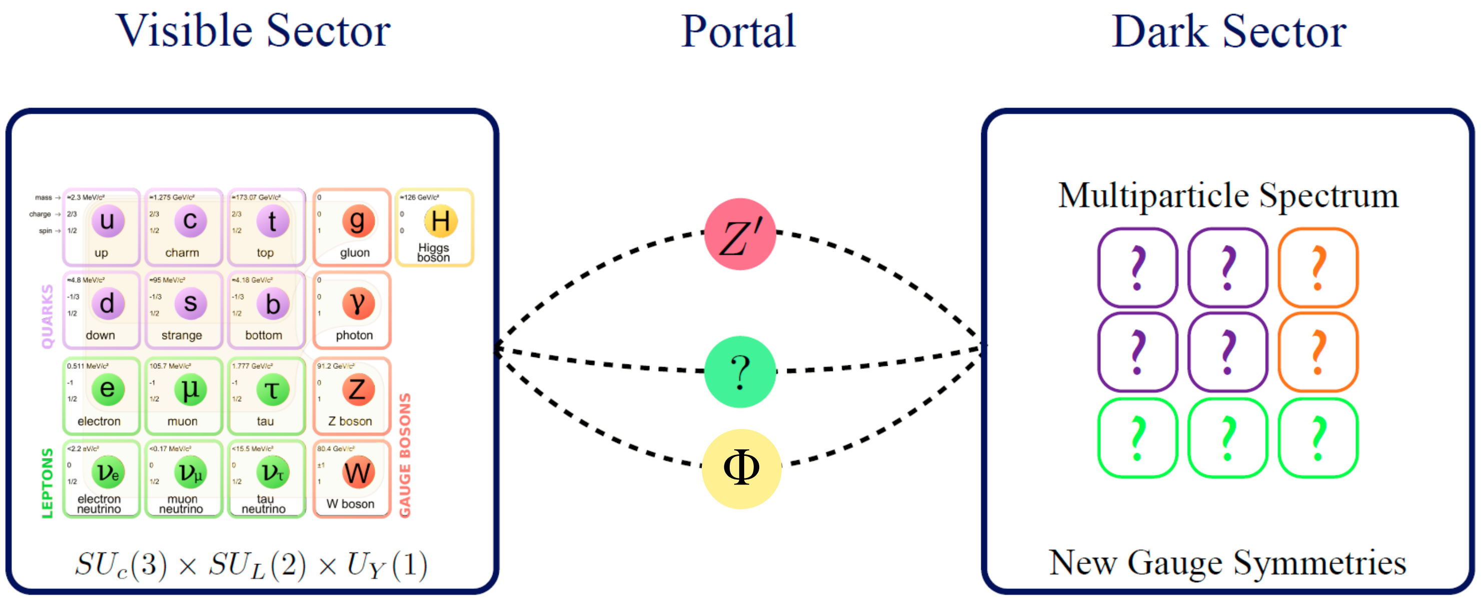
Semivisible Jets at CMS

Kevin Pedro
(Fermilab)
July 5, 2022

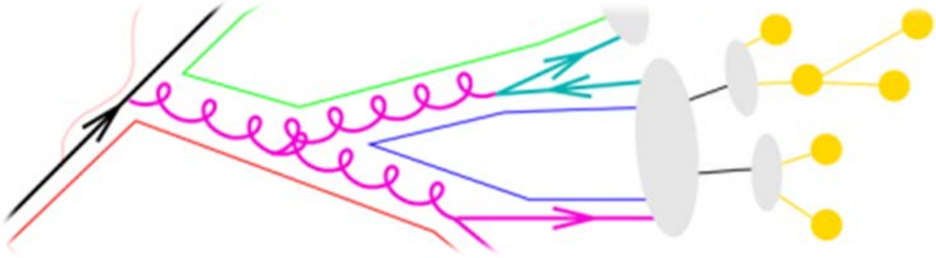
Hidden Sectors



- Simplest assumption: dark matter consists of a single species of weakly interacting massive particles
  - No observation of WIMPs → look for new models and phenomenology
- Dark matter may consist of multiple species of composite particles interacting via new, dark forces
  - Visible matter is mostly composite particles & has similar density to DM

Strongly Coupled Models

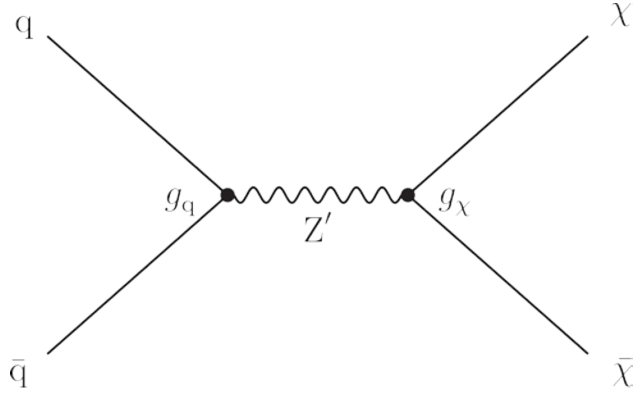
- New "dark QCD" force, $SU_{\text{dark}}(\boldsymbol{N_c^{\text{dark}}})$ (carried by dark gluons) with scale Λ_{dark}
- $\boldsymbol{N_f^{\text{dark}}}$ flavors of (fermionic) dark quarks χ_i (charged under $SU_{\text{dark}}(\boldsymbol{N_c^{\text{dark}}})$)
- Dark quarks *hadronize* to form dark mesons and baryons → "dark showers"



- Some dark hadrons may be *stable* because of conserved quantities
  - o Dark baryon number, dark isospin number, etc.
  - ➢ DM candidates!
- Other dark hadrons decay back to SM (through virtual mediators)
  - ➢ Leads to novel phenomenology

Production

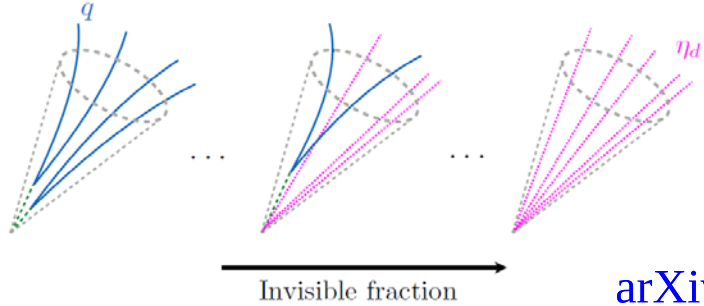
- Hidden sector couples to SM weakly via massive mediator: Z' from broken $U(1)$, vector, leptophobic, couplings g_q, g_χ



- Coupling choices aligned with LHC DM Working Group:
  - $g_q = 0.25$
  - $g_\chi = 1.0/\sqrt{(N_c^{\text{dark}} N_f^{\text{dark}})} = 0.5$
  - ➢ $B_{\text{dark}} = 47\%$, $\Gamma_{Z'}/m_{Z'} = 5.6\%$
    - Same as LHC DM models w/ $g_{DM} = 1.0$

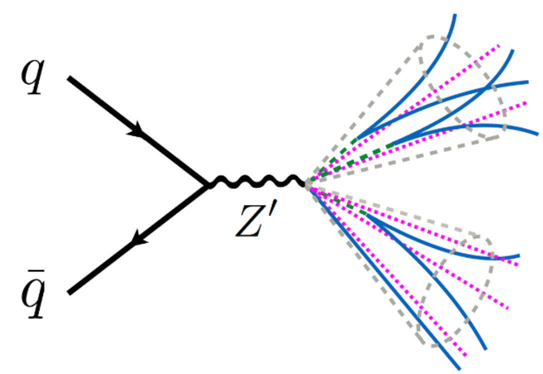
Decay

- Fraction of stable hadrons r_{inv} may vary from 0 to 1
  - Decreases w/ dark quark mass splitting, increases w/ N_f^{dark}
- ➢ Jets that contain *mix of visible and invisible particles* (prompt decays)
  - *Not covered* by existing searches for dijet resonances, p_T^{miss}+ISR



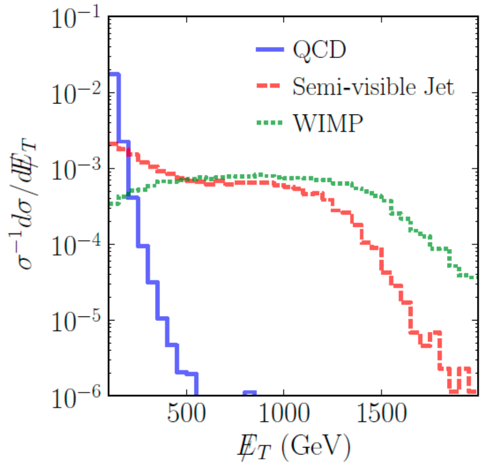
arXiv:1707.05326

- $Z' \to \chi\chi \to$ dark hadrons $\to$ SM quarks $\to$ SM hadrons
  - Decay to SM $\to$ two high-p_T, *wide* jets
  - ρ_{dark}: democratic decay
  - π_{dark}: mass insertion decay (prefer heavy flavor)
  - $N_c^{\text{dark}} = 2$, $N_f^{\text{dark}} = 2$, $m_\chi = \frac{1}{2} m_{\text{dark}}$

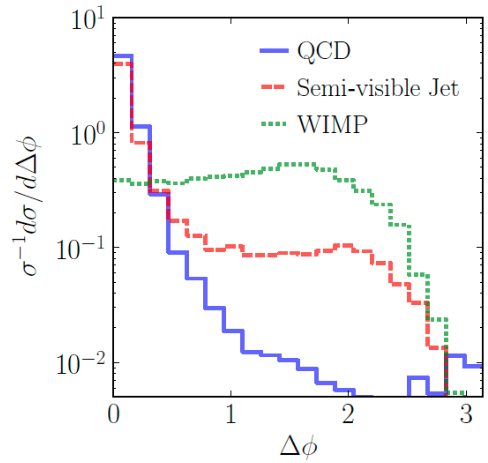


Resonant Search

- Kinematic signature: Less missing energy than WIMPs, aligned w/ jet

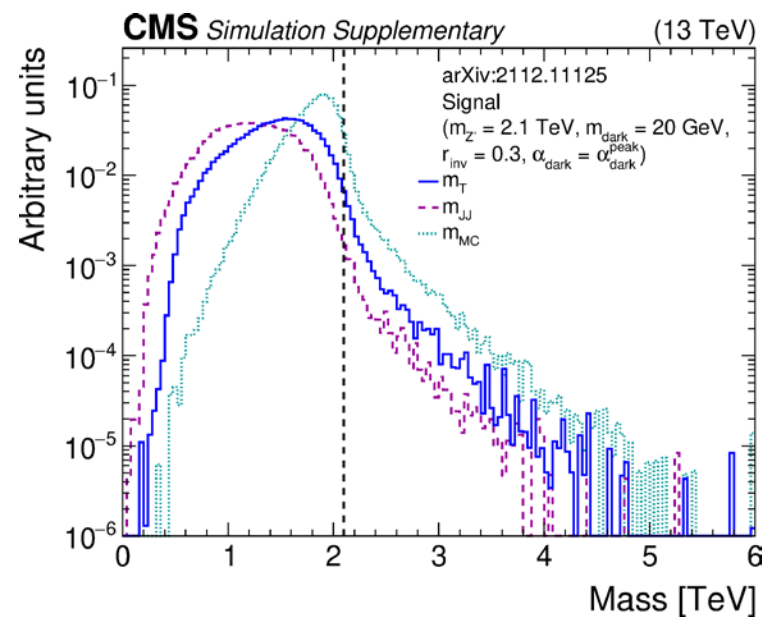


[arXiv:1503.00009](https://arxiv.org/abs/1503.00009)



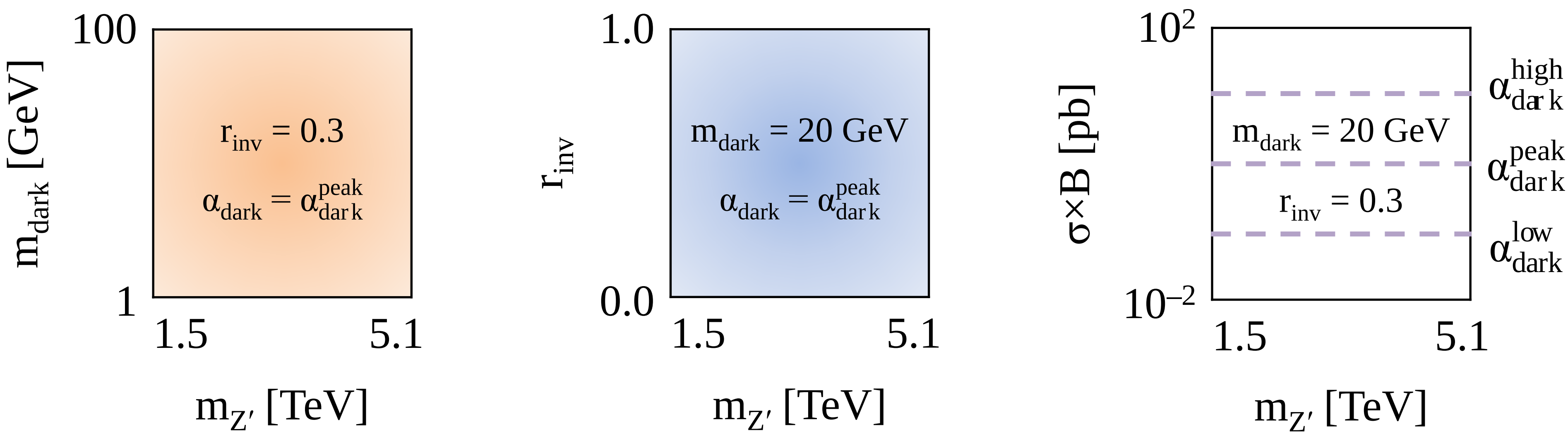
➢ Bump hunt in $\mathbf{m_T(JJ,p_T^{miss})}$

- o Kinematic edge at $\mathbf{m_{Z'}}$
- o Better resolution than $\mathbf{m_{JJ}}$
- o SM backgrounds have steeply falling distributions



Signal Models

- Parameters varied: $m_{Z'}$, m_{dark} (dark hadron mass scale), r_{inv}, α_{dark}
  - o α_{dark}: running coupling of dark QCD (alternate form of scale Λ_{dark})
  - o α_{dark}^{peak} maximizes dark hadron multiplicity (depends on m_{dark})
    - "Empirical" relationship derived from Pythia
    - Variations: $\alpha_{dark}^{high} = {}^{3}\!/_{2}\alpha_{dark}^{peak}$, $\alpha_{dark}^{low} = {}^{1}\!/_{2}\alpha_{dark}^{peak}$
- Three 2D scans ($m_{Z'}$ vs. m_{dark}, r_{inv}, α_{dark}) $\rightarrow$ 475 points
  - o Benchmark values: m_{dark} = 20 GeV, r_{inv} = 0.3, $\alpha_{dark} = \alpha_{dark}^{peak}$
- 4D scan with same grid of values would be 8208 points



Dual Strategy

- Dark QCD theories are very complicated
  - o Need to make choices about numerous parameters
  - o Plus modeling of hadronization/fragmentation, etc.
- First search for jets aligned with p_T^{miss} → maximize *generality* & *sensitivity*

"Inclusive" search

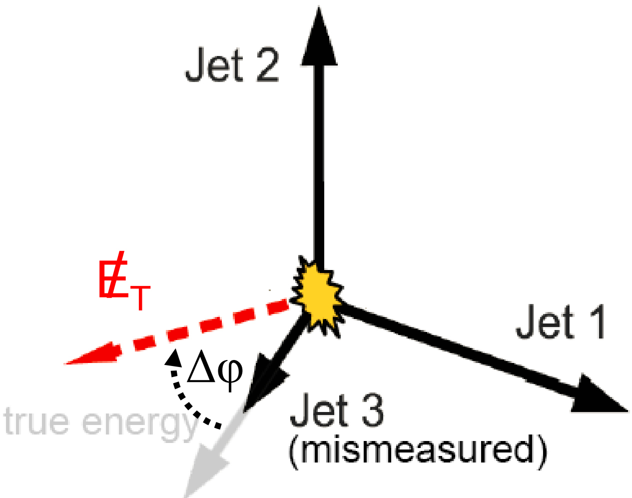
- Use only event-level kinematic variables
- Results apply to any model with similar kinematic behavior

"BDT-based" search

- Employ machine learning for optimized semivisible jet tagger
- Assumes chosen signal models are "correct"

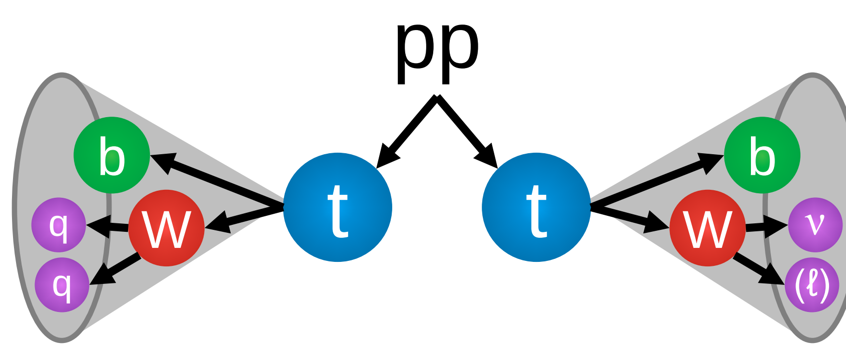
Backgrounds

QCD



- Jet mismeasurement induces $\not{E}_T$ aligned with jet
- Major background

$t\bar{t}$



- Wide, high-p_T jets: boosted tops
- "Lost" lepton ℓ: out of acceptance, can't veto (or hadronic τ)
- Neutrino aligned w/ wide jet: mimics semivisible jet

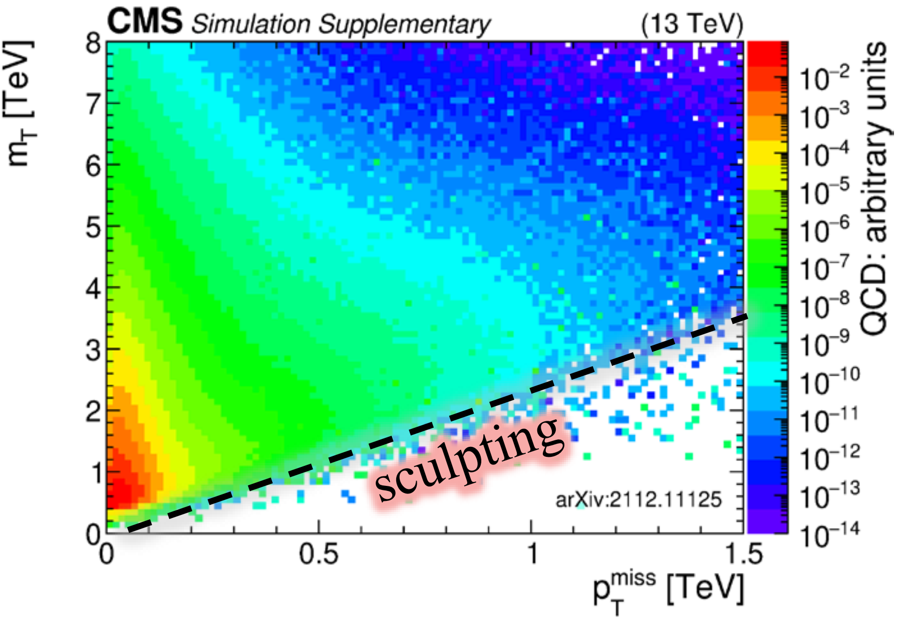
W(ℓν)+jets

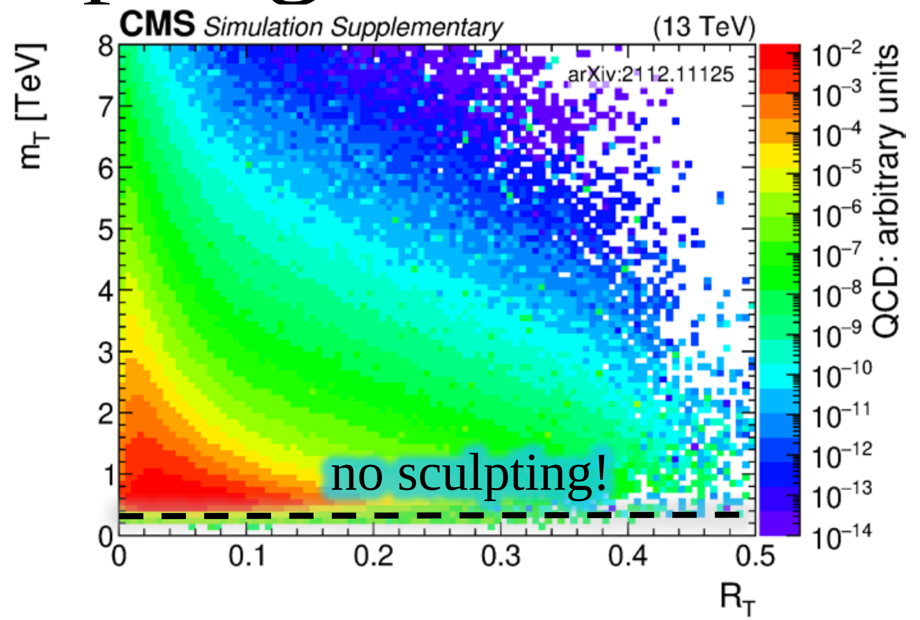
- Lost lepton or hadronic τ
- Less likely than $t\bar{t}$ to mimic semivisible jet, but higher σ

Z(νν)+jets

- Real $\not{E}_T$ from νν, but least likely to align with jet

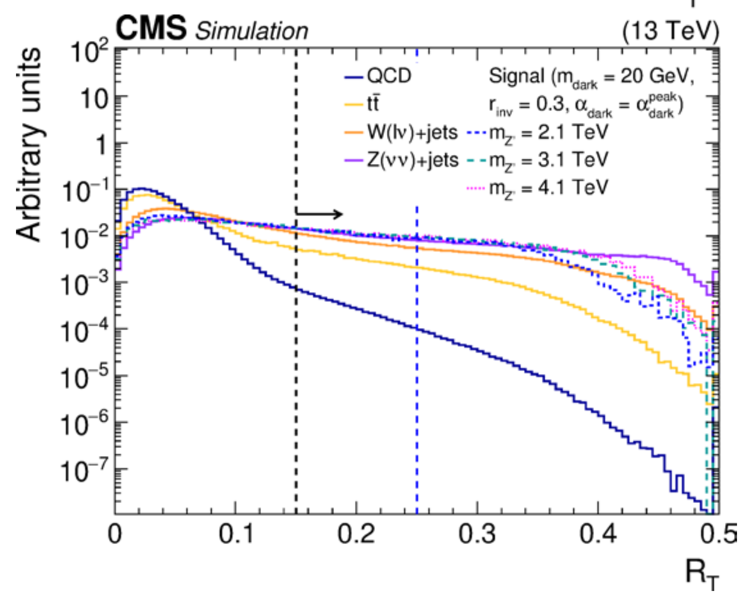
Mass Sculpting





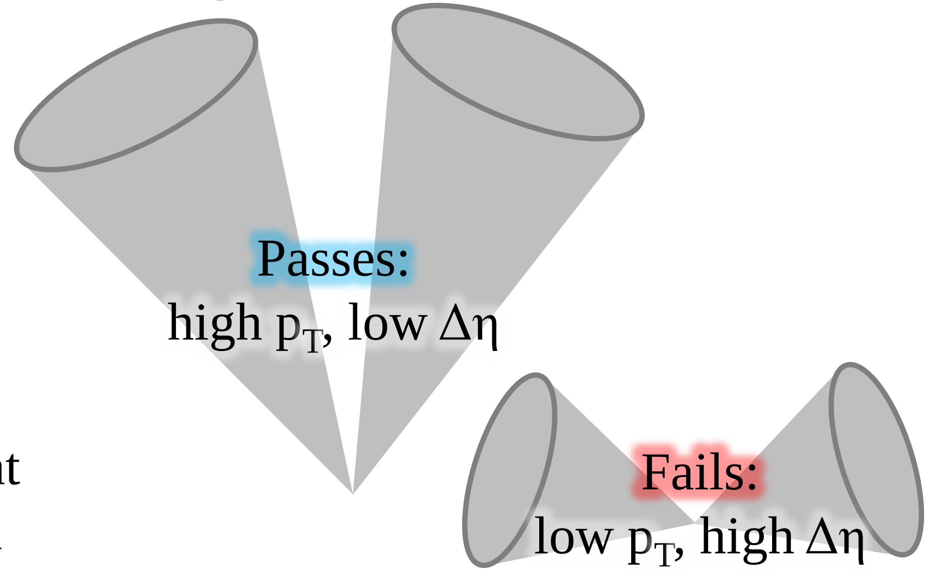
- Avoid/minimize direct cuts on m_T ingredients: p_T^{miss}, jet p_T
  - o Relative variable ("transverse ratio"): $R_T = p_T^{miss}/m_T$

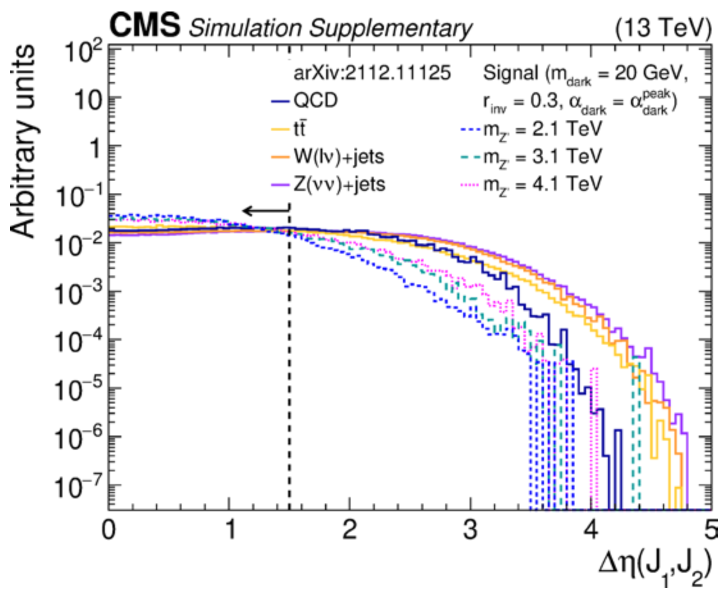
➢ Reject QCD background without shifting m_T peak

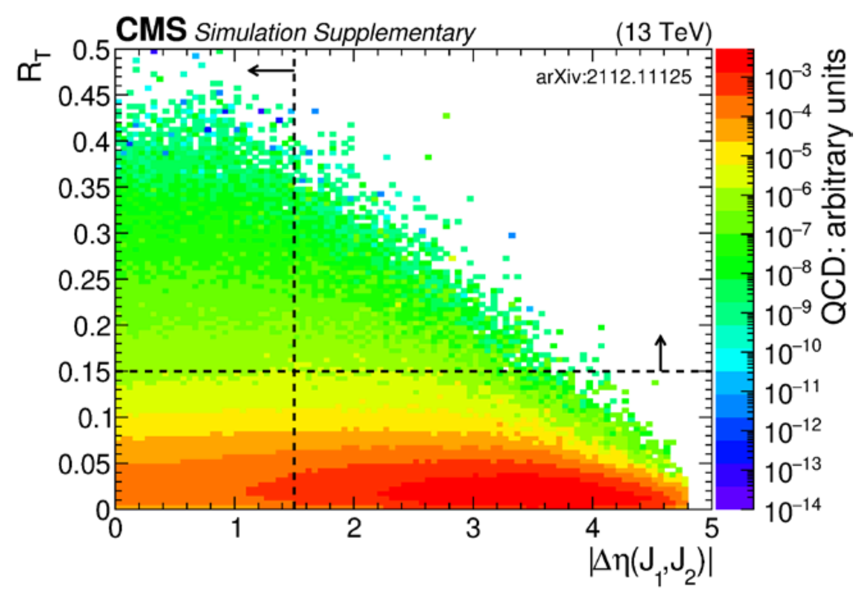


Triggering

- Trigger on jet p_T, H_T
  - ➢ Require low $\Delta\eta(J_1,J_2)$ for high efficiency
- Usually improves signal sensitivity
  - ➢ Most *t*-channel QCD events already rejected by R_T requirement
- m_T > 1500 GeV for trigger efficiency

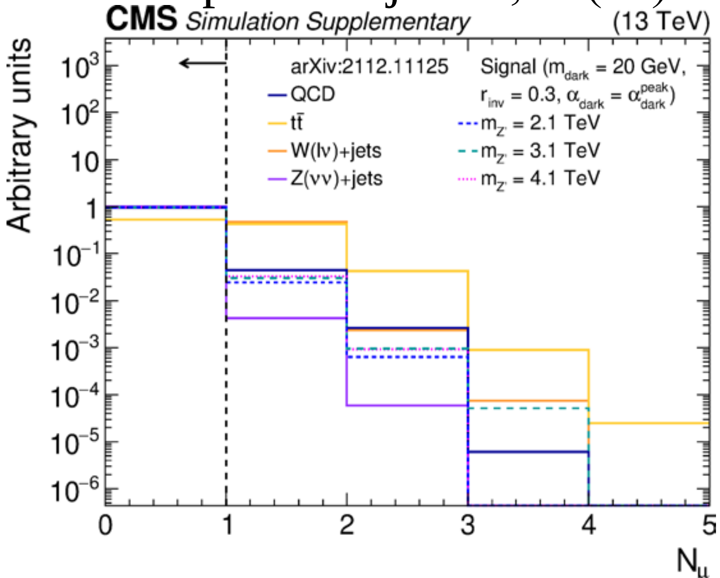


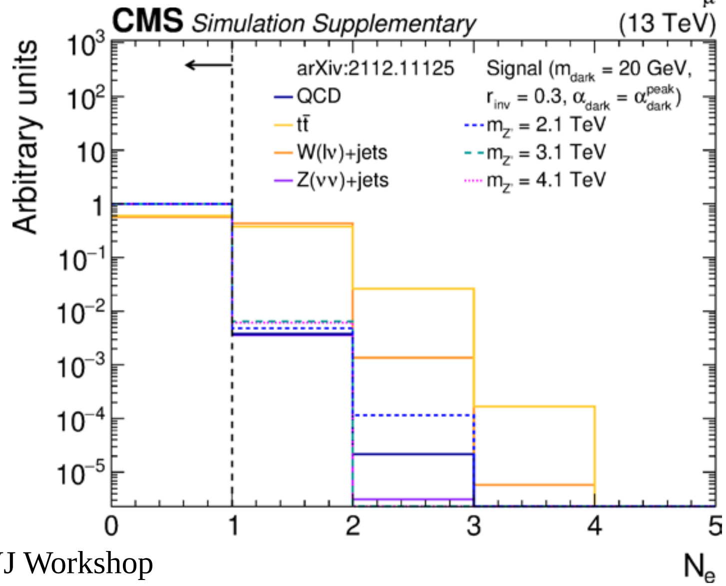




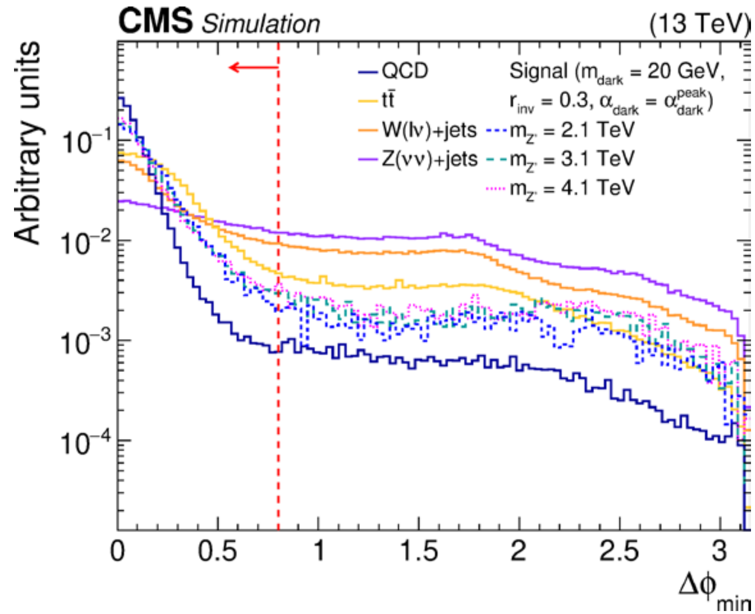
Electroweak Rejection

Veto leptons: reject $t\bar{t}$, $W(\ell\nu)$





Require low $\Delta\varphi_{min}(J_{1,2}, p_T^{miss})$:
Reject $t\bar{t}$, $W(\ell\nu)$, $Z(\nu\nu)$

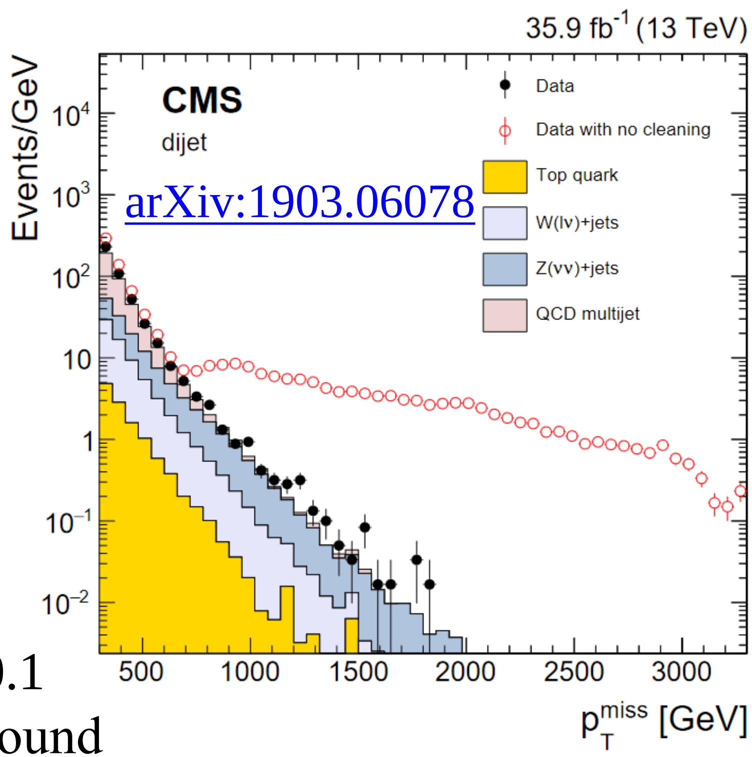


Instrumental Backgrounds

- Centrally-maintained filters reject *most* instrumental sources of artificial high-p_T^{miss} events
  - o But low-$\Delta\varphi$ region ignored by almost all analyses: filters not tuned here
- Major source of jet mismeasurement: nonfunctional ECAL readout channels ("dead" or "hot" cells)

➢ Custom filter vetoing events w/ narrow (AK4) jets w/ $\Delta R(j_{1,2}$, nonfunctional$) < 0.1$ → reject additional 40% of QCD background

  - o Signal efficiency 95%

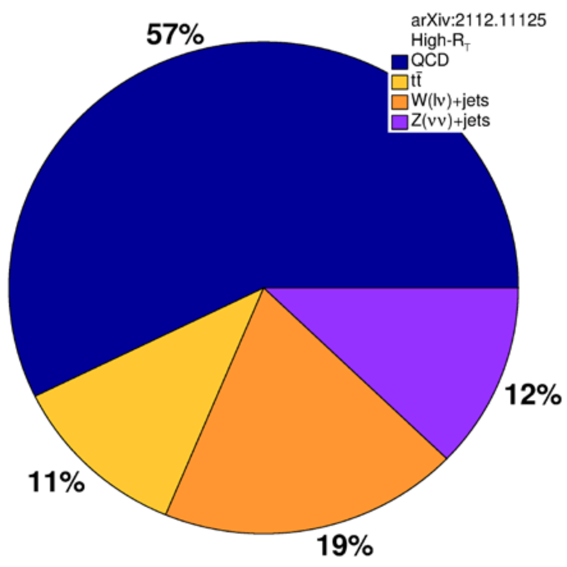


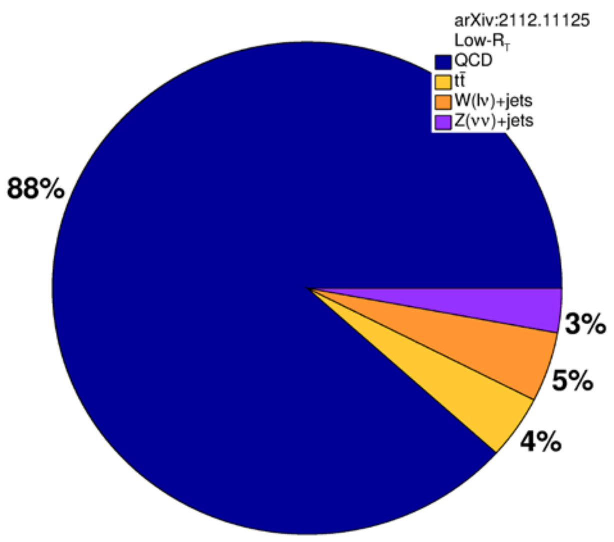
- Misreconstructed jets near barrel-endcap gap in ECAL
  - o Appear at high p_T^{miss} and high m_T
  - o Veto events w/ $p_T(j_1) > 1000$ GeV and $f_\gamma(j_1) > 0.7$

Inclusive Signal Regions

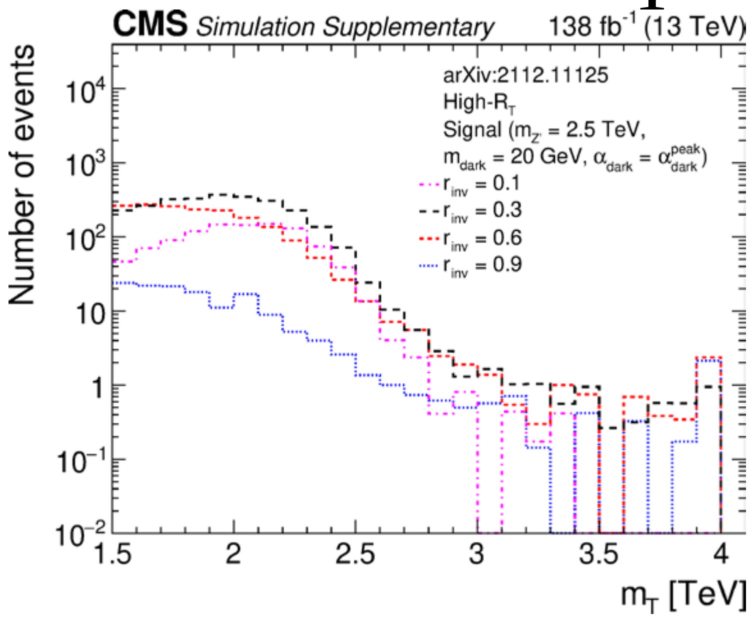
- With all inclusive selection requirements applied:
- If only one signal region were defined, high-R_T ($R_T > 0.25$) would have optimal significance
- Adding separate region low-R_T ($0.15 < R_T < 0.25$) improves expected performance

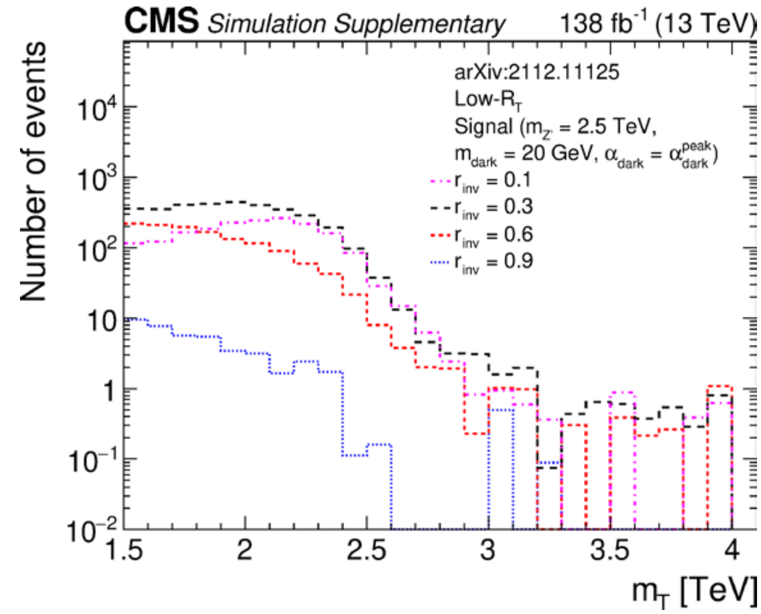
Process	Efficiency [%]
QCD	0.000016
$t\bar{t}$	0.0060
$W(\ell\nu)$+jets	0.0029
$Z(\nu\nu)$+jets	0.0085
signal	~17





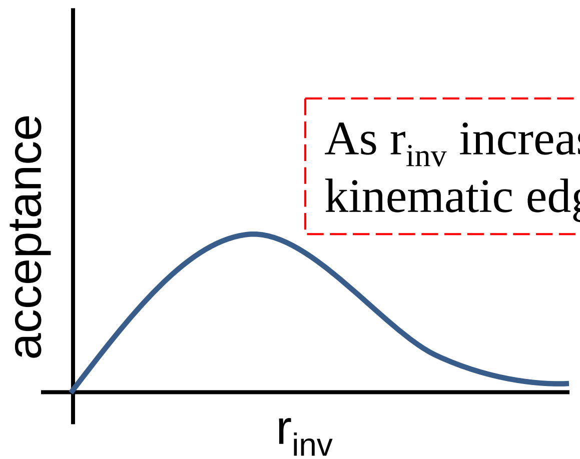
m_T Variations





- r_{inv} has largest impact on signal mass distributions
  - α_{dark} has minor impact; m_{dark} has very little impact

As r_{inv} increases, kinematic edge softens

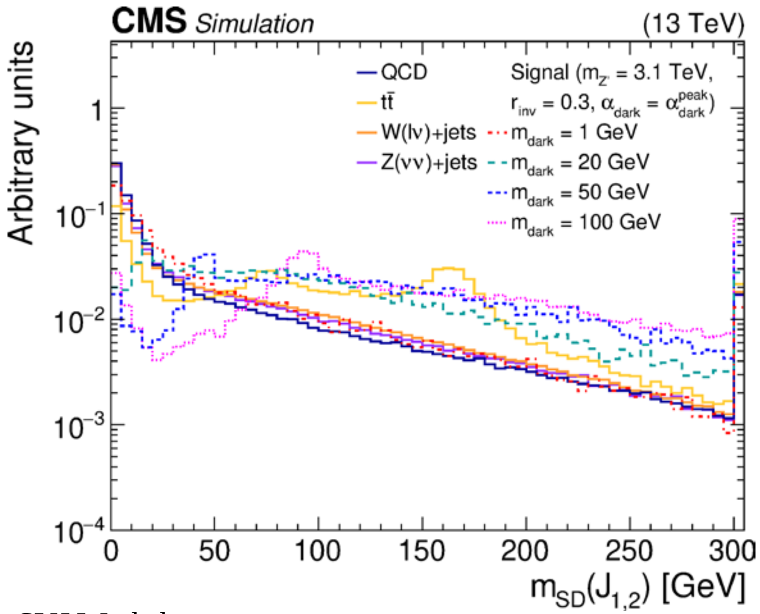


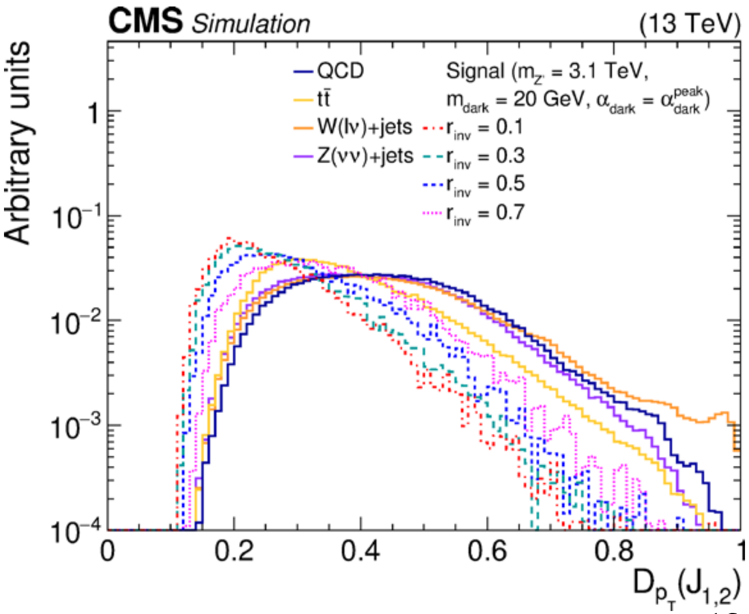
Low r_{inv} signals populate low-R_T signal region

Very high r_{inv} signals look like ISR

Tagging Semivisible Jets

- Various jet substructure variables (& $\Delta\varphi(J, p_T^{miss})$) can weakly discriminate between semivisible jets and SM background jets
  - o Heavy object tagging: m_{SD}, τ_{21}, τ_{32}, $N_2^{(1)}$, $N_3^{(1)}$
  - o Quark-gluon discrimination: D_{p_T}, σ_{major}, σ_{minor}, girth
  - o Flavor (energy fractions): f_γ, $f_{h\pm}$, f_{h0}, f_e, f_μ
- ➢ Combine useful variables into a BDT for strong discrimination!
  - o Background: equal mix of QCD and $t\bar{t}$; signal: mix of many models
  - o Reweight background jet p_T spectrum to match signal: avoid sculpting

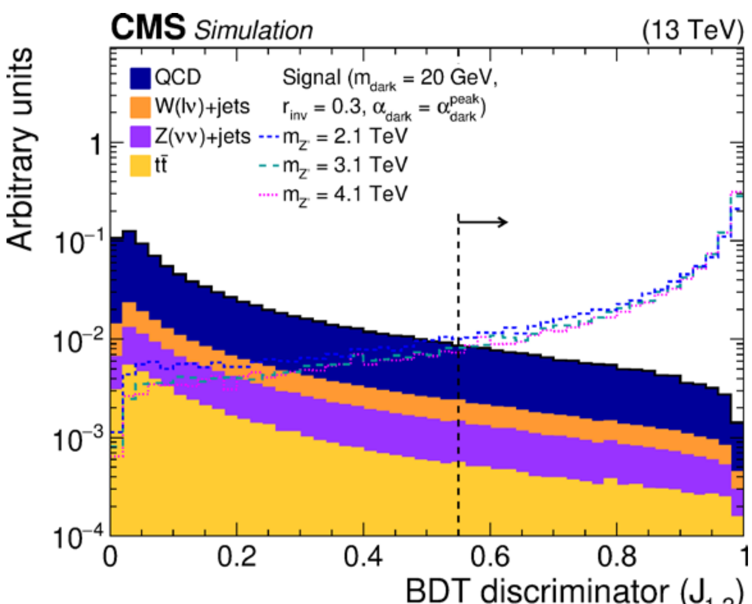


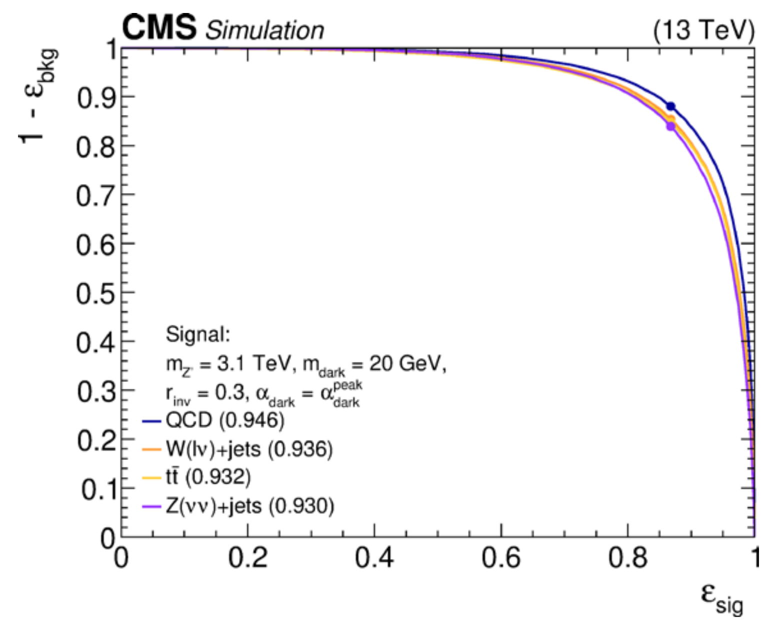


Tagger Performance

	$m_{Z'}$ = 3.1 TeV, m_{dark} = 20 GeV, r_{inv} = 0.3, $\alpha_{dark} = \alpha_{dark}^{peak}$		
	Acc (WP = 0.5)	AUC	$1/\varepsilon_B$ (ε_S = 0.3)
QCD	0.881	0.947	651.4
$t\bar{t}$	0.881	0.931	270.6
W(ℓν)+jets	0.881	0.936	441.5
Z(νν)+jets	0.881	0.930	420.7

- Strong and consistent performance
  - Training on only QCD ($t\bar{t}$) caused misclassification of $t\bar{t}$ (QCD) jets at rate of 10–20%
  - Some inefficiency for signals with high or low m_{dark}
- Working point 0.55 chosen based on background estimation

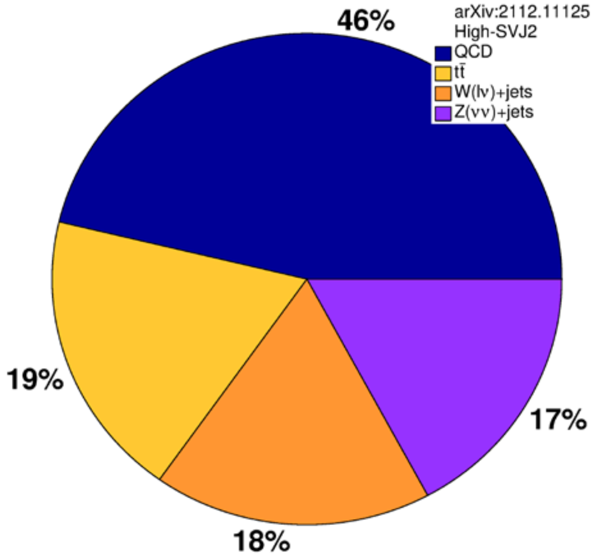


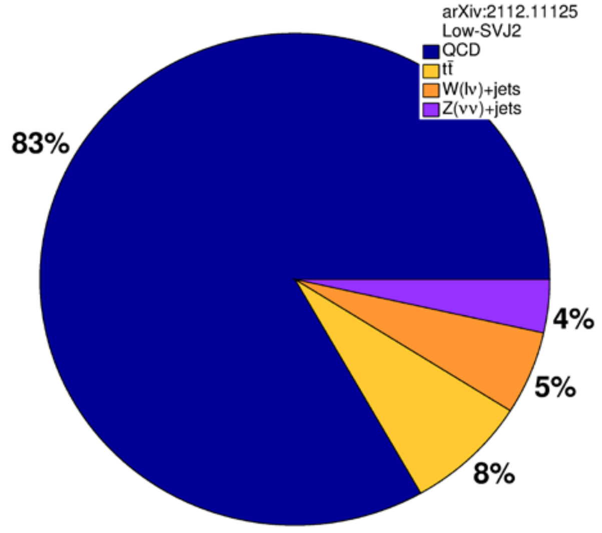


BDT-based Signal Regions

- Start from inclusive signal regions (high-R_T, low-R_T)
- Require both leading wide jets to be tagged as semivisible
  - high-SVJ2, low-SVJ2 regions: strict subsets of inclusive regions

➢ Reduce background by factor ~60 while preserving signal





Background Estimation

- Estimate smoothly-falling SM backgrounds via analytic fit to m_T data
- Primary fit function:

$$x = m_\mathrm{T}/\sqrt{s}$$

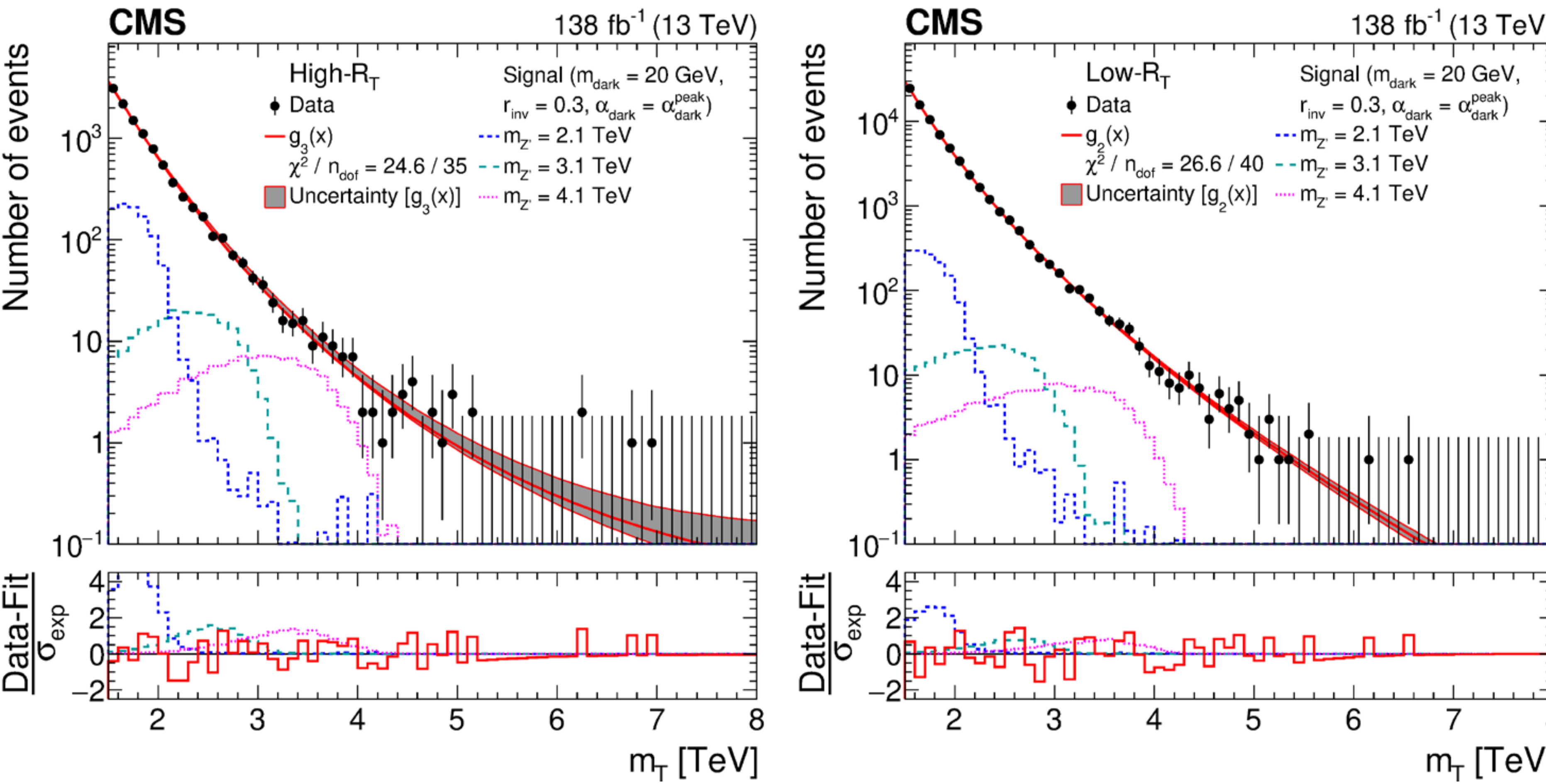
$$g(x) = \exp(p_1 x) x^{p_2(1+p_3\log(x)(1+p_4\log(x)(\cdots)))}$$

  - Perform fits varying sign and magnitude of initial parameter values (necessary to escape false minima)
  - Optimal # parameters for each signal region determined w/ Fisher test
- Several secondary functions (from other resonant searches) employed for bias studies:
  - Ensure that chosen function can fit different possible data distributions
  - Generate toy data with secondary functions, fit w/ primary function
  - $b = (\sigma_{ext} - \sigma_{inj})/\varepsilon_{\sigma_{ext}}$ should be normally distributed ($\mu = 0$, $\sigma = 1$)
  - $|\langle b\rangle| \leq 0.5$ in all cases → fits are sufficiently unbiased

Optimal # Parameters

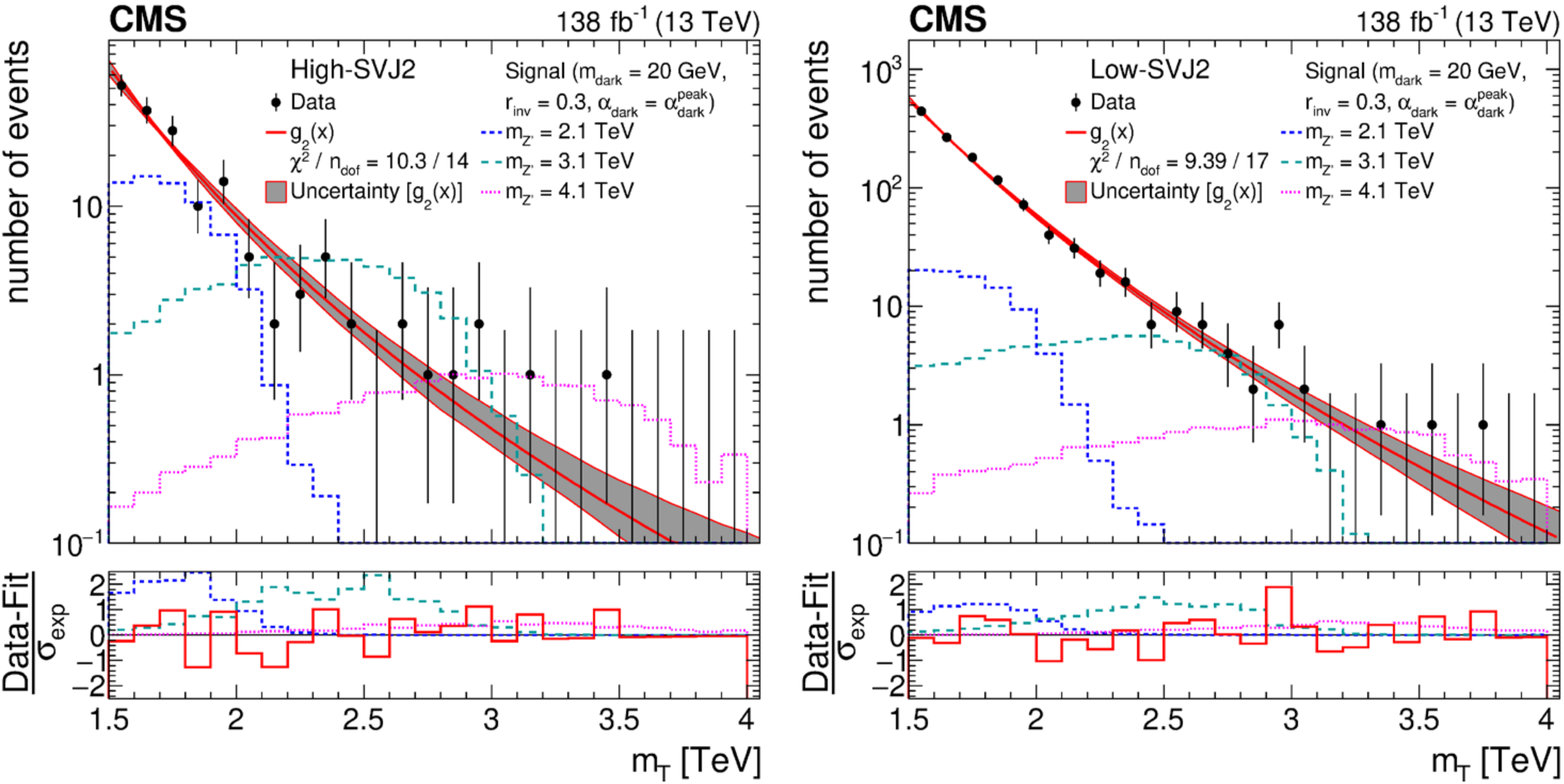
region	***g*(*x*)**
high-R_T	3
low-R_T	2
high-SVJ2	2
low-SVJ2	2

Background Fits (inclusive)



- No significant deviations from SM
  - Small pulls, few if any cases of several contiguous pulls > 0
- Signals shown w/ cross section at observed limit

Background Fits (BDT-based)

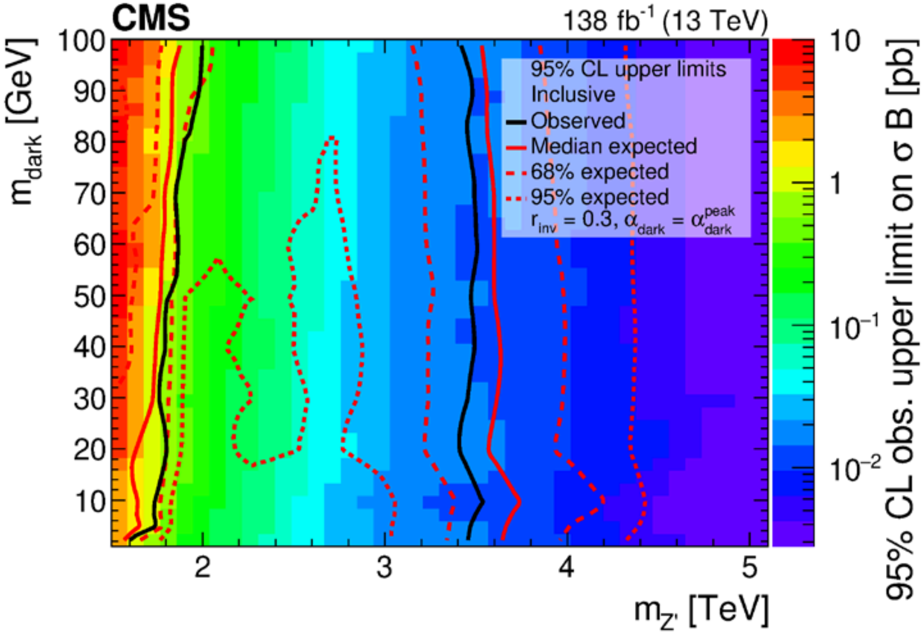


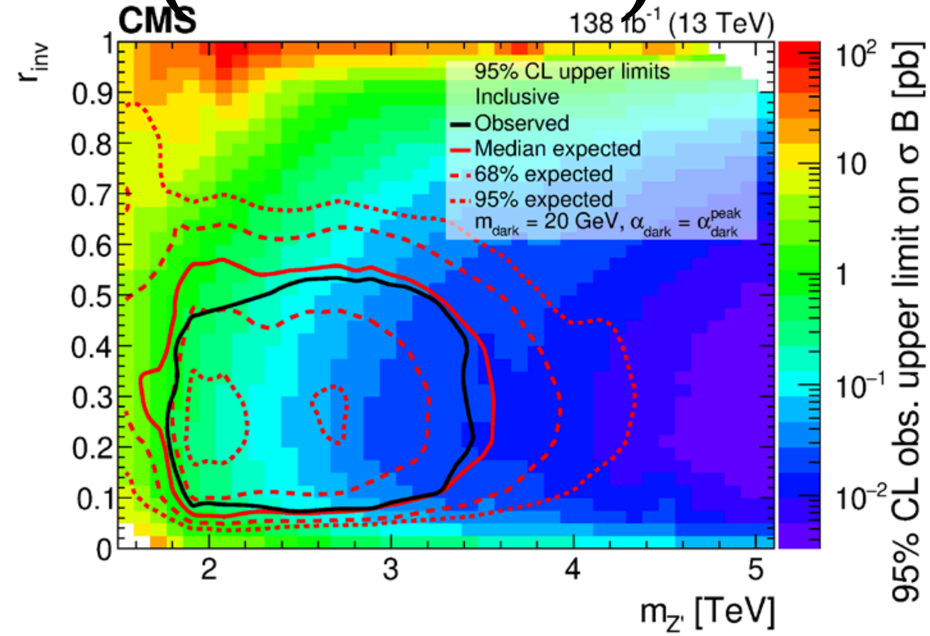
- No significant deviations from SM
  - Small pulls, few if any cases of several contiguous pulls > 0
- Signals shown w/ cross section at observed limit

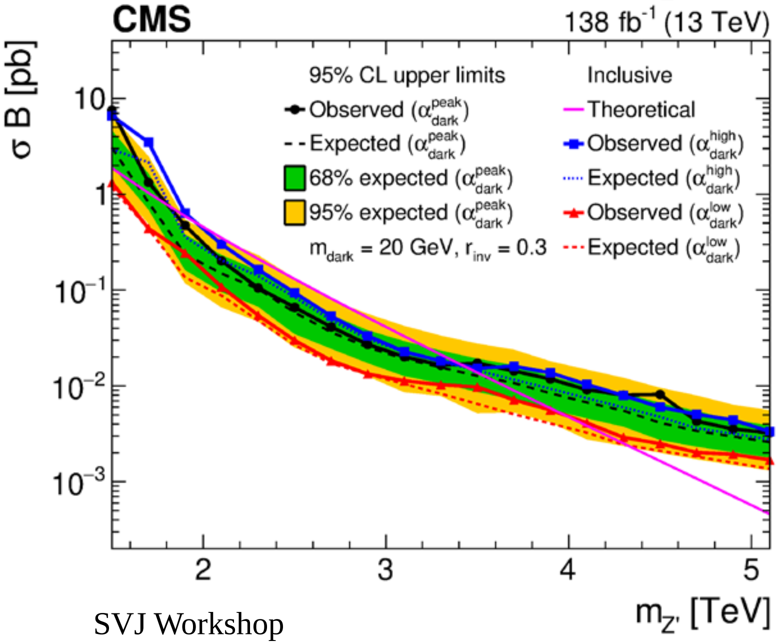
Systematic Uncertainties

- Signal:
  - Experimental: (uncorrelated between years of data-taking)
    - Luminosity, trigger efficiency, **jet energy corrections** (up to 12%), jet energy resolution, pileup, statistical uncertainties in simulated samples
  - Theoretical: (correlated between years of data-taking)
    - PDFs, renormalization/factorization scale, parton shower modeling (ISR/FSR), **jet energy scale/composition** (up to 21%)
- Background:
  - Fit parameters: freely floating, uncertainties arise from statistical uncertainty in data
  - **Fit normalizations**: also freely floating, can change by up to 10% $\rightarrow$ *most impactful uncertainty*

Observed Limits (inclusive)



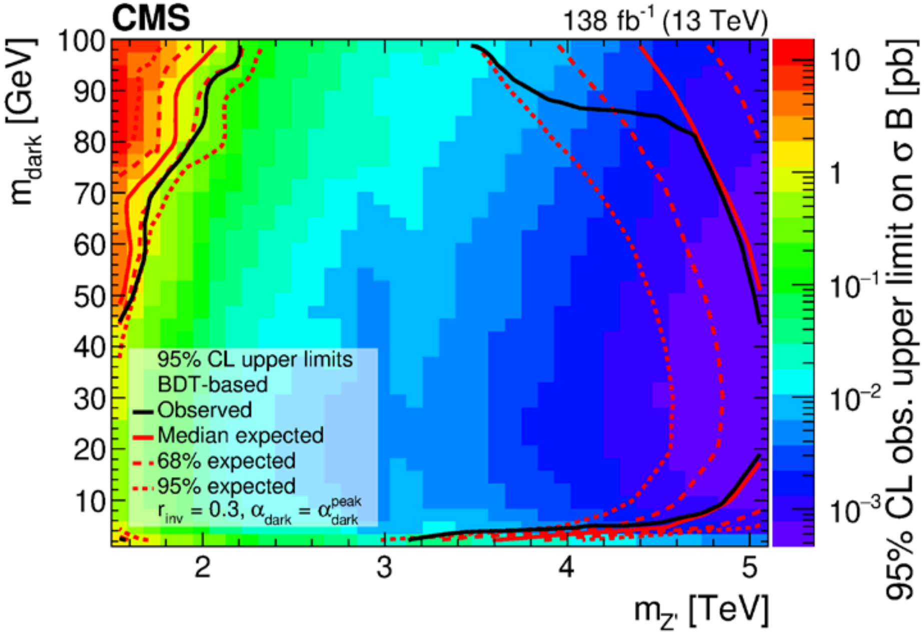


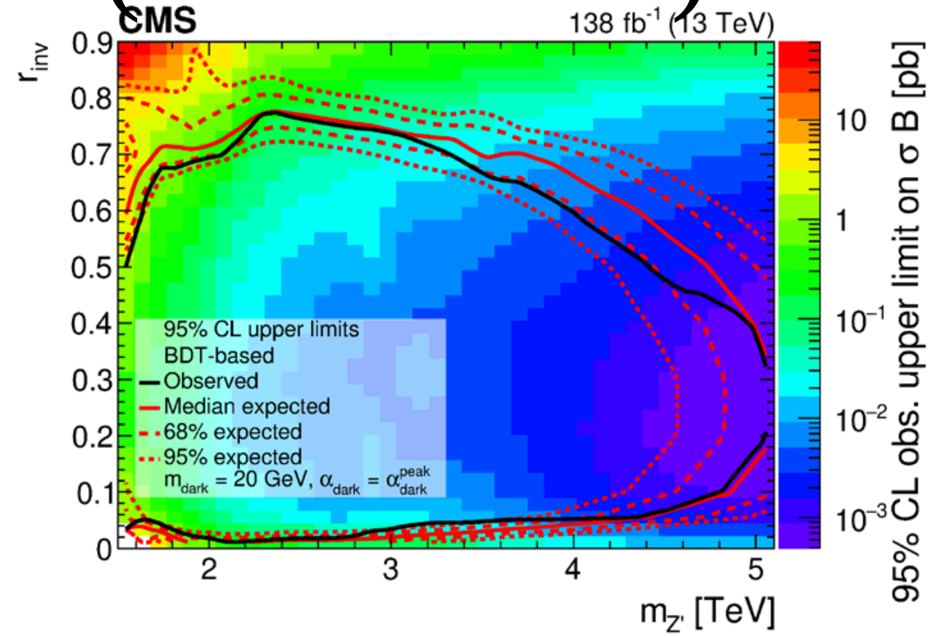


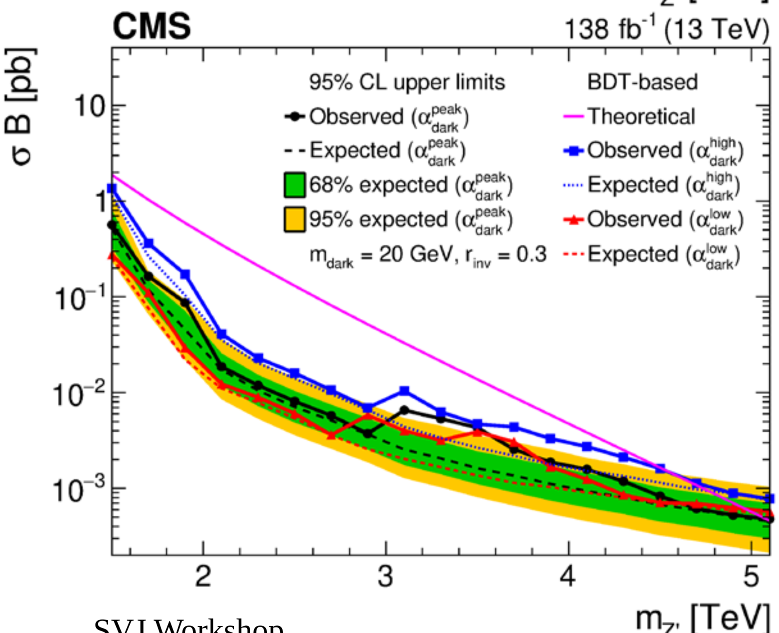
Observed (expected) exclusions:

- $1.5 < m_{Z'} < 4.0$ TeV ($1.5 < m_{Z'} < 4.3$ TeV)
- Depending on $m_{Z'}$:
  - $0.07 < r_{inv} < 0.53$ ($0.06 < r_{inv} < 0.57$)
  - All m_{dark}, α_{dark} variations

Observed Limits (BDT-based)







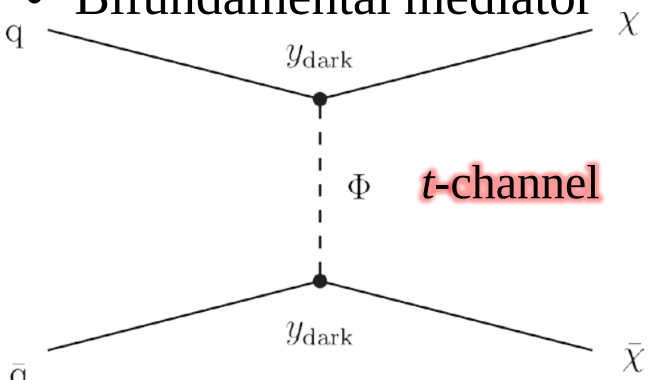
Observed (expected) exclusions:

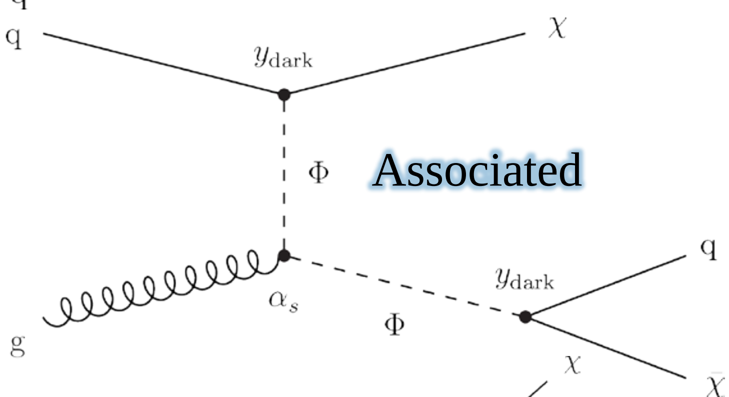
- $1.5 < m_{Z'} < 5.1$ TeV ($1.5 < m_{Z'} < 5.1$ TeV)
- Depending on $m_{Z'}$:
  - $0.01 < r_{inv} < 0.77$ ($0.01 < r_{inv} < 0.78$)
  - All m_{dark}, α_{dark} variations

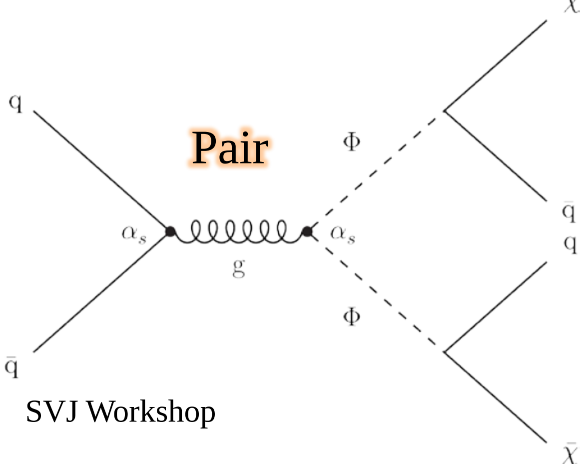
➢ Signal parameters excluded for wider range in $m_{Z'}$ vs. inclusive search

Future Semivisible Jet Searches

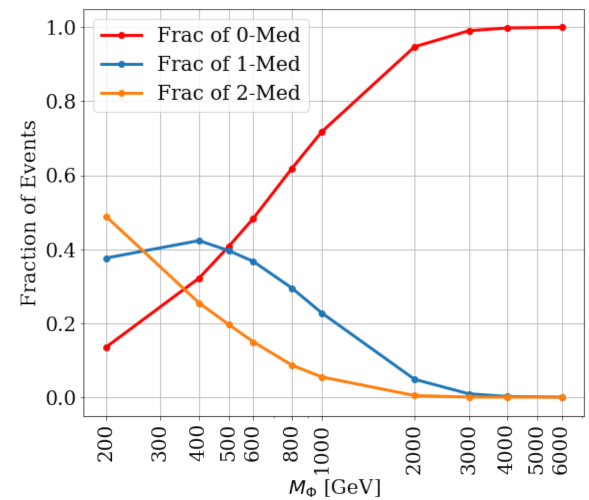
- Bifundamental mediator



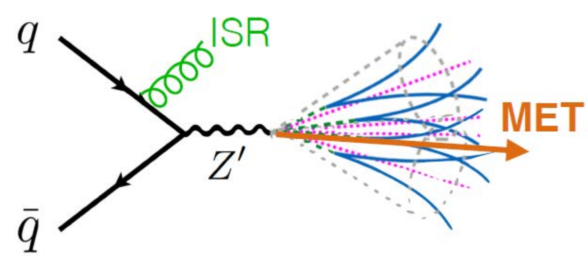




- Challenges:
  - Additional parameter y_{dark}
  - Non-resonant diagrams dominate as mediator mass increases

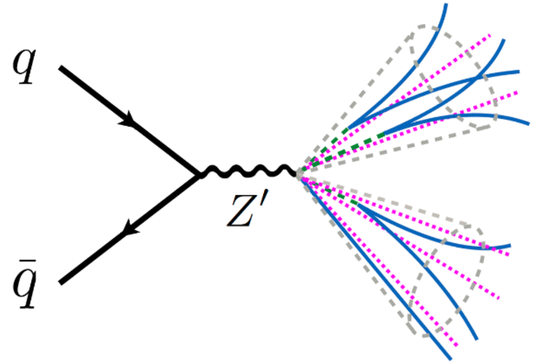


- Low-mass Z' mediators (boosted)



Conclusions

- CMS search directly excludes a large portion of semivisible jet model space for the first time
  - Sensitivity to a broad range of $m_{Z'}$, m_{dark}, r_{inv} values
- Dual strategy provides both generality and sensitivity
  - Inclusive search can be reinterpreted for any kinematically similar signal
  - BDT-based search improves background rejection by almost two orders of magnitude (first SVJ tagger applied to data)
- Ongoing search program targets different mediators, regions of model space
  - Will employ new techniques such as autoencoders to increase sensitivity and reinterpretability even further
  - Run 3 plans developing: opportunities to improve triggers, etc.
- Stay tuned for more SVJ results from CMS!



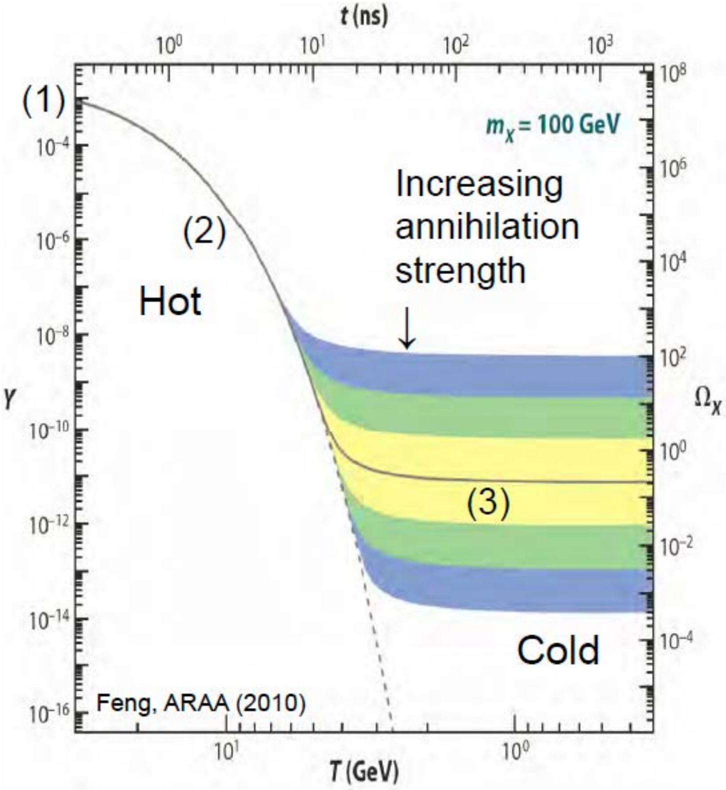
Backup

Strongly-Coupled Hidden Sector References

- M. J. Strassler and K. M. Zurek, "Echoes of a hidden valley at hadron colliders", Phys. Lett. B 651 (2007) 374, arXiv:hep-ph/0604261.
- Y. Bai and P. Schwaller, "Scale of dark QCD", Phys. Rev. D 89 (2014) 063522, arXiv:1306.4676.
- T. Cohen, M. Lisanti, and H. K. Lou, "Semivisible jets: Dark matter undercover at the LHC", Phys. Rev. Lett. 115 (2015) 171804, arXiv:1503.00009.
- P. Schwaller, D. Stolarski, and A. Weiler, "Emerging jets", JHEP 05 (2015) 059, arXiv:1502.05409.
- G. D. Kribs and E. T. Neil, "Review of strongly-coupled composite dark matter models and lattice simulations", Int. J. Mod. Phys. A 31 (2016) 1643004, arXiv:1604.04627.
- S. Knapen, S. Pagan Griso, M. Papucci, and D. J. Robinson, "Triggering Soft Bombs at the LHC", JHEP 08 (2017) 076, arXiv:1612.00850.
- T. Cohen, M. Lisanti, H. K. Lou, and S. Mishra-Sharma, "LHC searches for dark sector showers", JHEP 11 (2017) 196, arXiv:1707.05326.
- H. Beauchesne, E. Bertuzzo, G. Grilli di Cortona, and Z. Tabrizi, "Collider phenomenology of Hidden Valley mediators of spin 0 or 1/2 with semivisible jets", JHEP 08 (2018) 030, arXiv:1712.07160.
- S. Renner and P. Schwaller, "A flavoured dark sector", JHEP 08 (2018) 052, arXiv:1803.08080.
- CMS Collaboration, "Search for new particles decaying to a jet and an emerging jet", JHEP 02 (2019) 179, arXiv:1810.10069.
- H. Beauchesne, E. Bertuzzo, and G. Grilli Di Cortona, "Dark matter in Hidden Valley models with stable and unstable light dark mesons", JHEP 04 (2019) 118, arXiv:1809.10152.
- M. Park and M. Zhang, "Tagging a jet from a dark sector with jet-substructures at colliders", Phys. Rev. D 100 (2019) 115009, arXiv:1712.09279.
- T. Cohen, J. Doss, and M. Freytsis, "Jet substructure from dark sector showers", JHEP 09 (2020) 118, arXiv:2004.00631.
- J. Alimena et al., "Searching for long-lived particles beyond the Standard Model at the Large Hadron Collider", J. Phys. G 47 (2020) 090501, arXiv:1903.04497.
- E. Bernreuther, F. Kahlhoefer, M. Krämer, and P. Tunney, "Strongly interacting dark sectors in the early Universe and at the LHC through a simplified portal", JHEP 01 (2020) 162, arXiv1907.04346.
- E. Bernreuther, T. Finke, F. Kahlhoefer, M. Krämer, and A. Mück, "Casting a graph net to catch dark showers", SciPost Phys. 10 (2021) 046, arXiv:2006.08639.
- H. Mies, C. Scherb, and P. Schwaller, "Collider constraints on dark mediators", JHEP 04 (2021) 049, arXiv:2011.13990.
- C. Cesarotti, M. Reece, and M. Strassler, "Spheres To Jets: Tuning Event Shapes with 5d Simplified Models", JHEP 05 (2021) 096, arXiv:2009.08981.
- S. Knapen, J. Shelton, and D. Xu, "Perturbative benchmark models for a dark shower search program", Phys. Rev. D 103 (2021) 115013, arXiv:2103.01238.
- CMS Collaboration, "Search for resonant production of strongly-coupled dark matter in proton-proton collisions at 13 TeV", JHEP 06 (2022) 156, arXiv:2112.11125.
- F. Canelli, A. de Cosa, L. Le Pottier, J. Niedziela, K. Pedro, M. Pierini, "Autoencoders for Semivisible Jet Detection", JHEP 02 (2022) 074, arXiv:2112.02864.
- G. Albouy et al., "Theory, phenomenology, and experimental avenues for dark showers: a Snowmass 2021 report", arXiv:2203.09503.
- C. Cazzaniga, A. de Cosa, "Leptons lurking in semi-visible jets at the LHC", arXiv:2206.03909.

Dark Matter Relic Abundance

- Dark matter production, annihilation at equilibrium in early universe
- Universe expands and cools:
  stops DM production, then annihilation
- WIMPs imply fixed DM abundance (bottom)



$$\Omega_{\rm DM}h^2 = 0.12$$

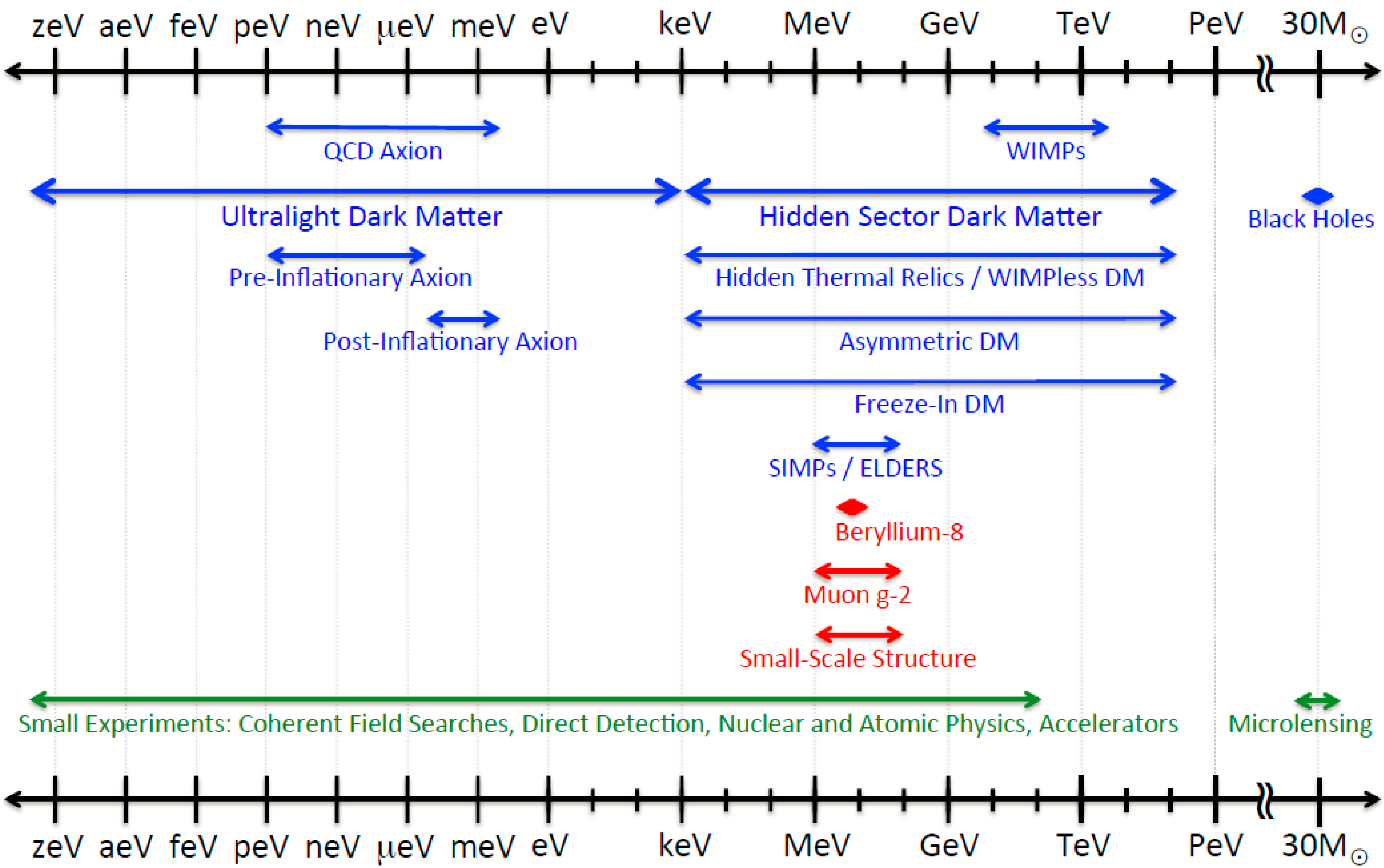
$$\Omega_{\rm DM}h^2 = \frac{0.2 \times 10^{-9}{\rm GeV}^{-2}}{\langle \sigma v \rangle}$$

$$\langle \sigma v \rangle \sim 10^{-9}{\rm GeV}^{-2}$$ (weak cross section)

$$\Omega_{\rm DM} \sim 0.2$$

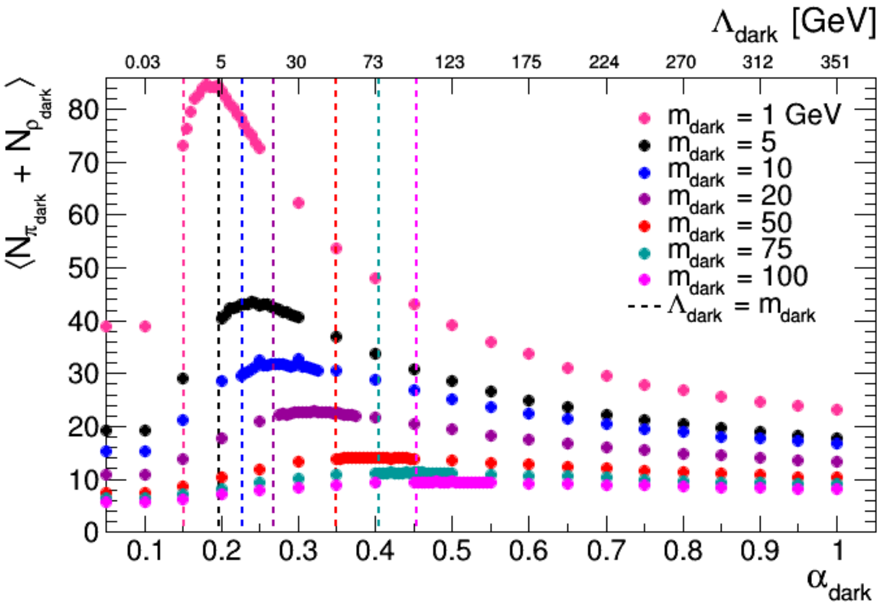
Dark Matter Landscape

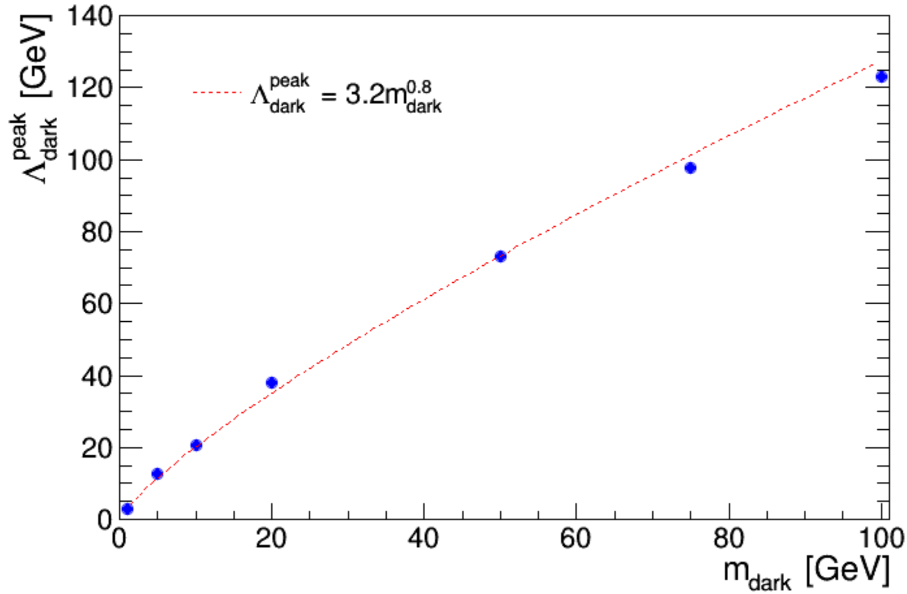
Dark Sector Candidates, Anomalies, and Search Techniques

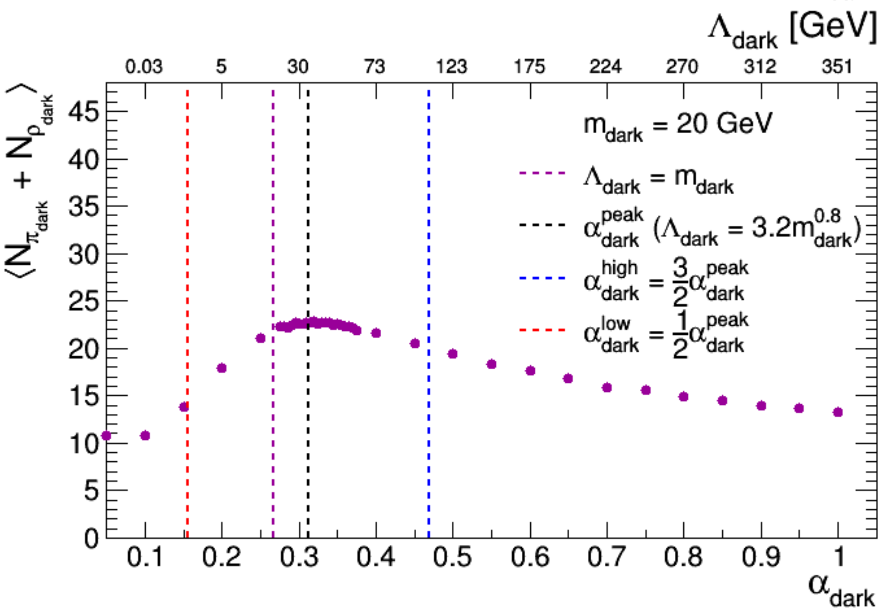


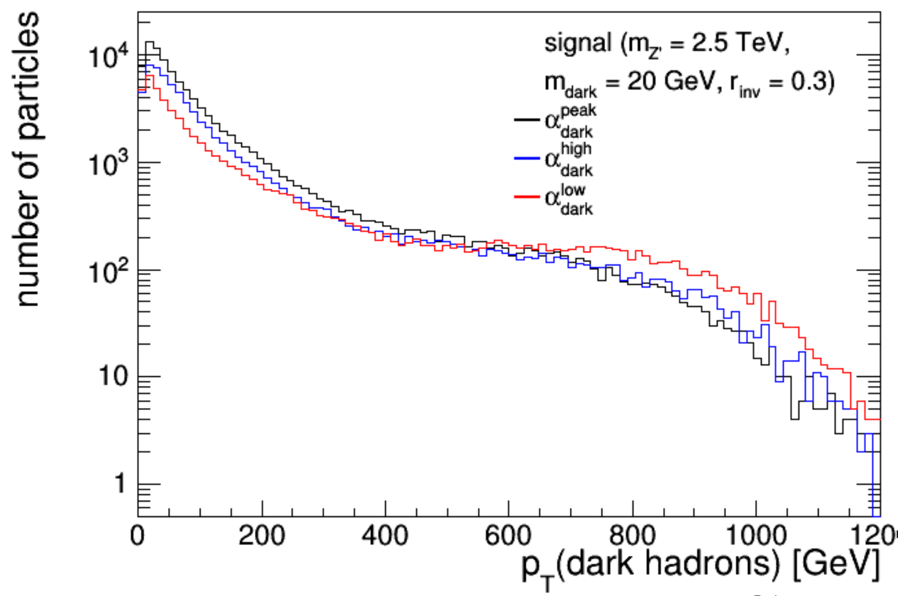
arXiv:1707.04591, G. Landsberg

α_{dark} variations

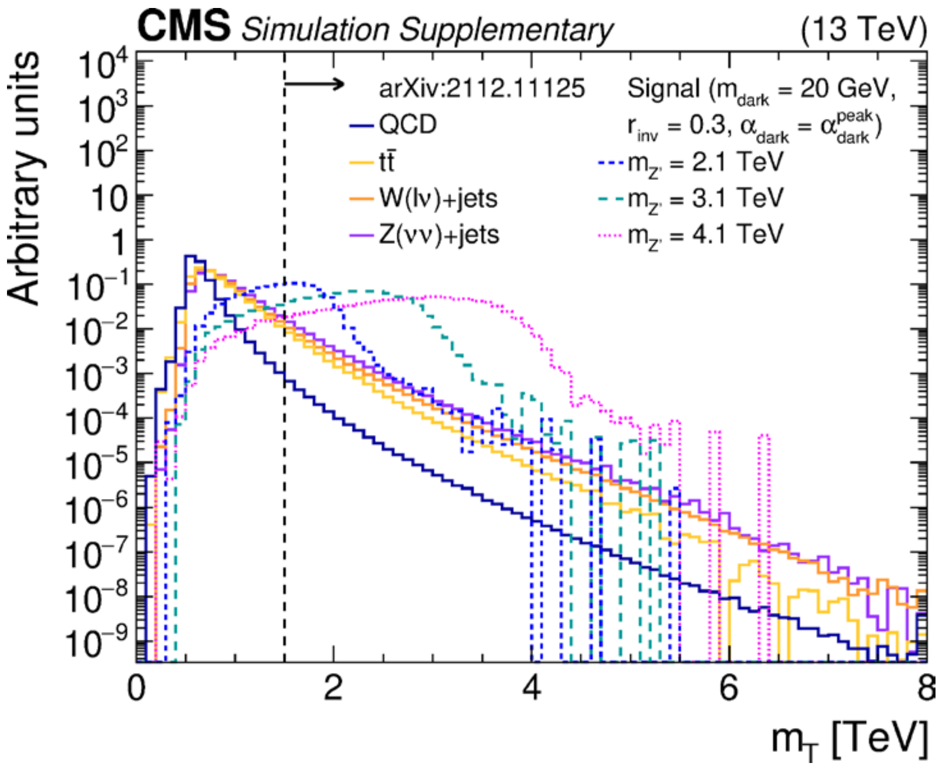


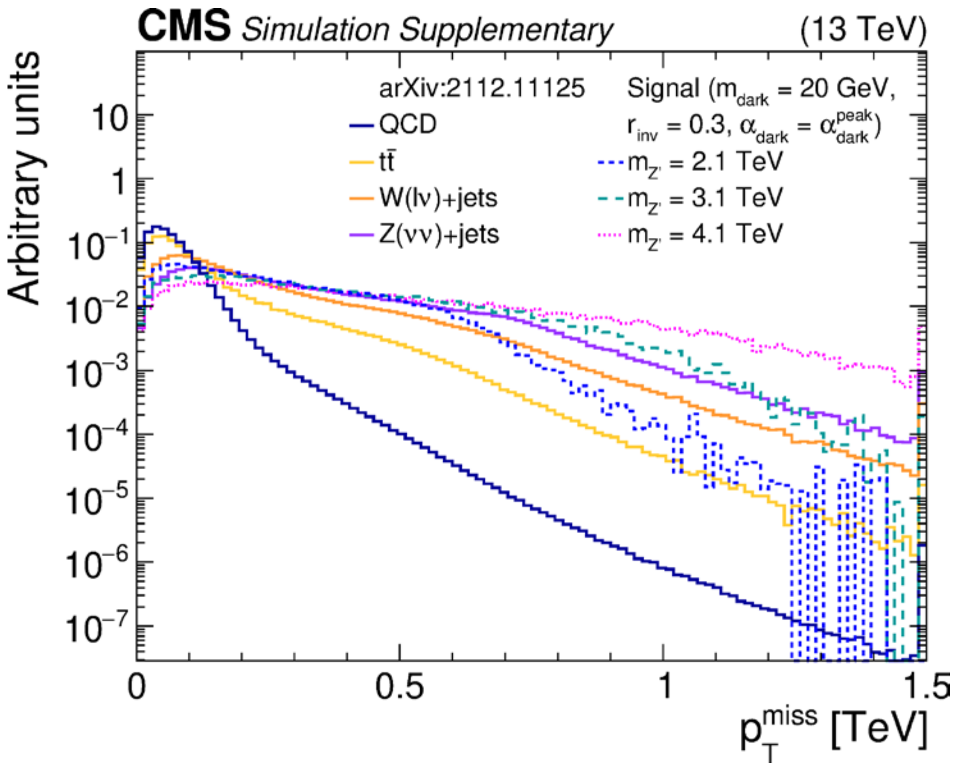






Semivisible Jet Kinematics





Event Selection

Preselection

- $N_J \geq 2$
- $p_T(J_1,J_2) > 200$ GeV, $|\eta(J_1,J_2)| < 2.4$, → Signal topology
  $J_{1,2}$ pass noise rejection → Data quality
- $R_T \equiv p_T^{miss}/m_T > 0.15$ → Reject QCD
- $\Delta\eta(J_1,J_2) < 1.5$
- $m_T > 1500$ GeV → Trigger efficiency
- e/μ veto ($p_T > 10$ GeV, $|\eta| < 2.4$) → Reject $t\bar{t}$, $W(\ell\nu)$
- p_T^{miss} filters
- Custom dead ECAL cell filter: veto events w/ $\Delta R(j_{1,2}, c_{nonfunctional}) < 0.1$
- Inactive HCAL filter (2018 only): veto events w/ $p_T(j) > 30$ GeV, $-3.05 < \eta(j) < -1.35$, $-1.62 < \varphi(j) < -0.82$ → Data quality

Final Selection

- Gap jet filter: veto events w/ $p_T(j_1) > 1000$ GeV, $f_\gamma(j_1) > 0.7$ → Data quality
- $\Delta\varphi_{min}(J_{1,2},p_T^{miss}) < 0.80$ → Reject $t\bar{t}$, $W(\ell\nu)$, $Z(\nu\nu)$

Control Regions

Single Muon

- $N_J \geq 2$
- $p_T(J_1,J_2) > 200$ GeV, $|\eta(J_1,J_2)| < 2.4$, $J_{1,2}$ pass noise rejection
- $R_T \equiv p_T^{miss}/m_T > 0.15$
- $\Delta\eta(J_1,J_2) < 1.5$
- e veto ($p_T > 10$ GeV, $|\eta| < 2.4$)
- $N_\mu \geq 1$ ($p_T > 50$ GeV, $|\eta| < 2.4$, medium ID, $I_{mini} < 0.2$, HLT match)

- Used for trigger efficiency measurement
- Corresponding Single Muon High-$\Delta\eta$ region used for CR trigger efficiency measurement

---

- Used for data quality studies
  - Statistically limited, but otherwise kinematically similar to signal region
- $\Delta\eta$ range maximizes data yield (in fully efficient region, $m_T > 1850$ GeV)
  - $1500 < m_T < 1850$ GeV can be used w/ trigger efficiency correction applied

High-$\Delta\eta$

Preselection, except:

- $1.5 < \Delta\eta(J_1,J_2) < 2.2$
- $m_T > 1850$ GeV

Cutflows

Selection	QCD	$t\bar{t}$	W+jets	Z+jets	$r_{inv} = 0.3$
$p_T(J_{1,2}) > 200\,GeV$, $\eta(J_{1,2}) < 2.4$	1.2	6.4	2.0	1.3	83.5
$R_T > 0.15$	1.3	12.1	18.5	34.6	39.7
$\Delta\eta(J_1, J_2) < 1.5$	94.9	88.0	85.1	78.8	80.0
$m_T > 1.5\,TeV$	0.20	3.1	4.0	5.6	81.8
$N_\mu = 0$	93.0	62.0	66.0	99.5	96.8
$N_e = 0$	99.6	59.8	57.3	99.6	99.4
p_T^{miss} filters	99.5	99.9	99.9	99.9	99.8
$\Delta R(j_{1,2}, c_{nonfunctional}) > 0.1$	60.6	95.1	95.2	95.6	95.2
veto $f_\gamma(j_1) > 0.7$ & $p_T(j_1) > 1.0\,TeV$	99.7	99.7	99.6	99.7	99.7
$\Delta\phi_{min} < 0.8$	94.8	81.7	61.8	44.7	87.7
Efficiency [%]	1.6e-05	0.0060	0.0029	0.0085	17
high-R_T	9.0	29.5	38.8	39.1	45.2
low-R_T	91.0	70.5	61.2	60.9	54.8
high-SVJ2	0.093	0.62	0.46	0.69	34.6
low-SVJ2	1.1	1.7	0.92	0.94	42.3

Variable Definitions

- Girth: $g = \sum_i \frac{p_{\mathrm{T},i}}{p_{\mathrm{T,jet}}} r_i$
- Major/minor axes:

$$\mathcal{M} = \begin{bmatrix} \sum_i p_{\mathrm{T},i}^2 \Delta\eta_i^2 & -\sum_i p_{\mathrm{T},i}^2 \Delta\eta_i \Delta\phi_i \\ -\sum_i p_{\mathrm{T},i}^2 \Delta\eta_i \Delta\phi_i & \sum_i p_{\mathrm{T},i}^2 \Delta\phi_i^2 \end{bmatrix}$$

$$\sigma_{\mathrm{major}} = \sqrt{\lambda_1 / \sum_i p_{\mathrm{T},i}^2}$$

$$\sigma_{\mathrm{minor}} = \sqrt{\lambda_2 / \sum_i p_{\mathrm{T},i}^2}$$

- p_TD: $p_{\mathrm{T}}D = \frac{\sqrt{\sum_i p_{\mathrm{T},i}^2}}{\sum_i p_{\mathrm{T},i}}$
- Nsubjettiness: $\tau_{21} = \tau_2/\tau_1$, $\tau_{32} = \tau_3/\tau_2$

$$\tau_N^{(\beta)} = \frac{1}{\sum_k p_{\mathrm{T},k} R_0} \sum_k p_{\mathrm{T},k} \min\{\Delta R_{1,k}^{(\beta)}, \Delta R_{2,k}^{(\beta)}, \ldots, \Delta R_{N,k}^{(\beta)}\}$$

- Energy correlation functions:

$${}_v e_n^{(\beta)} = \sum_{1 \le i_1 < \cdots < i_n \le n_{\mathrm{const.}}} z_{i_1} \ldots z_{i_n} \prod_{m=1}^{v} \min_{s<t \in \{i_1,\ldots,i_n\}}^{(m)} \left\{ \theta_{st}^{\beta} \right\}$$

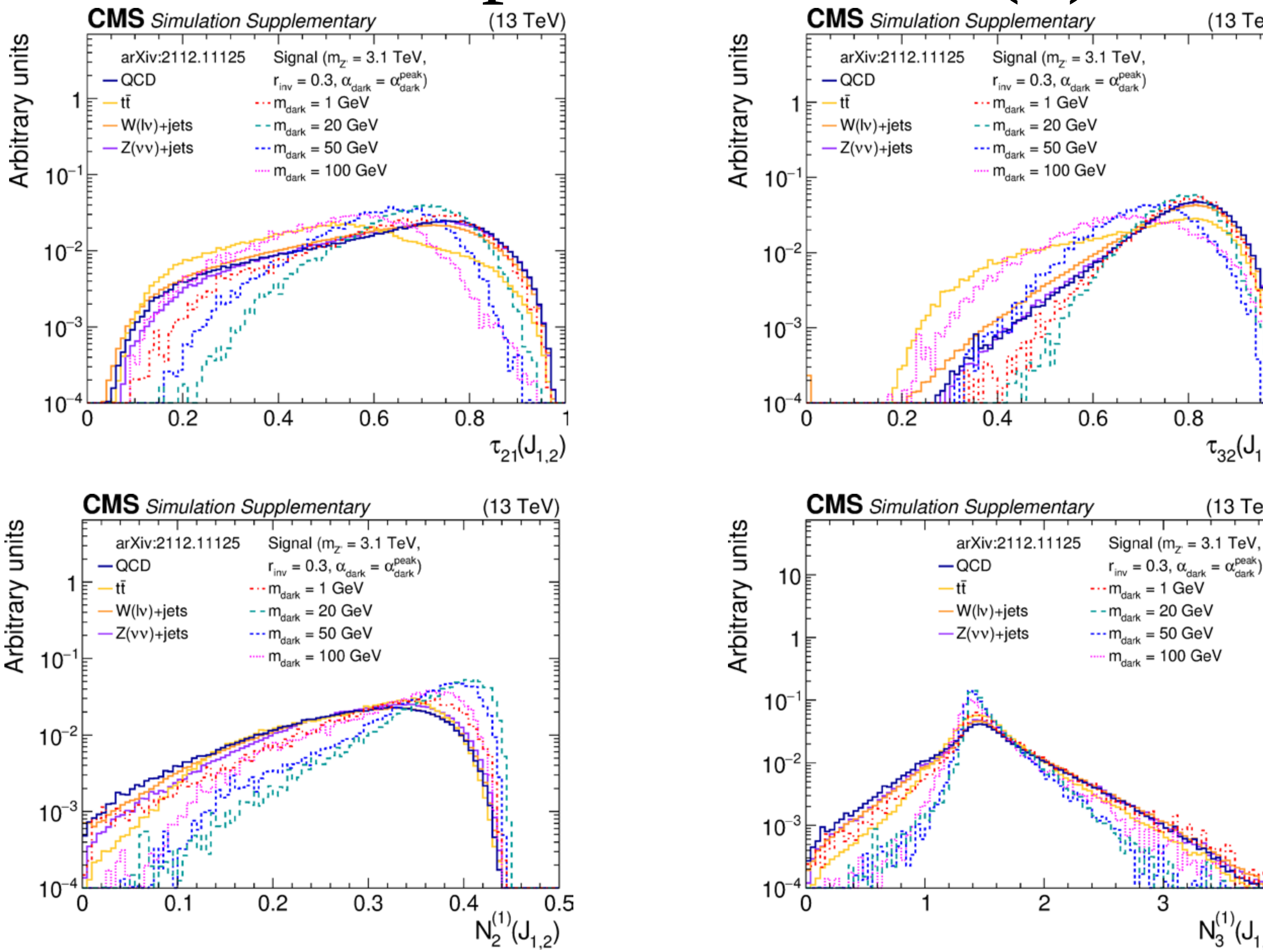
$$N_2^{(1)} = \frac{{}_2e_3^{(1)}}{\left({}_1e_2^{(1)}\right)^2} \qquad N_3^{(1)} = \frac{{}_2e_4^{(1)}}{\left({}_1e_3^{(1)}\right)^2}$$

$$m_{\mathrm{T}}^2 = \left[E_{\mathrm{T,JJ}} + E_{\mathrm{T}}^{\mathrm{miss}}\right]^2 - \left[\vec{p}_{\mathrm{T,JJ}} + \vec{p}_{\mathrm{T}}^{\mathrm{miss}}\right]^2$$
$$= m_{\mathrm{JJ}}^2 + 2p_{\mathrm{T}}^{\mathrm{miss}} \left[\sqrt{m_{\mathrm{JJ}}^2 + p_{\mathrm{T,JJ}}^2} - p_{\mathrm{T,JJ}} \cos(\phi_{\mathrm{JJ,miss}})\right]$$

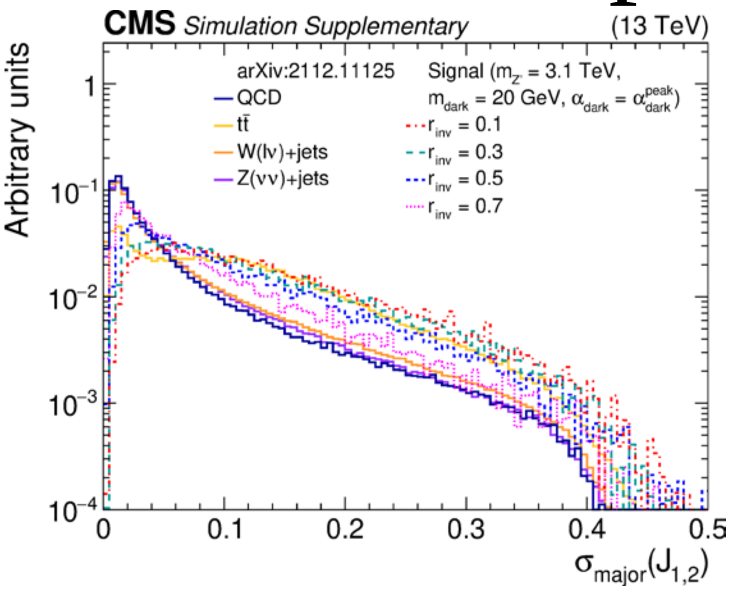
Soft Drop Mass

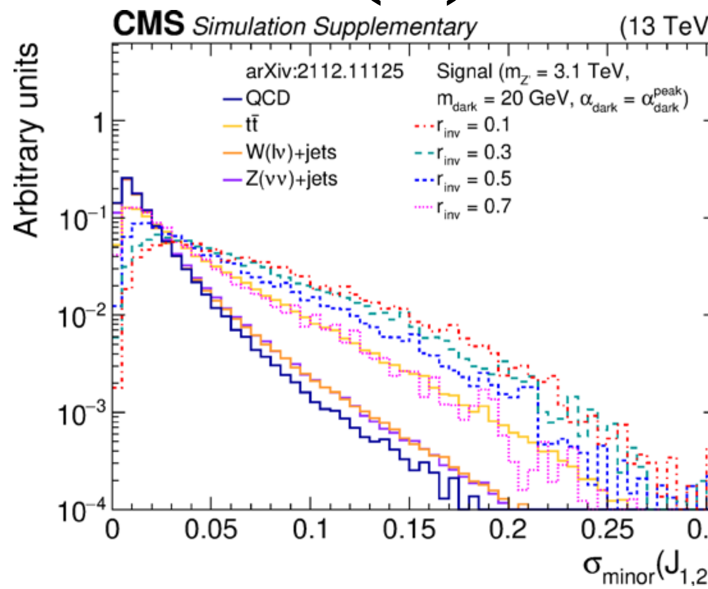
- Start w/ jet clustered by anti-k_t algorithm w/ R = 0.8
- Recluster jet constituents w/ Cambridge-Aachen algorithm
  - o Undo clustering one step at a time
  - o Get two subjets j_1, j_2
  - o Check condition: $\frac{\min(p_{Tj1}, p_{Tj2})}{p_{Tj1} + p_{Tj2}} > z_{cut} \times \left(\frac{\Delta R_{12}}{R_0}\right)^{\beta}$
  - o If met, then keep whole jet and stop
  - o If not met, keep higher p_T subjet and repeat
- CMS uses $z_{cut} = 0.1$, $\beta = 0$
- See arXiv:1402.2657
- Effect: drop soft constituents (at wide angles)
  → remove ISR, underlying event, pileup
- Mass calculation: find invariant mass from softdrop subjet 4-vectors

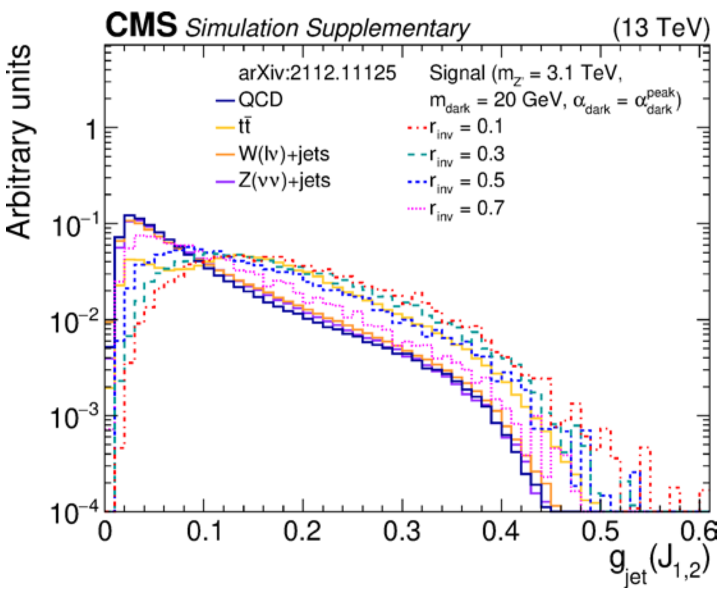
BDT Input Variables (1)

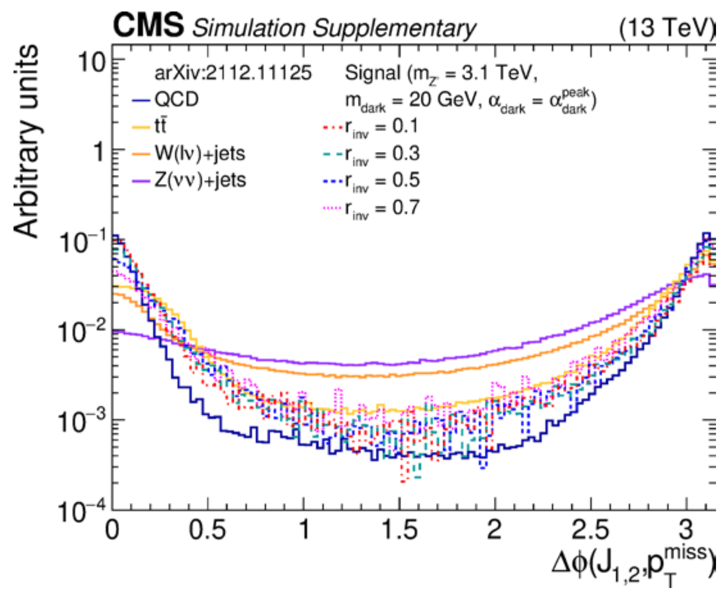


BDT Input Variables (2)

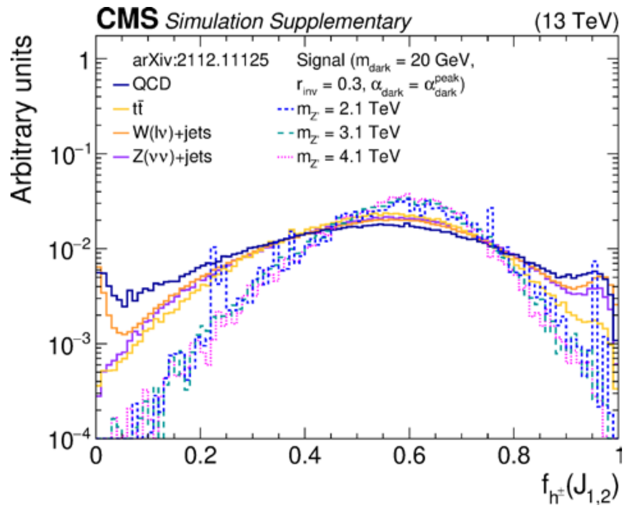


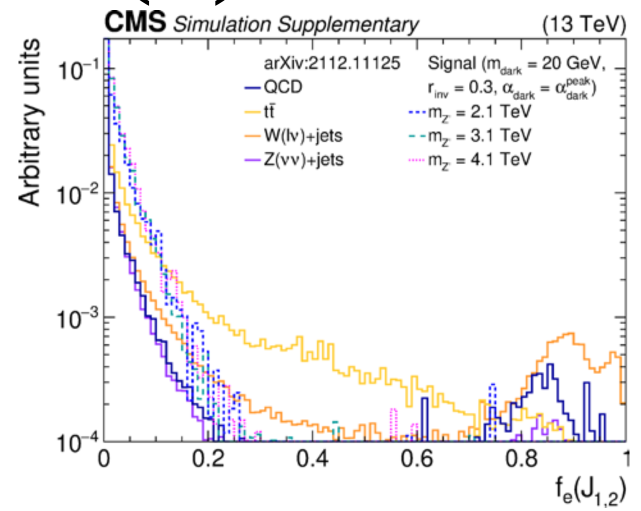


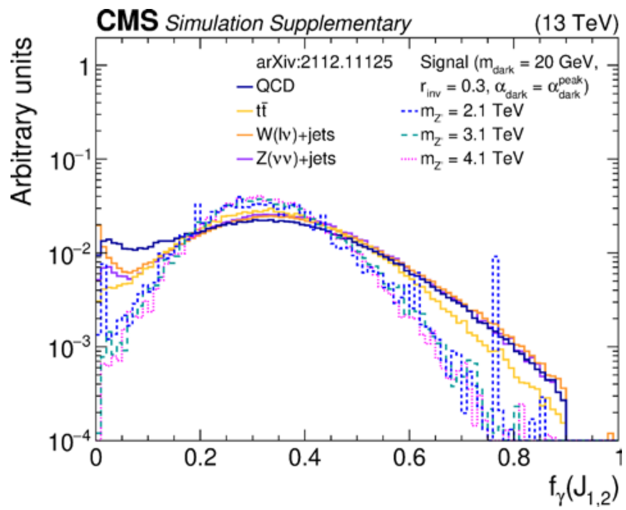


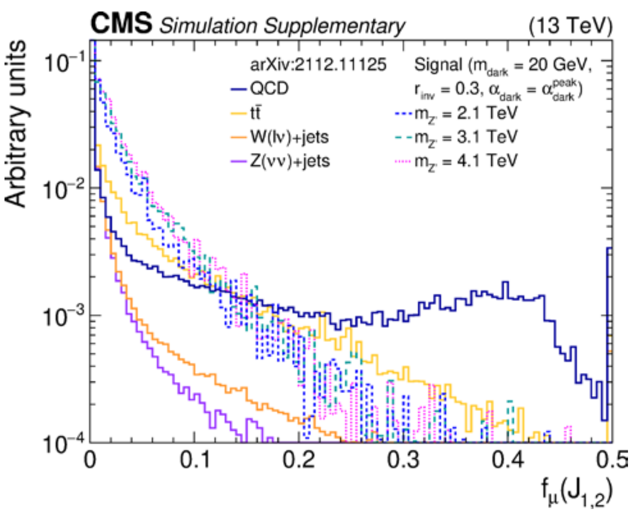


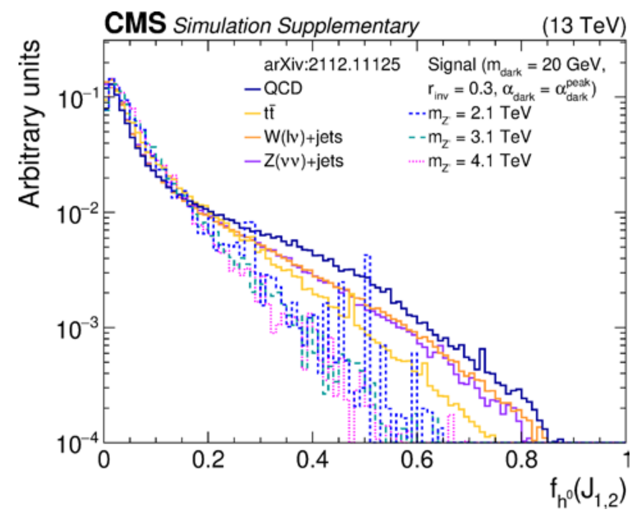
BDT Input Variables (3)



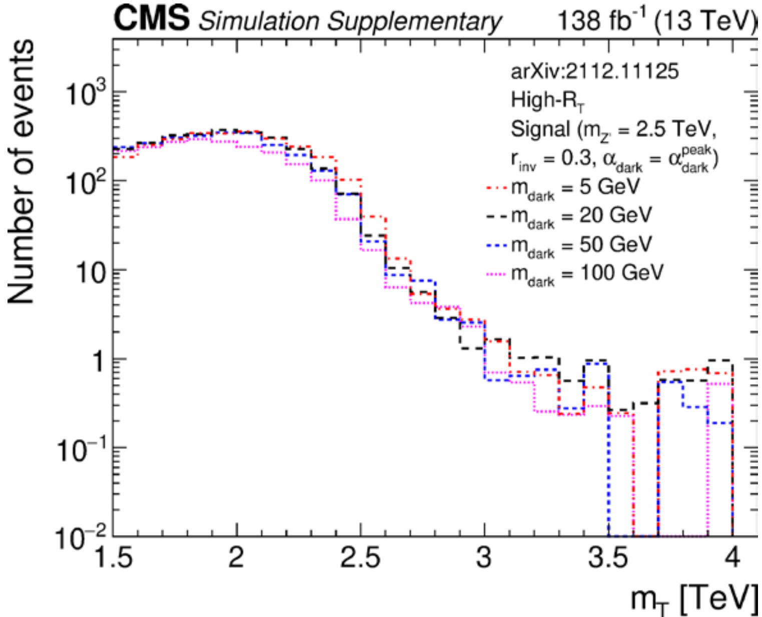




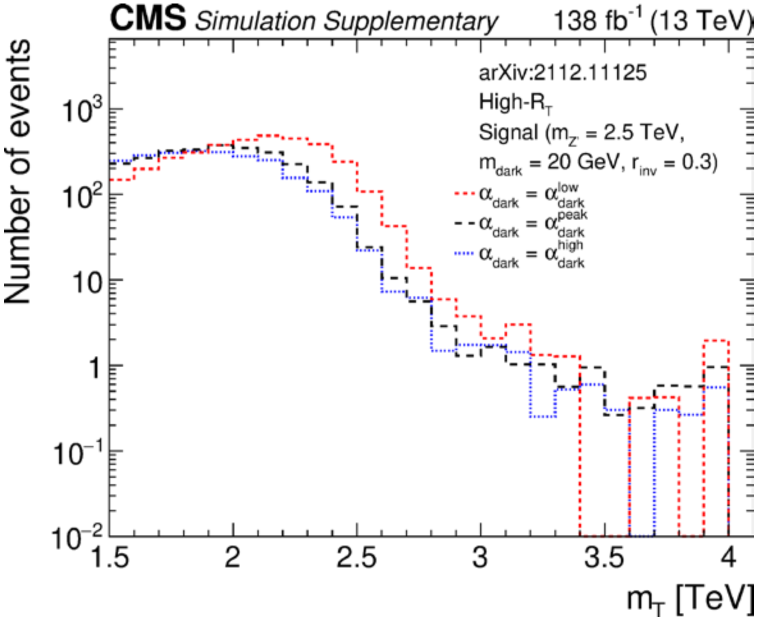


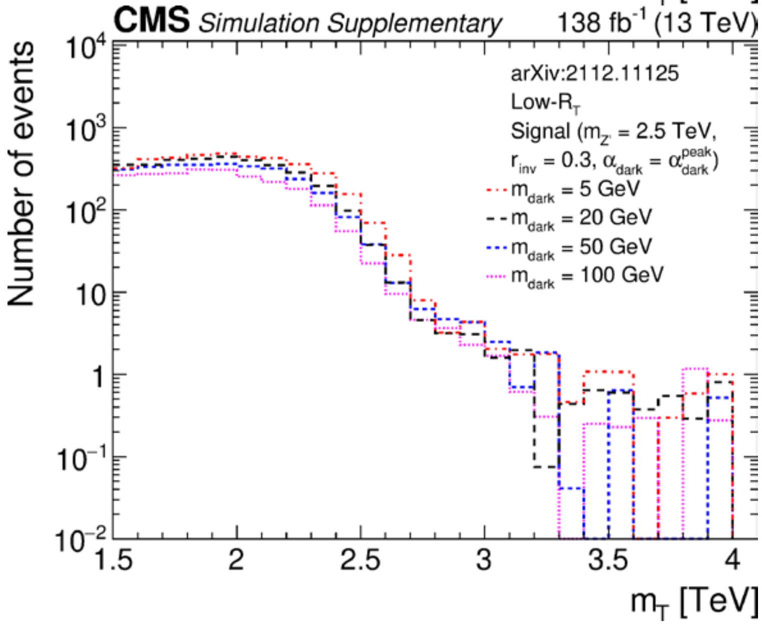


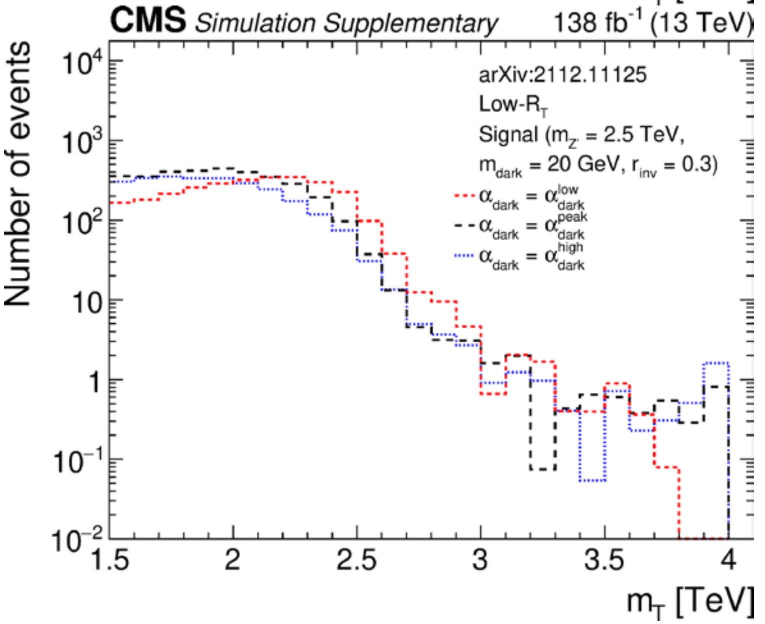
More m_T Variations



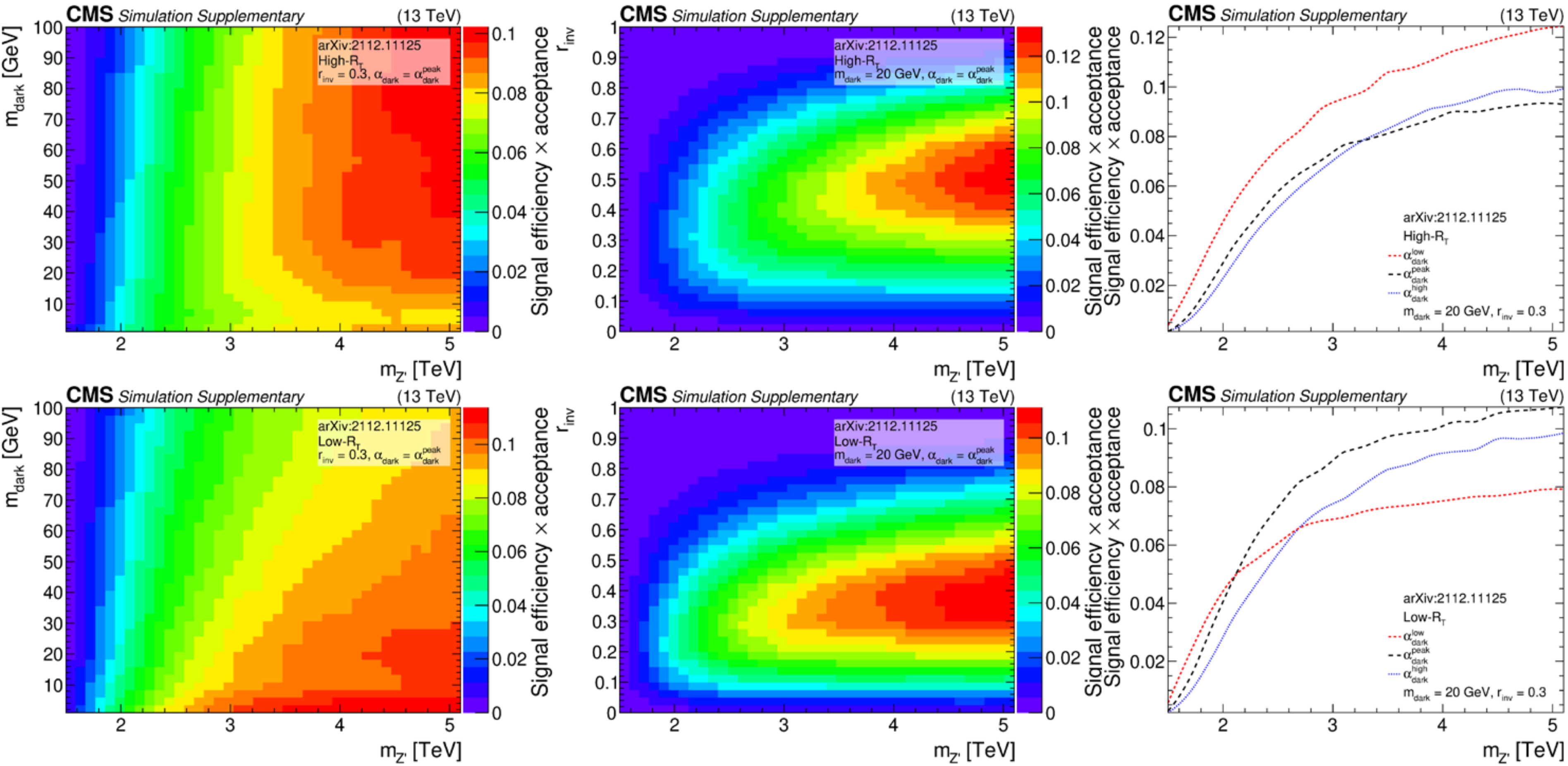
- α_{dark} has non-trivial impact
- m_{dark} has very little impact



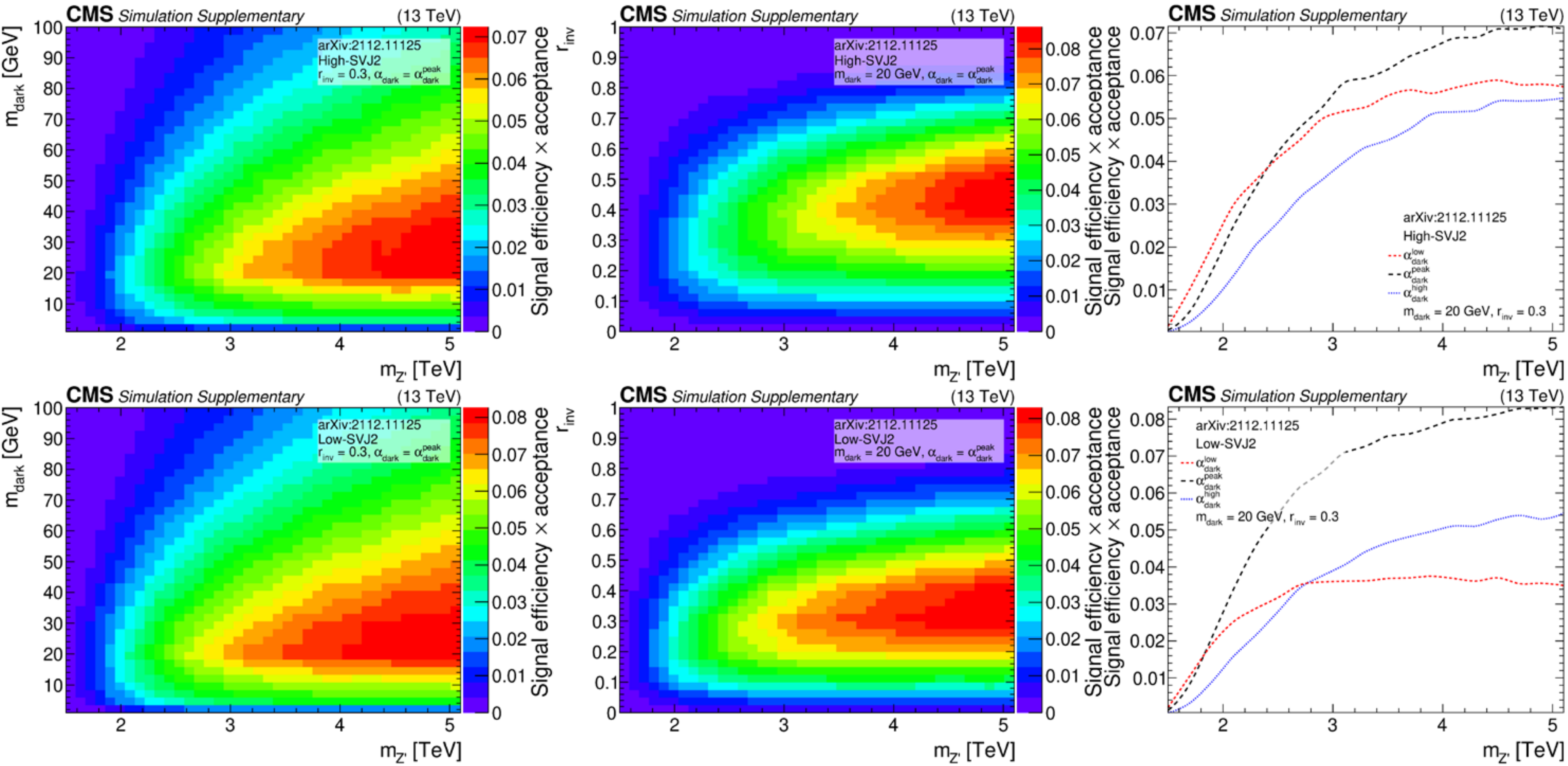




Signal Efficiency (inclusive)



Signal Efficiency (BDT-based)



Secondary Functions

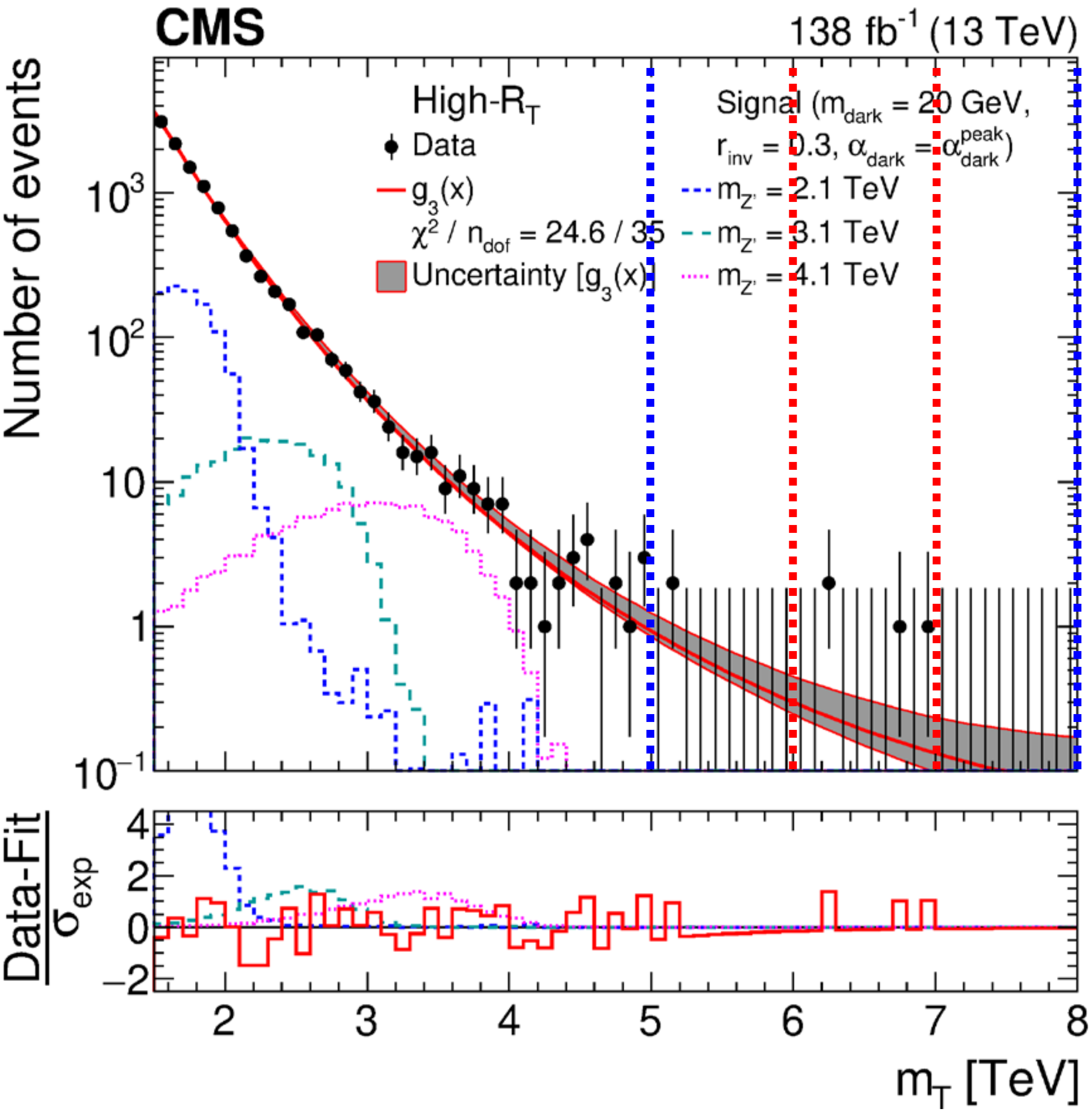
$$x = m_{\mathrm{T}}/\sqrt{s}$$

$$f(x) = (1-x)^{p_1}(x)^{p_2+p_3\log(x)}$$

$$h(x) = (x)^{-p_1}\exp(-p_2 x - p_3 x^2)$$

- *f*(*x*) from CMS dijet searches e.g. arXiv:1911.03947
- *h*(*x*) from UA2 dijet searches: Z. Phys. C **49** (1991) 17, Nucl. Phys. B **400** (1993) 3

High-R_T Tail Counts



m_T [TeV]	Obs.	Pred.
5–8	6	$8.4^{+2.1}_{-1.4}$
6–7	4	$2.0^{+0.6}_{-0.4}$

- Predicted counts and uncertainties obtained from integrating background fit