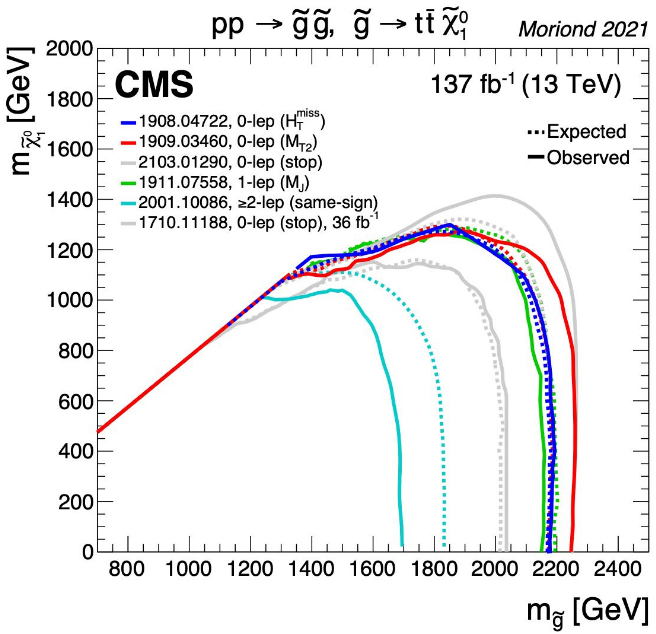
Constraining challenging regions of the SUSY parameter space with the CMS experiment

EPS-HEP Conference 2021

P.E. Meiring (UZH)
On behalf of the CMS Collaboration

Introduction

- No hint yet of New Physics at the Electroweak scale
  → Strong constraints on SUSY particle masses
- Traditional strategy of SUSY searches:
  - Search in events with **high p_T^{miss}** and **high p_T^{vis}**
- Not all phase-space covered
  → SUSY could still hide in corners of phase-space!
- Unconventional signatures → experimental challenges:
  - Particle reconstruction/identification
  - Process modeling
  - Event selection (online/offline)



EWK SUSY

Compressed Higgsinos

Challenge: Identification of (very) soft leptons

Strong SUSY

Light Stops in "Top Corridor"

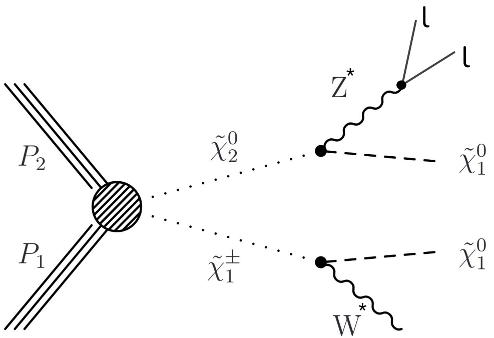
Challenge: Discrimination between signal and kinematically similar background

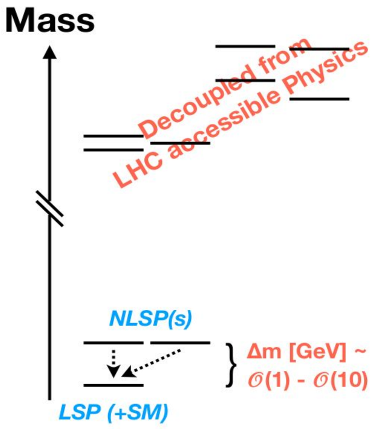
Strong SUSY

R-parity Violating Stops

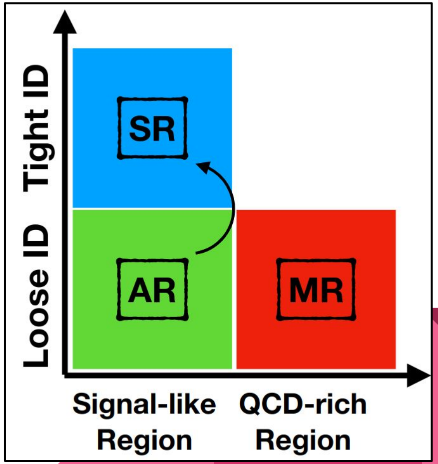
Challenge: Constrain the shape of a multivariate score from data

Compressed Higgsinos





- Natural SUSY most elegant solution to the Hierarchy Problem
  → Suggests $\underbrace{\text{light Higgsinos } (\sim m_H)}_{\text{Accessible at LHC!}}$ with $\underbrace{\text{small mass-splittings}}_{\text{Soft visible decay products}}$
- Moderate p_T^{miss} and soft leptons → **No handle to select events online**!
  - Require ISR jet → induce additional p_T^{miss}
  - Designed dedicated **soft μμ + p_T^{miss} trigger**
- **More challenges**:
  - Weak production → Low cross-section
  - Tail of p_T^{miss} distribution → Low statistics
  - Soft leptons → Many non-prompt lepton candidates
- Main background from non prompt leptons
  - Rely on data: Tight-to-loose method using fake rate (FR)

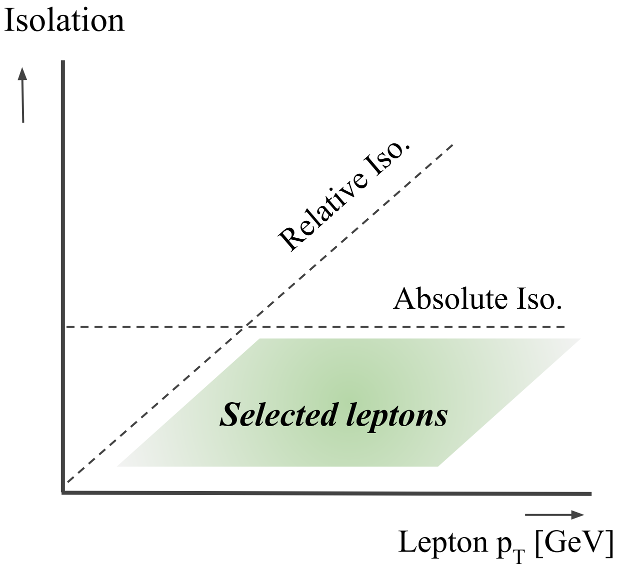


How to define Loose and Tight ID for soft leptons?

Compressed Higgsinos

- <u>Soft muons (electrons)</u>: 3.5 (5) < p_T < 30 GeV
  - Standard lepton ID tuned for harder objects → inadequate!

- Customized Loose/Tight lepton definition
  - **Absolute isolation** (Iso_{abs})
    - Energy sum of particles in $\Delta R < 0.3$
  - **Relative isolation** (Iso_{rel})
    - $Iso_{rel} = Iso_{abs}$ / lepton p_T
    - Tight cut for **softest leptons**
  - **3D impact parameter & significance**
    - Reduces non prompt contamination → **improves s/b**
  - **Electron MVA ID**
    - **p_T dependent working points**

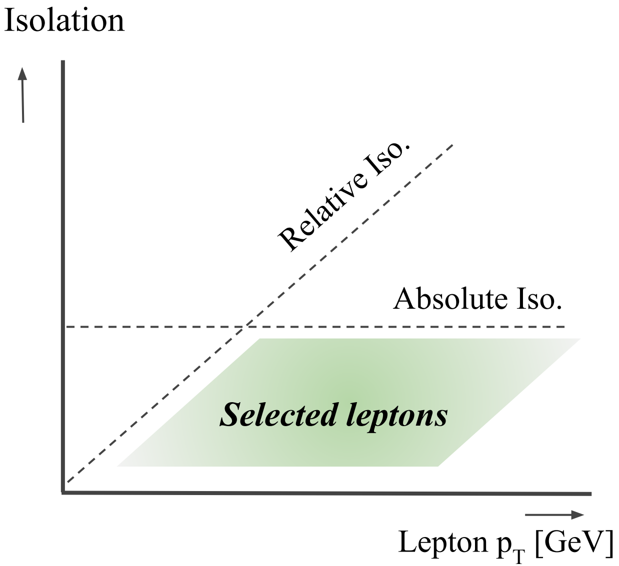


- Important: Jet flavor composition in Measurement Region not same as in Application / Signal Region!

  → <u>Flavor-independence</u> of FR established by careful tuning

Compressed Higgsinos

- **Soft muons (electrons)**: 3.5 (5) < p_T < 30 GeV
  - Standard lepton ID tuned for harder objects → inadequate!
- Customized Loose/Tight lepton definition
  - **Absolute isolation** (Iso_{abs})
    - Energy sum of particles in $\Delta R < 0.3$
  - **Relative isolation** (Iso_{rel})
    - $Iso_{rel} = Iso_{abs}$ / lepton p_T
    - Tight cut for **softest leptons**
  - **3D impact parameter & significance**
    - Reduces non prompt contamination → **improves s/b**
  - **Electron MVA ID**
    - **p_T dependent working points**



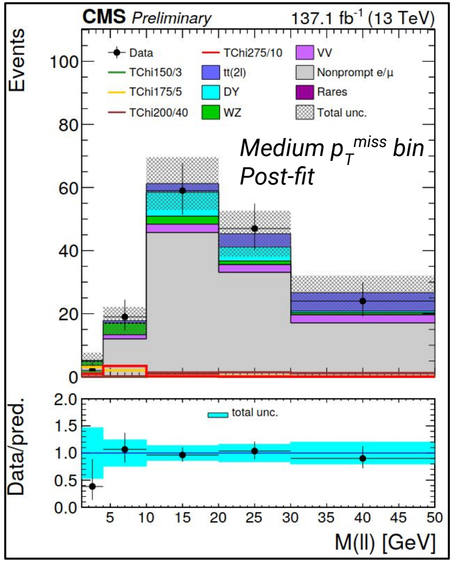
- Important: Jet flavor composition in Measurement Region not same as in Application / Signal Region!
  → Flavor-independence of FR established by careful tuning
- **Yet another challenge!**
  - Low statistics
  - Method attributes negative weights
    → Loss of statistical power

Large uncertainties make fully data-driven estimate unsuitable!

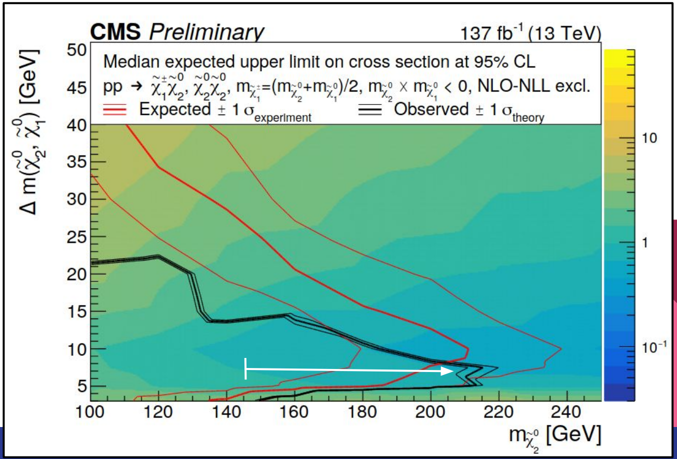
Compressed Higgsinos

- Data-driven **Tight-To-Loose method modified** (details in backup)
  - Both data and MC used → ***New semi-data-driven*** estimation
    → Transfer factor from data
    → Non-prompt m_{ll} shapes from MC templates (large statistics)
  - **Better shape agreement** and **smaller uncertainties**!
- Final fit performed in bins of p_T^{miss} and m_{ll}
  → m_{ll} as proxy for Δm (NLSP, LSP): good discriminator
- Exclusion limits of **Higgsino Simplified Model**
  - 215 GeV for $\Delta m = 7.5$ GeV
  - 150 GeV for $\Delta m = 3$ GeV

  → **Strong limits down to $\Delta m = 3$ GeV**!



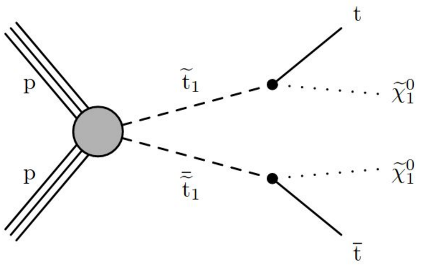
CMS-PAS-SUS-18-004

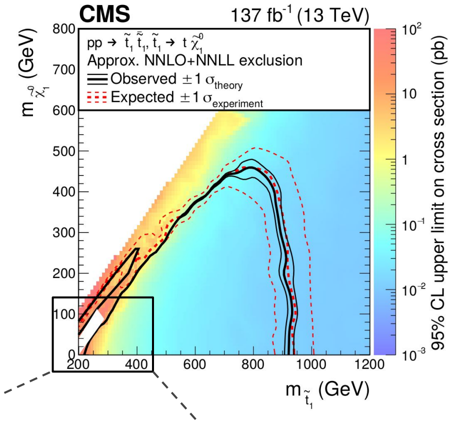


Light Stops in Top Corridor

- SUSY hidden in the **Top Corridor**?
  - $m(\tilde{t}) - m(N1) \simeq m(t)$
- **Pair produced stops** in Top Corridor implies:
  - Stable LSP neutralino as DM candidate
  - Uncovered phase-space, within reach of LHC!
- Phase-space **particularly challenging** if stops are light (<300 GeV)!
  - → Final states very similar to SM tt + low p_T^{miss}
  - → Signal acceptance strongly varying
- CMS analysis dedicated to Top Corridor with full Run 2 data
  - 2 opposite sign leptons
  - ≥2 (b-)jets
  - $p_T^{miss} > 50$ GeV

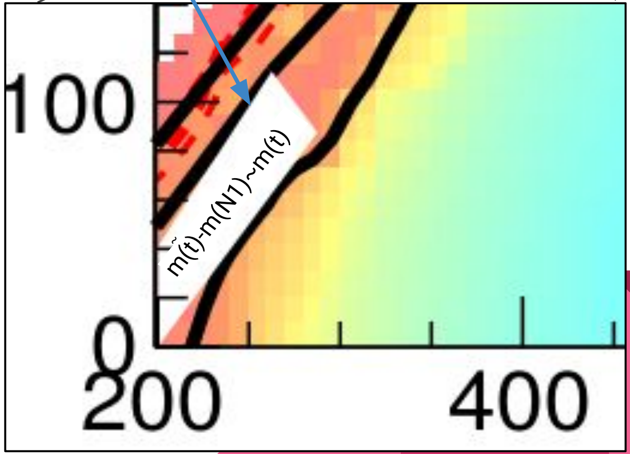
How to distinguish signal from kinematically similar SM background?





Light stops in Top Corridor

arXiv: 2008.05936



Light Stops in Top Corridor

- 90% of background in SR is **<u>irreducible SM tt</u>**
  → Very similar to signal!

- Maximally exploit differences of many kinematic variables
  → **Deep Neural Network** (DNN)

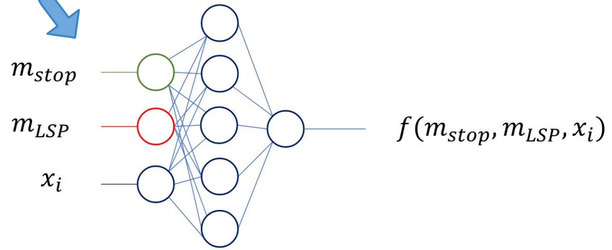
- Signal acceptance strongly varying in Top Corridor
  → <u>Discrimination power differs with</u> **(m_{stop}, m_{LSP})** !

- m_{stop} and m_{LSP} included as input → **<u>parametric</u> DNN**
  - Maximizes discrimination power per mass-hypothesis
  - Randomly assigned to background events

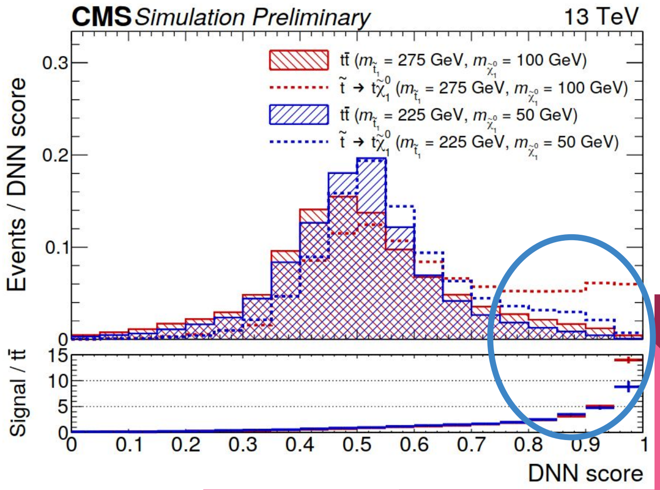
- **Good discrimination** observed at <u>large DNN score</u>!

7-hidden layer DNN with 11 input variables:

$$x_i \begin{cases} m_{T2},\ m_{ll},\ p_T^{miss},\ H_T, \\ p_T(e\mu),\ \Delta\eta(e\mu),\ \Delta\varphi(e\mu), \\ p_T(l_{1,2}),\ \eta(l_{1,2}) \end{cases}$$



CMS-PAS-SUS-20-002



Light Stops in Top Corridor

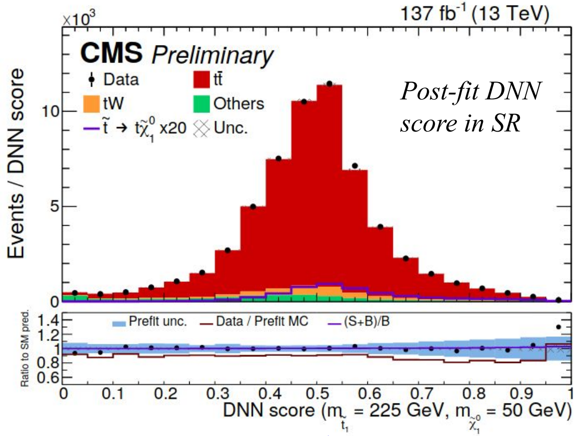
- **DNN score**: single variable with <u>maximum shape difference</u>
  → Fit variable for signal extraction!

- Other backgrounds (MC):
  - tW, DY, non-prompt, ttV, VV

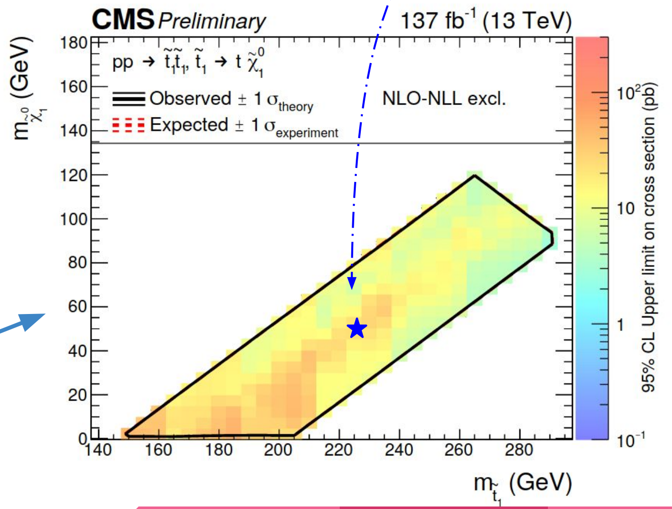
- Simultaneous fit to data performed with DNN scores in 9 channels
  - Year: 2016 / 2017 / 2018
  - Lepton flavors: ee / eμ / μμ

- Result included in the legacy **Run 2 CMS stop combination** analysis
  → Talk by Giulia Collura

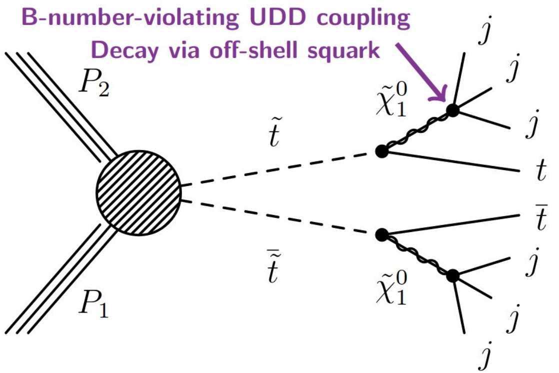
Light stops in Top Corridor excluded



CMS-PAS-SUS-20-002



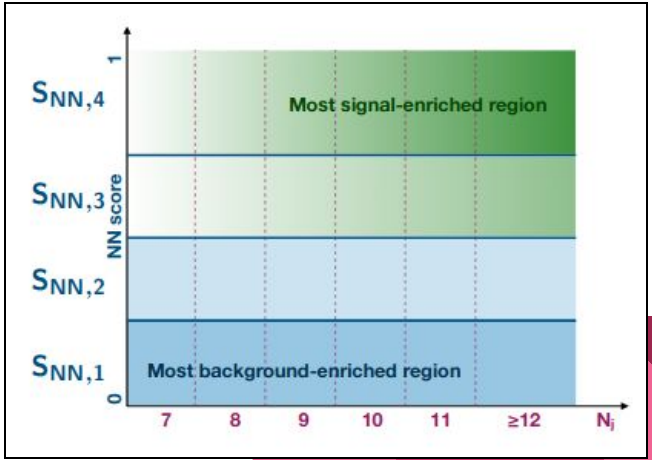
R-parity Violating Stops



- Several SUSY theories predict **R-parity violation** (RPV)
  - Unstable LSPs → decay to SM particles (no high p_T^{miss} signature) !
- **RPV** + **stop pair production** → Final states with:
  - Two SM top-particles
  - **high jet multiplicity (≥10)!**

  ⇒ tt+jets dominant background!
- N_{jets} poorly described by MC → Fit to data
  - But no signal-free CR possible!
- **NN trained** to discriminate signal from tt+jets
  - Extract N_{jets} shape from low NN score region

S_{NN} must be independent of N_{jets}



R-parity Violating Stops

- Two Neural Networks combined:
  - NN1: Classify events as signal or tt+jets background
  - NN2: Predict N_{jets}

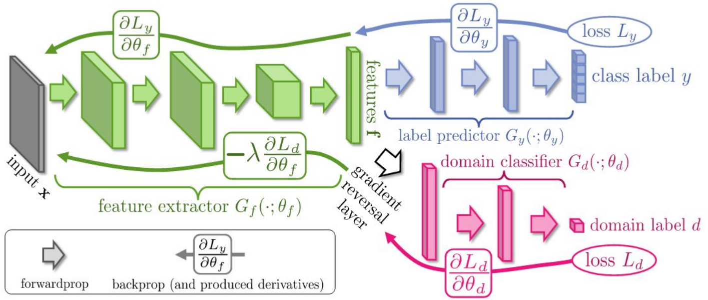
Total loss function to be minimized:

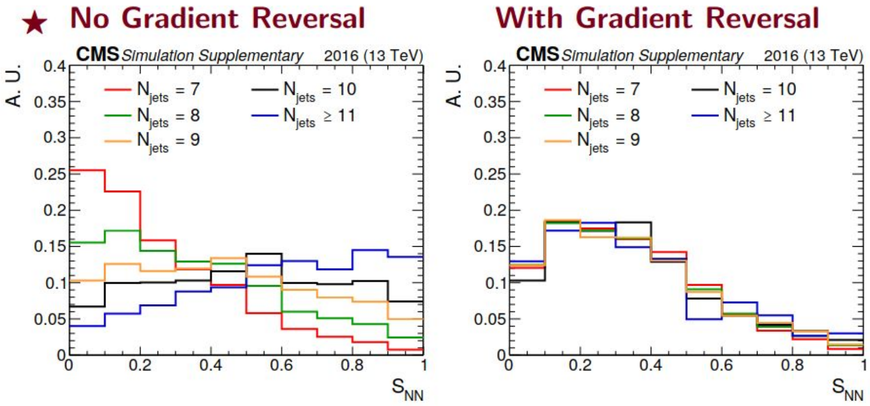
$$L_{tot} = L_{NN1} - \lambda L_{NN2}$$

Gradient reversal (GR) penalizes predictive power in N_{jets}!

- GR also used by other NN-based taggers
  - Example: Jet mass decorrelation in DeepAK8 jet classifier
- Inputs to the network:
  - 4-vectors of 7 highest p_T jets and lepton
  - Jet energy-momentum tensor eigenvalues
  - Jet angular variables

Schematic of gradient reversal
arXiv :1409.7495





arXiv: 2102.06976

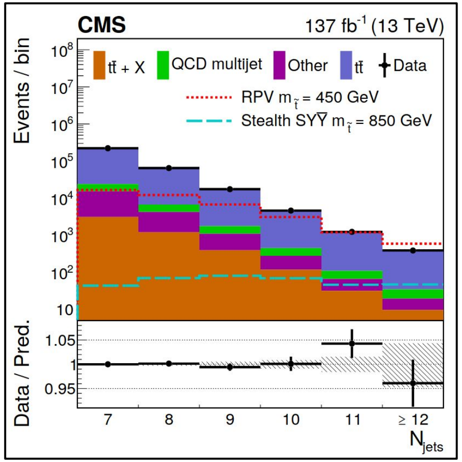
R-parity Violating Stops

- Gradient reversal cures dependency of N_{jets} and S_{NN}
  → Same N_{jets} shape in all S_{NN} bins!

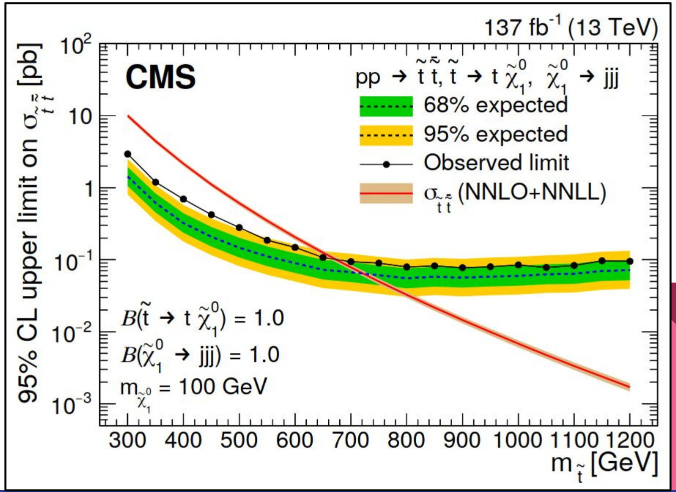
- Simultaneous **fit of N_{jets} in each S_{NN} bin**, using fit components:
  - tt+jets (90%) → $S_{NN,1}$ with **fixed shape across S_{NN}**
  - QCD → CR
  - tt+X → Simulation
  - Other backgrounds (tW, VV, etc.) → Simulation
  - Signal → Simulation

- Exclusion limits in **benchmark RPV model**:
  - **m_{stop} < 670 GeV excluded**
    (Even without requiring large p_T^{miss}!)

Most powerful limits ever set!



arXiv: 2102.06976

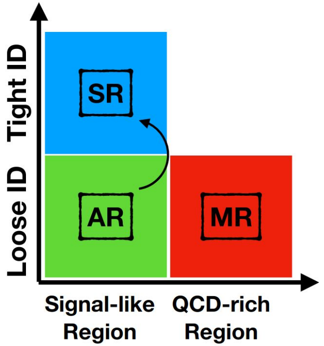


Summary

- Wide range of CMS analyses exploring **corners of SUSY parameter space**
- More unconventional signatures with increasing experimental challenges!
  - **Compressed Higgsinos**
    → Soft lepton identification and non-prompt background estimation
  - **Light Stop pairs in the Top Corridor**
    → Discrimination of signal from nearly identical background
  - **R-parity violating Stops**
    → Constraining the shape of a multivariate score from data
- Many searches with improved and innovative analysis strategies anticipated
  - Long-lived SUSY searches in Run 3!
    → New Physics parallel session on Tuesday!

Back-up

Semi-data-driven non-prompt estimation



- Fake rate measured in QCD enriched Measurement region (MR)
- Application Region (AR) divided into “side-bands” depending on nLeptons passing Loose ID but failing Tight ID
  - Per side-band a scale factor (SF) is derived as:

$$\text{Semi-DD SF} = \frac{\text{Data - MC (prompt bkg)}}{\text{MC (non-prompt bkg)}}$$

- Non-prompt MC events in each side-band of the AR are weighted by: $\boxed{\text{Weight} = \text{Semi-DD SF} \;\; \text{x} \;\; W_{FR}}$

nEvents with leptons passing (p) and/or failing (f) the tight ID

p_i (f_i) is the prompt (fake) rate given the p_T, η of lepton i

nEvents with leptons prompt (1) and/or non-prompt (0)

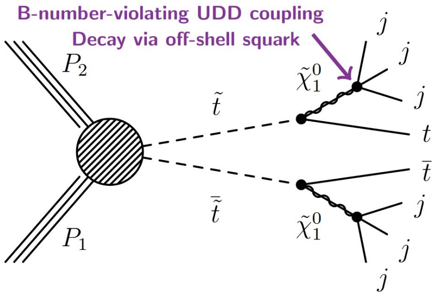
Weights to apply to events with at least 1 lepton failing the tight ID (i.e. in the AR)

$$\begin{pmatrix} N_{pp} \\ N_{pf} \\ N_{fp} \\ N_{ff} \end{pmatrix} = \begin{bmatrix} p_1 p_2 & p_1 f_2 & f_1 p_2 & f_1 f_2 \\ p_1(1-p_2) & p_1(1-f_2) & f_1(1-p_2) & f_1(1-f_2) \\ (1-p_1)p_2 & (1-p_1)f_2 & (1-f_1)p_2 & (1-f_2)f_2 \\ (1-p_1)(1-p_2) & (1-p_1)(1-p_f) & (1-f_1)(1-p_2) & (1-f_1)(1-f_2) \end{bmatrix} \begin{pmatrix} N_{11} \\ N_{10} \\ N_{01} \\ N_{00} \end{pmatrix}$$

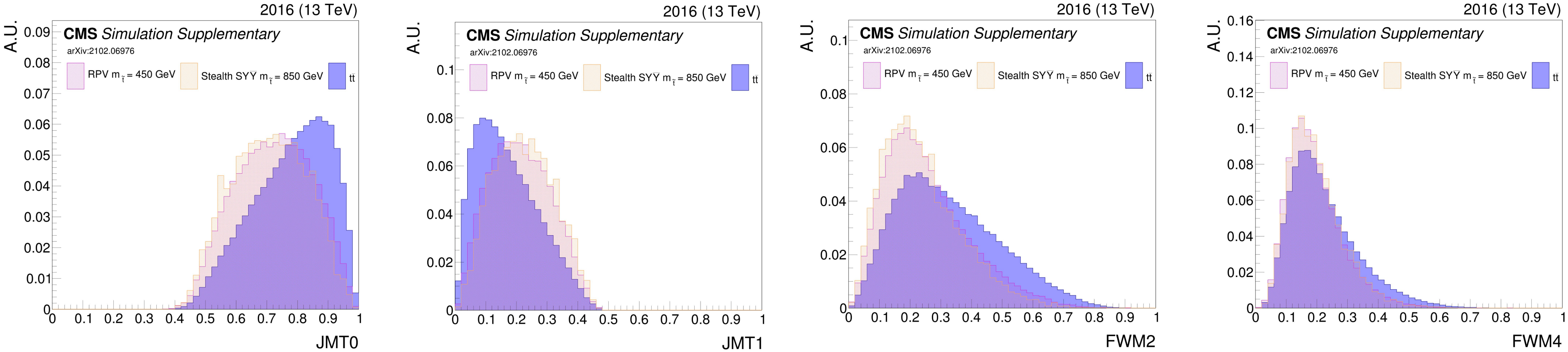
Invert matrix equation and assume p>>f

$$W_{pf} = \frac{f_2}{1-f_2} \qquad W_{fp} = \frac{f_1}{1-f_1} \qquad W_{ff} = \frac{-f_1 f_2}{(1-f_1)(1-f_2)}$$

R-parity Violating Stops



Example distributions (2016 MC) of input variables of the NN



Jet momentum tensor eigenvalues quantify the "flow" of energy and sphericity of the event. They are eigenvalues of the tensor defined as:

$$S^{\alpha,\beta} = \frac{\sum_i p_i^{\alpha} p_i^{\beta}}{\sum_j |p_j|^2}$$

Fox-Wolfram moments account for momenta and angles of the 7 highest p_T jets. Hence, a measure of the "event shape". Defined as:

$$H_l = \sum_{i,j} |p_i||p_j| P_l(\cos\theta_{ij})$$

R-parity Violating Stops

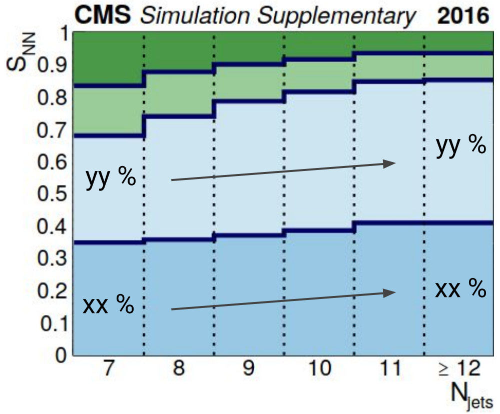
- After gradient reversal in the NN, only **residual dependencies** of N_{jets} and S_{NN} remain

  → Edges of the S_{NN} bins are adjusted per-N_{jets} to maximize signal significance while requiring equal tt+jets background fractions in each N_{jets} bin given a S_{NN} bin.

  Signal significance maximized with metric:

  $$\sum_{N_{jets}=7}^{12} \sum_{S_{NN,i=1}}^{4} N_{sig} / \sqrt{N_{t\bar{t}} + (0.20 \times N_{t\bar{t}})^2 + N_{QCD} + (0.20 \times N_{QCD})^2}$$

  → Also a dedicated systematic unc. applied



- During the final simultaneous fit, the tt+jets N_{jets} shape is parametrized as

  $$R(i) = a_2 + \left[\frac{(a_1 - a_2)^{i-7}}{(a_0 - a_2)^{i-9}} \right]^{1/2}$$

  where $i = N_{jets}$, $a_0 = \frac{M_8}{M_7}$, $a_1 = \frac{M_{10}}{M_9}$, $a_2 = \lim_{x \to \infty} \frac{M(x+1)}{M(x)}$

  and M_i is the number of events with N_{jets}=i

Analyses selections (Compressed Higgsinos)

Table 1: Definition of the MET bins of the SRs. The boundaries (in GeV) of p_T^{miss} and $p_T^{miss,corr}$ of every bin are described.

Search Region	Low MET		Med MET	High MET	Ultra MET
	p_T^{miss}	$p_T^{miss,corr}$	$p_T^{miss,corr}$	$p_T^{miss,corr}$	$p_T^{miss,corr}$
2ℓ Ewk	>125	$[125, 200]$	$[200, 240]$	$[240, 290]$	>290
2ℓ Stop	>125	$[125, 200]$	$[200, 290]$	$[290, 340]$	>340
3ℓ Ewk	>125	$[125, 200]$		>200	

Table 2: List of all criteria that events must satisfy in order to be selected in one of the SRs.

Variable	2ℓ Ewk Low	2ℓ Ewk Med/High/Ultra	2ℓ Stop Low	2ℓ Stop Med/High/Ultra	3ℓ Ewk Low	3ℓ Ewk Med/High/Ultra
N_{lep}	2	2	2	2	3	3
$p_T(\ell_1)$ [GeV] for μ(e)	[5,30]	[3.5(5),30]	[5,30]	[3.5(5),30]	[5,30]	[3.5(5),30]
$p_T(\ell_2)$ [GeV] for μ(e)	[5,30]	[3.5(5),30]	[5,30]	[3.5(5),30]	[5,30]	[3.5(5),30]
$p_T(\ell_3)$ [GeV] for μ(e)	-	-	-	-	[5,30]	[3.5(5),30]
1 OS pair	✓	✓	✓	✓	✓	✓
1 OSSF pair	✓	✓	-	-	✓	✓
$M(\ell\ell)_{SFOS}$ ($M(\ell\ell)^{min}_{SFOS}$ in 3ℓ) [GeV]	[4,50]	[1,50]	[4,50]	[1,50]	[4,50]	[1,50]
$M(\ell\ell)_{SFOS}$ ($M(\ell\ell)^{min}_{SFOS}$ in 3ℓ) [GeV]	veto[9,10.5]		veto[9,10.5]		veto[9,10.5]	
$M(\ell\ell)_{SFOS}$ ($M(\ell\ell)^{min}_{SFOS}$ in 3ℓ) [GeV]	veto[3,3.2]		veto[3,3.2]		veto[3,3.2]	
$M(\ell\ell)^{max}_{SFAS}$(AS= any sign) [GeV]	-	-	-	-	<60	
$p_T(\ell\ell)$ [GeV]	>3		>3		>3	
$M_T(\ell_i, p_T^{miss})$ [GeV] $(i = 1, 2)$	<70		-		-	
$\Delta R(\ell\ell)_{ij}$ (i,j=1,2,3)	-	>0.3	-	>0.3	-	>0.3
p_T^{miss} [GeV]	>125		>125		>125	
$p_T^{miss,corr}$ [GeV]	>125		>125		>125	
H_T [GeV]	>100		>100		>100	
p_T^{miss}/H_T	[2/3,1.4]		[2/3,1.4]		[2/3,1.4]	
Jet ID tight WP for leading jet	✓		✓		✓	
$N_b(p_T > 25$ GeV)	=0		=0		=0	
$M_{\tau\tau}$	veto[0,160]		veto[0,160]		veto[0,160]	

Electrons

Cut	LooseNotTight	Tight
p_T	≥ 5 GeV	≥ 5 GeV
$\lvert\eta\rvert$	≤ 2.5	≤ 2.5
IP_{3D}	< 0.0175 cm	< 0.01 cm
$\sigma_{IP_{3D}}$	< 2.5	< 2.0
d_{xy}	< 0.05 cm	< 0.05 cm
d_z	< 0.1 cm	< 0.1 cm
Iso_{abs}	$< 20.0\,\mathrm{GeV} + \frac{300.0\,\mathrm{GeV}^2}{p_T}$	< 5 GeV
Iso_{rel}	< 1.0	< 0.5
Deep CSV veto	–	custom loose WP
Electron MVA ID	custom ID	tight WP
No missing pixel hits	✓	✓
No conversion vertex	✓	✓

Muons

Cut	LooseNotTight	Tight
p_T	≥ 3.5 GeV	≥ 3.5 GeV
$\lvert\eta\rvert$	≤ 2.4	≤ 2.4
IP_{3D}	< 0.0175 cm	< 0.01 cm
$\sigma_{IP_{3D}}$	< 2.5	< 2.0
d_{xy}	< 0.05 cm	< 0.05 cm
d_z	< 0.1 cm	< 0.1 cm
Iso_{abs}	$< 20.0\,\mathrm{GeV} + \frac{300.0\,\mathrm{GeV}^2}{p_T}$	< 5 GeV
Iso_{rel}	< 1.0	< 0.5
Deep CSV veto	–	custom loose WP
loose ID	✓	✓
soft ID	✓	✓

Analyses selections (Light Stops in Top Corridor)

Baseline event selection

Opposite-sign dilepton pair

m_{ll} > 20 GeV

Leading lepton p_T > 25 GeV

At least 2 jets

At least 1 b-tagged jet

$|m_{(ee,\mu\mu)} - m_Z|$ > 15 GeV

Additional SR selections

p_T^{miss} > 50 GeV

m_{T2} > 80 GeV

Muons

Tight ID

p_T > 20 GeV and $|\eta|$ < 2.4

RelIso04 < 0.15

$d_z \leq 0.10$ cm and $d_{xy} \leq 0.05$ cm

Electrons

Tight cut-based ID (Fall-94X-V2)

p_T > 20 GeV and $|\eta|$ < 2.4

RelIso03 year/ η dependent

$d_z \leq 0.10$ (0.20) cm for barrel (endcaps)

$d_{xy} \leq 0.05$ (0.10) cm for barrel (endcaps)

Jets and b-jets

Tight or tightLepVeto ID

p_T > 30 GeV and $|\eta|$ < 2.4

ΔR(jet, lepton) > 0.4

b-tagged jets: medium WP of Deep Jet tagger

Analyses selections (R-Parity Violating Stops)

Selection	Jets	Electrons	Muons
Quality	AK4 PFJets with CHS DeepCSV medium WP for b jets	Tight cut-based ID Mini-isolation < 0.1	Medium ID Mini-isolation < 0.2
$\|\eta\|$	< 2.4		
p_T	> 30 GeV	2016: > 30 GeV 2017/2018: > 37 GeV	> 30 GeV

SR event selection

Require at least 7 jets
$H_T > 300$ GeV
At least one b jet
Exactly one lepton
$50 < M_{b,\ell} < 250$ GeV

Abstract

Supersymmetric models are characterized by a strong diversity of experimental signatures. Since general-purpose searches have not yet given any clear indication of new physics, dedicated methodologies and tools have been developed to target the regions of the parameter space where the analysis is most challenging and SUSY might still lie undetected. This presentation will describe relevant examples among searches performed by the CMS Collaboration using the full dataset of proton-proton collisions collected during the Run 2 of the LHC at a center-of-mass energy of 13 TeV.