 1 WW Control Region for each data period
  - Signal regions information entering in the fit: 1D histograms of BDT output score
  - Control regions information entering in the fit: 1-bin distributions. Top and W+Jets normalization freely float within the global fit

2016 BDT output



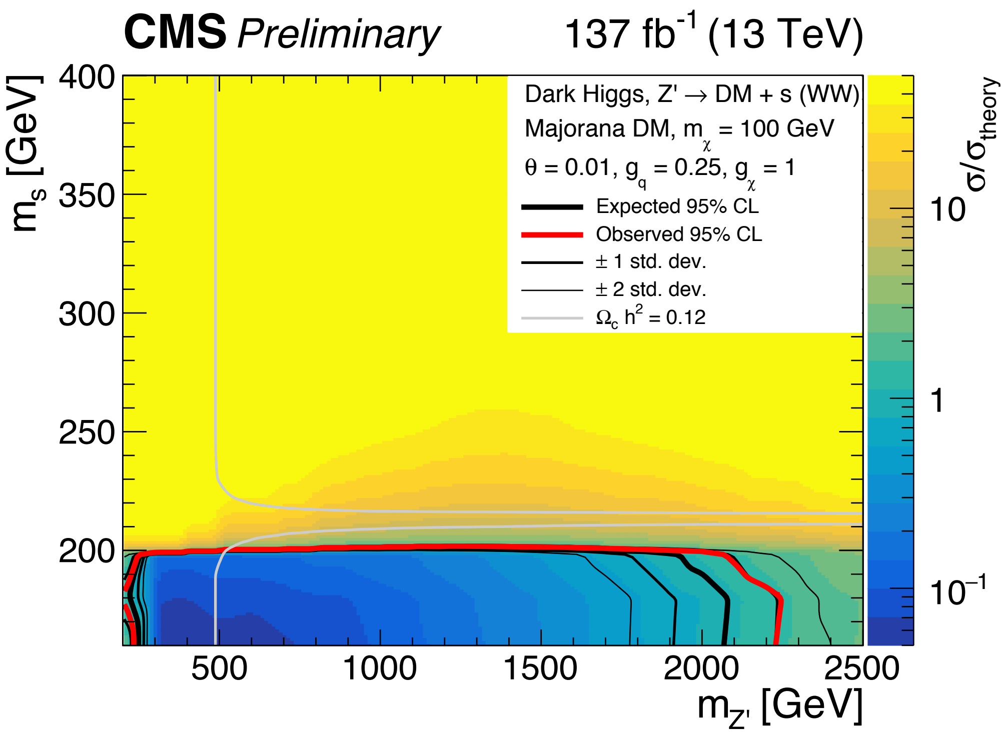
2017+2018 BDT output



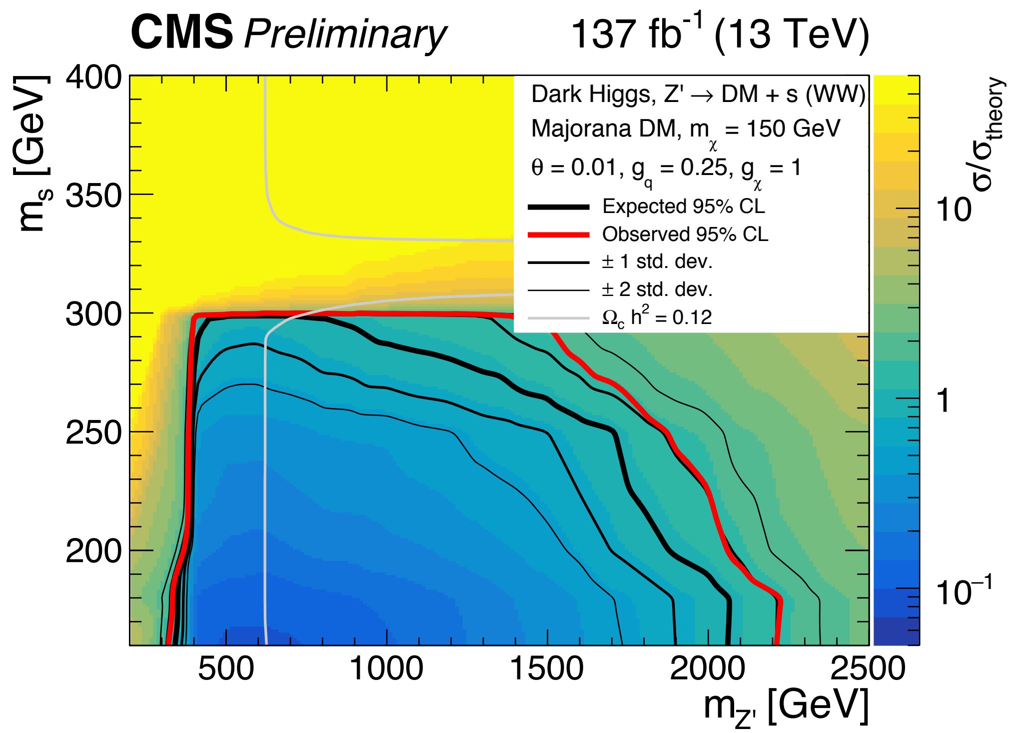
- Finer binning in 2017-2018 to squeeze the sensitivity

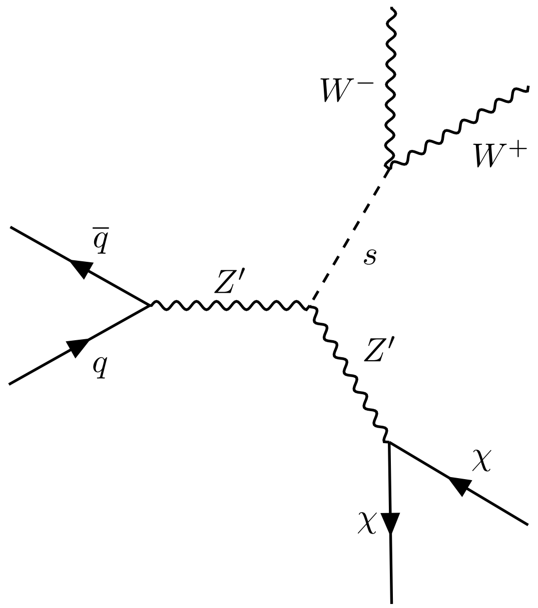
Results: combination

mχ = 100 GeV

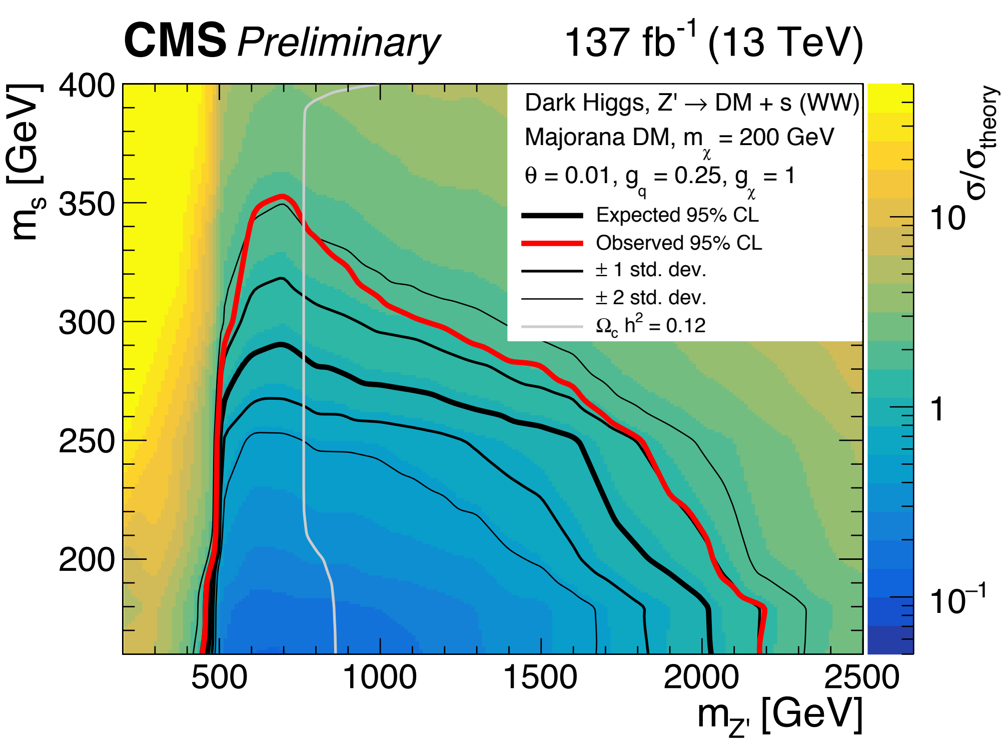


mχ = 150 GeV

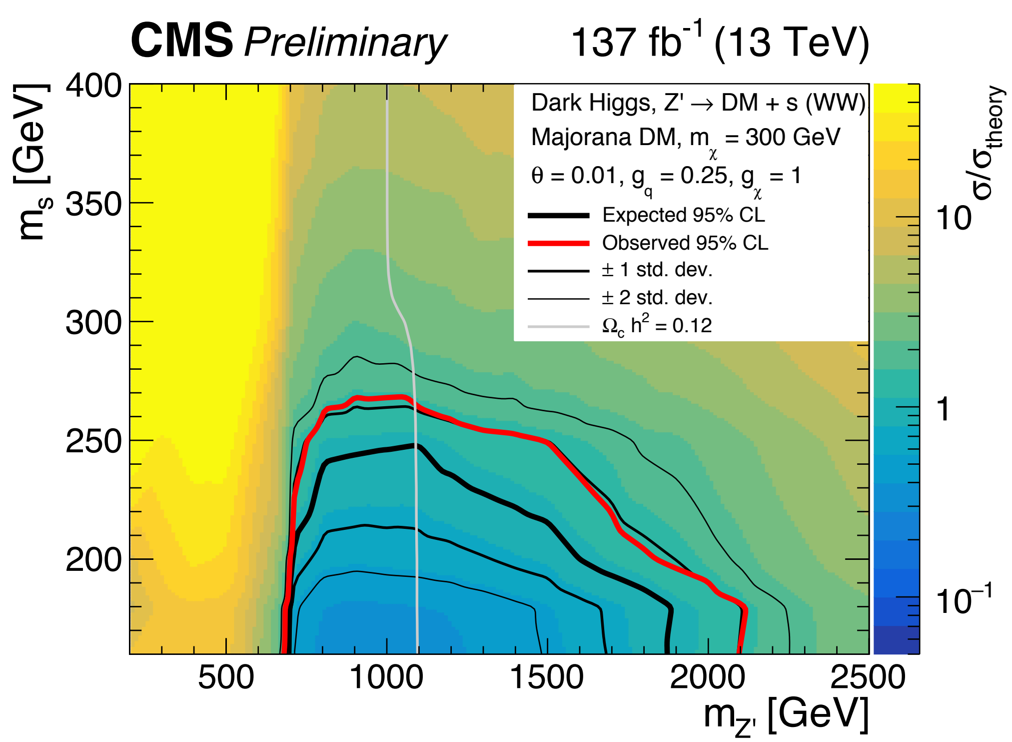




mχ = 200 GeV



mχ = 300 GeV



- Dominant $S \to \chi\chi$ decay mode (invisible) for $m_S \geq 2m_\chi$
- Observed limit is better than the expected due to slight data deficit in some of the sensitive bins

Conclusions

- **Most recent searches for Dark Matter performed with the CMS detector have been presented**
- Shown results on Run 2 data:
    - Search for inelastic Dark Matter with long-lived particles
    - Search for semi-visible jets
    - Search for dark Higgs
- **No significant discrepancies with SM predictions observed** → results are interpreted as limits on DM mediators, simplified model parameters, and on DM-nuclei interaction cross sections
- Large part of the CMS results exploit the full Run 2 dataset (138 fb^{-1}) so far
    - Already collected ~60 fb^{-1} on Run 3 at 13.6 TeV
    - New results to come and new exotic signatures to be inspected in Run 3. **Stay tuned!**