Searches for BSM in top final states in CMS

Victor Shang (University of Wisconsin-Madison)

on behalf of the CMS Collaboration

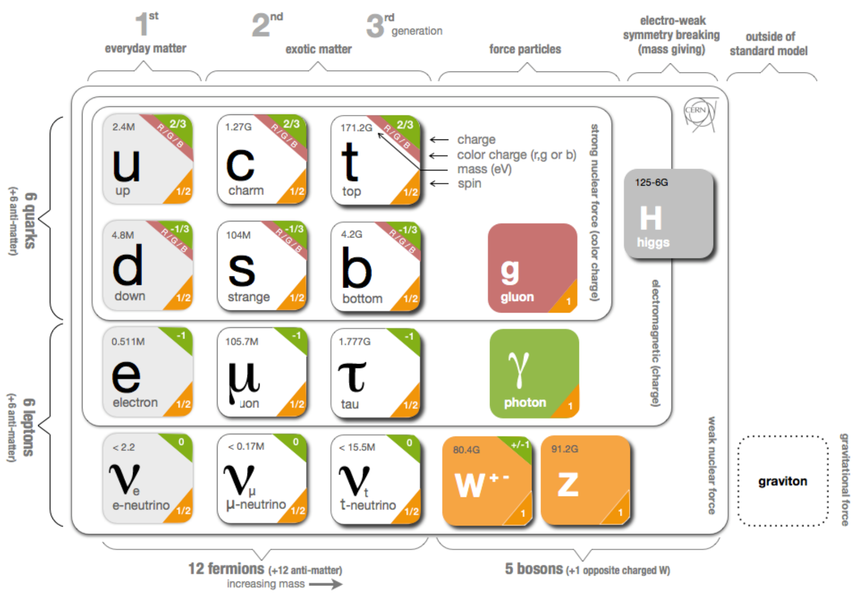
12th Large Hadron Collider Physics Conference (LHCP2024)
June 3-7, 2024

Beyond the Standard Model (BSM)

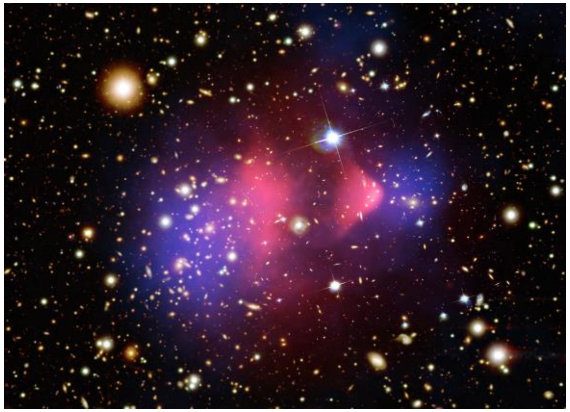
- Open questions: matter-antimatter asymmetry, gravity, neutrino masses, ***dark matter?***

- Large evidence for dark matter (DM) from astrophysical observations:
  - Mass of coma cluster, galactic rotation curves, gravitational lensing measurement of the bullet cluster, etc.

- Stable, interacts gravitationally, no color or charge
  - Possible candidate: Weakly Interacting Massive Particles (WIMPS)



Credit: David Galbraith

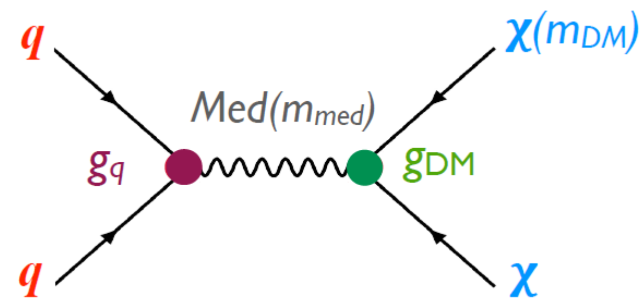


Credit: NASA/STScI, Magellan/U.Arizona/D.Clowe

Search for Dark Matter

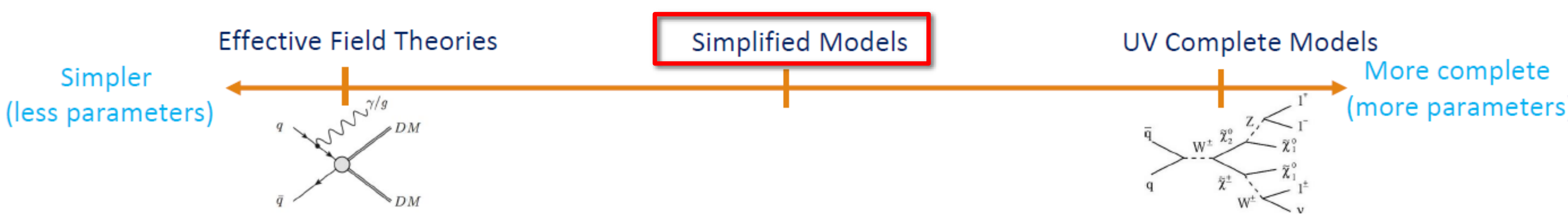
Ref. [CMS-PAS-EXO-22-014]

- Can study various models (EFT, **simplified models**, complete models)
  - Consider simplified model with single mediator between DM and SM (arXiv:1507.00966)
  - Minimal set of parameters: m_{med}, m_{DM}, g_q, g_{DM}
  - Assume spin-0 interaction: minimal flavor violation -> Yukawa coupling structure to SM



- Look for production of mediator from **top quark** that then decays into **two DM particles**
  - Infer DM in CMS from presence of **missing transverse momentum** (p_T^{miss})

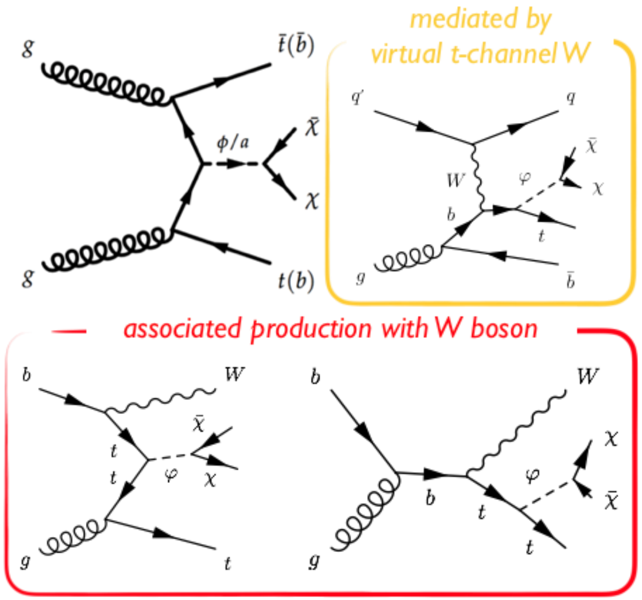
Spin-1	*vector* $g_q \sum_q V_\mu \bar{q}\gamma^\mu q$	*axial-vector* $g_q \sum_q A_\mu \bar{q}\gamma^\mu \gamma^5 q$
Spin-0	*scalar* $g_q \frac{\phi}{\sqrt{2}} \sum_f y_f \bar{f} f$	*pseudoscalar* $g_q \frac{iA}{\sqrt{2}} \sum_f y_f \bar{f}\gamma^5 f$

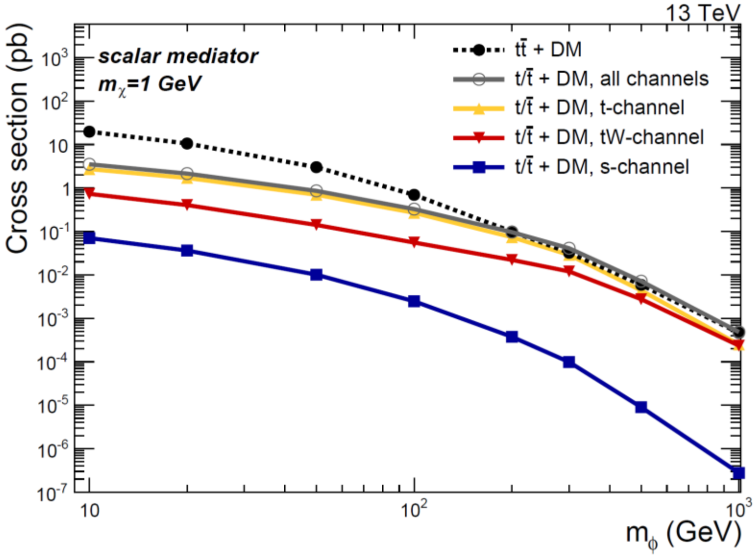


Dark Matter and Top Quarks

Ref. CMS-PAS-EXO-22-014

- ➢ Consider both **top quark pair ($t\bar{t}$)** and **single top ($t/\bar{t}$)** processes
  - ○ Only include t- and tW-channel processes for t+DM signal
- ➢ First time in CMS considering all three channels **All Hadronic (AH)**, **Semileptonic (SL)**, and **Dileptonic (DL)** for both tt+DM and t+DM
  - ○ Using full Run II data collected at CMS (2016-2018)



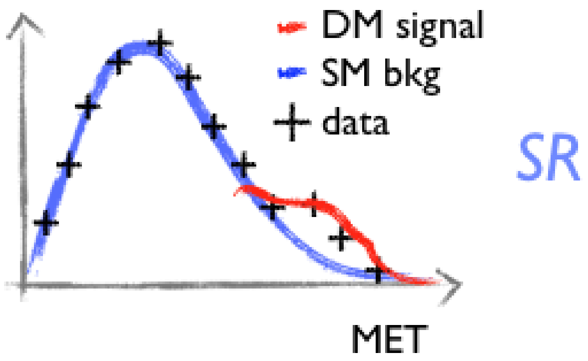


Credit: EPS-HEP 2017, 091

Search Strategy

Ref. [CMS-PAS-EXO-22-014]

- DM expected as excess in kinematic spectrum (i.e. p_T^{miss} or neural network output)
  - Dedicated ***signal regions (SR)*** for **AH**, **SL**, and **DL** with different discriminating variable selections applied to optimize signal sensitivity
  - Major backgrounds include **$t\bar{t}$** and **V+jets**



AH / SL

- Shape analysis of p_T^{miss} spectrum after selections for 0 and 1 lepton final states
  - Major SM background determined in simultaneous fit to data in orthogonal ***control regions (CR)***

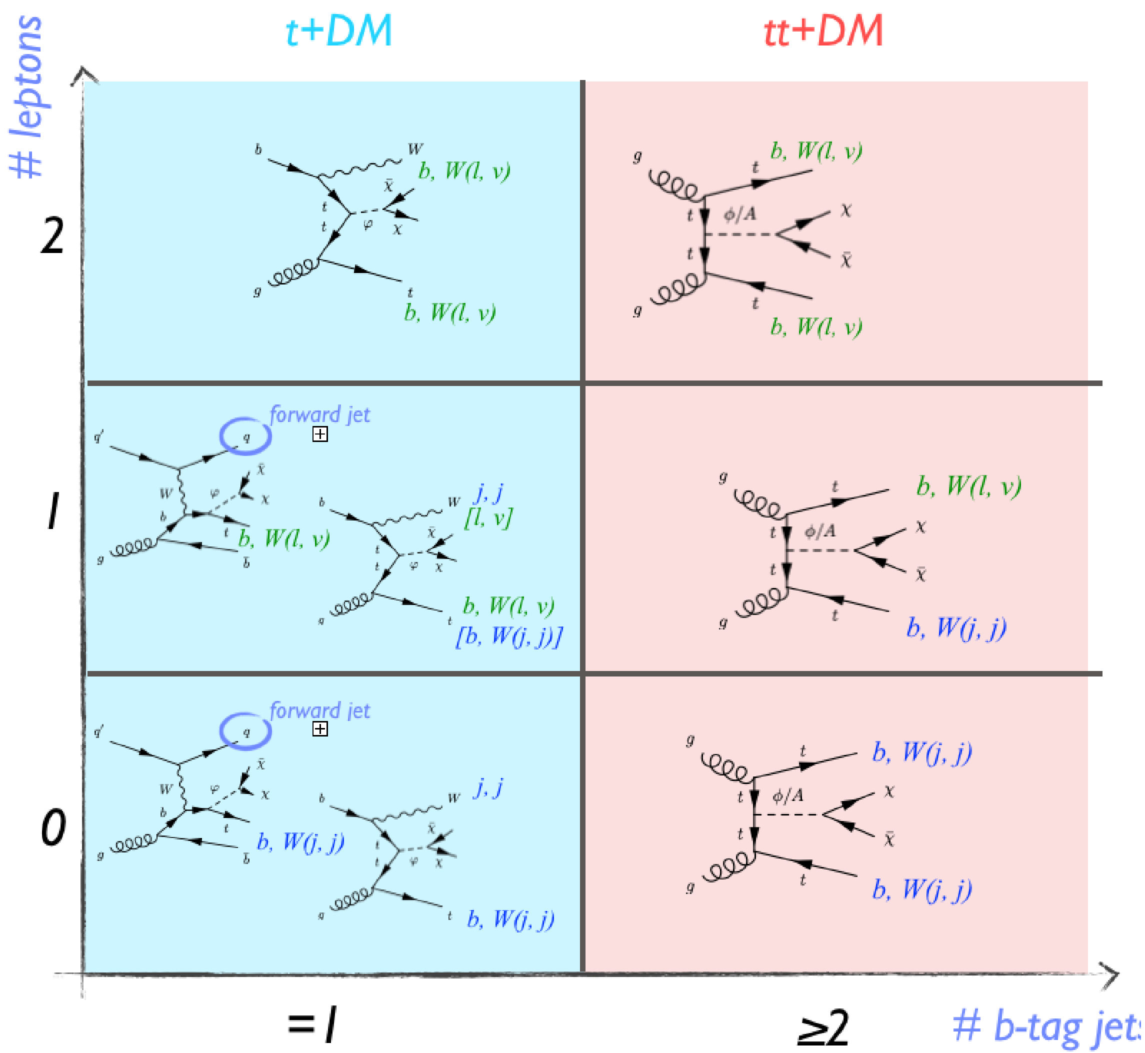
DL

- Shape analysis on neural network (NN) distributions after selections for 2 lepton final state
  - Various discriminating variables used to train NN
  - Background estimated predominantly from ***Monte Carlo (MC)*** simulations

Event Selection

Ref. CMS-PAS-EXO-22-014

- All channels: 1b and ≥2b regions targeting t+DM and tt+DM respectively

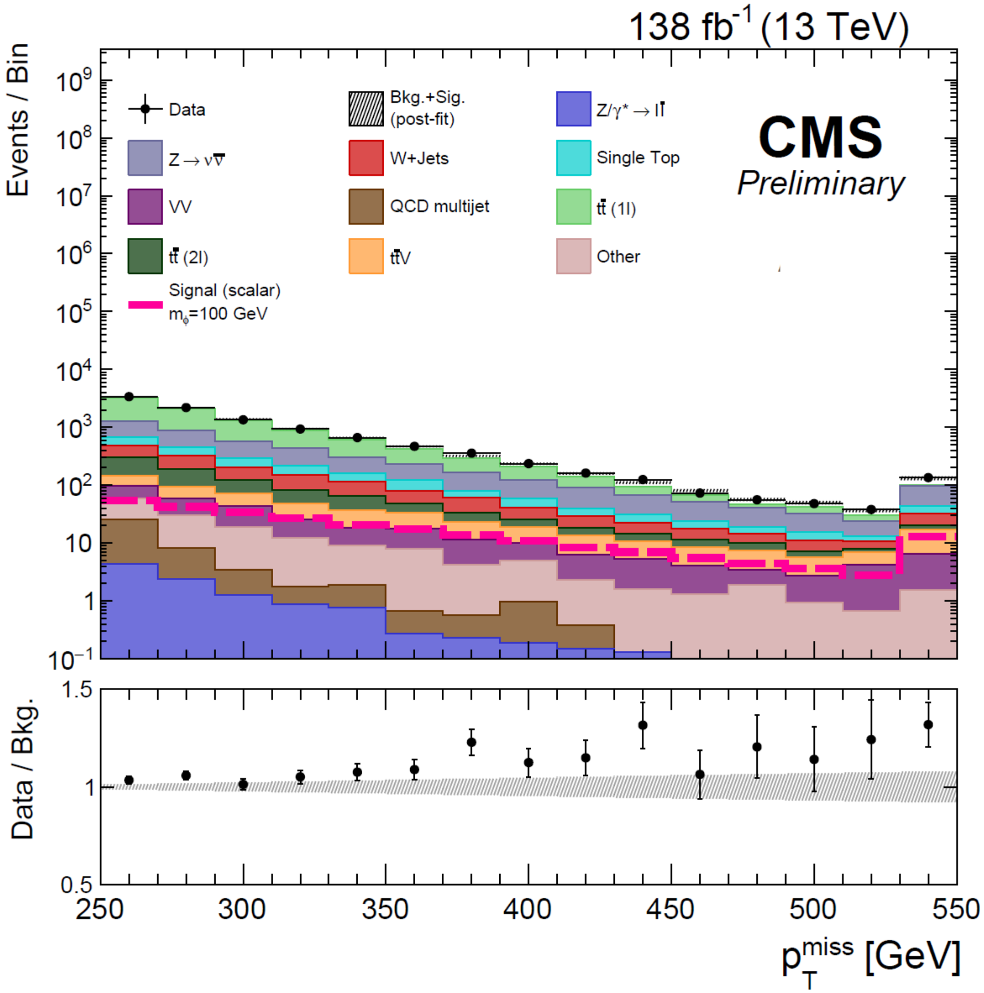


- **DL** channel:
  - ≥1 jets and pass tt kinematic reconstruction in ≥2b region
- **AH** and **SL** channels:
  - ≥1 forward jet categorization to target t-channel t+DM
  - p_T^{miss} > 250 GeV and ≥3 (AH) or ≥2 (SL) jets

Statistical Analysis

Ref. CMS-PAS-EXO-22-014

AH 2b SR

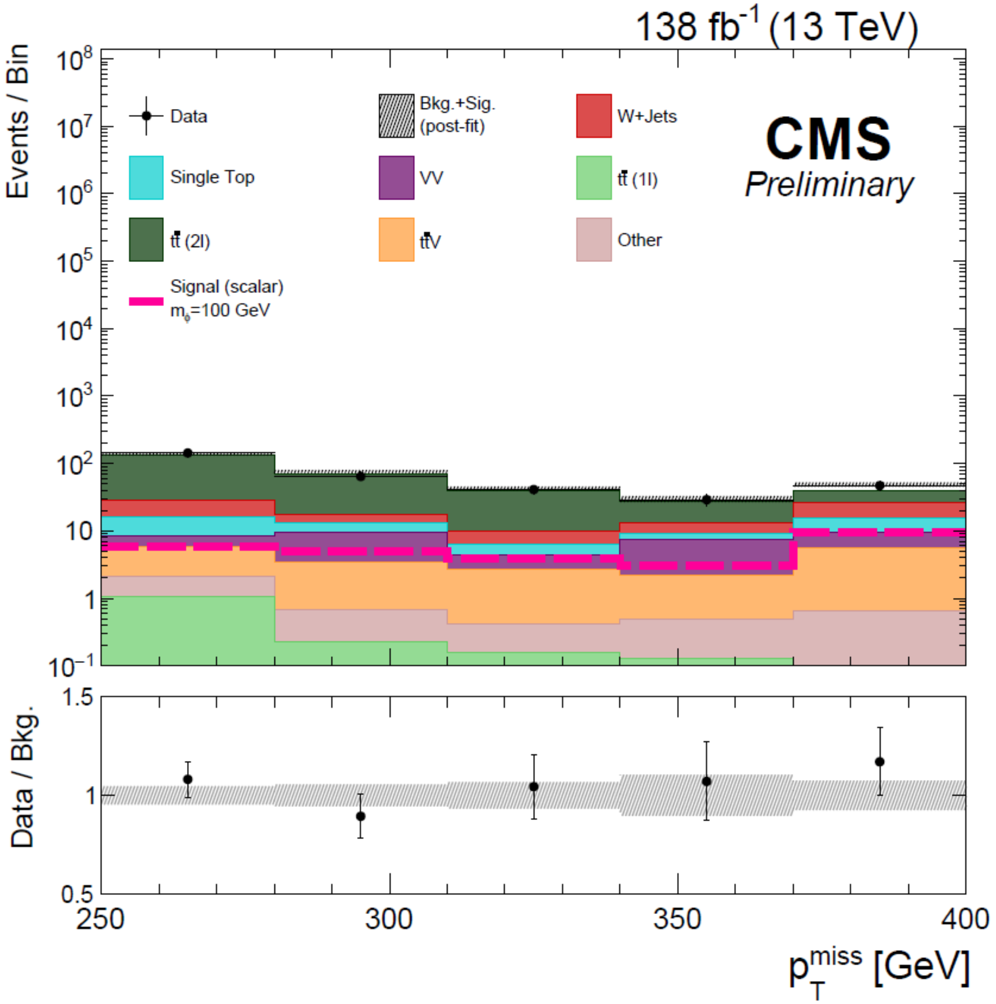


- Perform simultaneous maximum likelihood binned fit in SRs and CRs to p_T^{miss} (AH/SL) and NN output (DL) distributions
  - Consider normalization + shape effects of systematic uncertainties (**constrained nuisance parameters)**
  - Use CRs to estimate main backgrounds through rate parameters linked to SRs (**unconstrained parameters)**

Statistical Analysis

Ref. CMS-PAS-EXO-22-014

SL 1f-T2 SR

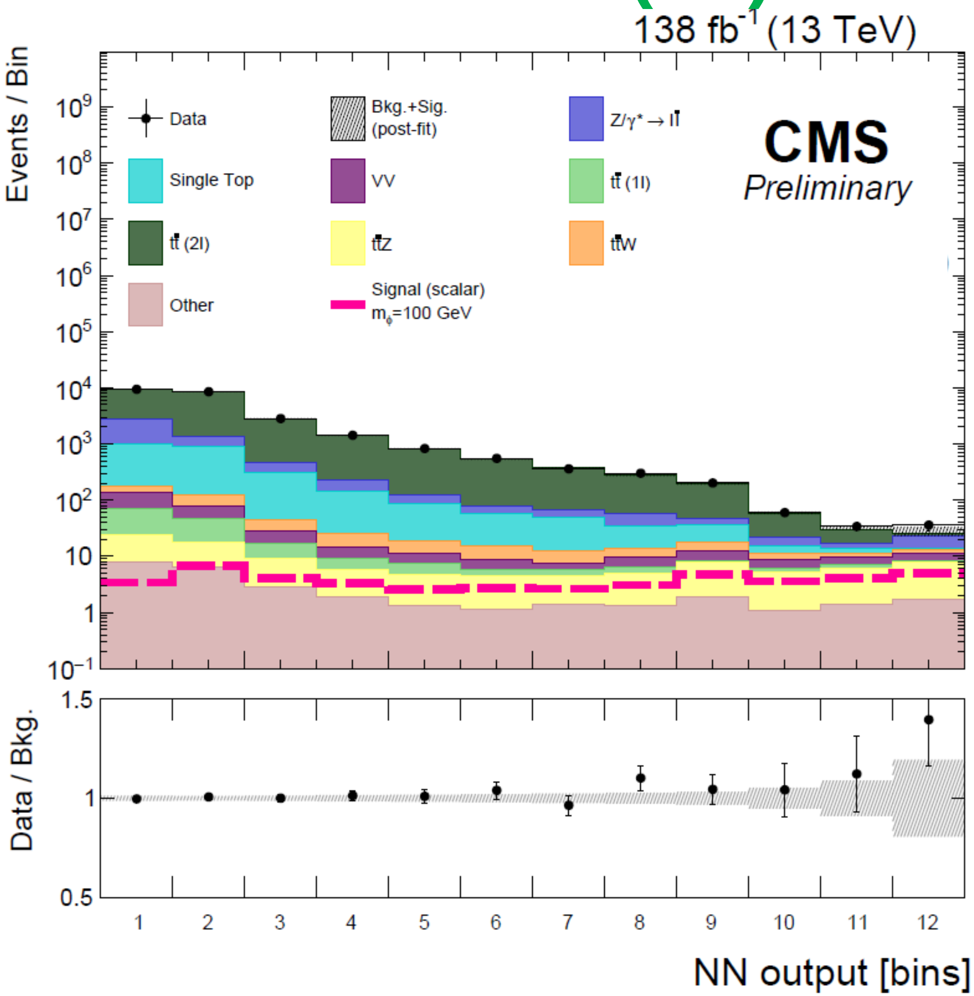


- Perform simultaneous maximum likelihood binned fit in SRs and CRs to p_T^{miss} (AH/SL) and NN output (DL) distributions
  - Consider normalization + shape effects of systematic uncertainties (**constrained nuisance parameters)**
  - Use CRs to estimate main backgrounds through rate parameters linked to SRs (**unconstrained parameters)**

Statistical Analysis

Ref. CMS-PAS-EXO-22-014

DL tW+DM SR (SF)

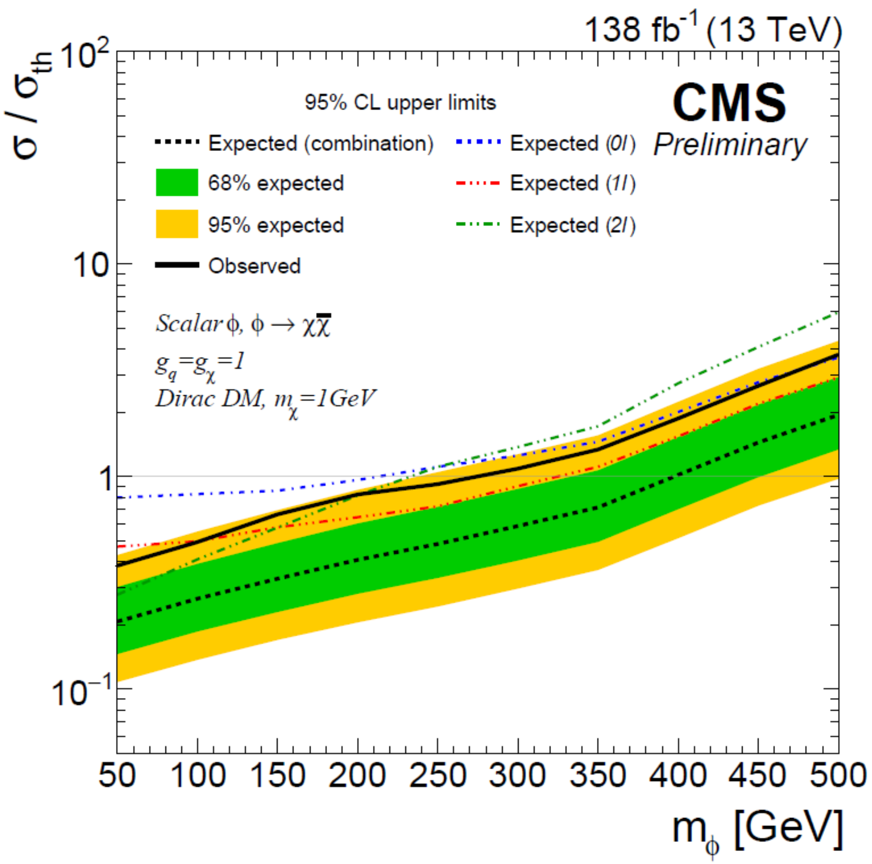


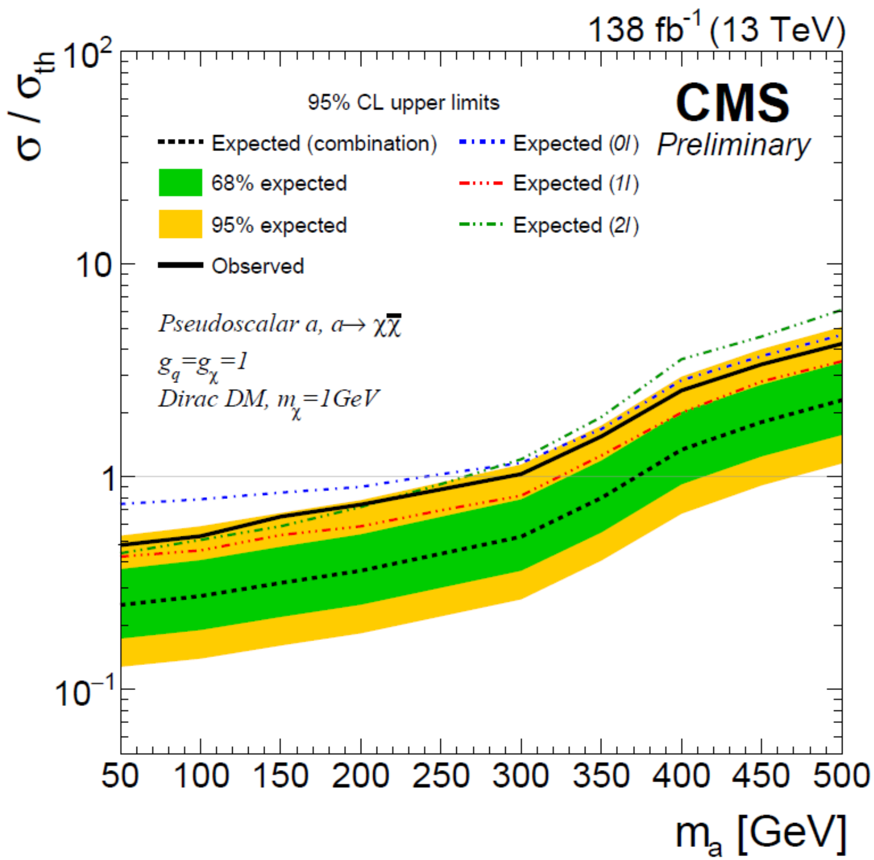
- Perform simultaneous maximum likelihood binned fit in SRs and CRs to p_T^{miss} (AH/SL) and NN output (DL) distributions
  - Consider normalization + shape effects of systematic uncertainties (**constrained nuisance parameters)**
  - Use CRs to estimate main backgrounds through rate parameters linked to SRs (**unconstrained parameters)**

Results

Ref. [CMS-PAS-EXO-22-014]

- Use estimated signal strength parameter from fit to set constraints on DM production cross sections relative to theory predictions
  - Scalar and pseudoscalar mediator masses excluded below 280(400) GeV and 290(380) GeV for obs(exp) limits, respectively

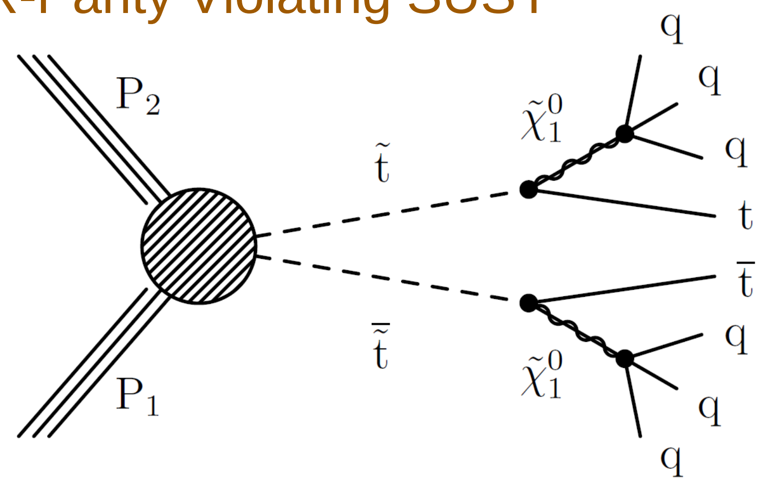




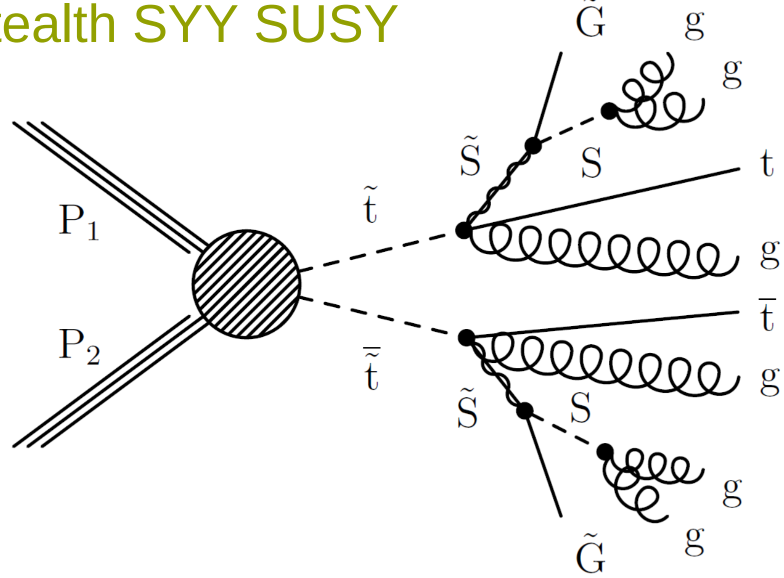
Stealth/RPV Stop Search

Ref.
SUS-23-001 (PAS TBD)

R-Parity Violating SUSY



Stealth SYY SUSY



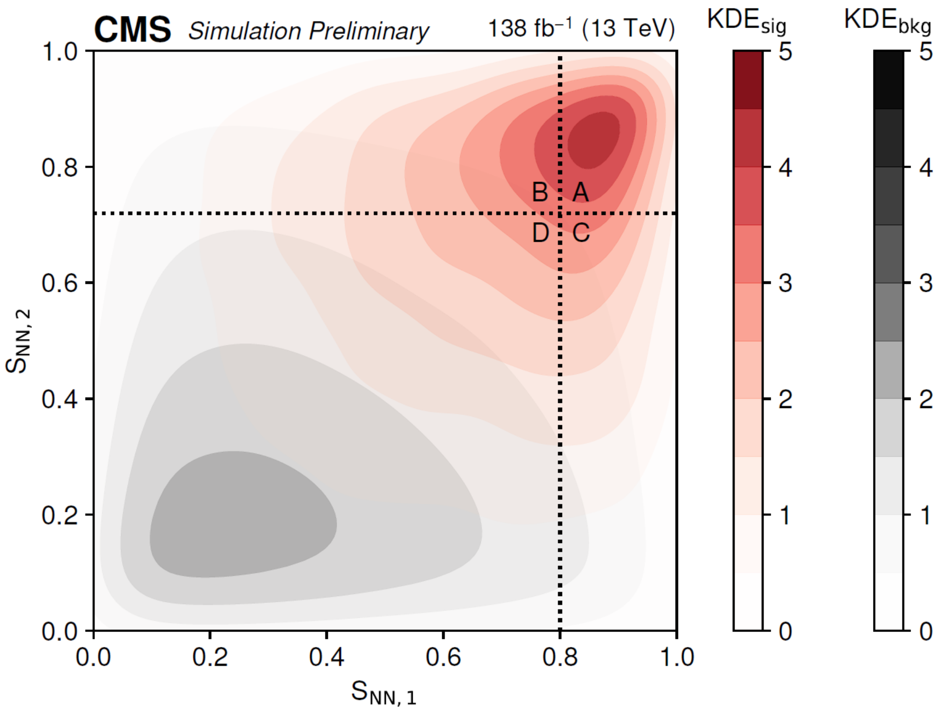
- Looking for both an R-parity violating (RPV) and a Stealth SUSY signature
  - Final state of both signal models is **$t\bar{t}$ + jets with little to no p_T^{miss}**
  - Primary observable is **N_{jets}**
  - Three channels: zero lepton (0l), one lepton (1l), and two lepton (2l)

Search Strategy

Ref.
SUS-23-001 (PAS TBD)

- Estimate main tt+jets background using novel **ABCDisCoTEC** method
  - Generates two independent signal vs. background discriminators to use as basis variables for ABCD background estimation
  - Signal and tt+jets estimated separately in each N_{jets} bin with simultaneous fit to data in four 'ABCD' bins (N_A, N_B, N_C, N_D) of $S_{NN,1}$ vs. $S_{NN,2}$ plane
  - 'ABCD' constraint: $N_A = \kappa \left(\frac{N_B N_C}{N_D} \right)$

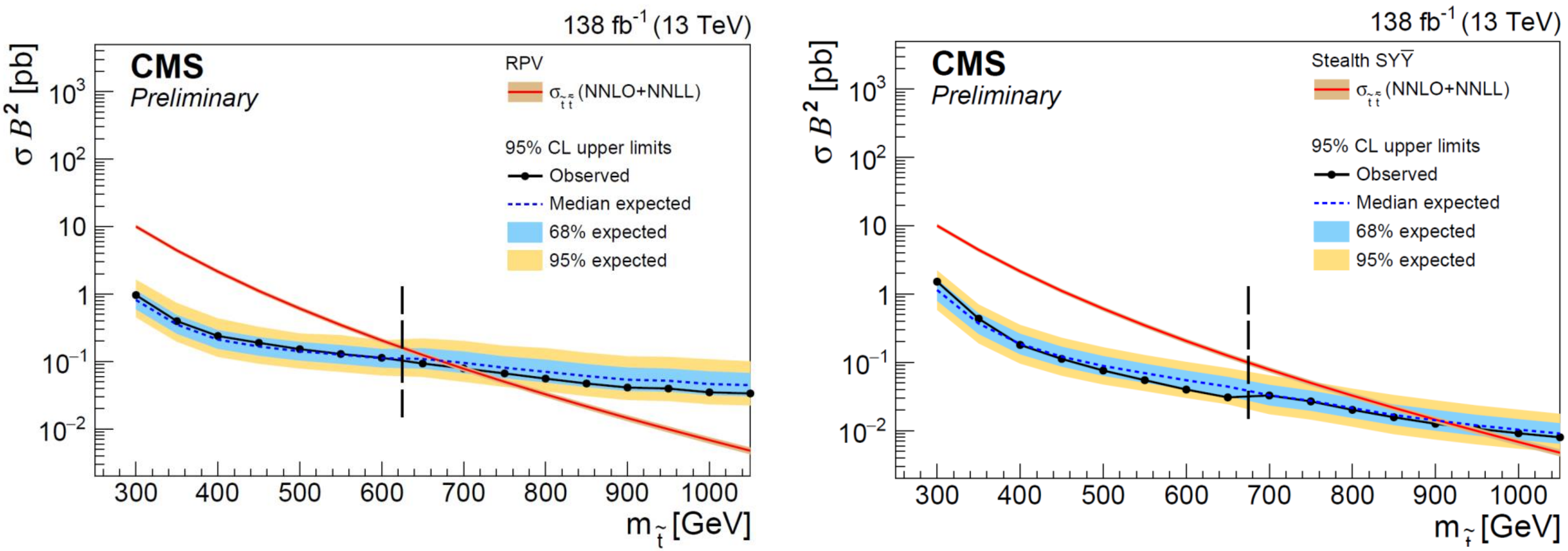
- Other backgrounds predicted either from MC simulation (tt+X, single top, Z+jets, Multiboson) or extracted from dedicated control region (QCD multijet)



Results

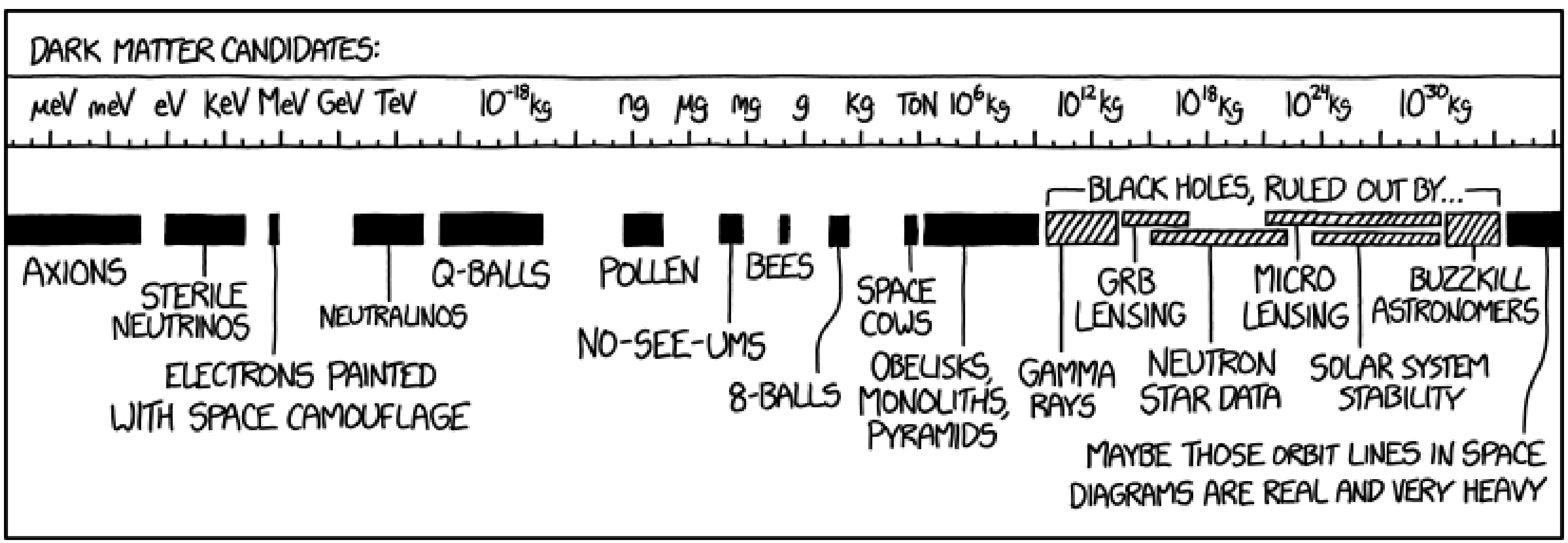
Ref.
SUS-23-001 (PAS TBD)

- Three channel combination limits shown for the RPV (left) and Stealth SYY (right) signal models
  - No significant excess observed above expected background for either model
- Mass exclusion limits set at 700 GeV (RPV) and 930 GeV (Stealth SYY)



Summary

- First time in CMS that search for DM with top quarks performed in all three channels **All Hadronic (AH)**, **Semileptonic (SL)**, and **Dileptonic (DL)** for **both** tt+DM and t+DM using full Run II data
- Search for **RPV** and **Stealth SUSY** signatures also performed at CMS using **novel ABCD-based method** for background estimation
- Improvements coming in recent future (Run 3), so stay tuned for future results!



Backup

SR and CR Definitions

Ref. CMS-PAS-EXO-22-014

	All-hadronic SRs			Single-lepton SRs			Dilepton SRs	
	0l, 1 b-tag, 0 FJ	0l,1 b-tag, 1 FJ	0l, 2 b-tag	1l, 1 b-tag, 0 FJ	1l, 1 b-tag, 1FJ	1l, 2 b-tag	2l, 1 b-tag	2l, 2 b-tag
n_{lep}		$=0$			$=1$		$=2$	
n_{jet}		≥ 3			≥ 2		≥ 1	
n_{b}	$=1$	$=1$	≥ 2	$=1$	$=1$	≥ 2	$=1$	≥ 2
Forward jets	$=0$	≥ 1	—	$=0$	≥ 1	—	—	—
$p_{\mathrm{T}}(\mathrm{j}_1)/H_{\mathrm{T}}$	—		<0.5		—		—	
$p_{\mathrm{T}}^{\mathrm{miss}}$		>250 GeV			>250 GeV		—	
m_{T}		—			>140 GeV		—	
$m_{\mathrm{T2}}^{\mathrm{W}}$		—			>180 GeV		—	
$\min\Delta\phi(\mathrm{j}_{1,2}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$		>0.8 rad.			>0.8 rad.		—	
$m_{\mathrm{T}}^{\mathrm{b}}$		>140 GeV			>140 GeV		—	
m_{ll}		—			—		>20 GeV	
$\lvert m_{\mathrm{ll}} - m_{\mathrm{Z}}\rvert$		—			—		> 15 GeV (ee and $\mu\mu$)	
$m_{\mathrm{T2}}^{\mathrm{ll}}$		—			—		>80 GeV	
Pass $\mathrm{t\bar{t}}$ reco		—			—		—	yes

	AH CRs				SL CRs			DL CRs	
	$\mathrm{t\bar{t}}$(1l) CR	W (l ν) CR	Z $\to$ ll CR	QCD CR	$\mathrm{t\bar{t}}$(2l) CR	W (l ν) CR	$\mathrm{t\bar{t}}$(2l) VR	DY CR	$\mathrm{t\bar{t}}$Z CR
n_{b}	≥ 1	$=0$	$=0$	≥ 1	≥ 1	$=0$	≥ 1	$=1$	≥ 1
n_{lep}	$=1$	$=1$	$=2$	$=0$	$=2$ (e, μ)	$=1$	$=2$	$=2$ (ee, $\mu\mu$)	$=3$
n_{jet}	≥ 3	≥ 3	≥ 3	≥ 3	≥ 2	≥ 2	≥ 1	≥ 1	≥ 3
$p_{\mathrm{T}}^{\mathrm{miss}}$ (GeV)	≥ 250	≥ 250	≥ 250 (had recoil)	≥ 250	≥ 250	≥ 250	—	—	—
M_{T} (GeV)	≤ 140	≤ 140	—	—	—	≥ 140	—	—	—
$\min\Delta\phi(\mathrm{j}_{1,2}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$	≥ 0.8	—	—	< 0.8	—	—	—	—	—
m_{ll} (GeV)	—	—	$[60, 120]$	—	—	—	> 20	—	—
$\lvert m_{\mathrm{ll}} - m_{\mathrm{Z}}\rvert$ (GeV)	—	—	—	—	—	—	> 15 (SF)	< 15	< 10 (OSSF)
$m_{\mathrm{T2}}^{\mathrm{ll}}$ (GeV)	—	—	—	—	≤ 80	≤ 80	≤ 80	≥ 80	—
Included in fit?	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes

Discriminating Variables (AH/SL)

Ref. CMS-PAS-EXO-22-014

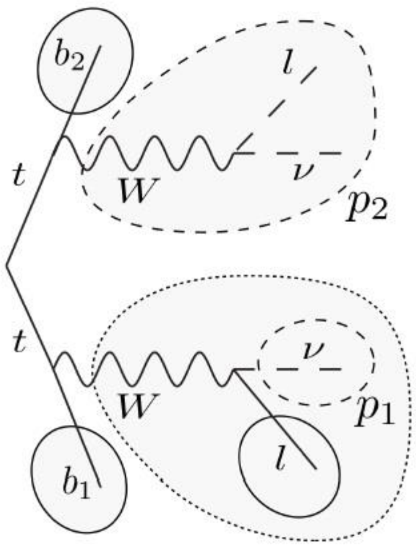
- **Modified topness (SL):** topness variable developed to suppress dileptonic tt background [https://arxiv.org/abs/1212.4495]

- **M_T (SL):** The transverse W mass defined at

$$M_T = \sqrt{2 p_T^{miss} p_T^{\ell} (1 - \cos(\Delta\phi))}$$

- **M_{T2}^W (SL):**

$$M_{T2}^{W} = \min_{\bar{p}_1, \bar{p}_2} \left\{ m_y : \left[\begin{array}{r} \vec{p}_1^{\,T} + \vec{p}_2^{\,T} = \vec{p}_T^{\text{miss}},\ \bar{p}_1^2 = 0,\ (\bar{p}_1 + \bar{p}_l)^2 = \bar{p}_2^2 = m_W^2, \\ (\bar{p}_1 + \bar{p}_l + \bar{p}_{b_1})^2 = (\bar{p}_2 + \bar{p}_{b_2})^2 = m_y^2 \end{array} \right] \right\}$$



- **$\min\Delta\phi(j_{1,2}, p_T^{miss})$ (SL and AH):** minimum opening angle between the first two leading p_T jets and p_T^{miss}

- **M_T^b (SL and AH):** the transverse mass of p_T^{miss} and b-tagged jet with highest b-tag discriminant

- **$jet_1\ p_T/H_T$ (AH):** ratio of leading p_T jet divided total hadronic energy in event

Neural Network Inputs (DL)

Ref. CMS-PAS-EXO-22-014

- **M_{T2}^{ll} (t+DM, tt+DM):** generalization of transverse mass defined as

$$M_{T2}(l,\bar{l}) = \min_{\vec{p}_{T,\bar{\nu}}+\vec{p}_{T,\nu}=\vec{p}_T^{\text{miss}}} [\max\{M_T(m_l, m_{\bar{\nu}}, \vec{p}_{T,l}, \vec{p}_{T,\bar{\nu}}), M_T(m_{\bar{l}}, m_{\nu}, \vec{p}_{T,\bar{l}}, \vec{p}_{T,\nu})\}],$$

with

$$M_T(m_1, m_2, \vec{p}_{T,1}, \vec{p}_{T,2}) = \sqrt{m_1^2 + m_2^2 + 2(E_{T,1}E_{T,2} - \vec{p}_{T,1}\cdot\vec{p}_{T,2})},$$

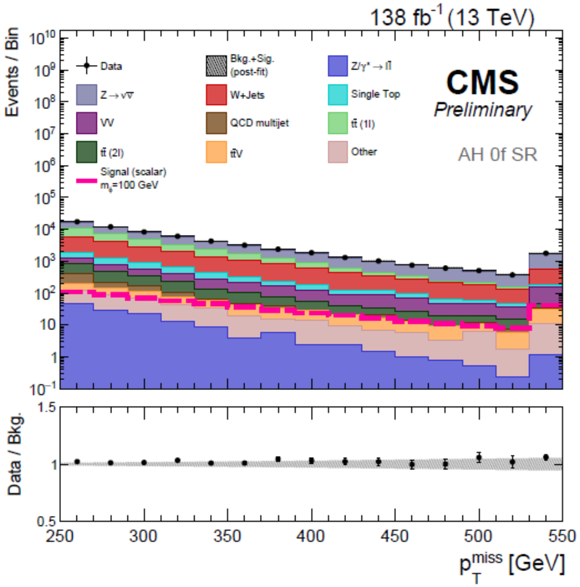
- **$\Delta\phi(\mathbf{l},\bar{\mathbf{l}})$ (t+DM, tt+DM):** absolute difference in azimuthal angle between two leptons
- **$\Delta\phi(\mathbf{p_T^{miss}}, \mathbf{l\bar{l}b})$ (t+DM, tt+DM):** absolute difference in azimuthal angle between p_T^{miss} and other visible objects (leptons, b-tagged jet)
- **p_T^{dark} (tt+DM):** subtraction of neutrino momentum from total p_T^{miss} defined as

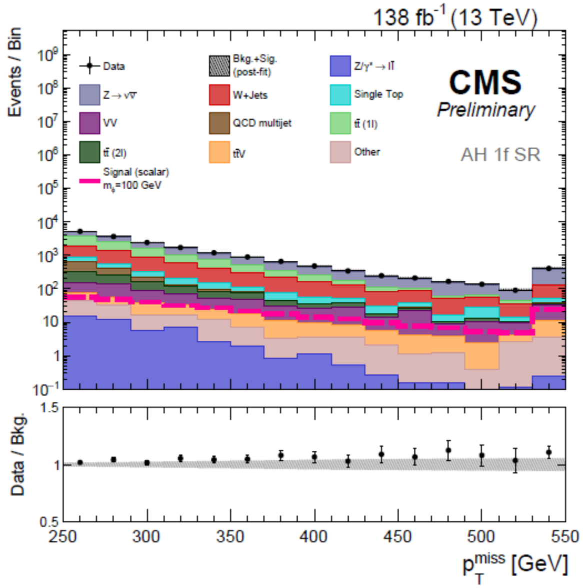
$$p_T^{\text{dark}} \doteq \sqrt{(\not{E}_x - p_{x,\nu} - p_{x,\bar{\nu}})^2 + (\not{E}_y - p_{y,\nu} - p_{y,\bar{\nu}})^2}$$

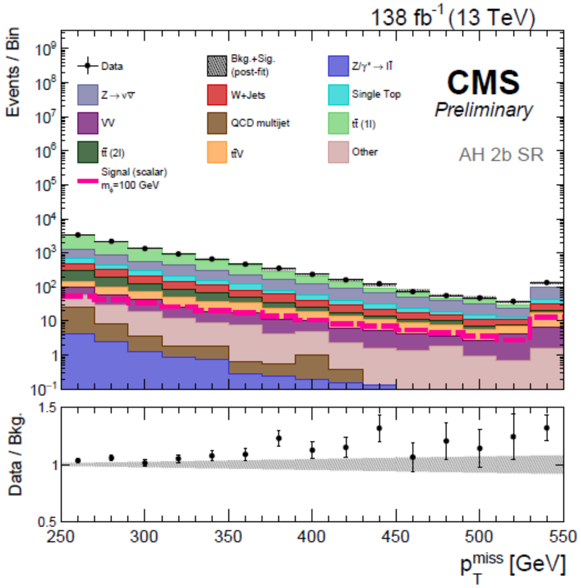
- **c_{hel} (tt+DM):** spin correlation variable defined as dot product of the two lepton directions measured in their parent top quark and antiquark rest frames
- **$\Delta\phi(\mathbf{t},\bar{\mathbf{t}})$ (tt+DM):** absolute difference in azimuthal angle between top quark and antiquark

SR Post-fit Distributions (AH)

Ref. CMS-PAS-EXO-22-014

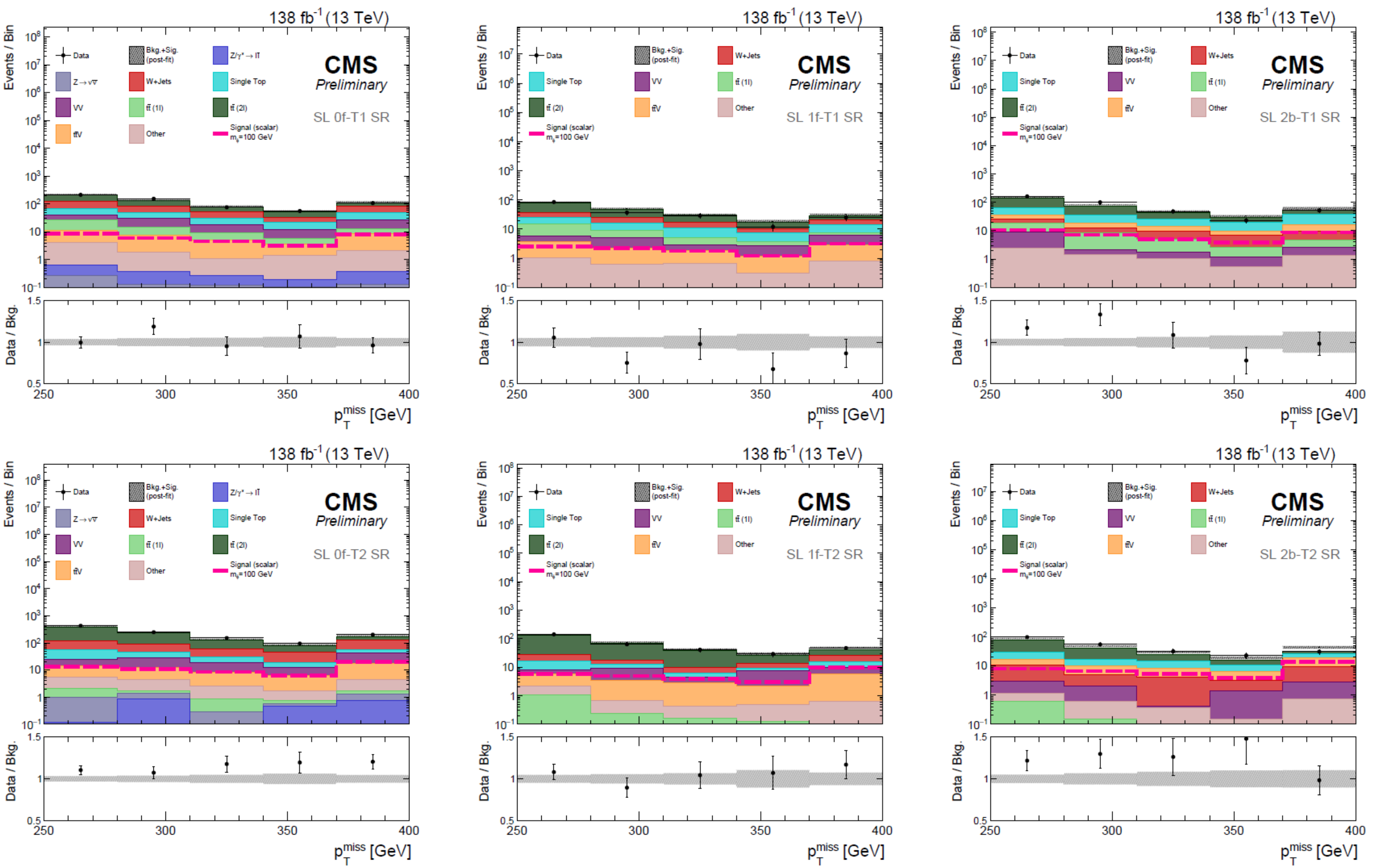






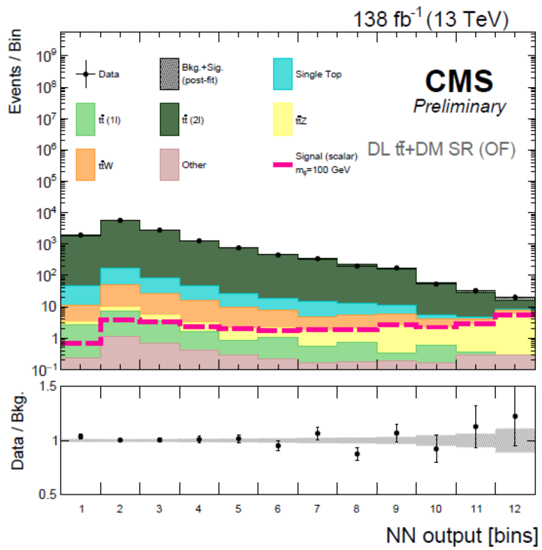
SR Post-fit Distributions (SL)

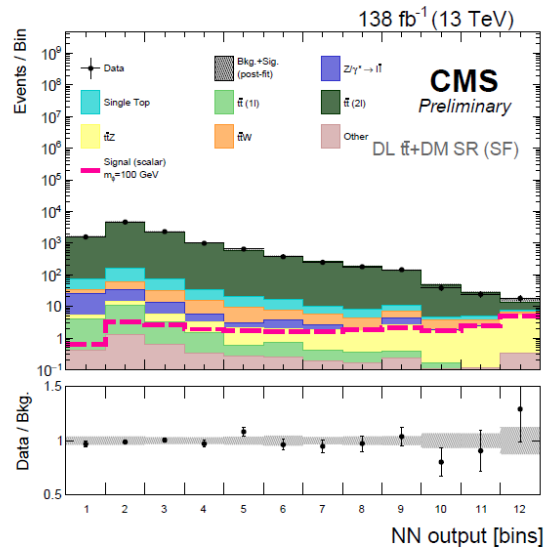
Ref. CMS-PAS-EXO-22-014

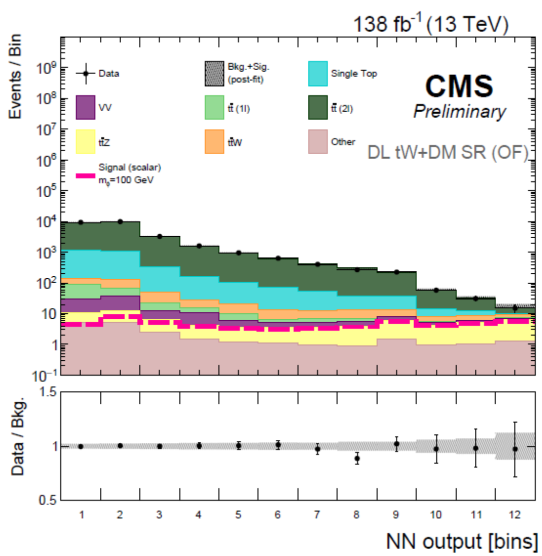


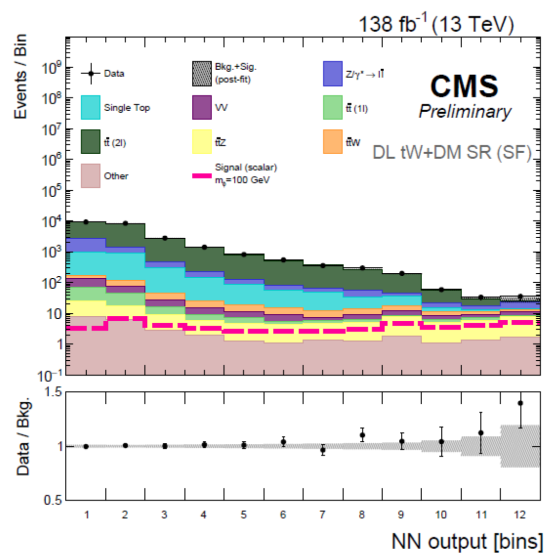
SR Post-fit Distributions (DL)

Ref. [CMS-PAS-EXO-22-014]









Systematic Uncertainties

Ref. CMS-PAS-EXO-22-014

Systematic	Name in datacard	# nuisances	Shape effect?	Correlated years?	Relevant processes
PDF	CMS_pdf	1	Yes	Yes	All MC (not data-driven QCD)
ME scales	QCDscale_{ren/fac}_{dataset}	$2 * n_{\text{dataset}}$	Yes	Yes	All (but uncorrelated between datasets)
Parton shower	CMS_PS{isr/fsr}	2	Yes	Yes	VV, ST, t/tt+DM (AH+SL), All (DL)
Luminosity	lumi_{year}	3	No	No	All
Pileup	CMS_scale_pu	3	Yes	Yes	All
Trigger	CMS_eff_{met/lep/dilep}_trigger	3	Yes	Yes	All
Lepton efficiency	CMS_eff_{e/m}{0/1}	4	Yes	Yes	All
Jet Energy Scale	CMS_scale{source}_j	27	Yes	Yes	All
Jet Energy Resolution	CMS_res_j_{year}	3	Yes	No	All
Unclusterred MET	CMS_UncMET_{year}	3	Yes	No	All
b-tagging efficiency (corr)	CMS_eff_b{_light}	2	Yes	Yes	All (uncorrelated between DL and AH/SL)
b-tagging efficiency (uncorr)	CMS_eff_b{_light}_{_year}	2	Yes	No	All (uncorrelated between DL and AH/SL)
W/Z + heavy-flavor	CMS_HF_{W/Z}	2	Yes	Yes	W/Z + jets (AH and SL)
EWK and QCD scale factors	CMS_{W/Z}{qcd/ewk}Weight{Ren/Fac}	6	Yes	Yes	W/Z + jets (AH and SL)
b-tagged jet mult. norm.	nbjet_{W/Z}	2	Yes	Yes	W/Z + jets (AH and SL)

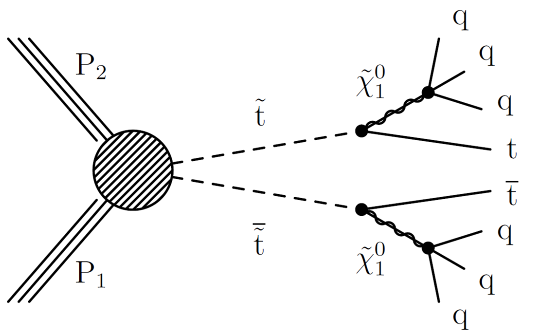
- Most important systematics include b-tagging efficiency, jet energy resolution, W/Z + heavy-flavor, ME scales, and ttZ normalization

Signal Models

Ref.
SUS-23-001 (PAS TBD)

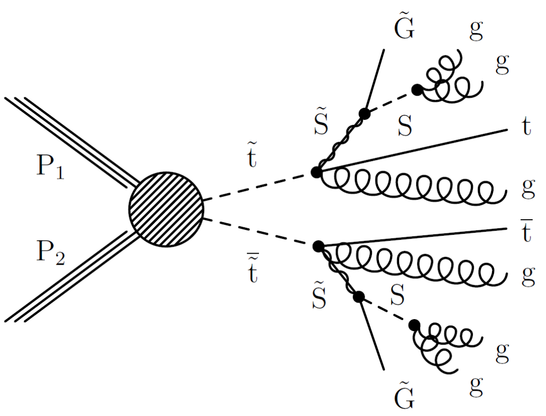
R-Parity Violating SUSY

- R-parity is not conserved, allowing for an additional interaction via the UDD coupling
- LSP neutralino ( ) no longer stable and decays to SM products via
- Mass of set to 100 GeV for this analysis



Stealth SYY SUSY

- Introduces a “hidden” sector comprised of a mass degenerate singlino ($M_{\tilde{S}}$ = 100 GeV) and singlet (M_S = 90 GeV)
- The singlino decays via a hidden sector to the singlet and a light, low-energy gravitino ($M_{\tilde{G}}$ = 1 GeV)
- The gravitino is too light to leave a noticeable p_T^{miss} signature



SR Selections

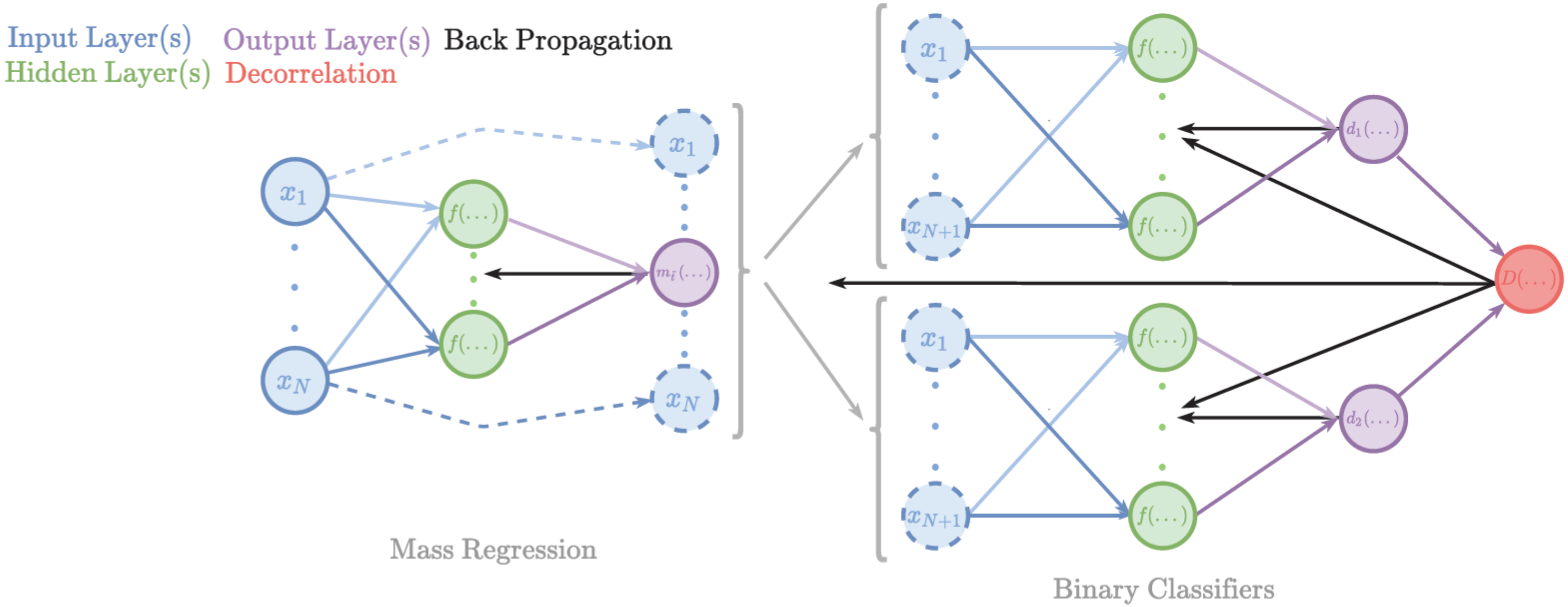
Ref.
SUS-23-001 (PAS TBD)

Selection Criteria		0ℓ	1ℓ	2ℓ
$N_{\text{leptons}}^{\text{iso}}$		0	1	2, oppositely charged
$N_{\text{muon}}^{\text{non-iso}}$		0	0	0
H_{T} (GeV)		> 500	> 500	> 500
N_{jets}	$(p_{\mathrm{T}} > 30\,\text{GeV})$	≥ 8	≥ 7	≥ 6
	$(p_{\mathrm{T}} > 45\,\text{GeV})$	≥ 6	—	—
$N_{\text{b jets}}$	$(p_{\mathrm{T}} > 30\,\text{GeV})$	≥ 2	≥ 1	≥ 1
	$(p_{\mathrm{T}} > 45\,\text{GeV})$	≥ 1	—	—
N_{t}		≥ 2	—	—
$M_{b\ell}$ (GeV)		—	$> 50, < 250$	—
$M_{\ell\ell}$ (GeV)		—	—	< 81 or > 101
$\Delta R_{\text{b jets}}$		≥ 1	—	—

ABCDisCoTEC Method

Ref.
SUS-23-001 (PAS TBD)

$$L_{Total} = \lambda_{BCE} L_{BCE} + \lambda_{DisCo} L_{DisCo} + \lambda_{Closure} L_{Closure} + \lambda_{MR} L_{MR}$$



- Train two **binary classifiers** to identify signal from background
- DisCo Loss: minimize **distance correlation** between $S_{NN,1}$, $S_{NN,2}$
- Closure Loss: directly minimize **non-closure** of ABCD constraint used in fit:

$$L_{Closure} = \frac{N_A N_D - N_B N_C}{N_A N_D + N_B N_C}$$

- **Mass regression** layer to infer $M_{\tilde{t}}$ and use as classification input

ABCDisCoTEC Details

Ref.
SUS-23-001 (PAS TBD)

Signal model	channel	low-mass boundaries	high-mass boundaries
RPV	0ℓ	(0.52, 0.54)	(0.74, 0.80)
RPV	1ℓ	(0.84, 0.42)	(0.80, 0.72)
RPV	2ℓ	(0.52, 0.58)	(0.50, 0.50)
SY$\overline{\text{Y}}$	0ℓ	(0.76, 0.70)	(0.54, 0.56)
SY$\overline{\text{Y}}$	1ℓ	(0.44, 0.42)	(0.68, 0.82)
SY$\overline{\text{Y}}$	2ℓ	(0.40, 0.42)	(0.48, 0.48)

- Lower bounds on ($S_{NN,1}$, $S_{NN,2}$) defining the A region for the low- and high-mass optimization for each signal model and search channel

$$\kappa = \frac{N_{\mathbf{A},\mathrm{MC}}}{N_{\mathbf{A},\mathrm{Pred.}}} = \frac{N_{\mathbf{A},\mathrm{MC}} N_{\mathbf{D},\mathrm{MC}}}{N_{\mathbf{B},\mathrm{MC}} N_{\mathbf{C},\mathrm{MC}}}$$

- Correction factor for each N_{jets} category to account for discrepancies between number of tt+jets events predicted and observed in region **A** in simulation

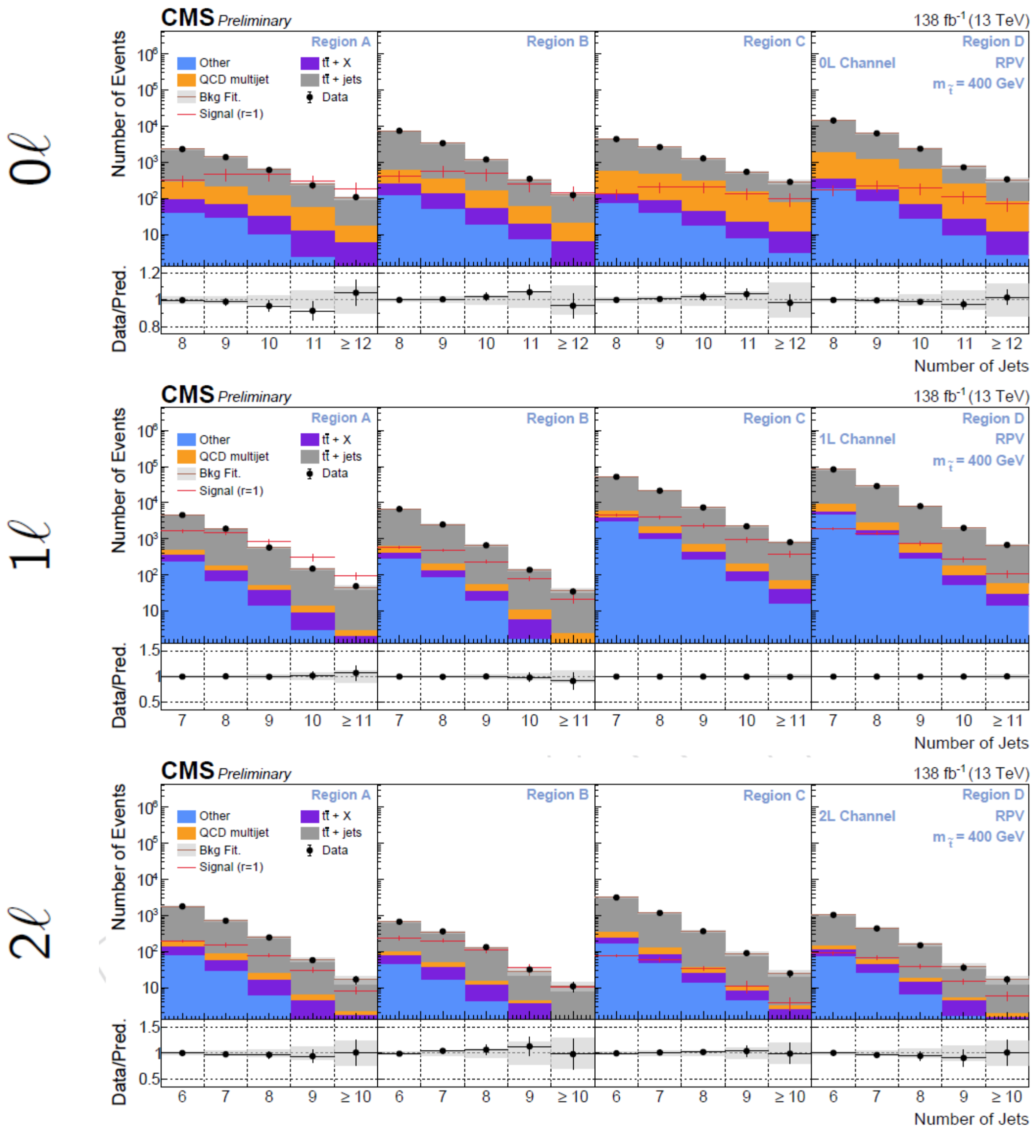
Systematic Uncertainties

Ref. SUS-23-001 (PAS TBD)

Source of uncertainty	0ℓ $t\bar{t}$	0ℓ Other	0ℓ RPV	1ℓ $t\bar{t}$	1ℓ Other	1ℓ RPV	2ℓ $t\bar{t}$	2ℓ Other	2ℓ RPV
PDF	—	0–0 (1)	0–0 (0)	—	0–0 (1)	0–0 (0)	—	0–0 (2)	0–0 (0)
(μ_R, μ_F) scales	—	1–8 (16)	0–1 (2)	—	0–5 (13)	0–1 (2)	—	0–7 (19)	0–1 (2)
FSR	1–3 (3)	1–14 (56)	1–12 (18)	0–1 (3)	1–12 (26)	1–7 (15)	0–6 (9)	2–21 (100)	1–9 (16)
ISR	—	0–10 (17)	1–4 (5)	—	0–8 (15)	1–4 (5)	—	3–11 (17)	0–4 (5)
Pileup	—	0–2 (12)	0–1 (3)	—	0–1 (22)	0–0 (3)	—	0–9 (26)	0–1 (9)
Non-Closure	3	—	—	5	—	—	7	—	—
κ Stat. Unc.	1–4 (7)	—	—	1–7 (8)	—	—	5–15 (19)	—	—
QCD TF	—	1–4 (16)	—	—	0–0 (1)	—	—	0–0 (0)	—
JES	—	4–18 (100)	0–10 (27)	—	1–18 (100)	1–14 (19)	—	4–100 (100)	1–15 (25)
JER	—	0–8 (23)	0–2 (4)	—	0–5 (35)	0–1 (4)	—	0–20 (45)	0–3 (9)
b tagging	—	0–1 (7)	1–2 (3)	—	0–1 (3)	0–0 (0)	—	0–1 (7)	0–0 (2)
t tagging	—	26–33 (42)	26–31 (34)	—	—	—	—	—	—
Jet trigger	—	0–1 (1)	0–0 (1)	—	—	—	—	—	—
Lepton ID	—	—	—	—	3–3 (4)	3–3 (3)	—	5–6 (6)	5–5 (6)
Prefiring	—	0–2 (7)	0–2 (3)	—	0–3 (6)	0–3 (4)	—	0–3 (11)	0–2 (4)
Integrated Luminosity	—	1.6	1.6	—	1.6	1.6	—	1.6	1.6
Theoretical Cross Section	—	20	—	—	20	—	—	20	—

Combination Post-fit Plots

Ref.
SUS-23-001 (PAS TBD)



- Background-only post-fit plots shown for three channels (RPV $M_{\tilde{t}}$=400 GeV)