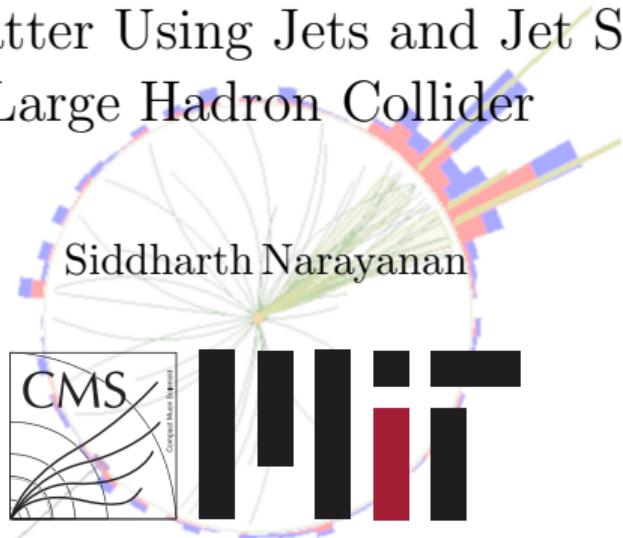


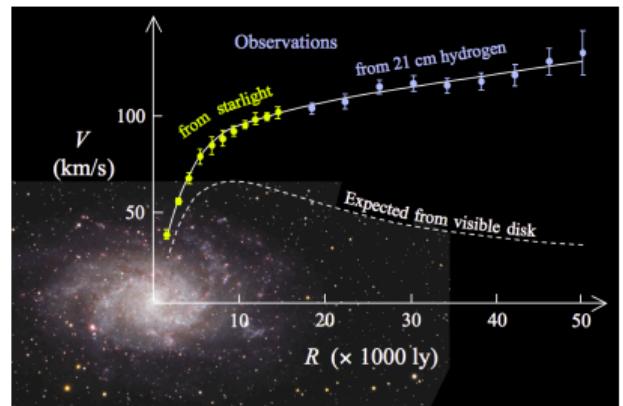
Searching for Dark Matter Using Jets and Jet Substructure at the Large Hadron Collider



Ph.D. Thesis Defense - 2019/01/25

Dark matter - in space

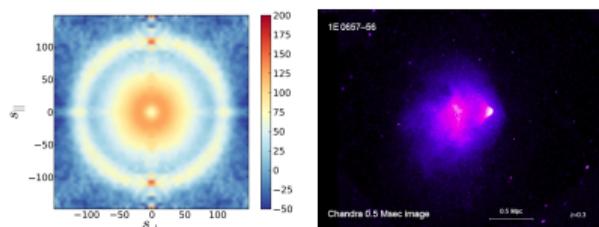
Strong astrophysical evidence for DM:



[1,2]



[3]

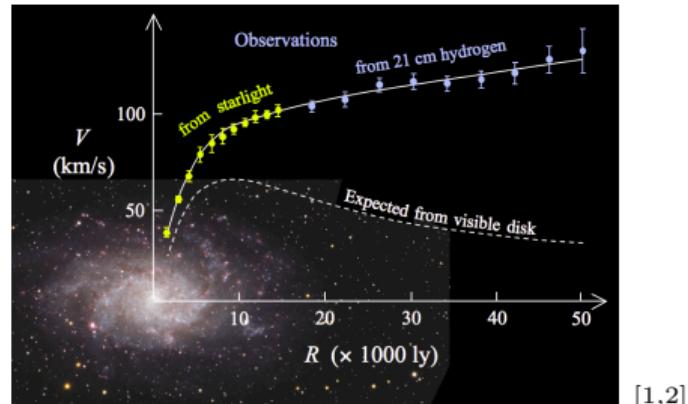


[4]

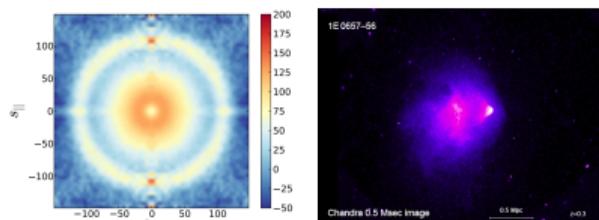
[5]

Dark matter - in space

Strong astrophysical evidence for DM:



[3]



Weakly Interacting Massive Particles

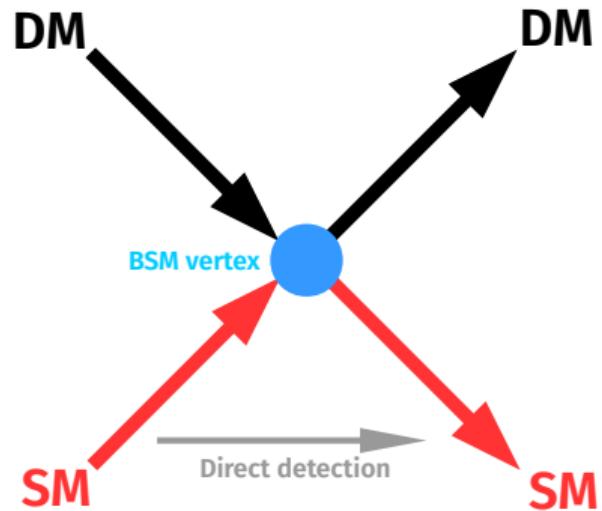
- ▶ Weakly: DM-SM coupling $g_\chi \sim g$
- ▶ Massive: mass $m_\chi \sim 100$ GeV
- ▶ The “WIMP miracle”:

$$\Omega \propto \frac{\rho}{\rho_c} \propto \frac{1}{\langle \sigma v \rangle}$$

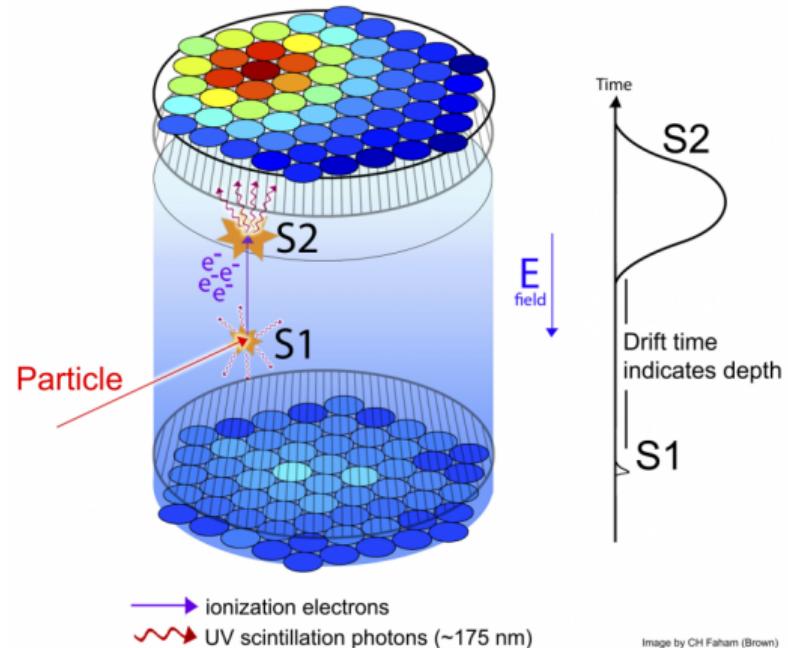
$$\Omega_{\text{meas.}} = 0.12, \quad \Omega_\chi \sim 0.1$$

- ▶ Almost any DM model probed by a collider has a WIMP
- ▶ Many other models exist (axions, MACHOs, ...)

Dark matter - in a laboratory

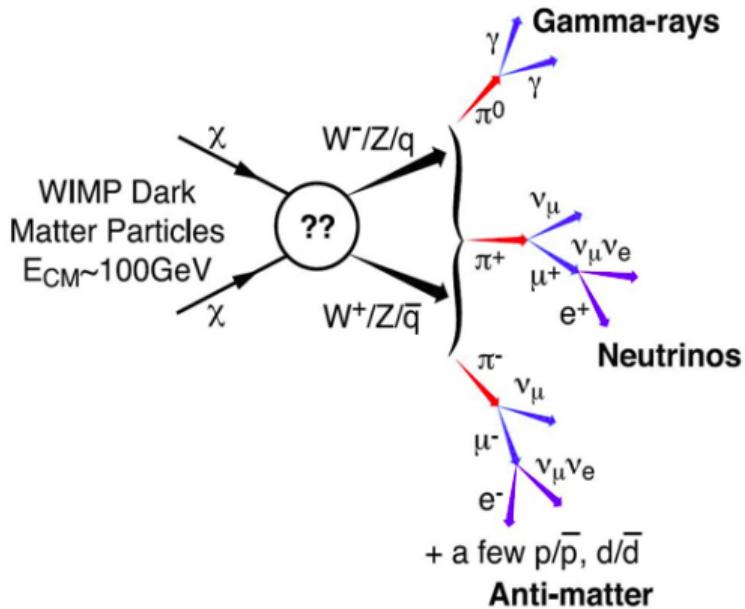
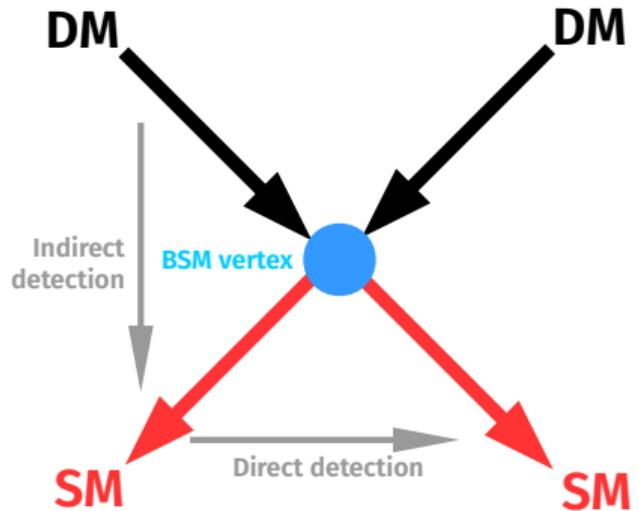


Search for DM-SM interactions as Earth moves through DM halo



LUX (shown), PandaX, CDMS, Xenon1T,...

Dark matter - in a laboratory

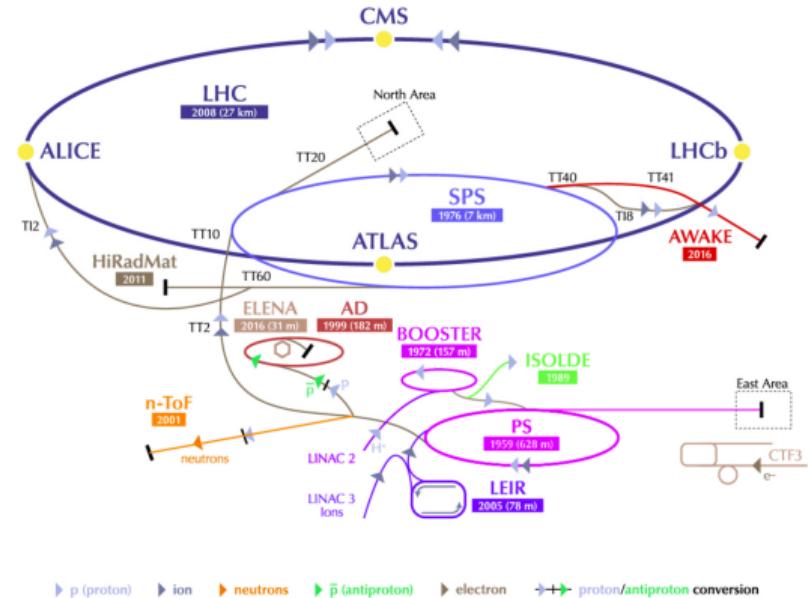
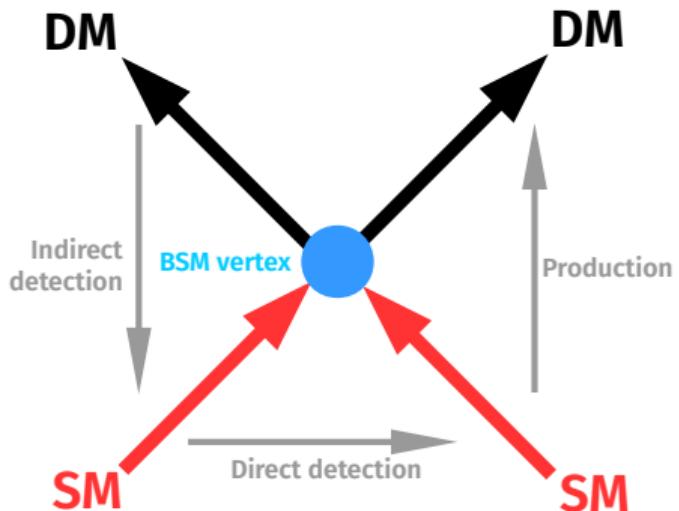


[7]

Look for SM remnants of DM-DM
annihilation in space

FermiLAT, AMS

Dark matter - in a laboratory

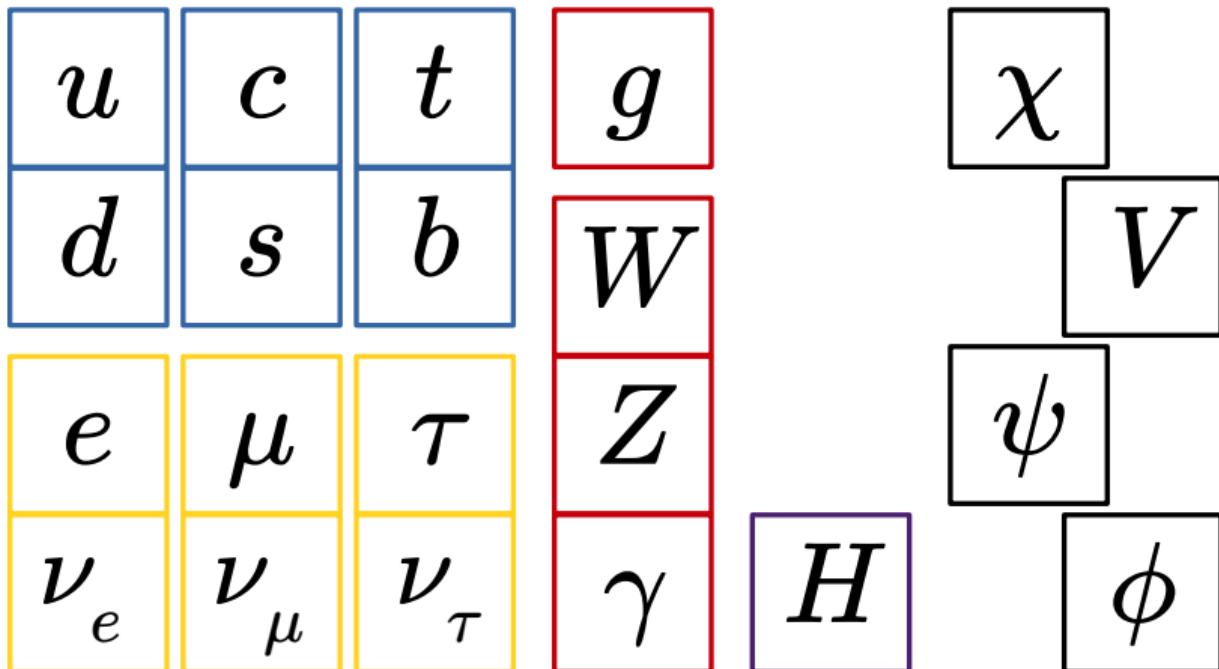


Exploit DM-SM interaction to produce DM
in a laboratory

Production of WIMPs at a hadron collider

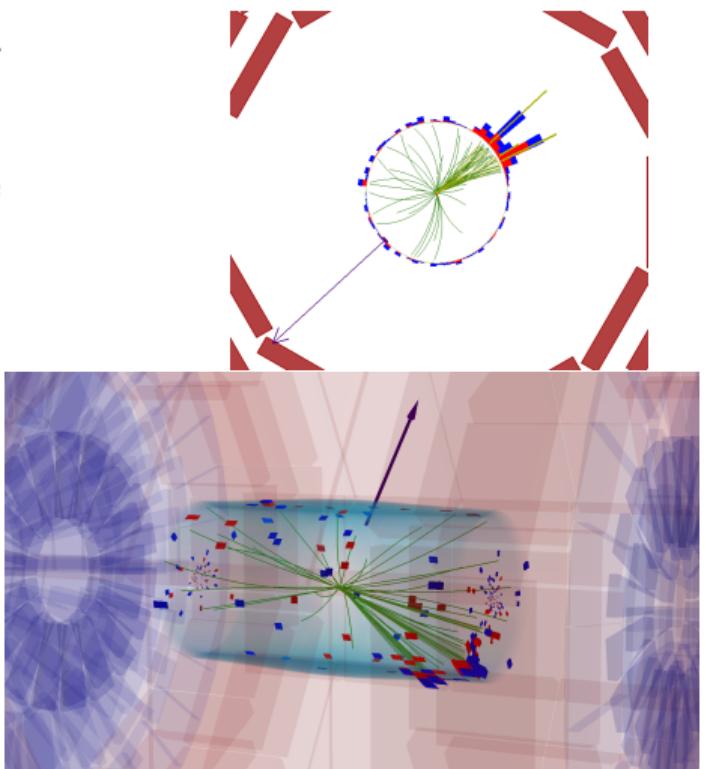
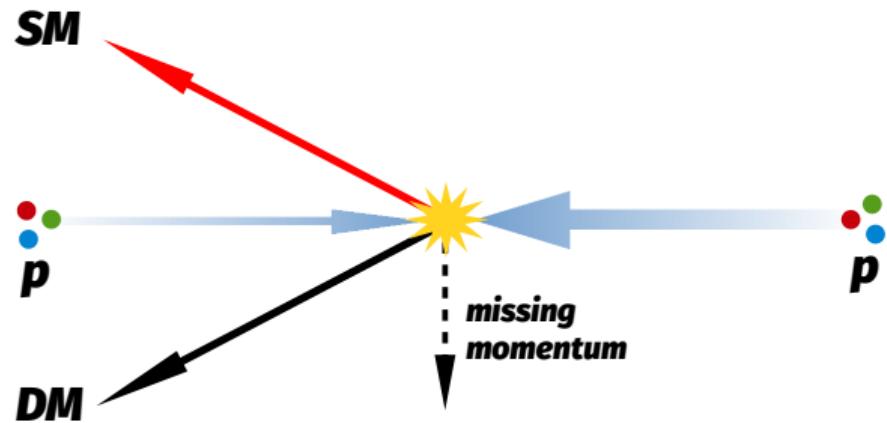
Effective coupling to quarks/gluons $\gtrsim 10^{-4}$

Masses $\lesssim \sqrt{s}$



Seeing the invisible at a hadron collider

- ▶ By definition, DM will not interact with a detector
- ▶ Look for production of DM with SM particle(s)
- ▶ Key observable is transverse momentum imbalance



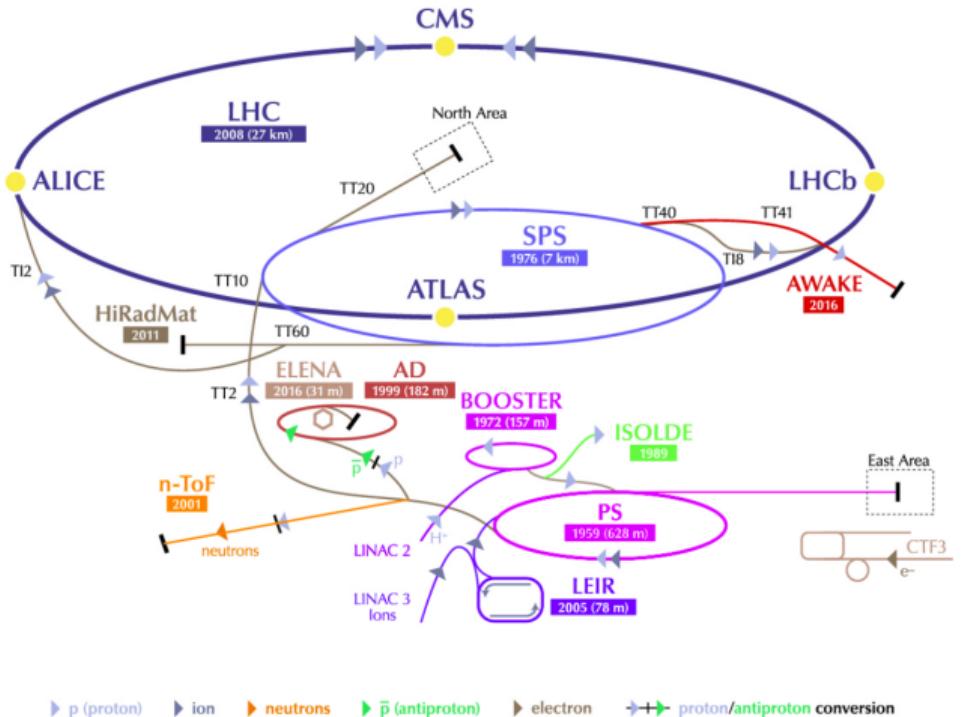
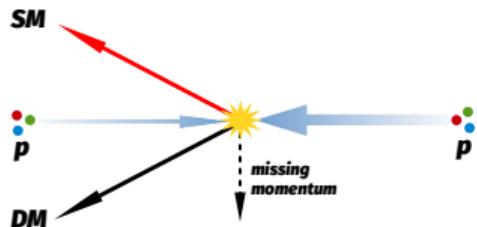
Which SM particle(s)?

Many to choose from. SM particle choice \Leftrightarrow type of DM you can look for

	SM particle	Minimal extension	Higgs-like	Extra dimensions	Extended Higgs sector	Flavor violation
Quarks	$q(g)$					
	t					
	qq'					
	$t\bar{t}$					
	$b/b\bar{b}$					
Gauge bosons	γ					
	$V \rightarrow q\bar{q}'$					
	$Z \rightarrow \ell^+\ell^-$					
Higgs	$H \rightarrow x\bar{x}$					

p_T^{miss} at the LHC

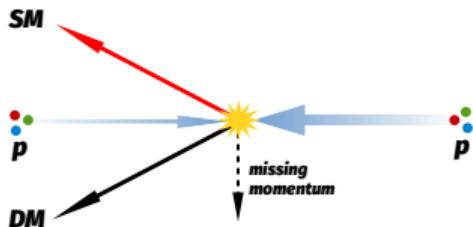
- CMS records collisions from the LHC
 - Today: $\sqrt{s} = 13 \text{ TeV}$ pp collision data from 2016
- Missing momentum:



p_T^{miss} at the LHC

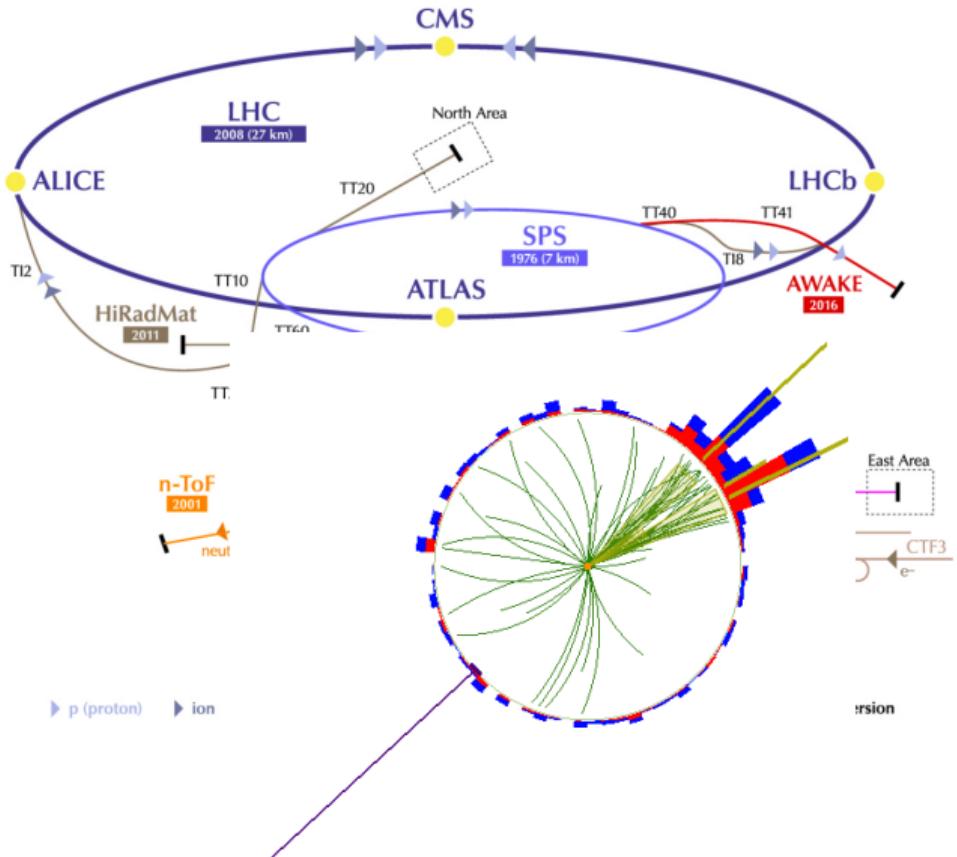


- ▶ CMS records collisions from the LHC
 - ▶ Today: $\sqrt{s} = 13 \text{ TeV}$ pp collision data from 2016
 - ▶ Missing momentum:



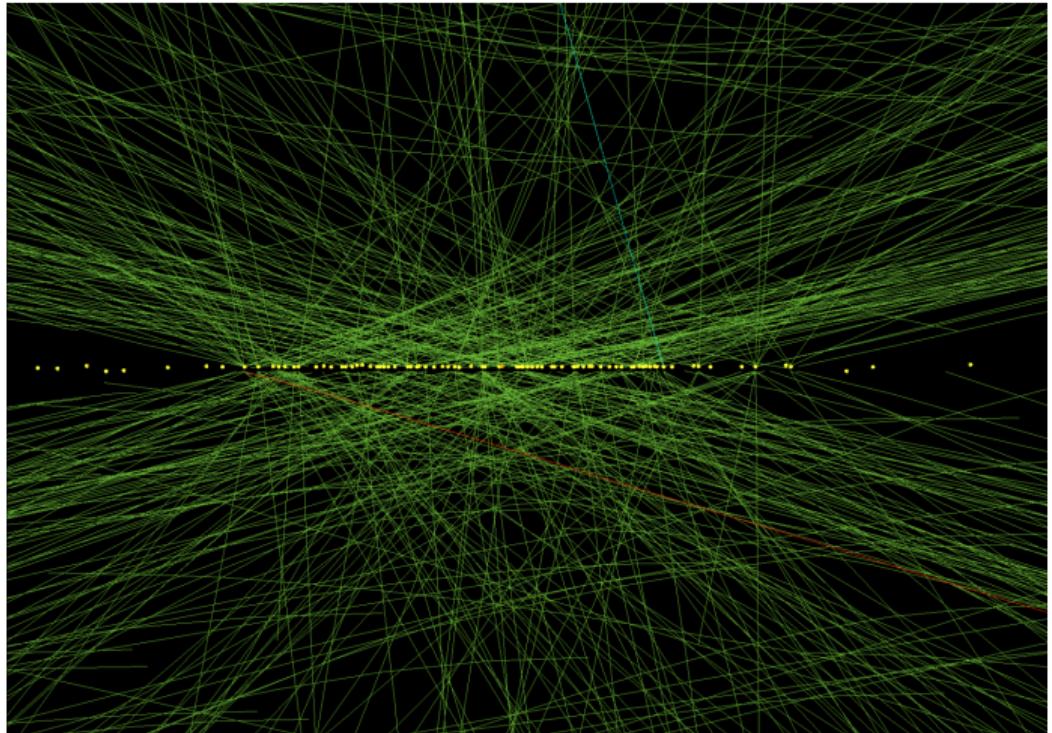
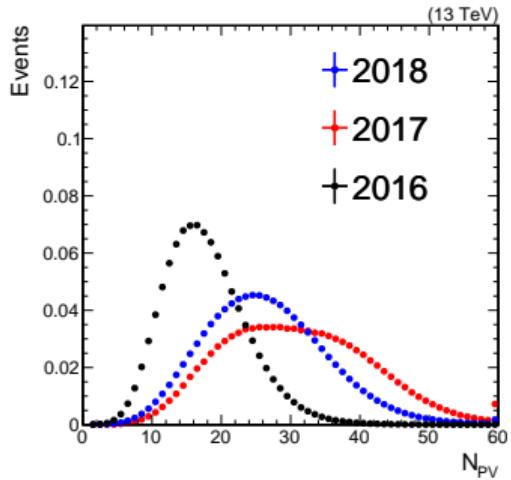
- turns into:

$$\vec{p}_T^{\text{miss}} = - \sum_{i \in \text{particles}} \vec{p}_{T,i}$$



Proton collisions

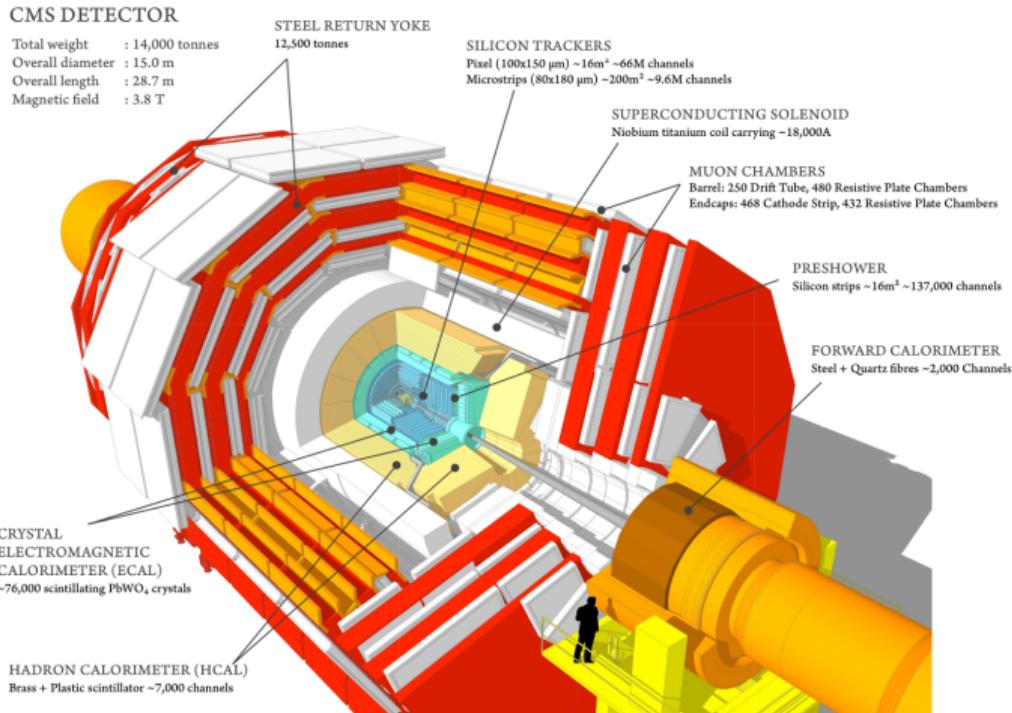
- ▶ Two 6.5 TeV proton beams
- ▶ $\sim 10^{11}$ protons per bunch
- ▶ Average collision has 10-25 pile-up interactions



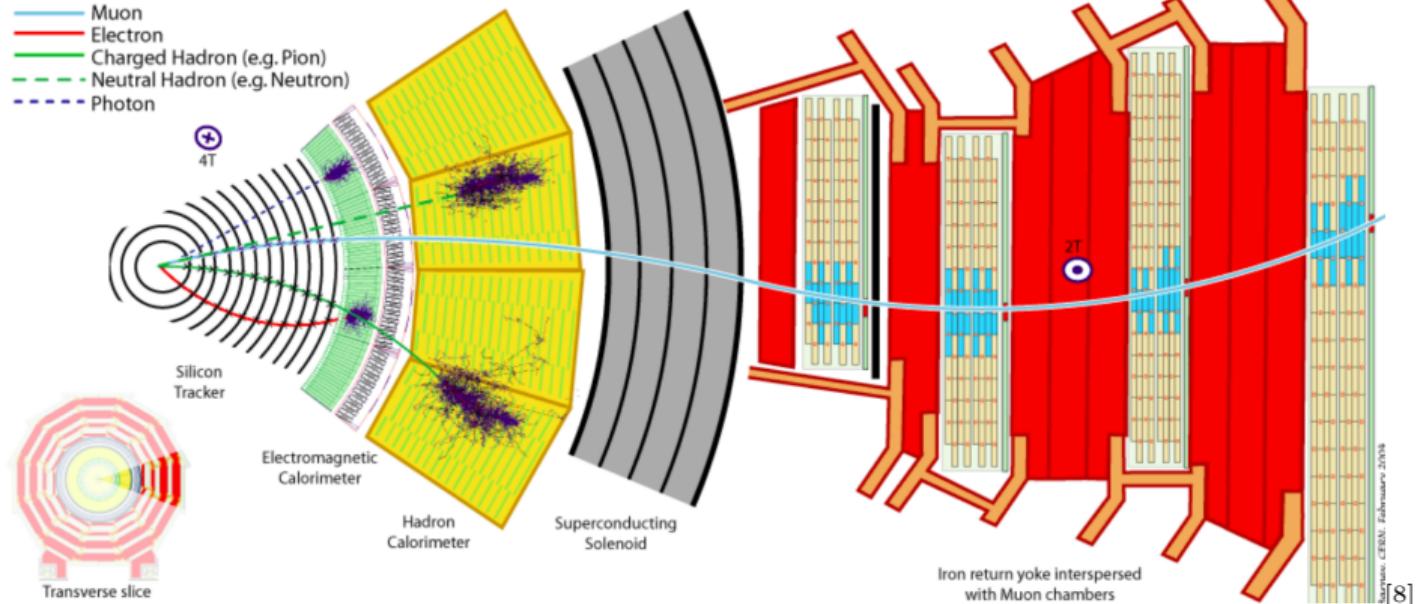
Compact Muon Solenoid

All particles in sum \Rightarrow
all subdetectors help measure p_T^{miss} !

- ▶ Solenoidal magnet
 - ▶ 3.8 T B field
- ▶ Silicon tracker
 - ▶ Charged particles' \vec{p}
 - ▶ Track vertices
- ▶ Calorimeters
 - ▶ EM and hadronic
 - ▶ Good energy resolution
 - ▶ Large coverage
- ▶ Muon chambers
 - ▶ ID muons
 - ▶ Help measure \vec{p}



Particle reconstruction



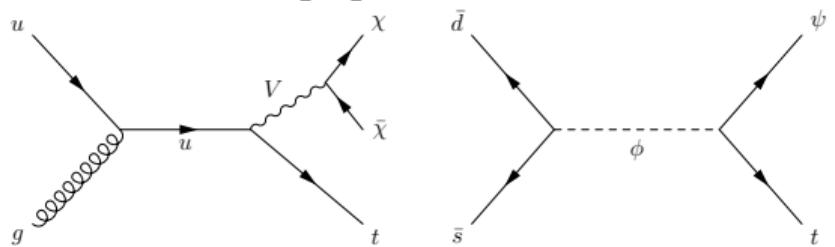
Tracker	ECAL	HCAL	Muon
$ \eta < 2.5$ $\frac{0.015\% \cdot p_T}{\text{GeV}} \oplus 0.5\%$	$ \eta < 3$ $\frac{3\%}{\sqrt{E/\text{GeV}}} \oplus \frac{12\%}{E/\text{GeV}} \oplus 0.3\%$	$ \eta < 5$ $\frac{85\%}{\sqrt{E/\text{GeV}}} \oplus 7.4\%$	$ \eta < 2.4$ 3% at 100 GeV

Outline of this talk

DM production mode

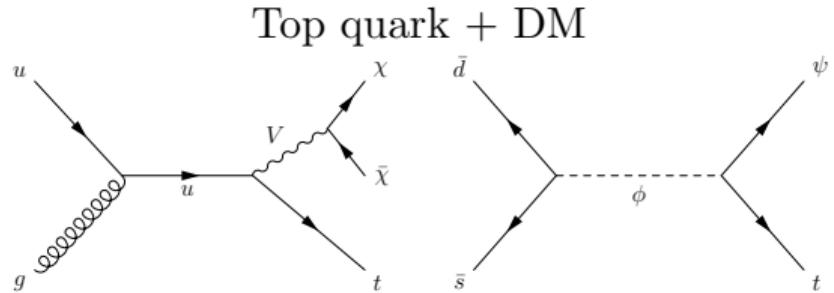
Highlights

Top quark + DM



Outline of this talk

DM production mode



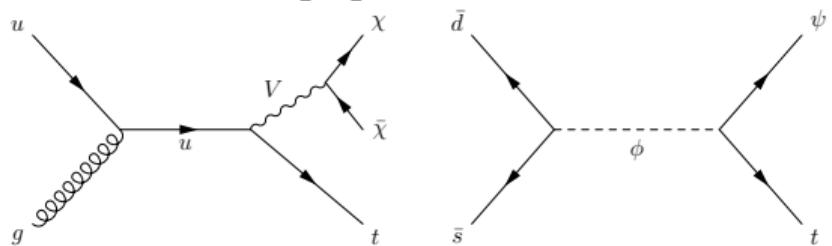
Highlights

Jet substructure
Invisible background estimation

Outline of this talk

DM production mode

Top quark + DM

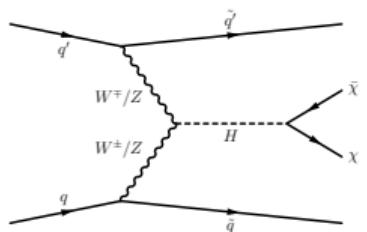


Highlights

Jet substructure

Invisible background estimation

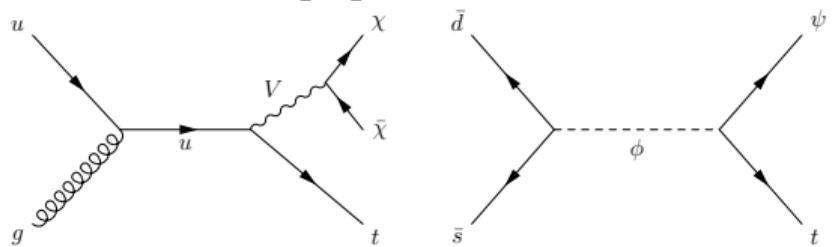
Higgs \rightarrow DM



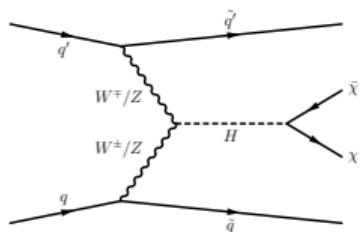
Outline of this talk

DM production mode

Top quark + DM



Higgs \rightarrow DM



Highlights

Jet substructure
Invisible background estimation

Forward jets

Electroweak SM backgrounds

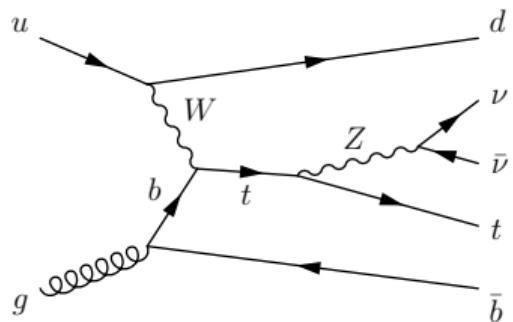
Mono-Top

Hallmarks of top quark+ p_T^{miss}

- ▶ Final state violates flavor conservation
 - ▶ SM will have b quark in the final state
- ▶ Excess mono-top production \Rightarrow flavor-changing BSM

Hallmarks of top quark+ p_T^{miss}

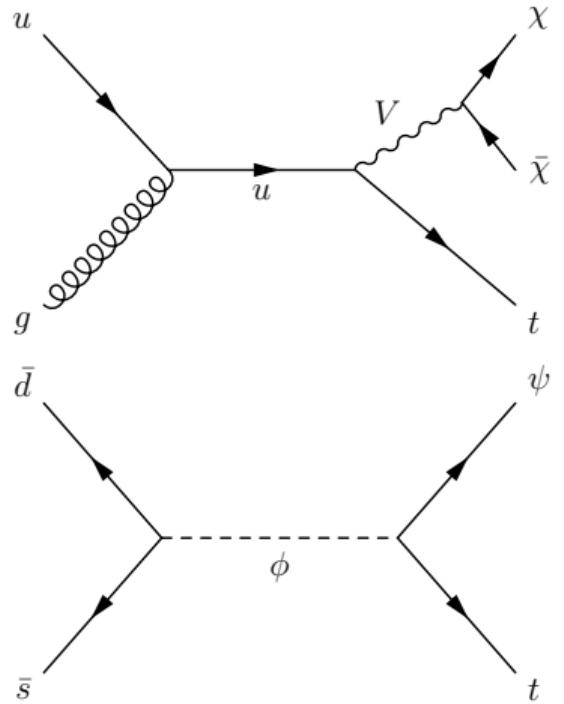
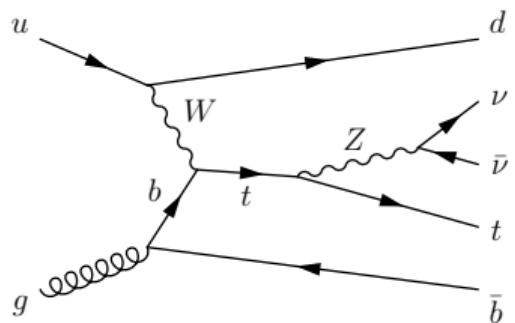
- ▶ Final state violates flavor conservation
 - ▶ SM will have b quark in the final state
- ▶ Excess mono-top production \Rightarrow flavor-changing BSM
- ▶ Leading SM process: 0.14 pb \Rightarrow 5000 events in 36/fb



Hallmarks of top quark+ p_T^{miss}

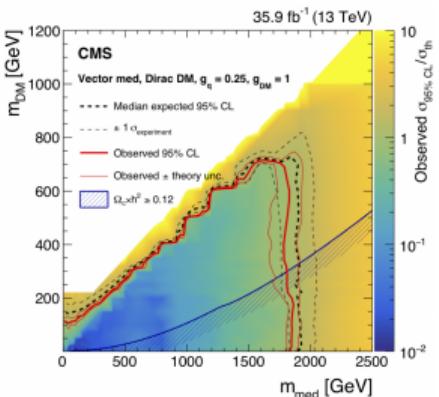
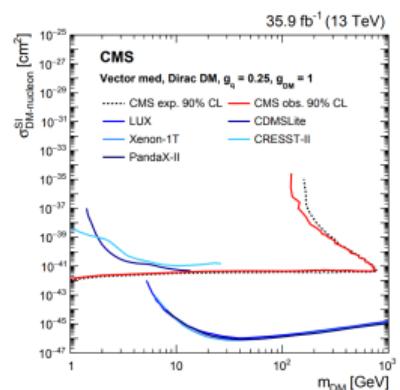
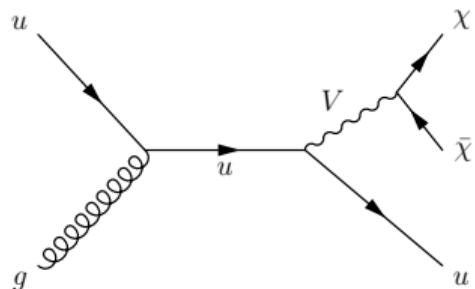


- ▶ Final state violates flavor conservation
 - ▶ SM will have b quark in the final state
 - ▶ Excess mono-top production \Rightarrow flavor-changing BSM
 - ▶ Leading SM process: $0.14 \text{ pb} \Rightarrow 5000 \text{ events in } 36/\text{fb}$

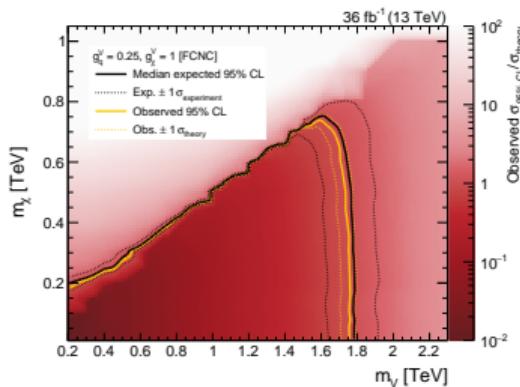
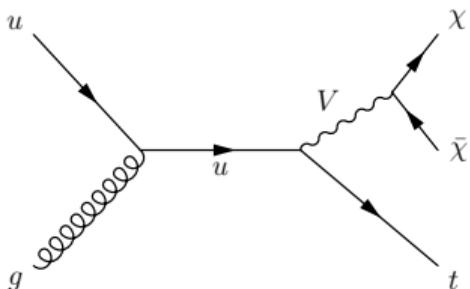


Connection to other DM models

Flavor conserving: diagonal g_u^V

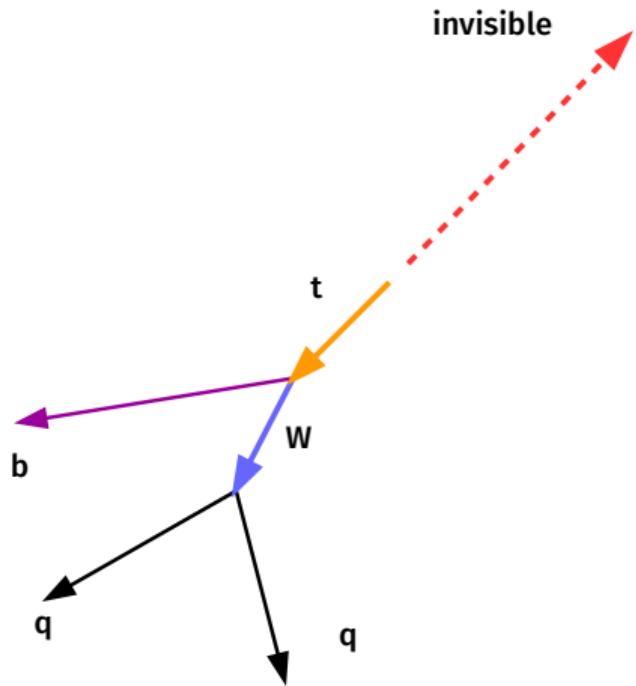


Flavor violating: off-diagonal g_u^V



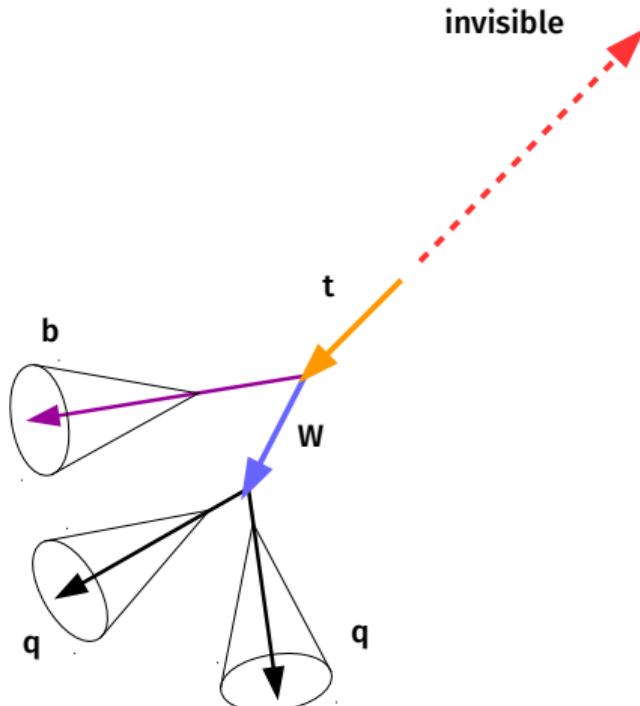
Anatomy of a mono-top event

Hadronic decay \Rightarrow larger \mathcal{B} , no p_T^{miss}



Anatomy of a mono-top event

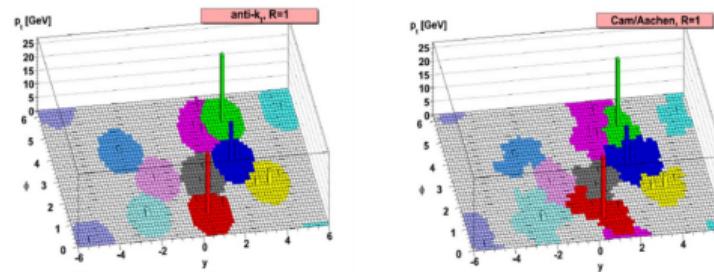
Quarks shower and hadronize into jets



- Particles are clustered into jets based on a distance metric:

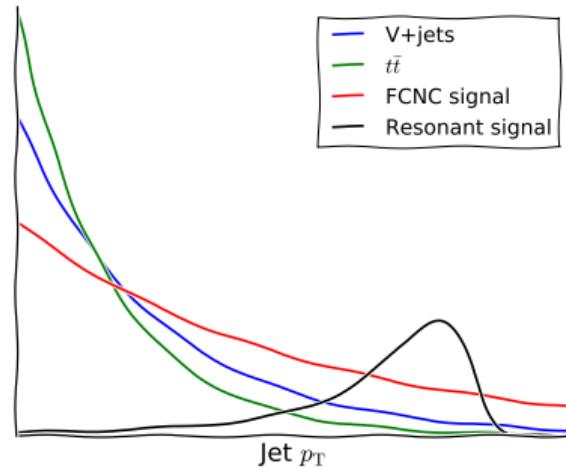
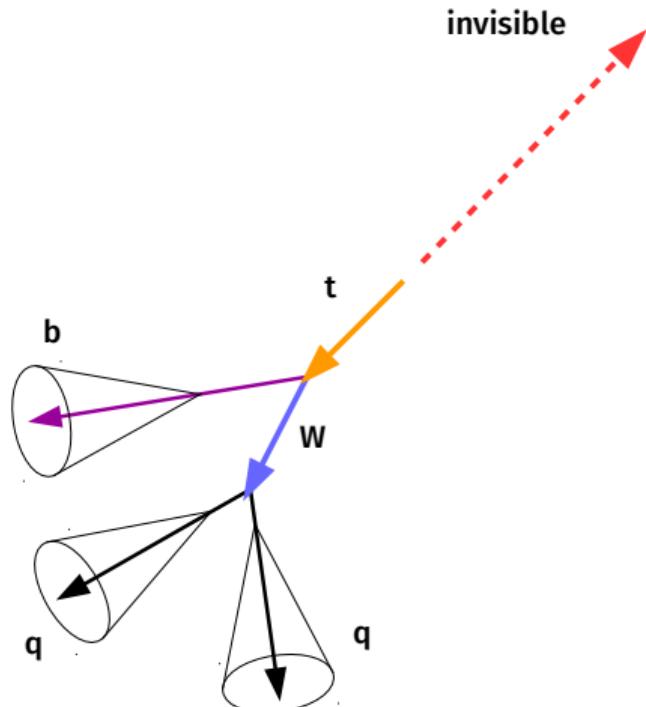
$$d_{ij} = \min\{p_{T,i}^{2q}, p_{T,j}^{2q}\} \frac{\Delta R(p_i^\mu, p_j^\mu)^2}{R}$$

- $q = -1$: anti- k_T (AK)
- $q = 0$: Cambridge-Aachen (CA)
- Single-parton jet: AK $R = 0.4$



Anatomy of a mono-top event

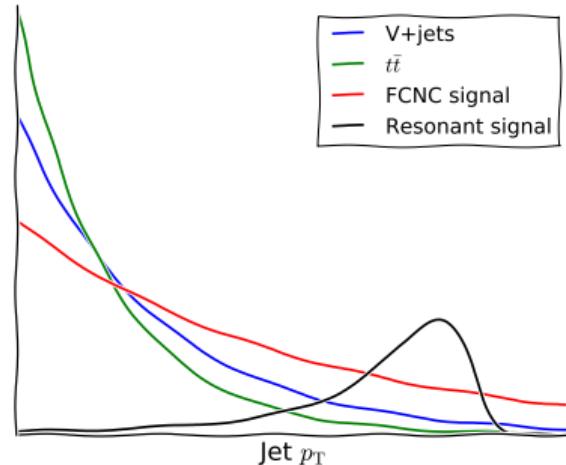
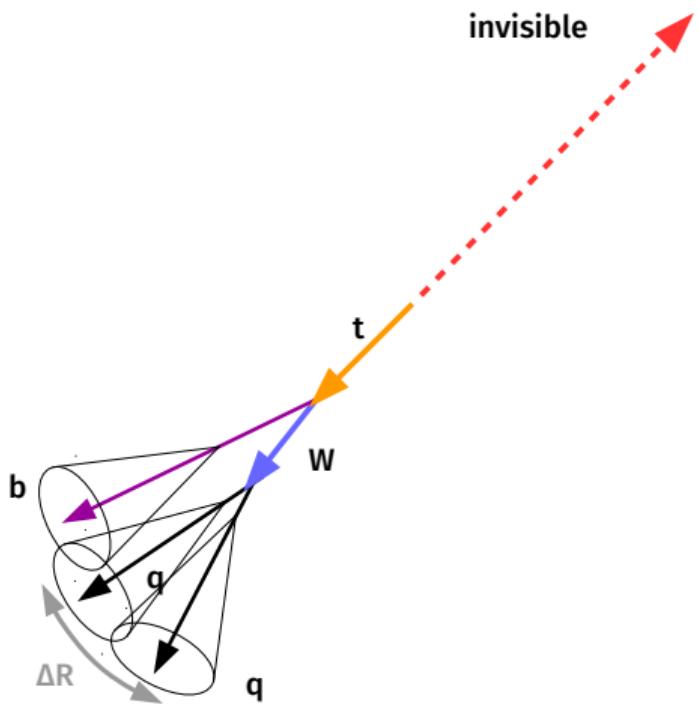
Quarks shower and hadronize into jets



- ▶ Signal more energetic than SM
- ▶ Maximize S/B \Rightarrow large jet p_T

Anatomy of a mono-top event

Decay products collimate

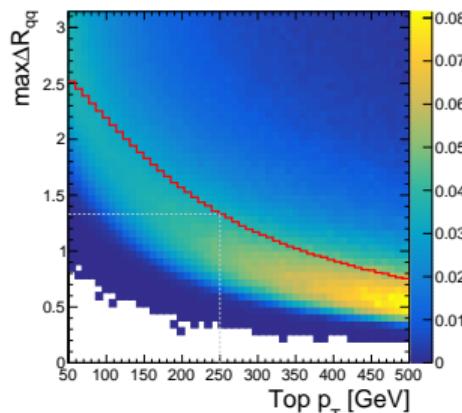


- ▶ Signal more energetic than SM
- ▶ Maximize S/B \Rightarrow large jet p_T
- ▶ Separation between jets: $\Delta R \sim 2m_t/p_T$
 - ▶ $p_T > 250 \text{ GeV} \Rightarrow$ jets ($R = 0.4$) overlap
 - ▶ $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$

Reconstruction of top quark jet

Clustering

- ▶ Three AK $R = 0.4$ jets \rightarrow single CA $R = 1.5$ jet

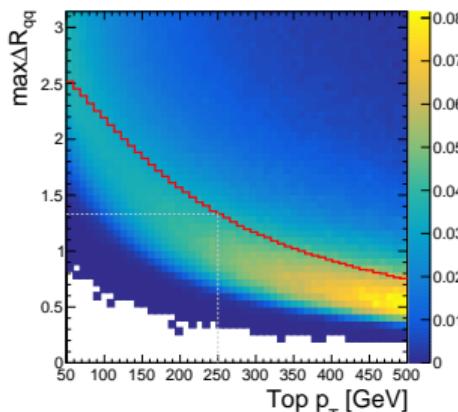


- ▶ These are huge jets: half the detector!
- ▶ Many extra particles

Reconstruction of top quark jet

Clustering

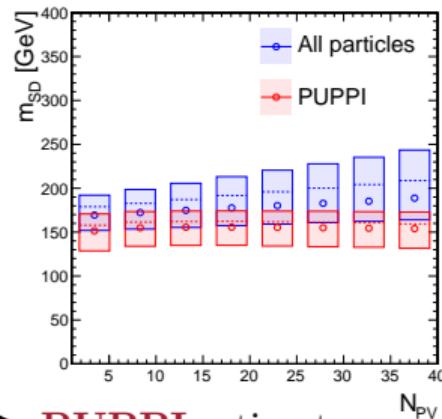
- ▶ Three AK $R = 0.4$ jets \rightarrow single CA $R = 1.5$ jet



- ▶ These are huge jets: half the detector!
- ▶ Many extra particles

Pileup particles

- ▶ 10-25 vertices per collision
- ▶ PU particles are isotropic

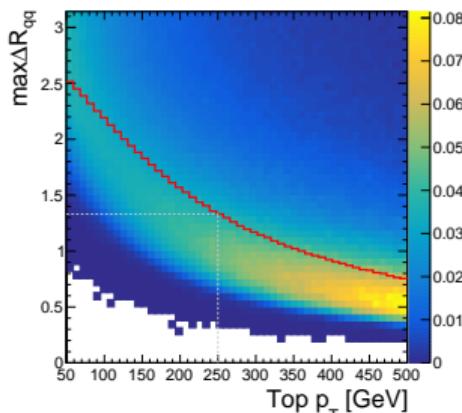


- ▶ **PUPPI** estimates
 $P(\text{PU}|p_T, \eta, \phi)$ from proximity to PV particles

Reconstruction of top quark jet

Clustering

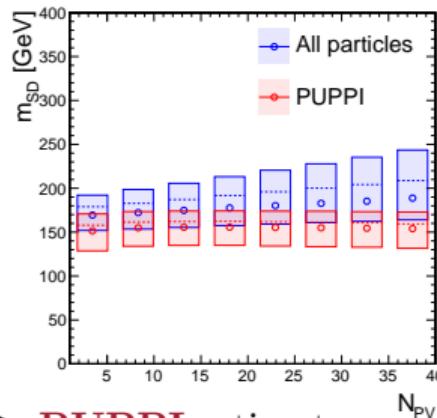
- ▶ Three AK $R = 0.4$ jets \rightarrow single CA $R = 1.5$ jet



- ▶ These are huge jets: half the detector!
- ▶ Many extra particles

Pileup particles

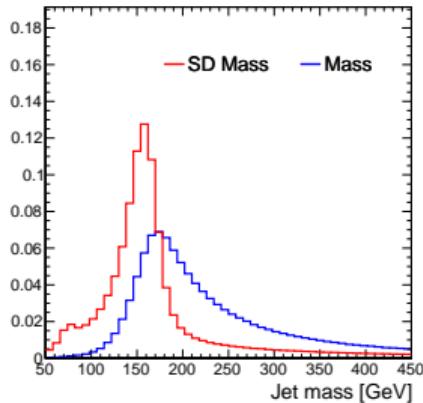
- ▶ 10-25 vertices per collision
- ▶ PU particles are isotropic



- ▶ **PUPPI** estimates $P(\text{PU}|p_T, \eta, \phi)$ from proximity to PV particles

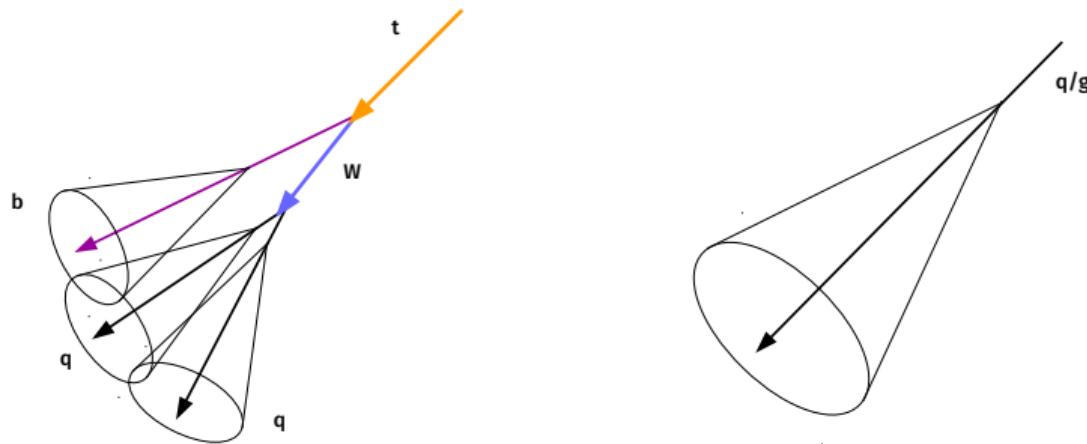
Non-PS radiation

- ▶ ISR, UE, MPI



- ▶ **Soft drop** removes wide-angle and soft radiation from jet

- Top quark $\rightarrow 3q \Rightarrow$ top jet has 3 “prongs”: regions of correlated radiation



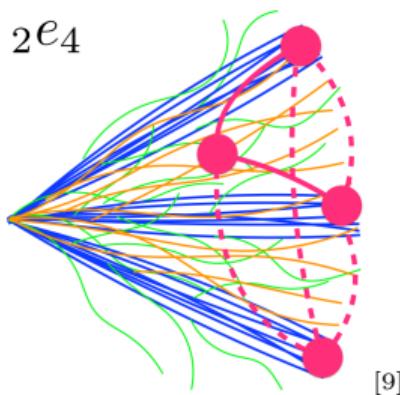
- **Substructure** observables are sensitive to such features
 - N -subjettiness, subjet algorithms, ECFs,...

Energy correlation functions

ECFs are **N**-point distance-weighted correlation functions among particles of the jet

$$e(a, \mathbf{N}, \alpha) \sim \sum_{\mathbf{N} \text{ particles } \in J}$$

sets of **N** particles

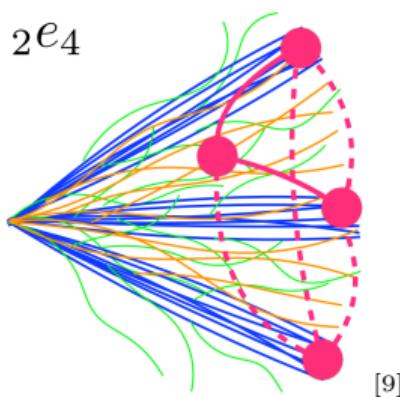


Energy correlation functions

ECFs are **N**-point distance-weighted correlation functions among particles of the jet

$$e(a, \mathbf{N}, \alpha) \sim \sum_{\mathbf{N} \text{ particles } \in J} \left[\prod_{p \in \text{particles}} \frac{E_p}{E_J} \right]$$

sets of **N** particles energy fractions

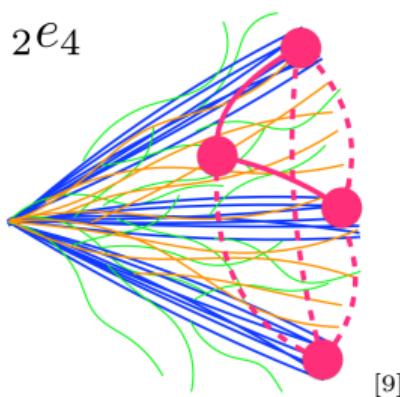


Energy correlation functions

ECFs are **N**-point distance-weighted correlation functions among particles of the jet

$$e(a, \mathbf{N}, \alpha) \sim \sum_{\mathbf{N} \text{ particles } \in J} \left[\prod_{p \in \text{particles}} \frac{E_p}{E_J} \right] \times \min \left\{ \prod_{p,q \in \text{particles}}^a \theta(p, q) \right\}^\alpha$$

sets of **N** particles
energy fractions
opening angle

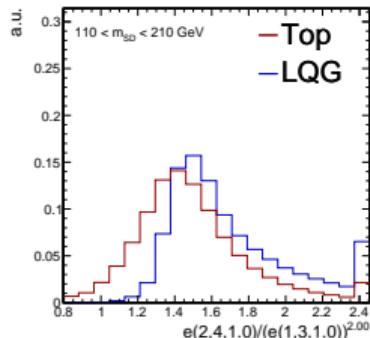
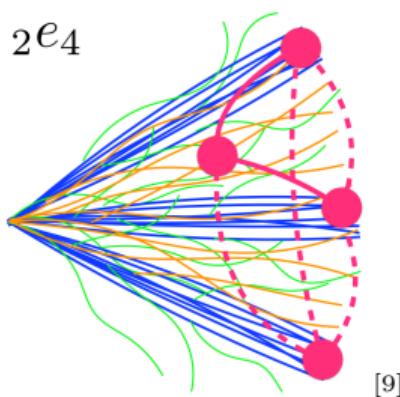


Energy correlation functions

ECFs are **N**-point distance-weighted correlation functions among particles of the jet

$$e(a, \mathbf{N}, \alpha) \sim \sum_{\mathbf{N} \text{ particles } \in J} \left[\prod_{p \in \text{particles}} \frac{E_p}{E_J} \right] \times \min \left\{ \prod_{p,q \in \text{particles}}^a \theta(p, q) \right\}^\alpha$$

sets of **N** particles
energy fractions
opening angle



$$e(4)/e(3)$$

Energy correlation functions

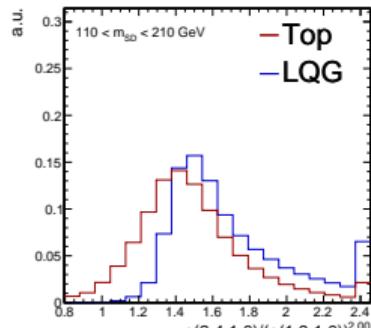
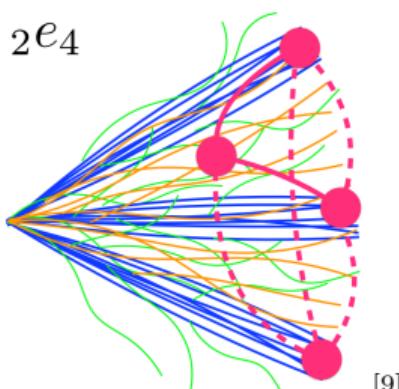
ECFs are **N**-point distance-weighted correlation functions among particles of the jet

$$e(a, \mathbf{N}, \alpha) \sim \sum_{\mathbf{N} \text{ particles } \in J} \left[\prod_{p \in \text{particles}} \frac{E_p}{E_J} \right] \times \min \left\{ \prod_{p,q \in \text{particles}}^a \theta(p, q) \right\}^\alpha$$

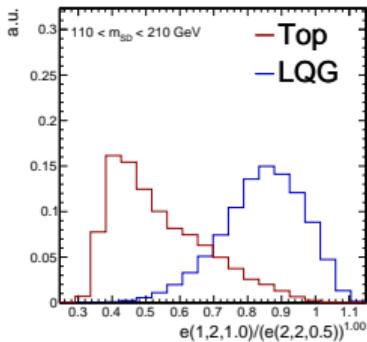
sets of **N** particles

energy fractions

opening angle



$e(4)/e(3)$



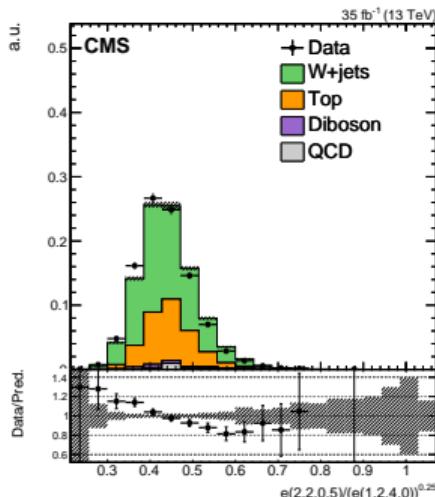
$e(2)/e(2)$

Building a combined tagger

Modeling issues

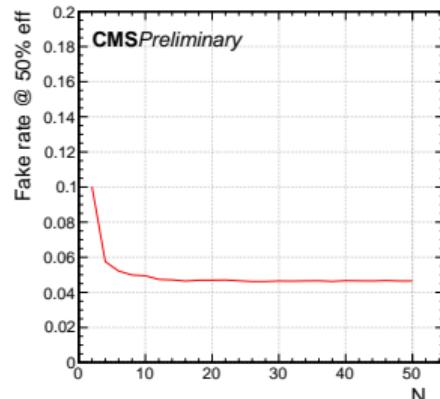
$$\frac{e(a, N, \alpha)}{e(b, M, \beta)^x}, \quad M \leq N, \quad x = \frac{a\alpha}{b\beta}$$

- ▶ Some ECF ratios poorly described by PS model



Dimensional reduction

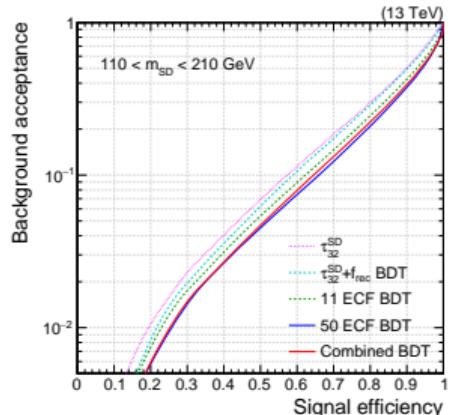
- ▶ Large ECF ratio space
- ▶ Expensive to compute



- ▶ Embed dimensional reduction into boosted decision tree training

Combined tagger

- ▶ Final discriminator combines ECFs, τ_{32} , f_{rec}



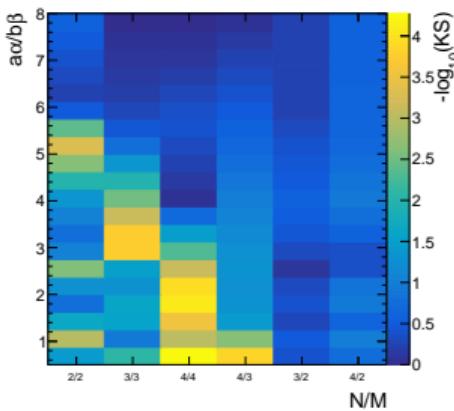
- ▶ 30% better background rejection than τ_{32}

Building a combined tagger

Modeling issues

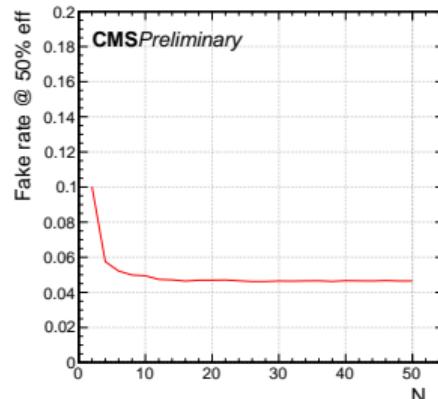
$$\frac{e(a, N, \alpha)}{e(b, M, \beta)^x}, \quad M \leq N, \quad x = \frac{a\alpha}{b\beta}$$

- ▶ Some ECF ratios poorly described by PS model



Dimensional reduction

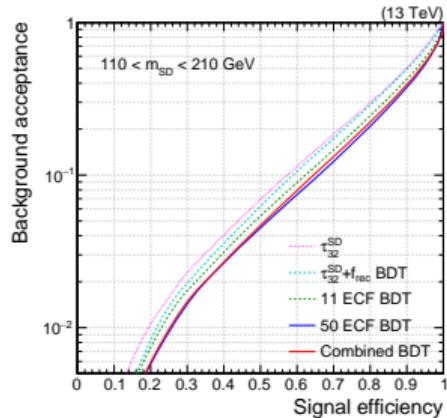
- ▶ Large ECF ratio space
- ▶ Expensive to compute



- ▶ Embed dimensional reduction into boosted decision tree training

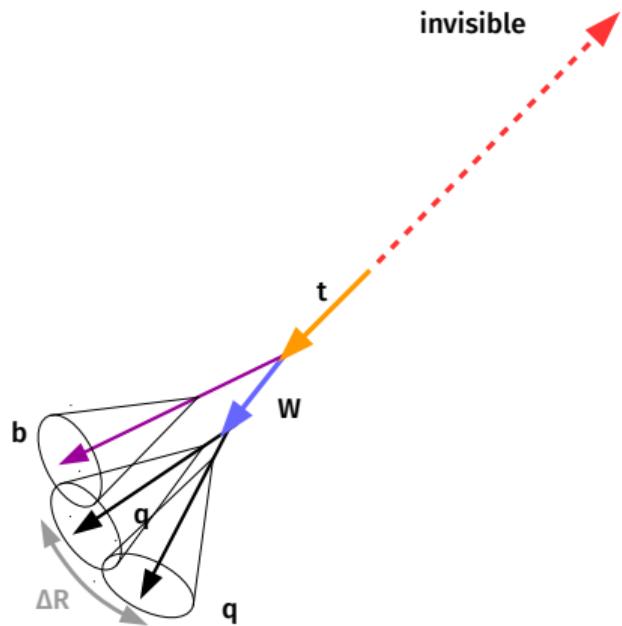
Combined tagger

- ▶ Final discriminator combines ECFs, τ_{32} , f_{rec}



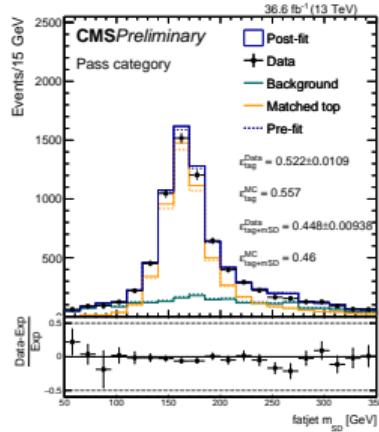
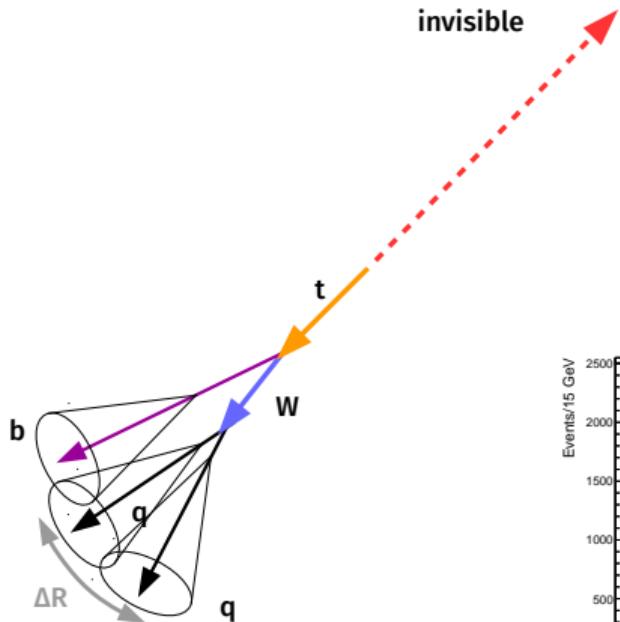
- ▶ 30% better background rejection than τ_{32}

Selecting mono-top events

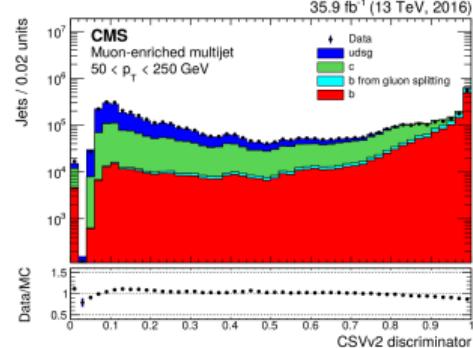


- ▶ $p_T^{\text{miss}} > 250 \text{ GeV}$ (trigger)
- ▶ No e, μ, τ_h , and AK4 b jets
- ▶ One CA15 jet, $p_T > 250 \text{ GeV}$
- ▶ Selected by BDT

Selecting mono-top events

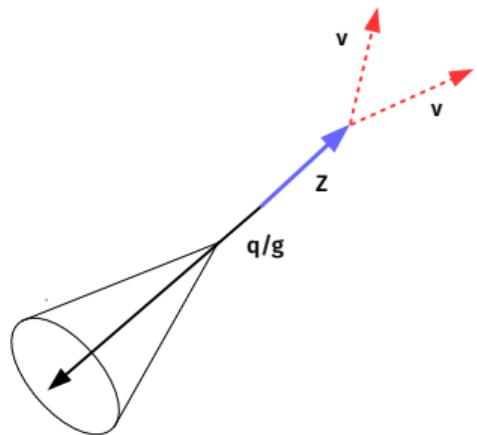


- $p_T^{\text{miss}} > 250 \text{ GeV}$ (trigger)
- No e, μ, τ_h , and AK4 b jets
- One CA15 jet, $p_T > 250 \text{ GeV}$
 - Selected by BDT
 - Mass consistent with m_t
 - Signature of B meson decay inside jet
- Lab frame $c\tau \sim \mathcal{O}(\text{mm})$



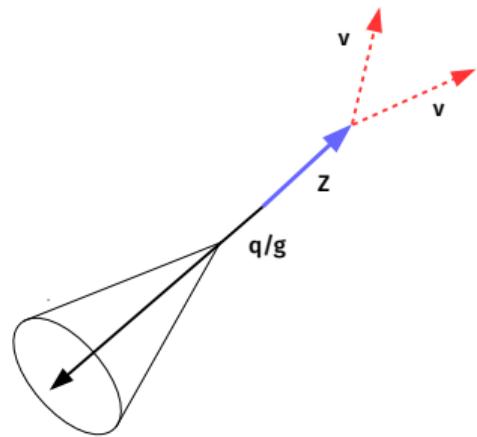
SM backgrounds

$Z \rightarrow \nu\nu$ (30%)

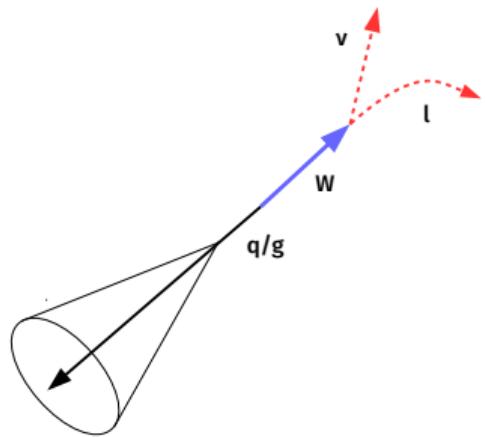


SM backgrounds

$Z \rightarrow \nu\nu$ (30%)

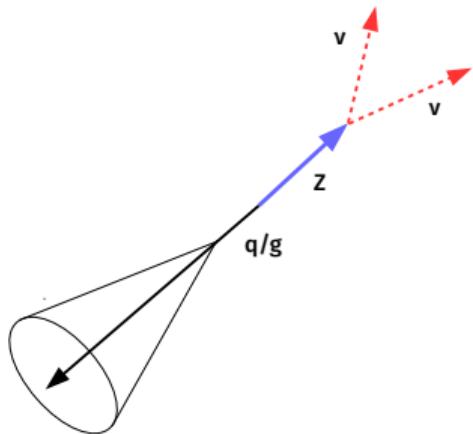


$W \rightarrow (\ell)\nu$ (15%)

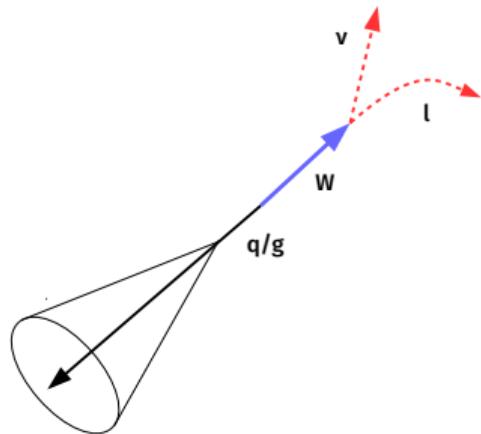


SM backgrounds

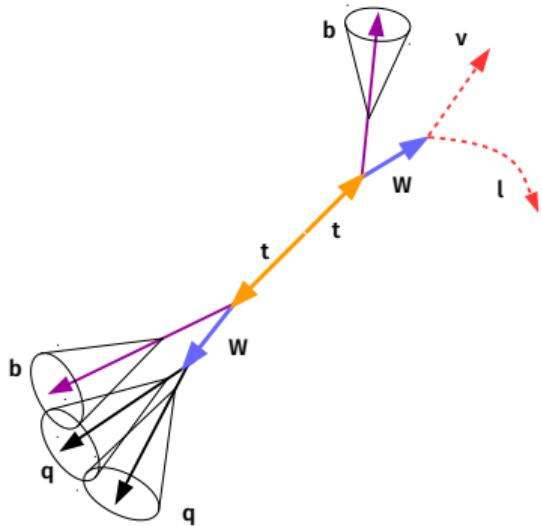
$Z \rightarrow \nu\nu$ (30%)



$W \rightarrow (\ell)\nu$ (15%)

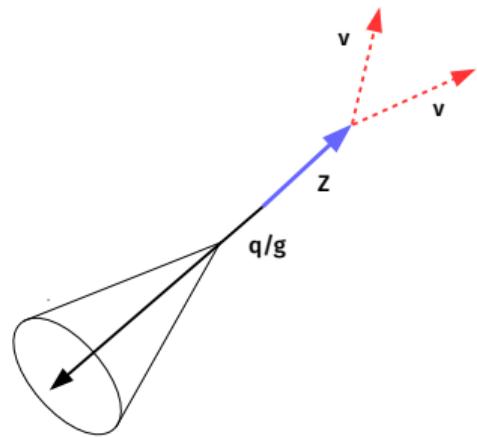


t quark pair (50%)

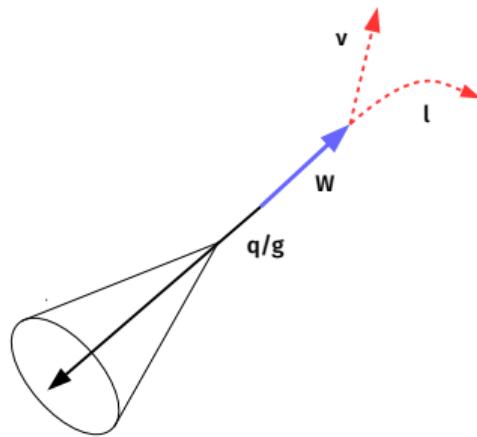


SM backgrounds

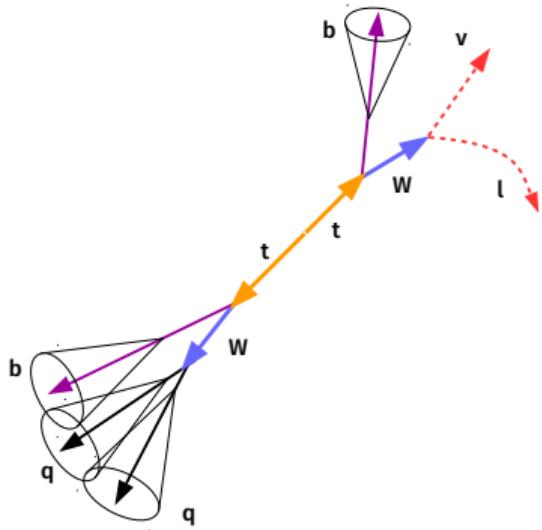
$Z \rightarrow \nu\nu$ (30%)



$W \rightarrow (\ell)\nu$ (15%)



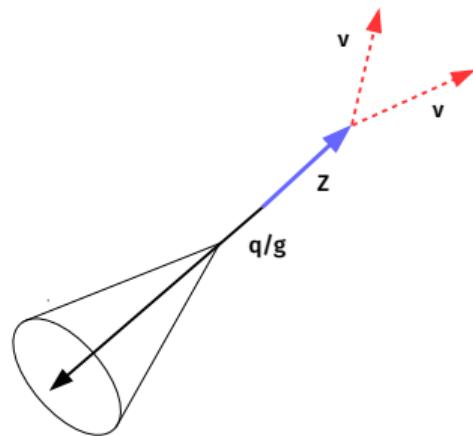
t quark pair (50%)



Note that p_T^{miss} is the transverse momentum of the **vector boson**

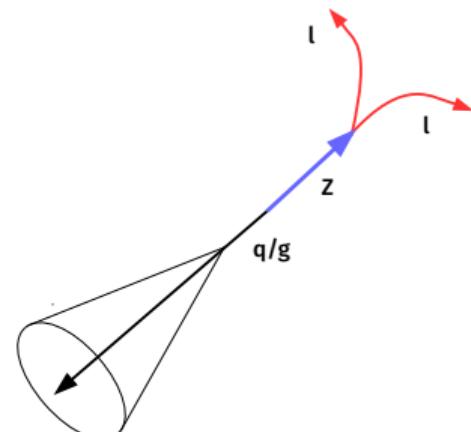
Prediction of Z p_T^{miss} spectrum

- MC predictions have $\sim 30\%$ uncertainties \Rightarrow use data to constrain
- p_T^{miss} is driven by jet measurement
- **Hadronic recoil (U)** \equiv momentum imbalance if we pretend ℓ^\pm, γ are invisible.

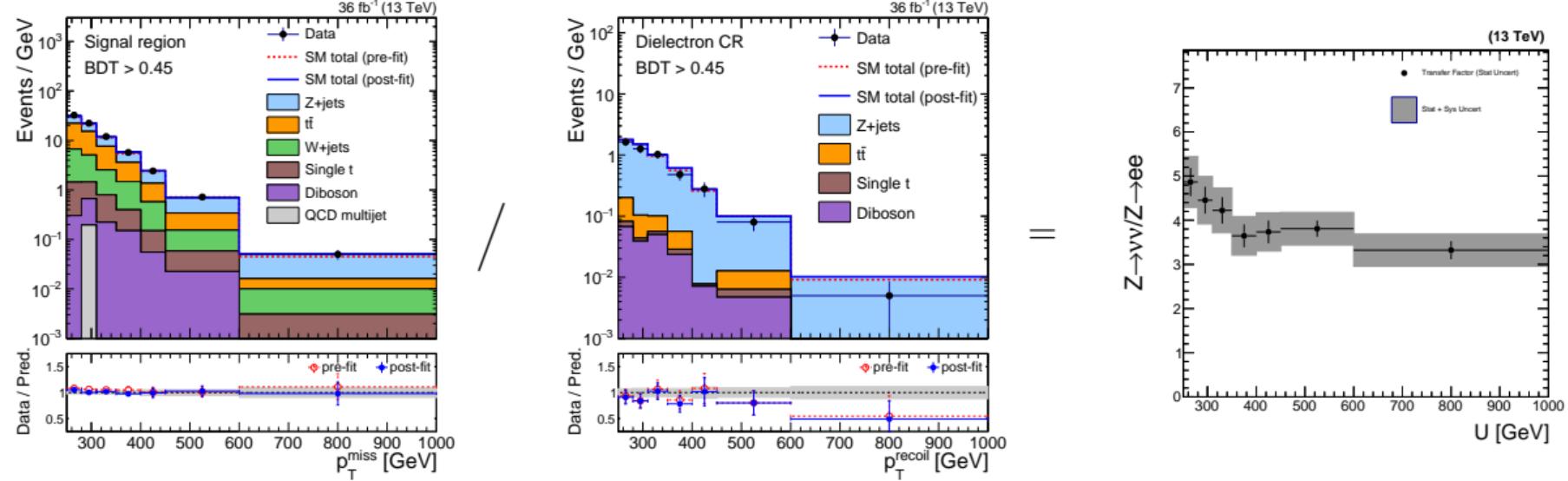


$$p_T^{\text{miss}} \longleftrightarrow p_T^{\text{miss}}(\text{no } \ell)$$

In both cases:
 $U \approx p_T^Z$

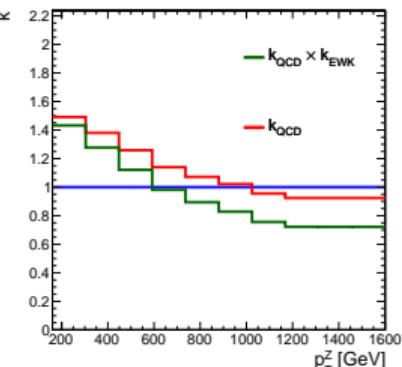
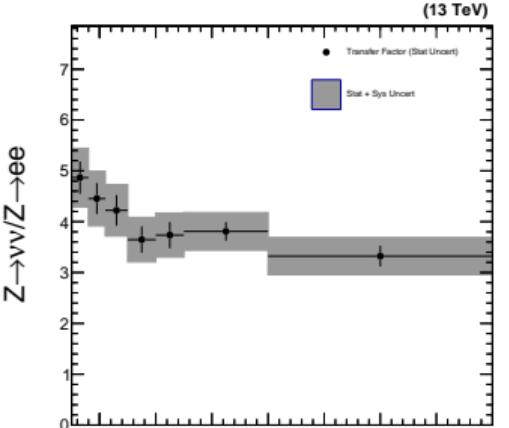
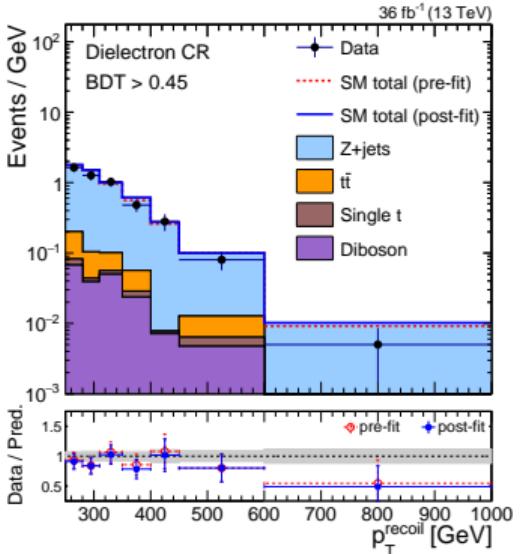
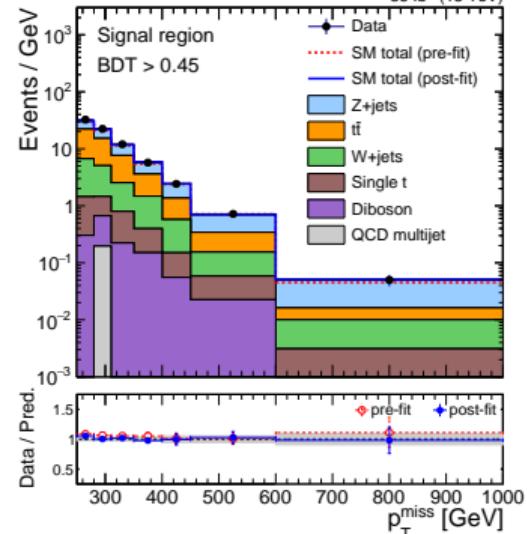


The real observable: $T = Z(\rightarrow \nu\nu)/Z(\rightarrow \ell\ell)$



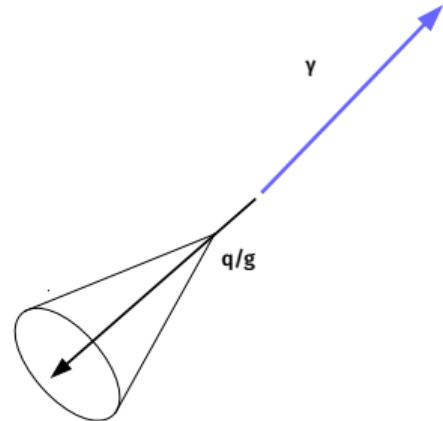
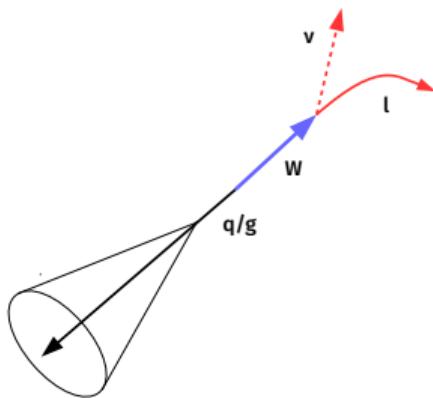
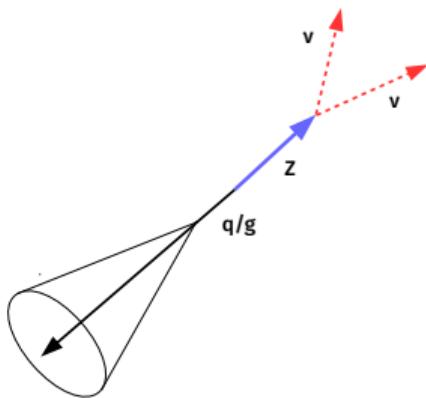
- ▶ Exp. and MC stat. uncertainties are small ($\leq 7\%$)
- ▶ Jet-related experimental uncertainties cancel
- ▶ Data stat. uncertainties are large at high U

The real observable: $T = Z(\rightarrow \nu\nu)/Z(\rightarrow \ell\ell)$



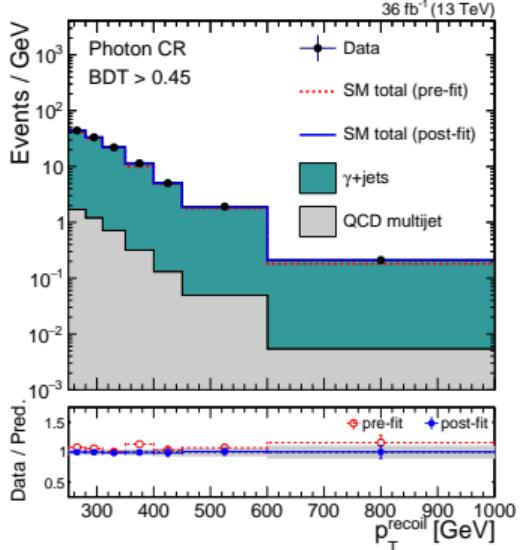
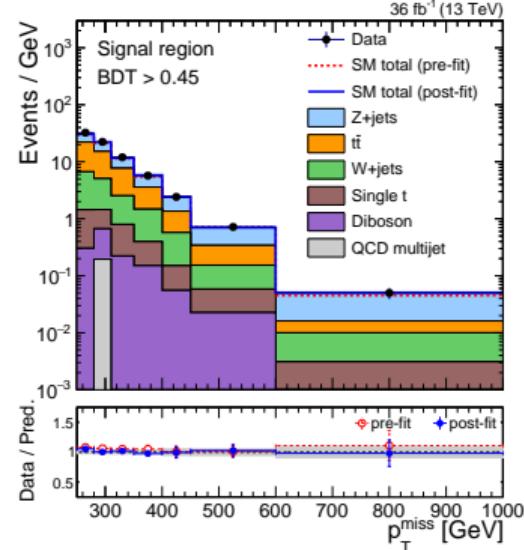
- ▶ Exp. and MC stat. uncertainties are small ($\leq 7\%$)
- ▶ Jet-related experimental uncertainties cancel
- ▶ Data stat. uncertainties are large at high U
- ▶ Compare to LO-to-NLO effect (+40% to -30%)

Additional constraints: Z/γ , Z/W

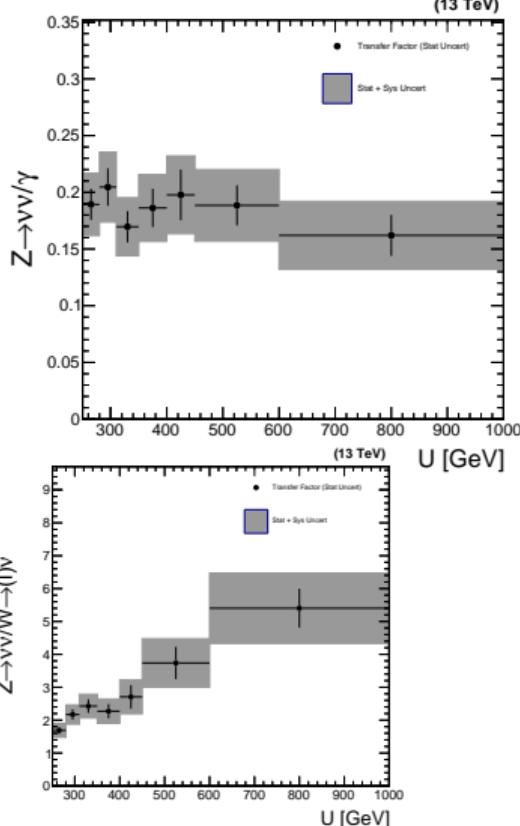


- ▶ Higher-order corrections needed for Z/γ and Z/W predictions
- ▶ Total (partial) cancellation of jet (theory) uncertainties

Additional constraints: Z/γ , Z/W

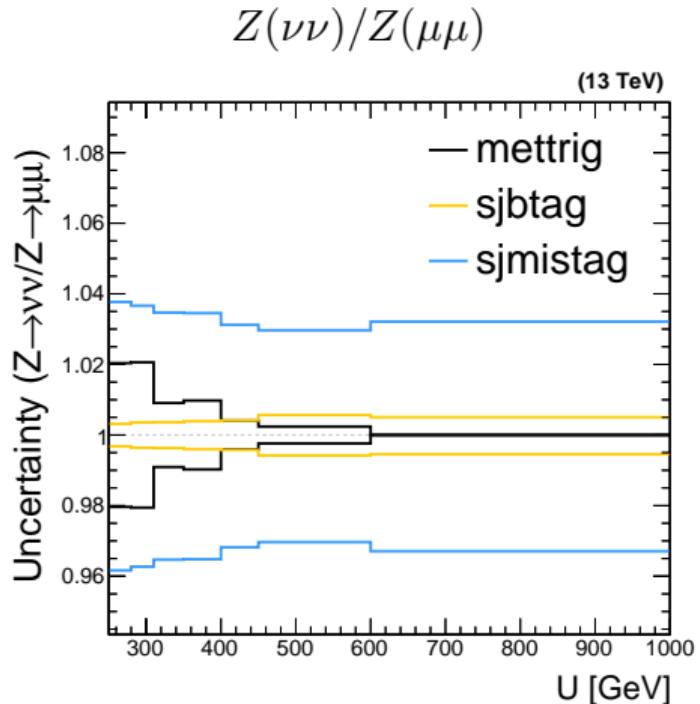


=

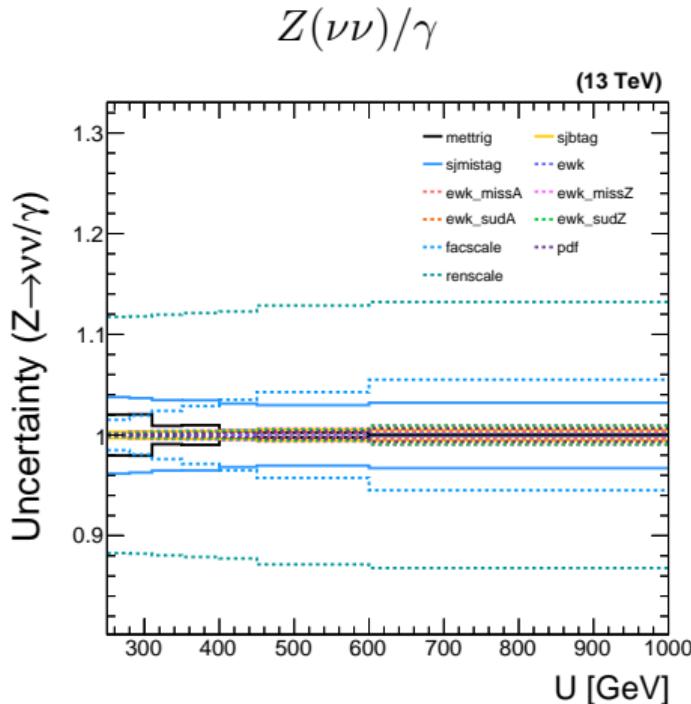


- ▶ Higher-order corrections needed for Z/γ and Z/W predictions
- ▶ Total (partial) cancellation of jet (theory) uncertainties
- ▶ Small data stat. uncertainties

Extrapolation uncertainties



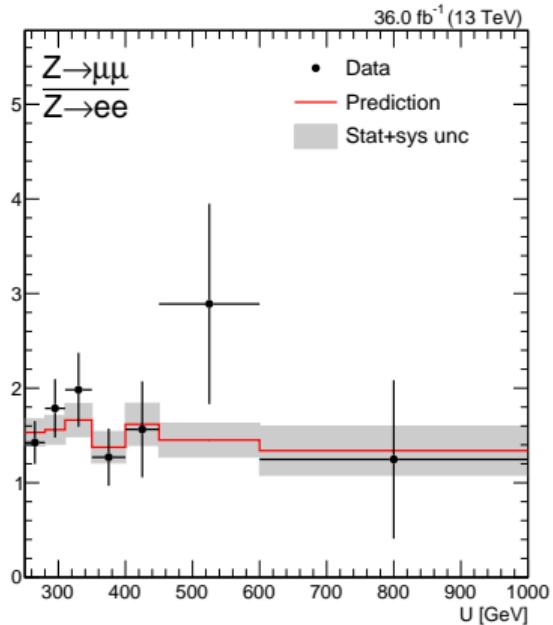
Small experimental uncertainties



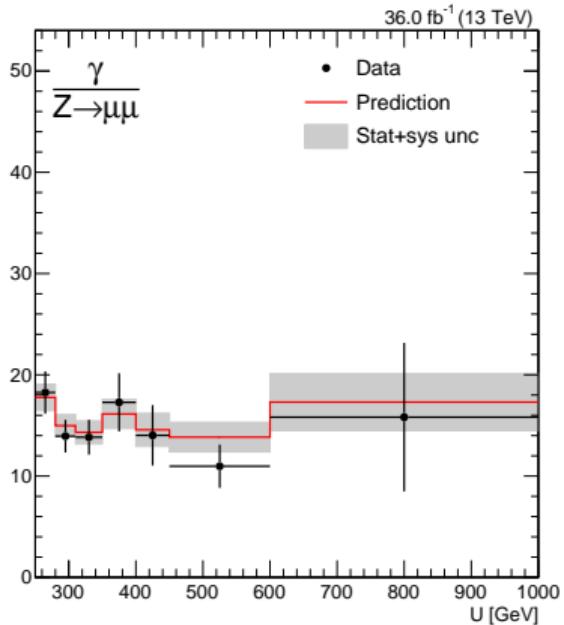
Large theoretical uncertainties

Extrapolation uncertainties

$Z(\mu\mu)/Z(ee)$



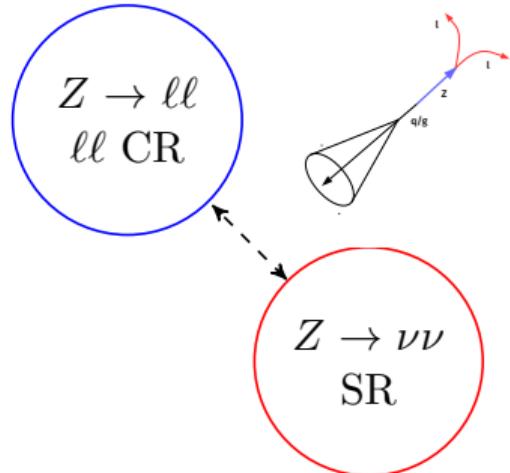
$Z(\mu\mu)/\gamma$



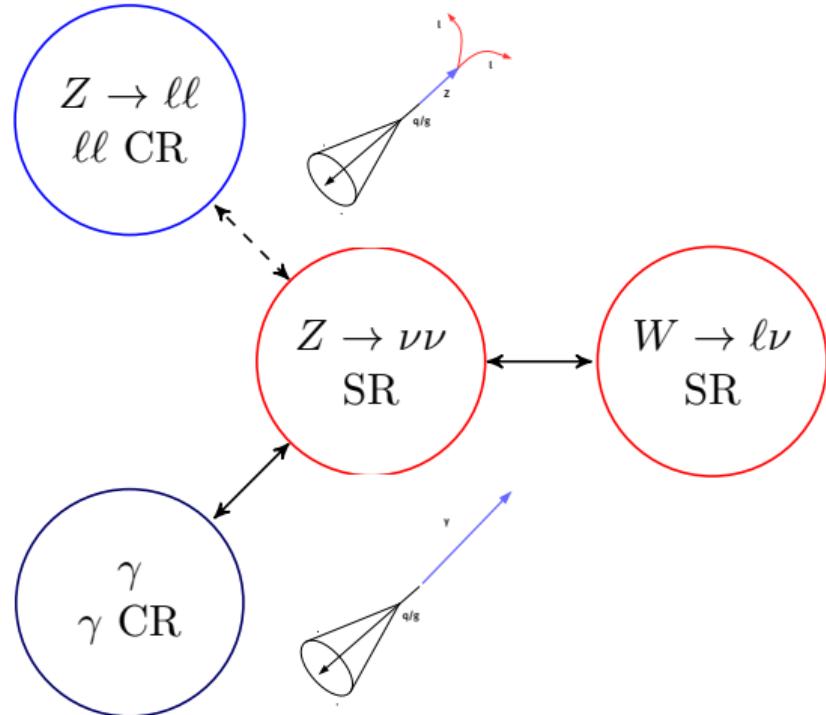
Small experimental uncertainties

Large theoretical uncertainties

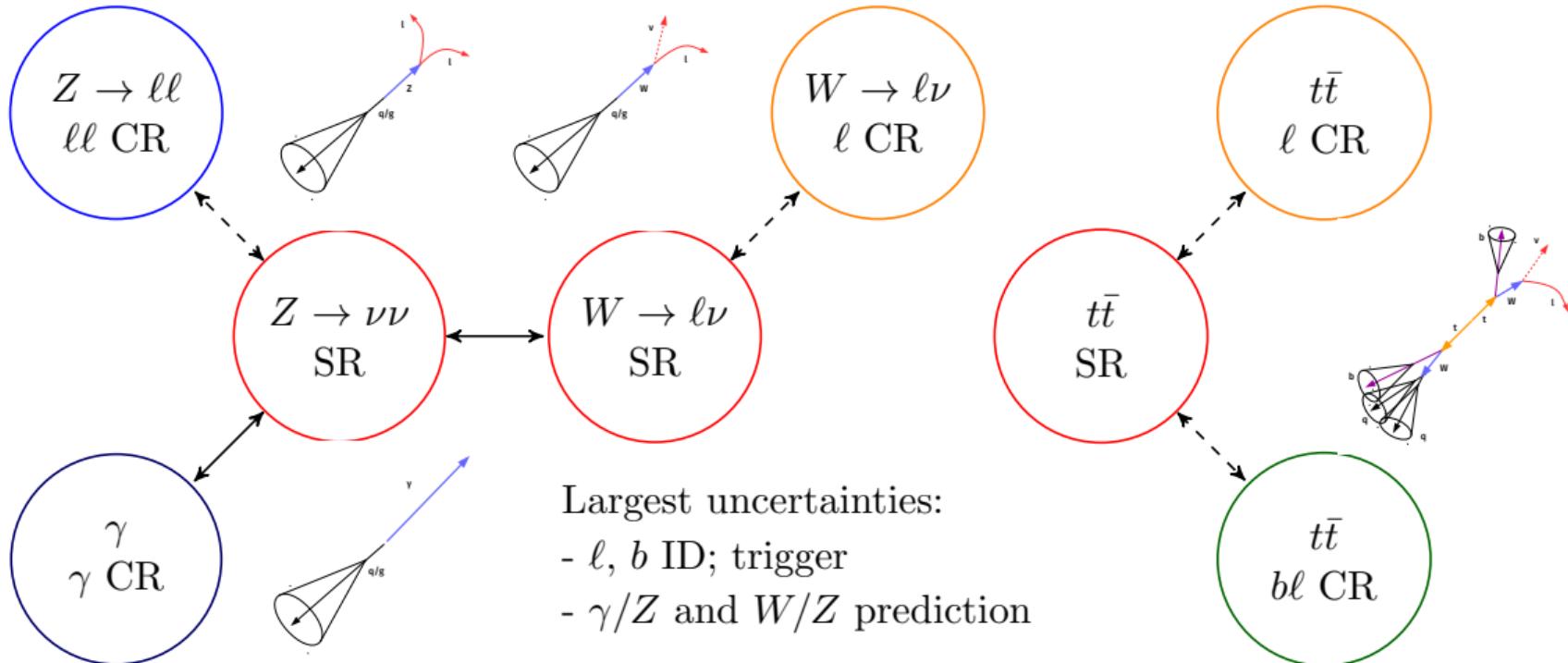
Background estimation summary



Background estimation summary



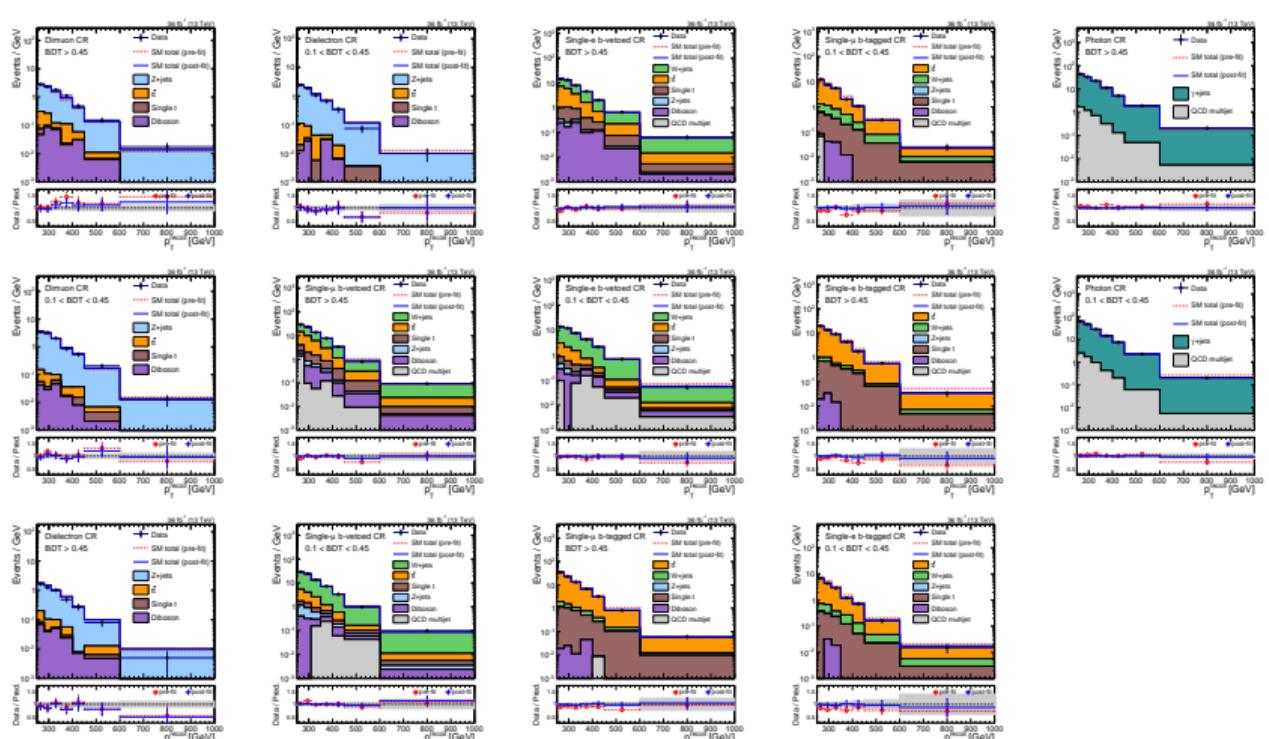
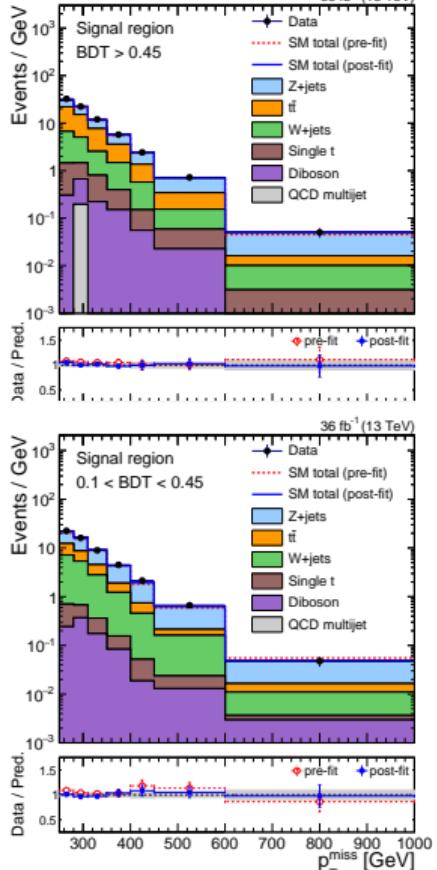
Background estimation summary



Likelihood assuming one major background ($Z \rightarrow \nu\nu$) and one CR :

$$\begin{aligned}\mathcal{L}(\mathbf{d} \mid \mu, \boldsymbol{\mu}^Z, \boldsymbol{\theta}) = & \prod_i \text{Pois} \left(d_i^{\text{SR}} \mid \mu S_i(\boldsymbol{\theta}) + \mu_i^Z + B_i^{\text{SR}}(\boldsymbol{\theta}) \right) \\ & \times \prod_i \text{Pois} \left(d_i^{\text{CR}} \mid \frac{\mu_i^Z}{T_i(\boldsymbol{\theta})} + B_i^{\text{CR}}(\boldsymbol{\theta}) \right) \\ & \times \prod_j p(\theta_j)\end{aligned}$$

MLE recoil distributions (SM hypothesis)

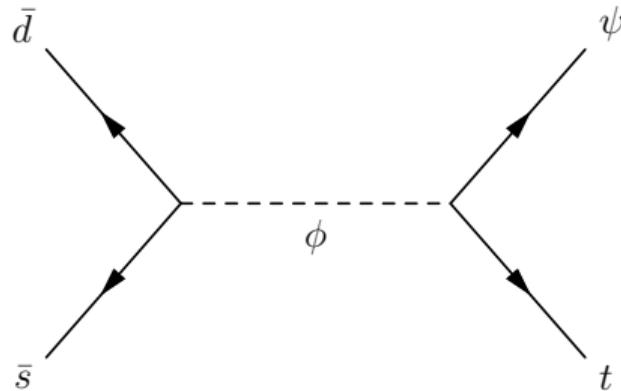


Limits on mono-top benchmark models

Benchmark models probe range of mono-top kinematics.

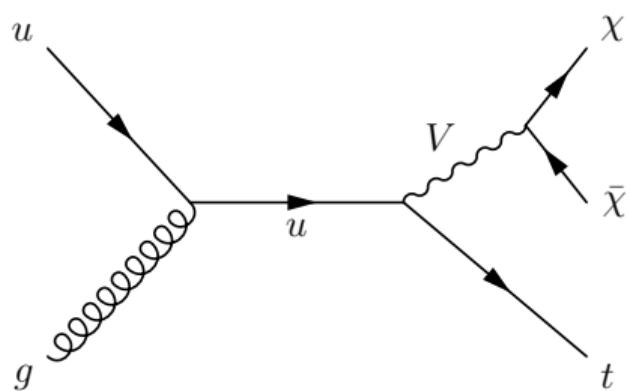
Resonant scalar

- p_T of top quark increases with m_ϕ
- Better selection efficiency at high m_ϕ



FCNC

- Falling p_T^{miss} spectra \Rightarrow worse signal eff.
- Constraints on m_V and couplings g_χ, g_q

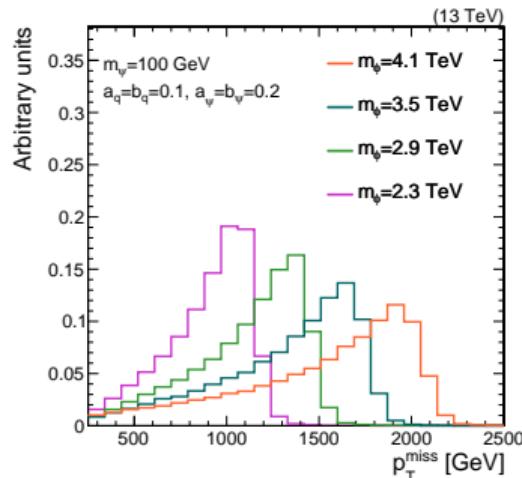


Limits on mono-top benchmark models

Benchmark models probe range of mono-top kinematics.

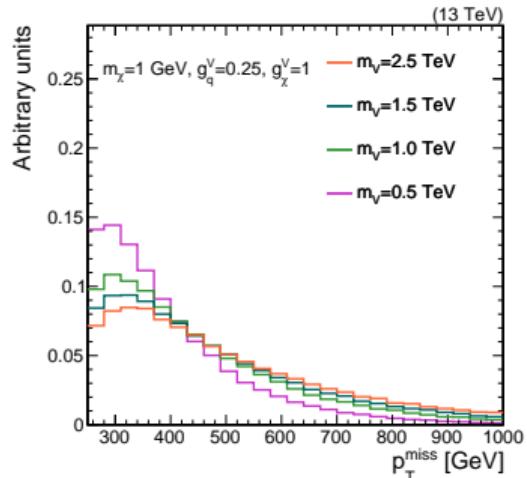
Resonant scalar

- p_T of top quark increases with m_ϕ
- Better selection efficiency at high m_ϕ



FCNC

- Falling p_T^{miss} spectra \Rightarrow worse signal eff.
- Constraints on m_V and couplings g_χ, g_q



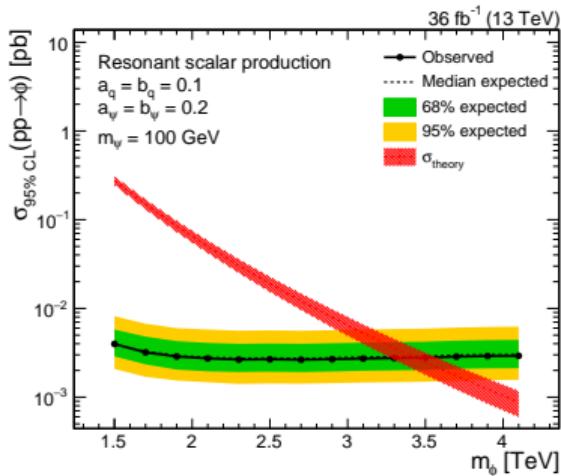
Limits on mono-top benchmark models



Benchmark models probe range of mono-top kinematics.

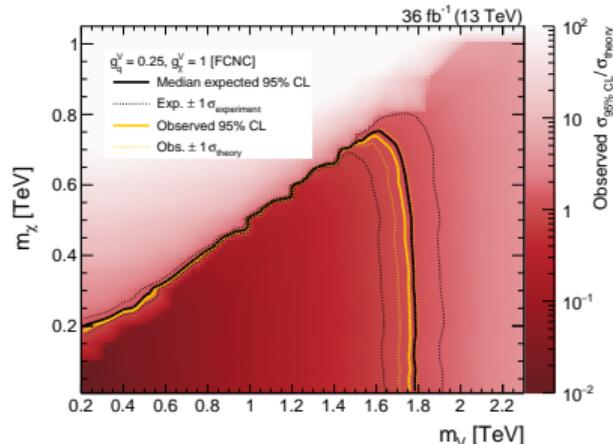
Resonant scalar

- p_T of top quark increases with m_ϕ
- Better selection efficiency at high m_ϕ



FCNC

- Falling p_T^{miss} spectra \Rightarrow worse signal eff.
- Constraints on m_V and couplings g_χ, g_q

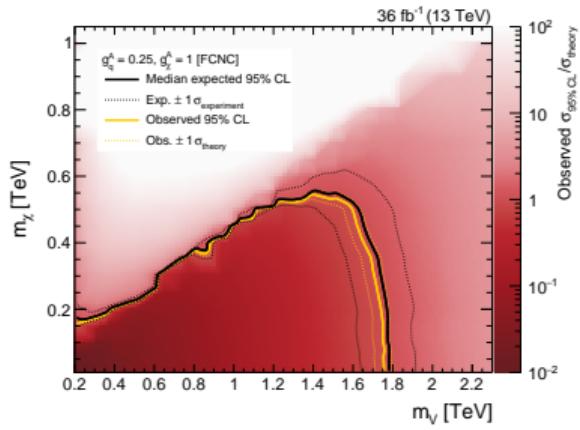


FCNC parameter space

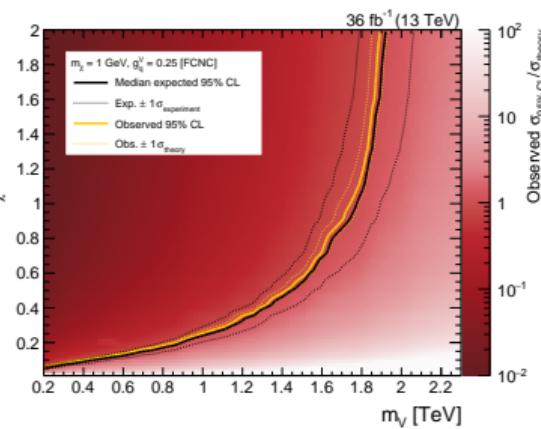


6 free parameters: m_χ , m_V , g_q^V , g_q^A , g_χ^V , g_χ^A

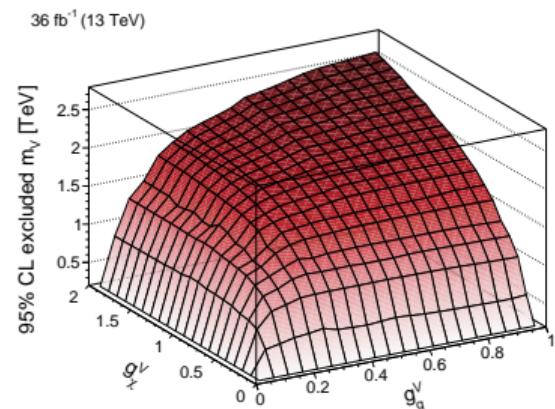
Axial-vector current



m_V vs g_q^V



m_V vs g_q^V vs g_χ^V



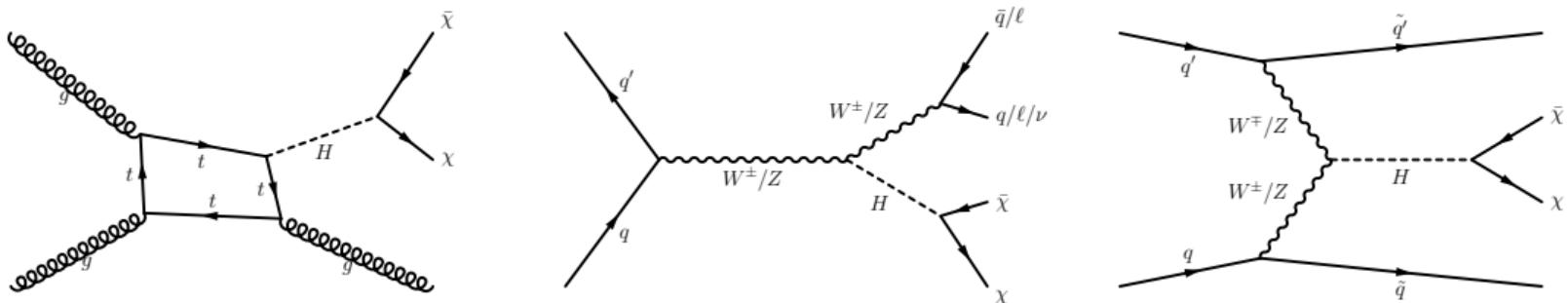
Comparison to previous results

	CMS Run 1	This result
Dataset size	20/fb	36/fb
Collision energy	8 TeV	13 TeV
$\sigma(pp \rightarrow \phi \rightarrow t\psi)$, $m_\phi = 1.5$ TeV	0.04 pb	0.19 pb
Reconstruction	3 resolved jets	1 merged jet
Excluded m_ϕ	0.35 TeV	0.8-3.4 TeV
Excluded resonant $\sigma \times \mathcal{B}$	0.6 pb	0.002 pb
Excluded m_V	0.65 TeV	0.2-1.8 TeV
Excluded FCNC $\sigma \times \mathcal{B}$	0.2 pb	0.018 pb

Vector boson fusion $H \rightarrow$ invisible

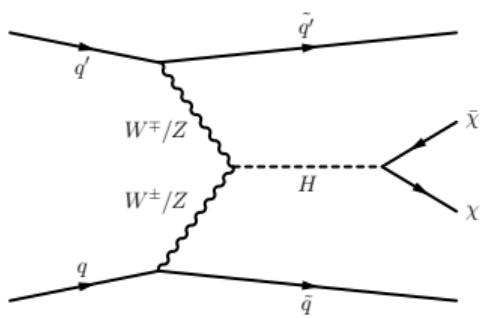
Invisible Higgs

- ▶ Core assumptions:
 - ▶ DM mass generation is through Higgs mechanism
 - ▶ $2m_\chi < m_H$
- ▶ Production mode \Rightarrow mono-**X** channels
 - ▶ $gg \rightarrow H + \text{jet(s)} \Rightarrow \text{mono-jet}$
 - ▶ $VH \Rightarrow \text{mono-}V(qq') \text{ and mono-}Z(\ell\ell)$
 - ▶ Vector boson fusion $\Rightarrow \text{VBF}+H \rightarrow \text{inv}$

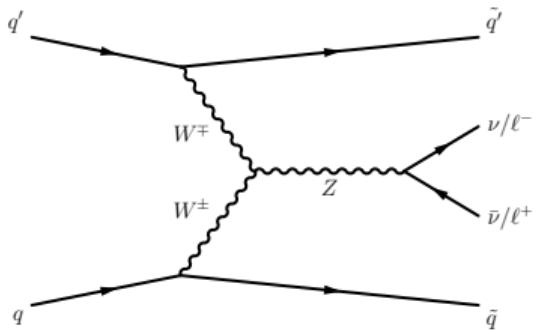


Production of bosons with two jets

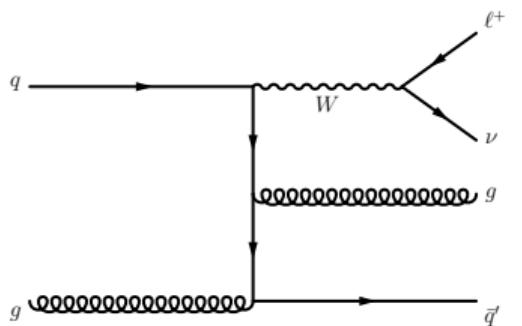
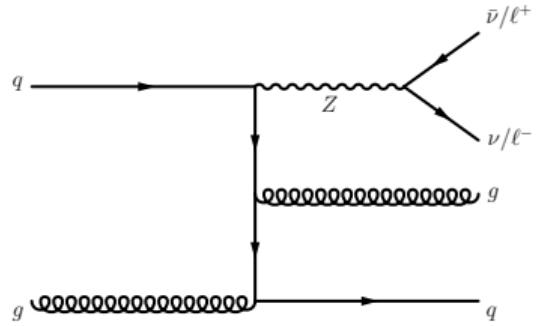
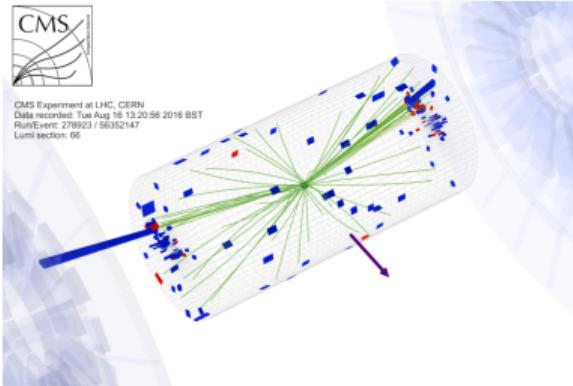
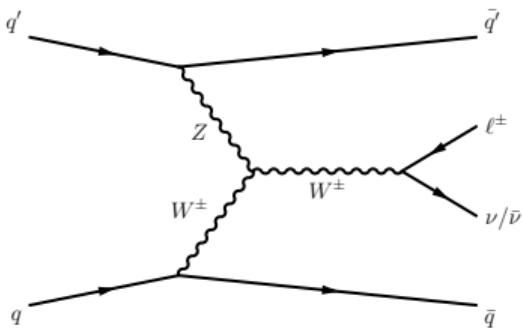
$$H \rightarrow \chi\bar{\chi}$$



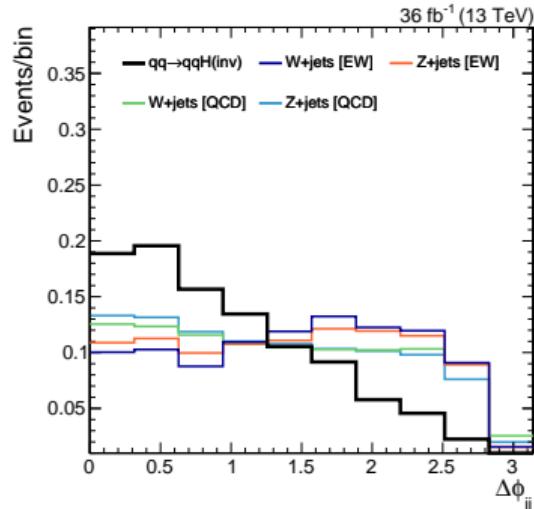
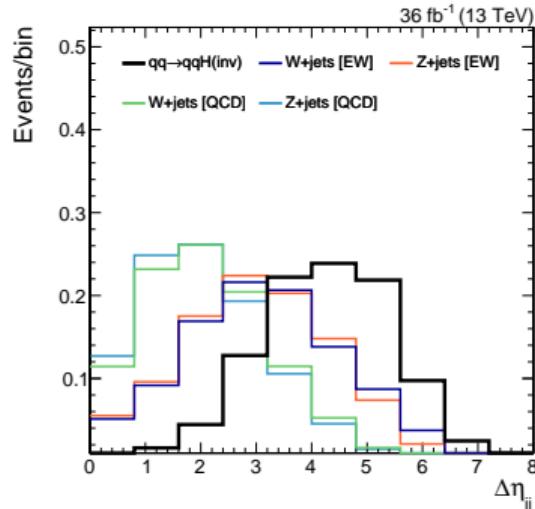
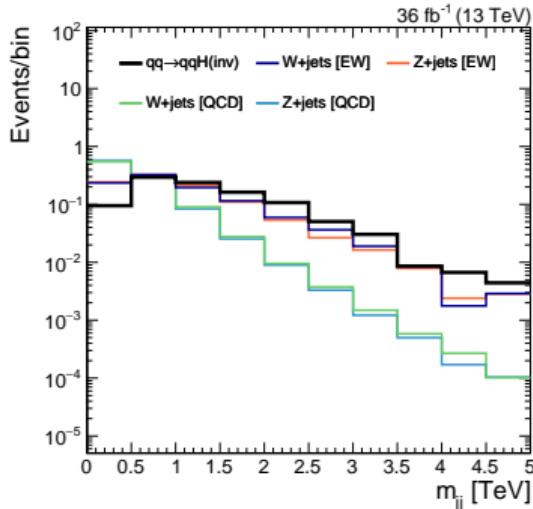
$$Z$$



$$W$$



Dijet kinematics in QCD/EW/VBF processes



$\Delta\eta$ and m_{jj} strongly correlated

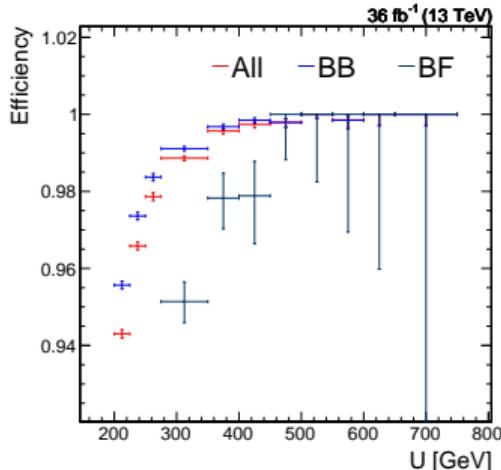
Use m_{jj} distribution to extract signal. Require $\Delta\eta > 1$ and $\Delta\phi < 1.5$.

Forward jets and L1 trigger

Resolution of p_T^{miss} strongly depends on location of jets

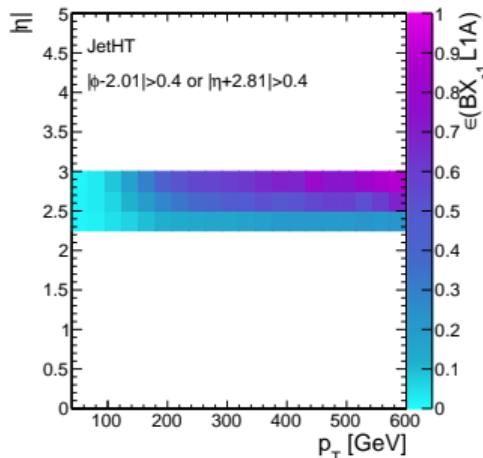
L1 p_T^{miss} calculation

- ▶ High rate from noise and multijet events
- ▶ Only $|\eta| < 3$ particles



L1 pre-firing

- ▶ Collisions are 25 ns apart
- ▶ Mis-timed signals block interesting events

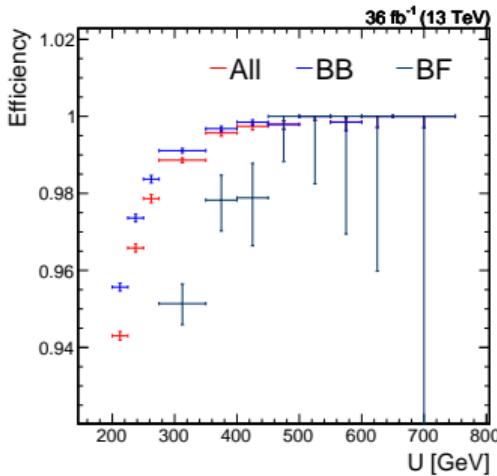


Forward jets and L1 trigger

Resolution of p_T^{miss} strongly depends on location of jets

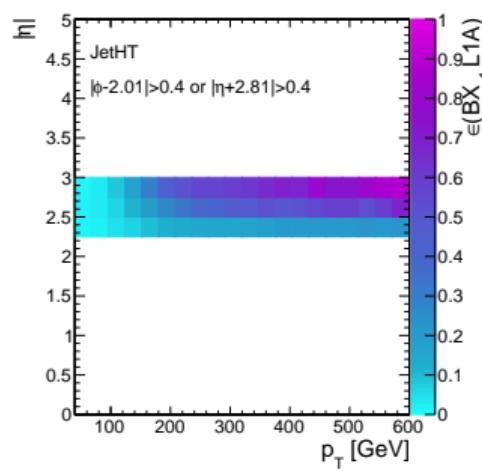
L1 p_T^{miss} calculation

- High rate from noise and multijet events
- Only $|\eta| < 3$ particles



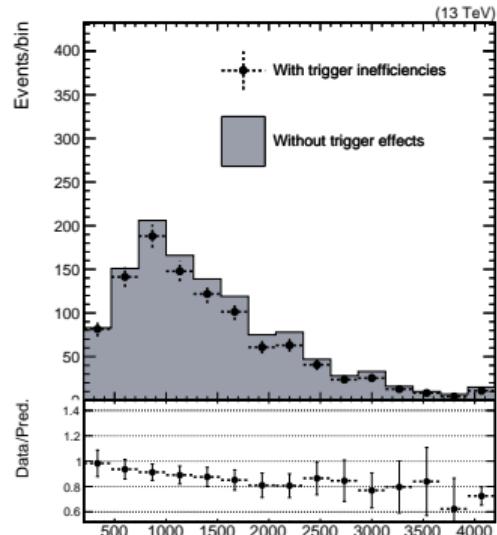
L1 pre-firing

- Collisions are 25 ns apart
- Mis-timed signals block interesting events



Net effect

- 20% of signal is lost when data is collected
- Fixed in 2018 data

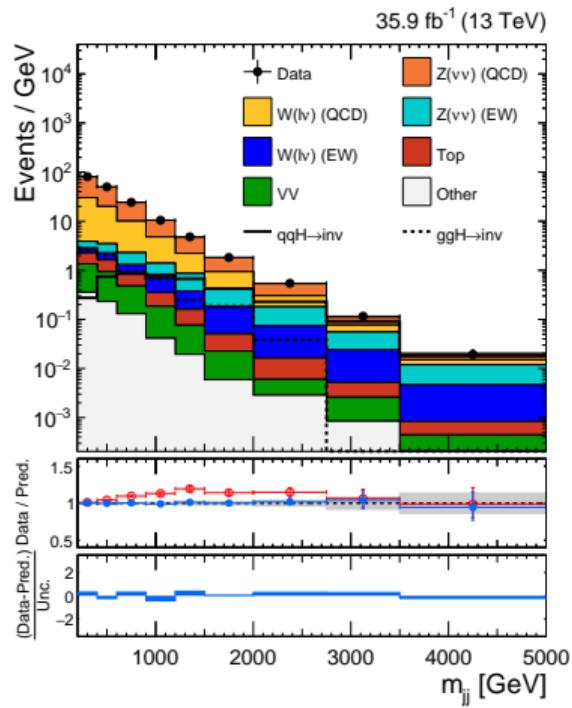
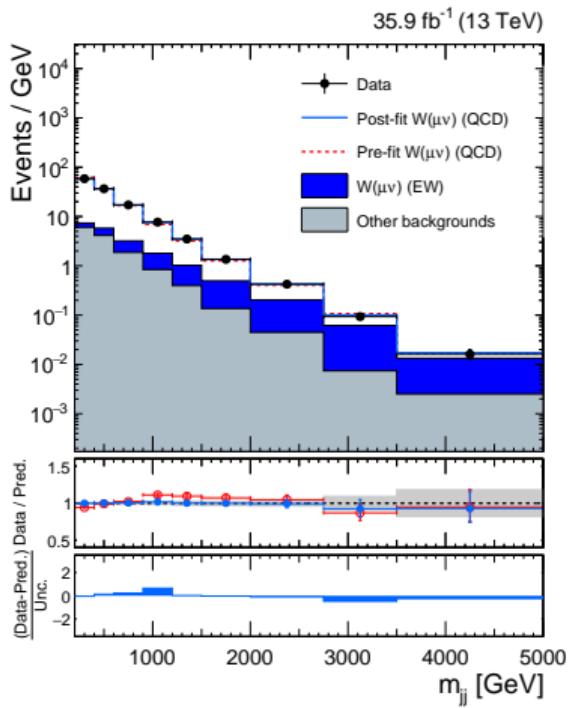


V +jet estimation



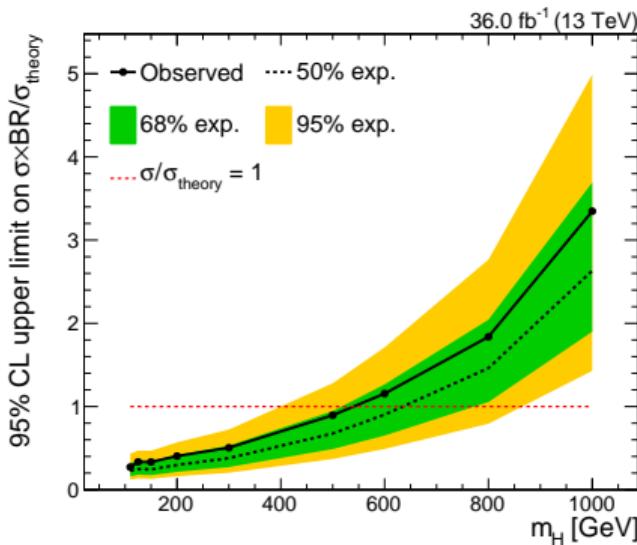
- Precision estimation of QCD and EW V +jets
- Similar strategy as mono-top, but no γ CR

Uncertainty	Size
W/Z QCD	15%
W/Z EW	15%
Trigger	2%
Lepton ID	2-3%

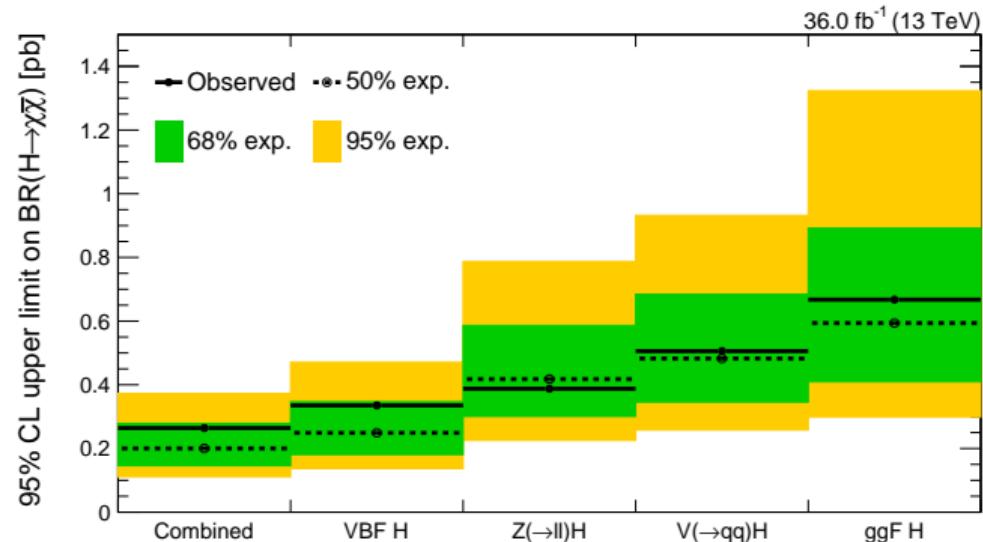


Constraints on $\mathcal{B}(H \rightarrow \text{inv})$

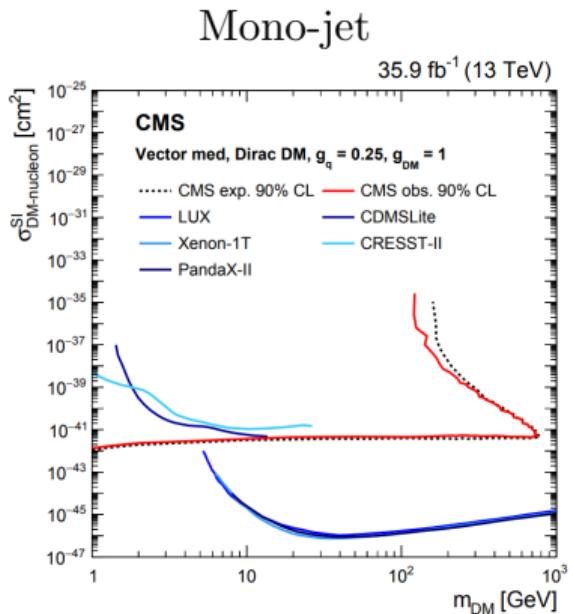
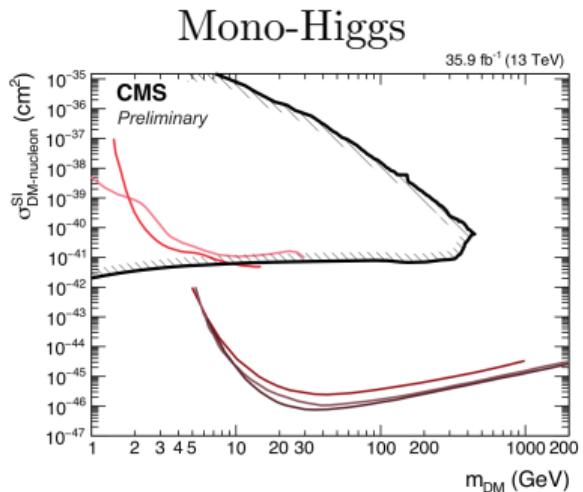
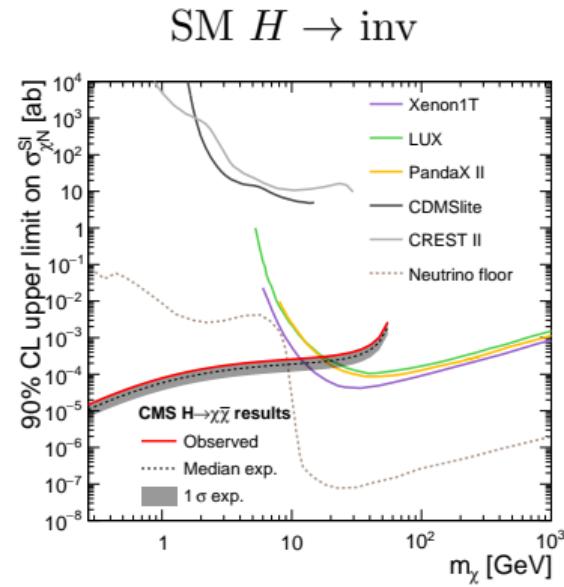
VBF-only



All production modes (SM Higgs)



Comparison of LHC and direct detection constraints



- LHC constraints strongest at low DM mass
- Constraints depend strongly on choice of model

Impact of analysis strategy and L1 losses



	Previous CMS result	This result
Collision energy	7,8,13 TeV	13 TeV
Dataset size	5,20,2/fb	36/fb
$\sigma(qq' \rightarrow qq'H)$	1.2,1.6,3.7 pb	3.7 pb
Exp. limit $\mathcal{B}(H \rightarrow \text{inv})$	0.30	0.25
Non-differential	-	0.30
No L1 effect	-	0.21

- ▶ Same sensitivity from old analysis technique on new dataset
 - ▶ Driven by losses in L1
- ▶ Improvements driven by precise measurement of EW+QCD $V+\text{jet}$ spectra

Improving DM searches at LHC: prediction of $V+jets$



VBF		
Uncertainty	Size	Effect
W/Z EW	15%	50%
W/Z QCD	15%	25%
Trigger	2%	20%
Lepton ID	2-3%	15%

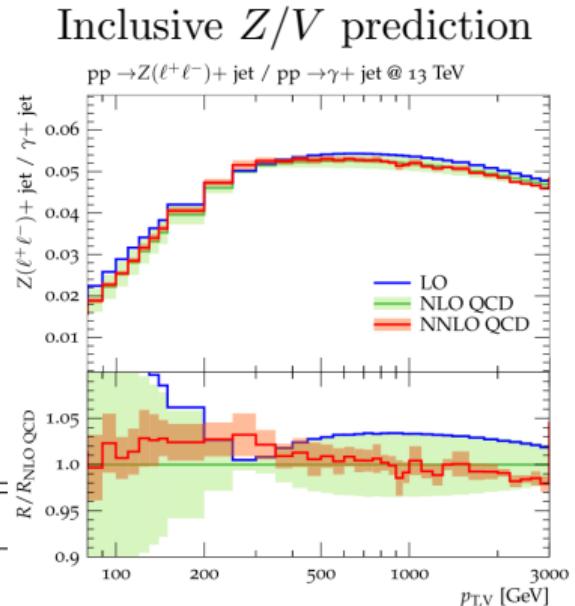
Mono-top		
Uncertainty	Size	Effect
V/Z QCD	15%	30%
b ID	2-10%	30%
Lepton ID	2-3%	15%

Improving DM searches at LHC: prediction of V +jets

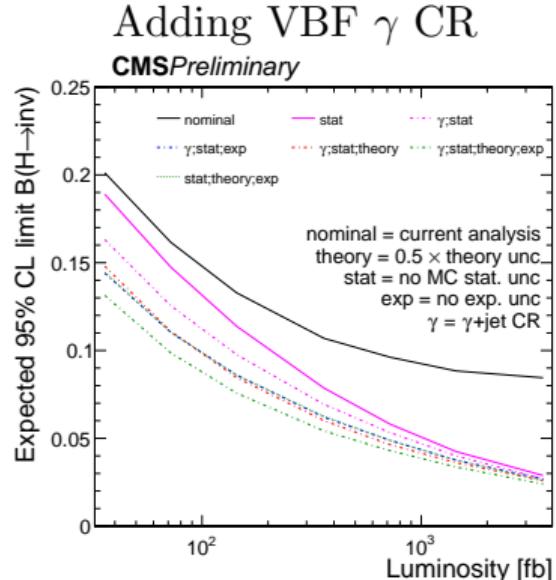


Uncertainty	VBF	Size	Effect
W/Z EW		15%	50%
W/Z QCD		15%	25%
Trigger		2%	20%
Lepton ID		2-3%	15%

Uncertainty	Mono-top	Size	Effect
V/Z QCD		15%	30%
b ID		2-10%	30%
Lepton ID		2-3%	15%



[10]



Conclusions

Mono-top

- ▶ Boosted top reconstruction extends reach
- ▶ Jet substructure is critical
- ▶ Strong constraints on flavor-changing DM models
- ▶ Substructure methods can be applied broadly

VBF $H \rightarrow$ invisible

- ▶ Very different set of jet challenges: triggering, backgrounds, energy calibration
- ▶ Precise measurement of QCD and EW V spectra

Conclusions

Mono-top

- ▶ Boosted top reconstruction extends reach
- ▶ Jet substructure is critical
- ▶ Strong constraints on flavor-changing DM models
- ▶ Substructure methods can be applied broadly

VBF $H \rightarrow$ invisible

- ▶ Very different set of jet challenges: triggering, backgrounds, energy calibration
- ▶ Precise measurement of QCD and EW V spectra

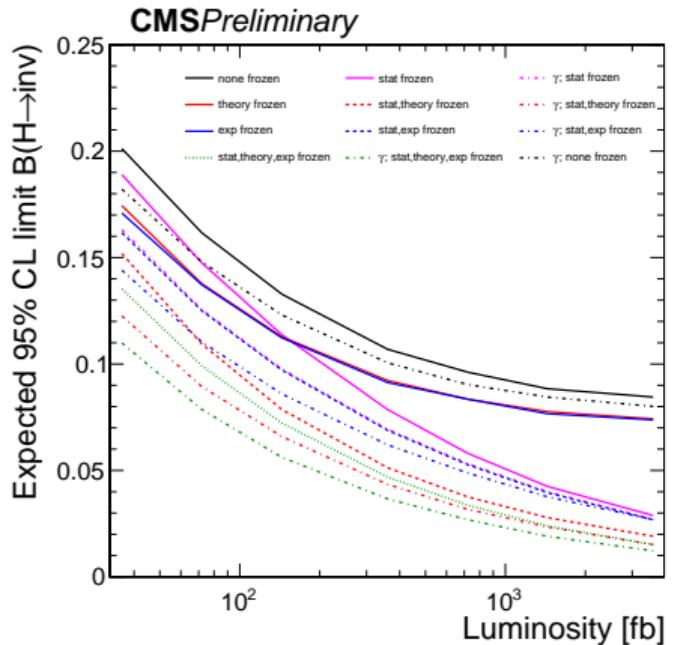
Conclusions

Mono-top

- ▶ Boosted top reconstruction extends reach
- ▶ Jet substructure is critical
- ▶ Strong constraints on flavor-changing DM models
- ▶ Substructure methods can be applied broadly

VBF $H \rightarrow$ invisible

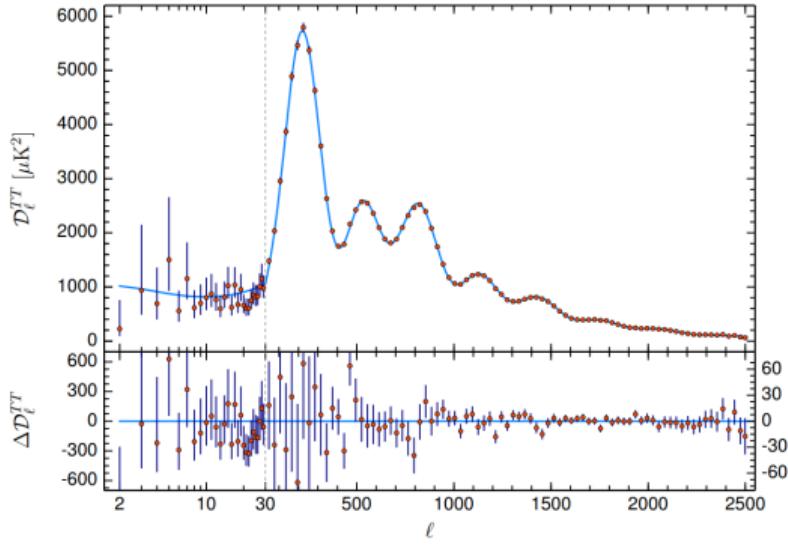
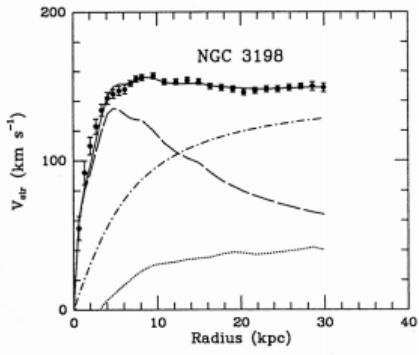
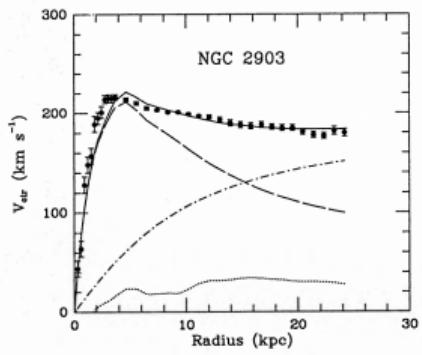
- ▶ Very different set of jet challenges: triggering, backgrounds, energy calibration
- ▶ Precise measurement of QCD and EW V spectra

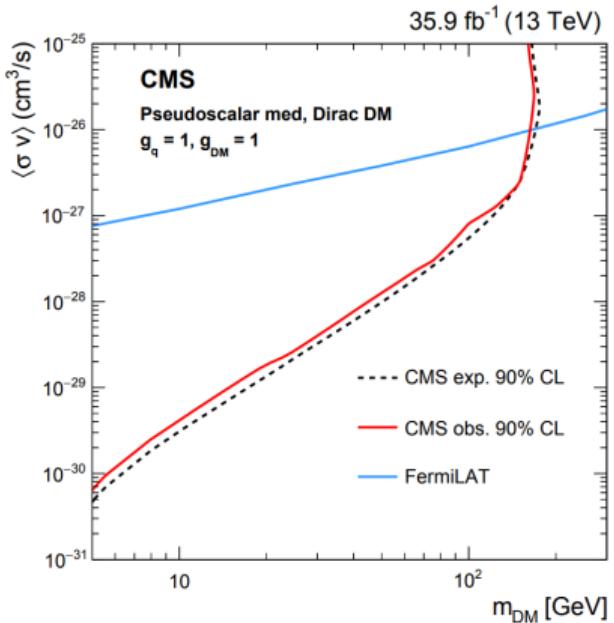
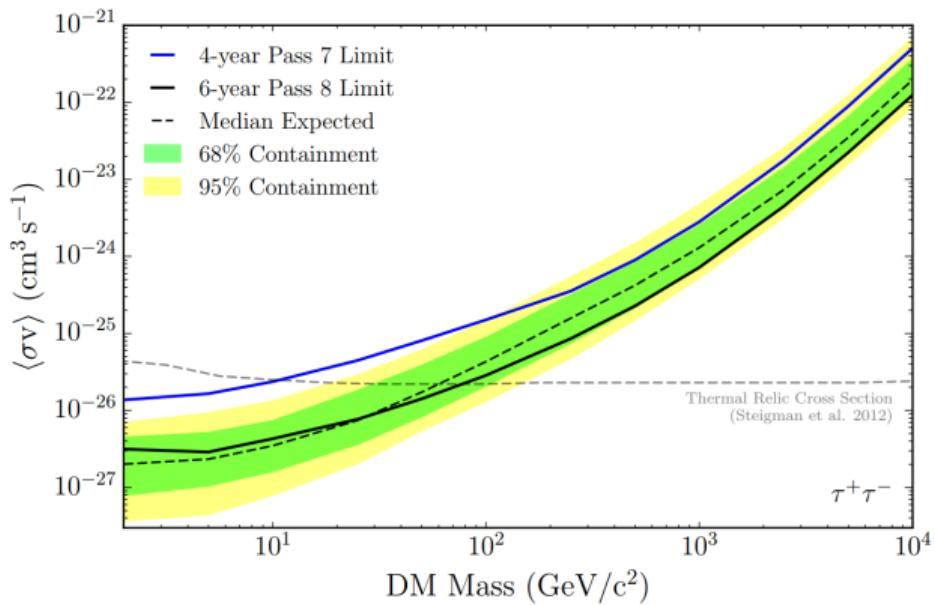


BACKUP

References

- [1] https://en.wikipedia.org/wiki/Galaxy_rotation_curve
- [2] arXiv:astro-ph/9909252
- [3] hubblesite.org/image/1276/news_release/2003-01
- [4] chandra.harvard.edu/photo/2006/1e0657/more.html
- [5] arXiv:astro-ph/1202.0090
- [6] [https://kipac.stanford.edu/highlights/
let-there-be-light-upon-dark-digging-deeper-dark-matter-lux](https://kipac.stanford.edu/highlights/let-there-be-light-upon-dark-digging-deeper-dark-matter-lux)
- [7] <https://fermi.gsfc.nasa.gov/science/eteu/dm/>
- [8] <https://cms-docdb.cern.ch/cgi-bin/PublicDocDB>ShowDocument?docid=4172>
- [9] arXiv:1609.07473
- [10] arXiv:1705.04664

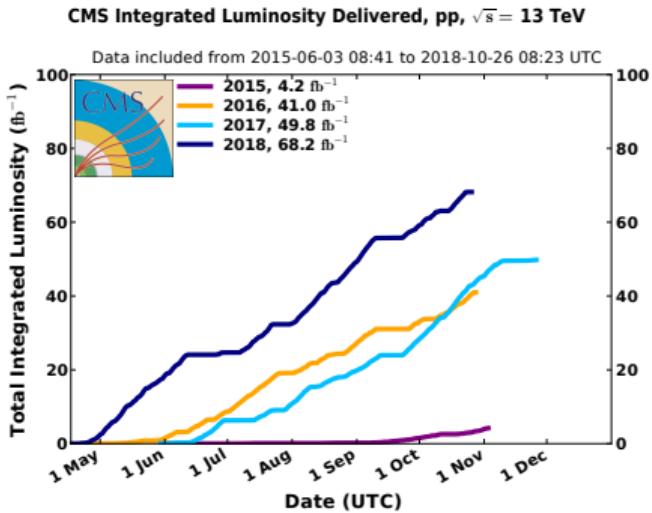




LHC luminosity

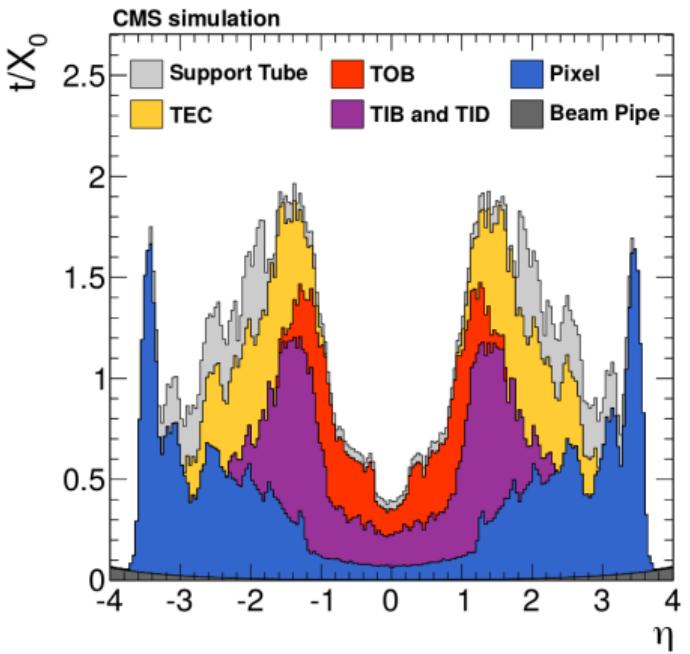
$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma F}{4\pi \epsilon \beta^*}$$

- ▶ N_b = particles per bunch
- ▶ n_b = bunches per beam
- ▶ f_{rev} = frequency of revolution
- ▶ $\gamma = E/m$ of beam
- ▶ ϵ = emittance of beam
- ▶ β^* = beta function at collision point
- ▶ F = factor accounting for beam intersection geometry



Tracking

- ▶ Hit resolution: $10\text{-}30 \mu\text{m}$ (pixels) and $10\text{-}500 \mu\text{m}$ (strips)
- ▶ Vertex resolution: $10\text{-}12 \mu\text{m}$
- ▶ Track efficiency: $> 98\%$ above 1 GeV



Calorimeters



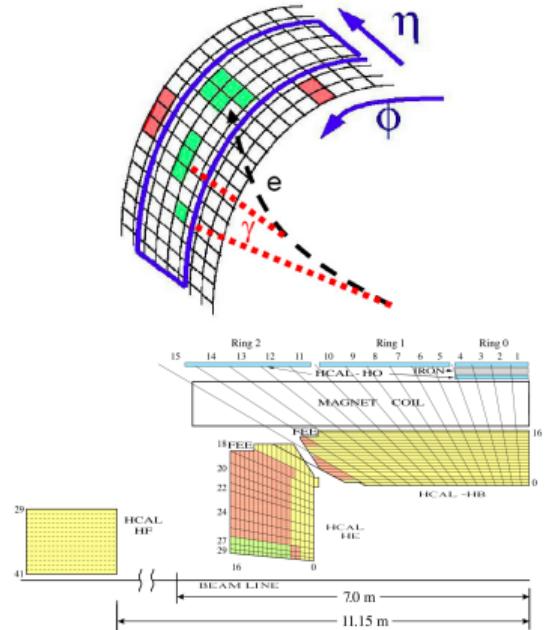
ECAL

- ▶ Barrel ($|\eta| < 1.44$) and ($|\eta| < 3$) endcap structures
- ▶ Lead tungstate crystals, $2.2 \times 2.2 \times 23$ cm 3
- ▶ $r_M = 2.19$ cm, $X_0 = 0.89$ cm

HCAL

- ▶ HB ($|\eta| < 1.4$), HE ([1.3, 3]), HF ([3, 5]), HO (1.3)
- ▶ Brass absorber ($\lambda_I = 1.5$ cm) and plastic scintillator]
- ▶ Thickness is 5.8-10.6 λ_I
- ▶ Tower segmentation is 0.087×0.087 in $\eta \times \phi$

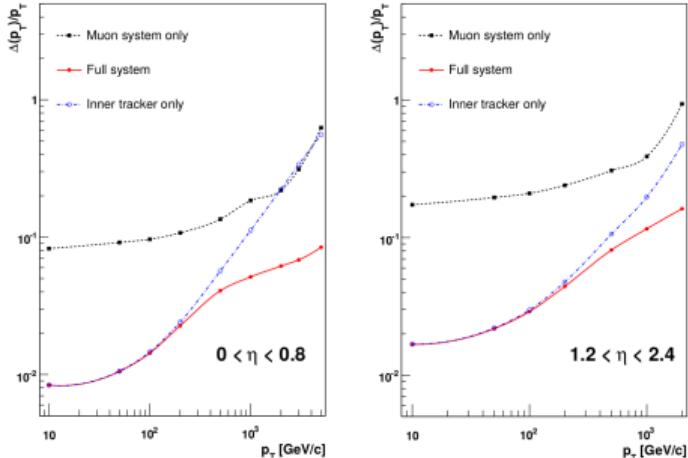
$$\frac{\sigma_E}{E} = \frac{1.98}{\sqrt{E/\text{GeV}}} \oplus 0.09$$



Muon chambers



- ▶ Drift tubes (barrel) and cathode strip chambers (endcaps)
- ▶ Hit resolution $78 - 152 \mu\text{m}$
- ▶ Efficiency is over 95%

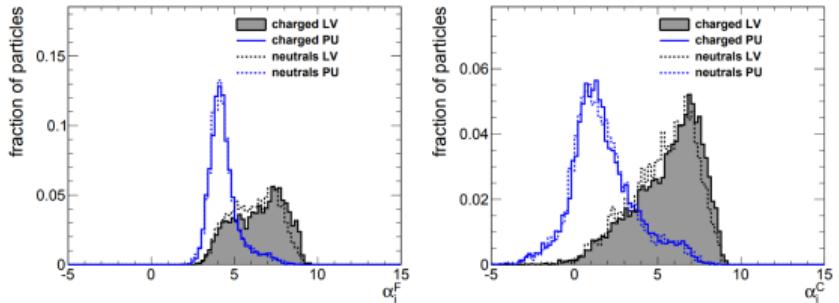


$$\alpha_i = \log \sum_{j \neq i} \frac{p_{T,j}}{\Delta R_{ij}} H(\Delta R_{ij} - R_{\min}) H(R_0 - \Delta R_{ij})$$

- ▶ Compute α_i for PV charged, PU charged, and neutrals
- ▶ Extrapolate in η and charge
- ▶ Compare neutral α_i to PV and PU distributions
- ▶ Rescale neutrals by probability:

$$w_i = P(\chi^2 < x_i | N_{\text{dof}} = 1)$$

$$x_i = H(\alpha_i - \bar{\alpha}) \frac{(\alpha_i - \bar{\alpha})^2}{\sigma^2}$$



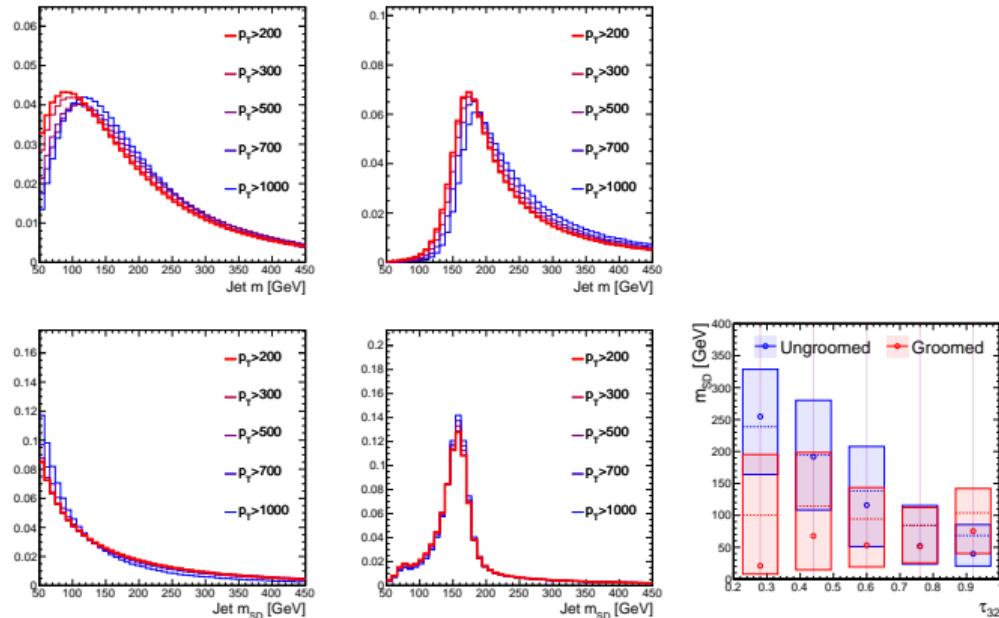
Soft drop

- ▶ Step through clustering tree, starting root node
- ▶ If the softer of the two branches $j = 1, 2$ satisfies:

$$\frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} < \left(\frac{\Delta R_{12}}{R} \right)^\beta$$

then drop it

- ▶ Re-compute quantities as a function of the groomed jet



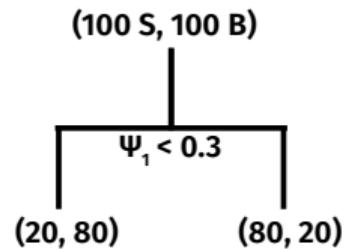
Classification (and regression) trees

- ▶ Sample the ECF ratio space with finite (but large) grid
- ▶ Greedily train a decision tree $T_0(\psi) \in \{0, 1\}$
- ▶ At each node n , choose a ψ_j and a boundary d_n
- ▶ Minimize some cost function, e.g.:

$$\ell(X, y; j, d_n) = -\hat{\pi}_B \ln \hat{\pi}_B - \hat{\pi}_S \ln \hat{\pi}_S$$

$$\hat{\pi}_c = \hat{P}(y = c | \psi_j < d_n)$$

- ▶ Proceed until stopping condition
 - ▶ Perfect classification
 - ▶ Maximum depth/nodes
 - ▶ Small change in ℓ



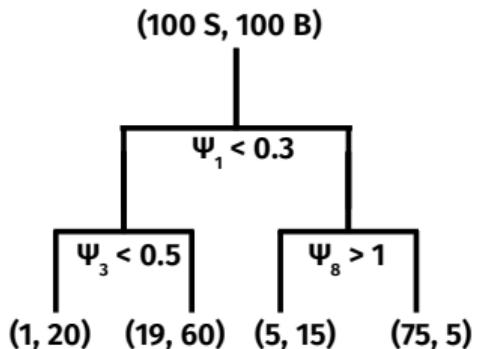
Classification (and regression) trees

- ▶ Sample the ECF ratio space with finite (but large) grid
- ▶ Greedily train a decision tree $T_0(\psi) \in \{0, 1\}$
- ▶ At each node n , choose a ψ_j and a boundary d_n
- ▶ Minimize some cost function, e.g.:

$$\ell(X, y; j, d_n) = -\hat{\pi}_B \ln \hat{\pi}_B - \hat{\pi}_S \ln \hat{\pi}_S$$

$$\hat{\pi}_c = \hat{P}(y = c | \psi_j < d_n)$$

- ▶ Proceed until stopping condition
 - ▶ Perfect classification
 - ▶ Maximum depth/nodes
 - ▶ Small change in ℓ



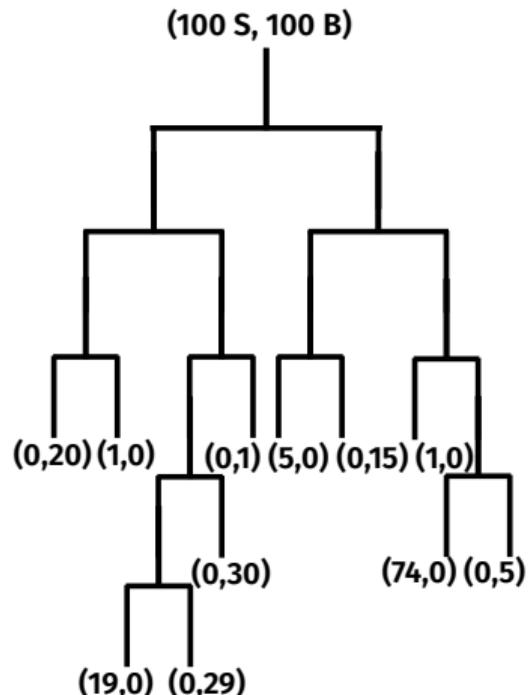
Classification (and regression) trees

- ▶ Sample the ECF ratio space with finite (but large) grid
- ▶ Greedily train a decision tree $T_0(\psi) \in \{0, 1\}$
- ▶ At each node n , choose a ψ_j and a boundary d_n
- ▶ Minimize some cost function, e.g.:

$$\ell(X, y; j, d_n) = -\hat{\pi}_B \ln \hat{\pi}_B - \hat{\pi}_S \ln \hat{\pi}_S$$

$$\hat{\pi}_c = \hat{P}(y = c | \psi_j < d_n)$$

- ▶ Proceed until stopping condition
 - ▶ Perfect classification
 - ▶ Maximum depth/nodes
 - ▶ Small change in ℓ



Boosted trees

- ▶ Decision trees prone to overfitting
 - ▶ Need to be very large for complex decision boundaries
- ▶ Boosting is a standard method to combine an ensemble of shallow trees
- ▶ Need to define a global loss function L , e.g.

$$L(y_i; f_i) = \ln(1 + \exp(-y_i f_i))$$

$$L(y_i; f_i) = \ln(1 + \exp(-y_i f_i))$$

1. Initialize classifier $f_0 = T_0$
2. Until some stopping condition (index $m = 1, \dots$):
 - 2.1. Compute the “residual”

$$r_{mi} = -\nabla_f L(y_i; f)|_{f=f_{m-1}(\psi_i)}$$

$$L = (y - f_m(\psi))^2 \Rightarrow r_m = y - f_m(\psi)$$

- 2.2. Fit a regression tree T_m to predict r_{mi} as a function of x_i :

$$\ell(X, r_m; j, d, \hat{r}) = \sum_{i|\psi_j i < d} (r_{mi} - \hat{r})^2$$

- 2.3. Update $f_m = f_{m-1} + \nu T_m$

Dimensional reduction

- ▶ ECF ratio space is much bigger than the problem we're solving
 - ▶ Large subsets are useless/correlated
 - ▶ Strong non-linear relationships among useful ratios
- ▶ Inference speed is important
 - ▶ Reduce number of ψ_i to compute
 - ▶ Make trees smaller
- ▶ Standard dimensional reduction techniques aren't ideal
 - ▶ PCA
 - ▶ Kernel PCA
 - ▶ Autoencoders, etc.

- ▶ Can we embed dimensional reduction into the classification?

1. Train a BDT with trees T_1, \dots, T_n
2. For each ψ_i , define a score:

$$s_i = \sum_{m=1}^n \nu^{m-1} \sum_{\text{nodes using } \psi_i \text{ in } T_m} N_{\text{samples}}(\text{node}) \times (\ell(\text{node}) - \ell(\text{parent}))^2$$

3. Remove (one or more) ψ_i with smallest s_i and repeat

- ▶ Iterative training is expensive, but can significantly improve inference speed
- ▶ Grid computing resources makes training semi-parallelizable

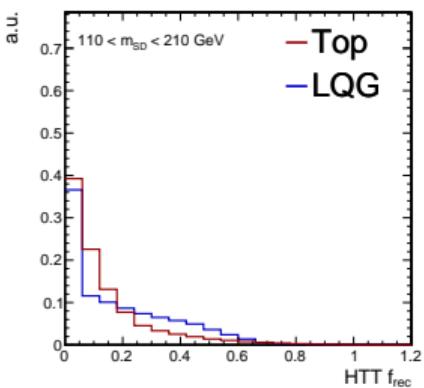
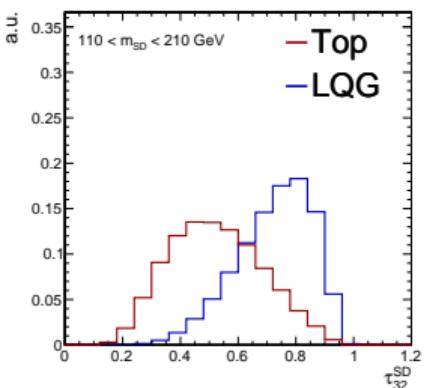
τ_{32} and f_{rec}

- $\tau_{32} = \tau_3/\tau_2$, where:

$$\tau_N = \frac{\sum_{i \in \text{jet}} p_{T,i} \min\{\Delta R_{ia} | a \in A\}}{\sum_{i \in \text{jet}} p_{T,i} R}$$

- Compute the filtered subjets of the CA15 jet, and choose the three (1, 2, 3) most consistent with top decay products. Then define:

$$f_{\text{rec}} = \min_{0 \leq i < j \leq 2} \left| \frac{m_{ij}/m_{123}}{m_W/m_t} - 1 \right|$$



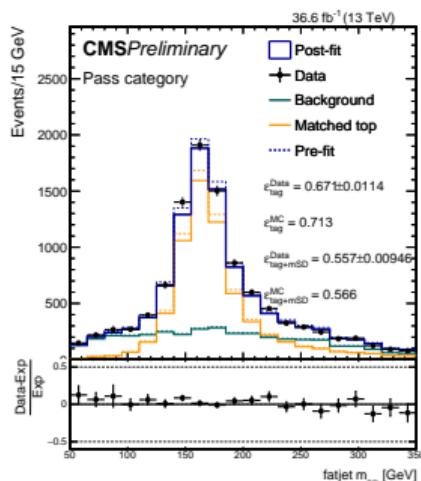
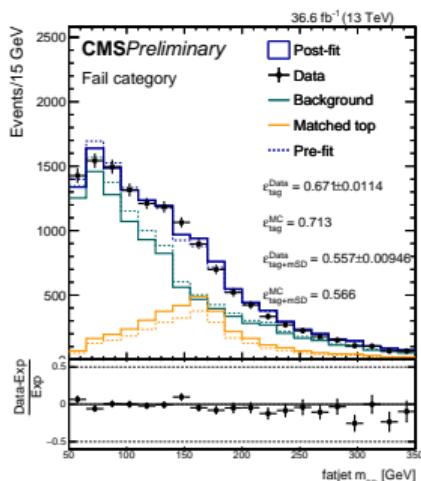
Efficiency in data

$$SF_{top}(0.1) = 1.08 \pm 0.04(\text{total})$$

$\pm 0.03(\text{statistical}) \pm 0.02(\text{JES} + \text{JER}) \pm 0.02(\text{merging}) \pm 0.002(b)$

$$SF_{top}(0.45) = 1.07 \pm 0.06(\text{total})$$

$\pm 0.03(\text{statistical}) \pm 0.02(\text{JES} + \text{JER}) \pm 0.014(\text{merging}) \pm 0.000(b)$



Monotop Lagrangians



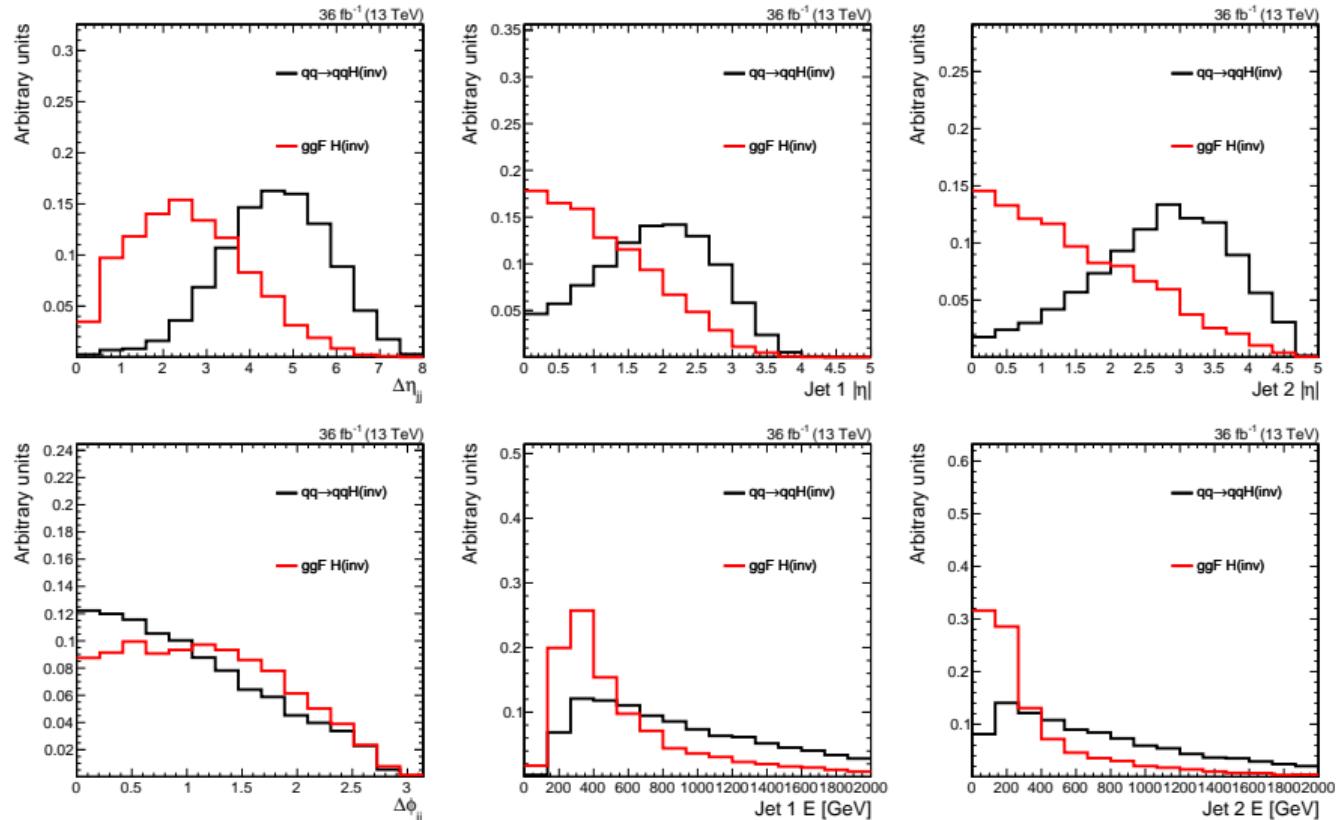
$$\mathcal{L}_{\text{int}} = V_\mu \bar{\chi} \gamma^\mu (g_\chi^V + g_\chi^A \gamma_5) \chi + \bar{q}_u \gamma^\mu (g_u^V + g_u^A \gamma_5) q_u V_\mu + \bar{q}_d \gamma^\mu (g_d^V + g_d^A \gamma_5) q_d V_\mu + \text{h.c.},$$

$$\mathcal{L}_{\text{int}} = \phi \bar{d}_i^C [(a_q)^{ij} + (b_q)^{ij} \gamma^5] d_j + \phi \bar{t} [a_\psi + b_\psi \gamma^5] \psi + \text{h.c.}$$

Monotop likelihood

$$\begin{aligned}
 \mathcal{L} \left(\mathbf{d} \mid \mu, \boldsymbol{\mu}_{\text{SR}}^{Z \rightarrow \nu\nu}, \boldsymbol{\mu}_{\text{SR}}^{t\bar{t}}, \boldsymbol{\theta} \right) = \\
 \prod_{i \in \text{bins}} \left[\text{Pois} \left(d_i^{\text{SR}} \mid \mu S_i^{\text{SR}}(\boldsymbol{\theta}) + \mu_{\text{SR},i}^{Z \rightarrow \nu\nu} + \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{Z/W,i}^{\text{SR}}(\boldsymbol{\theta})} + \mu_{\text{SR},i}^{t\bar{t}} + B_i^{\text{SR}}(\boldsymbol{\theta}) \right) \right. \\
 \times \prod_{X=\mu\mu, ee} \text{Pois} \left(d_i^X \mid \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{Z,i}^X(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta}) \right) \\
 \times \prod_{X=b\mu, be} \text{Pois} \left(d_i^X \mid \frac{\mu_{\text{SR},i}^{t\bar{t}}}{T_{t\bar{t},i}^X(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta}) \right) \\
 \times \prod_{X=\mu, e} \text{Pois} \left(d_i^X \mid \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{W,i}^X(\boldsymbol{\theta}) T_{Z/W,i}^{\text{SR}}(\boldsymbol{\theta})} + \frac{\mu_{\text{SR},i}^{t\bar{t}}}{T_{t\bar{t},i}^X(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta}) \right) \\
 \left. \times \text{Pois} \left(d_i^\gamma \mid \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{\gamma,i}^\gamma(\boldsymbol{\theta})} \right) \right] \times \prod_{j=0}^{n_\theta} p_j(\theta_j)
 \end{aligned}$$

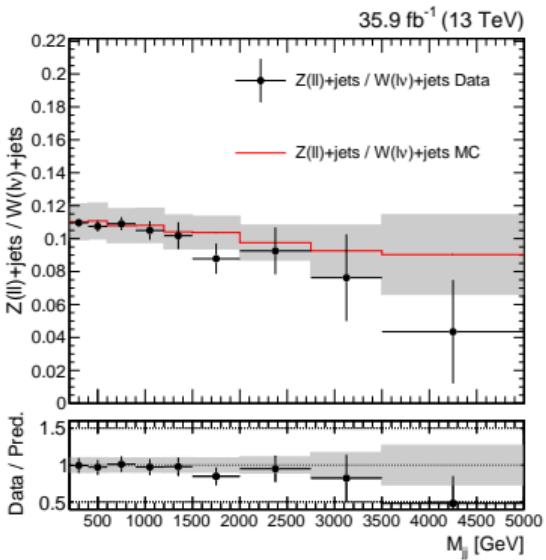
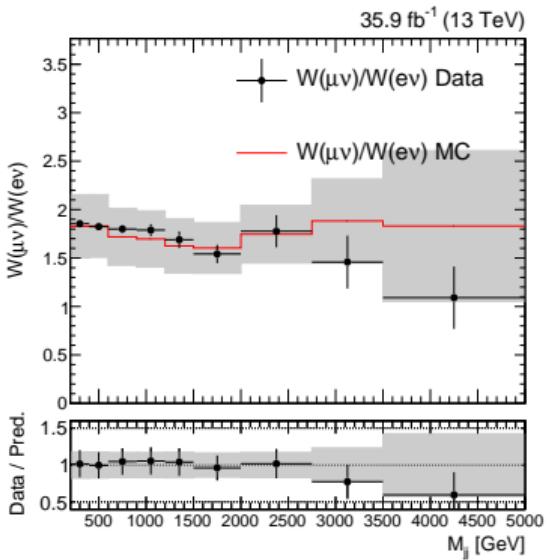
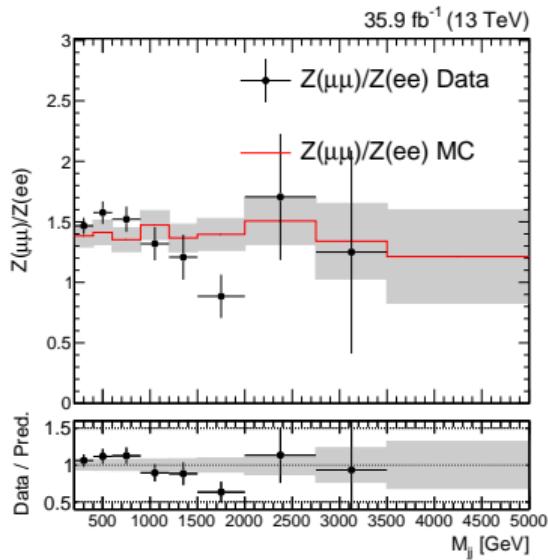
VBF vs gluon fusion



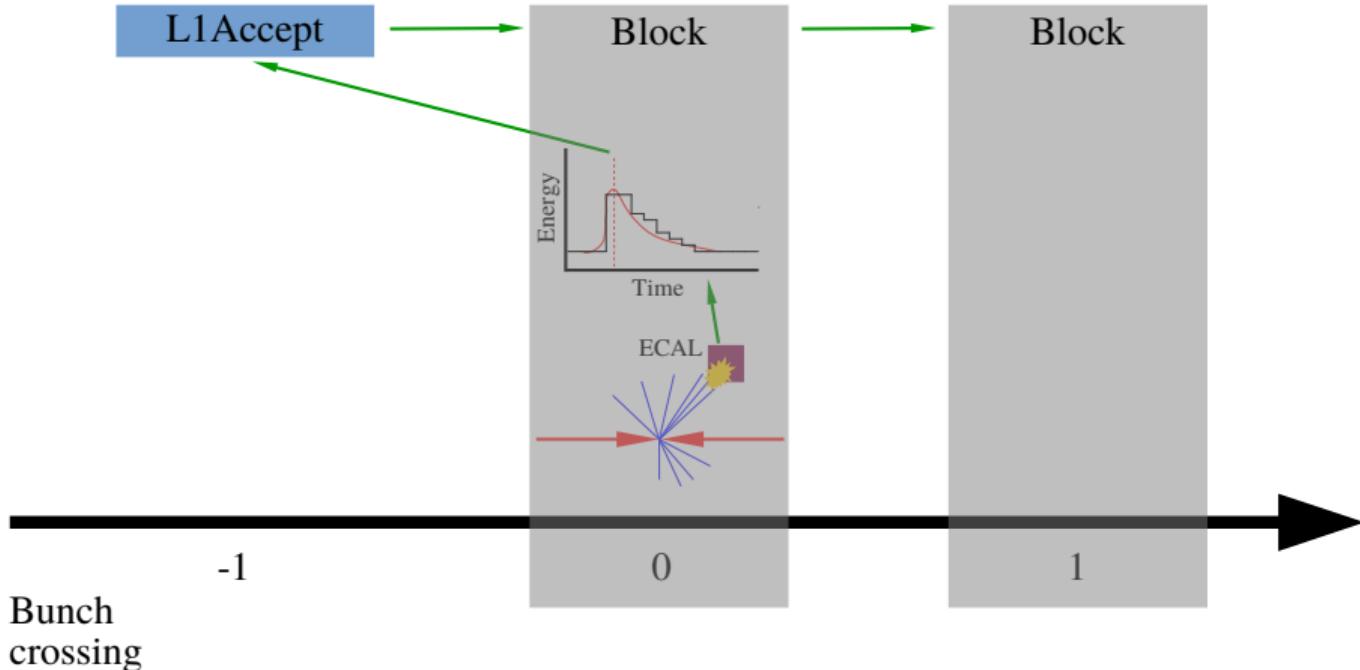
VBF likelihood

$$\begin{aligned}
 \mathcal{L}(\mathbf{d} | \mu, \boldsymbol{\mu}_{\text{SR}}^{Z \rightarrow \nu\nu}, \boldsymbol{\theta}) = \\
 \prod_{i \in \text{bins}} \left[\text{Pois} \left\{ d_i^{\text{SR}} \mid \mu S_i^{\text{SR}}(\boldsymbol{\theta}) + \left(1 + \frac{1}{T_{Z,i}^{\text{QE}}(\boldsymbol{\theta})}\right) \left(1 + \frac{1}{T_{Z/W,i}^{\text{SR}}(\boldsymbol{\theta})}\right) \mu_{\text{SR},i}^{Z \rightarrow \nu\nu} + B_i^{\text{SR}}(\boldsymbol{\theta}) \right\} \right. \\
 \times \prod_{X=\mu,e} \text{Pois} \left\{ d_i^X \mid \left(1 + \frac{1}{T_{Z,i}^{\text{QE}}(\boldsymbol{\theta})}\right) \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{W,i}^X(\boldsymbol{\theta}) T_{Z/W,i}^{\text{SR}}(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta}) \right\} \\
 \left. \times \prod_{X=\mu\mu,ee} \text{Pois} \left\{ d_i^X \mid \left(1 + \frac{1}{T_{Z,i}^{\text{QE}}(\boldsymbol{\theta})}\right) \frac{\mu_{\text{SR},i}^{Z \rightarrow \nu\nu}}{T_{Z,i}^X(\boldsymbol{\theta})} + B_i^X(\boldsymbol{\theta}) \right\} \right] \times \prod_{j=0}^{n_\theta} p_j(\theta_j)
 \end{aligned}$$

VBF T validation



L1 pre-firing mechanism



Un-pre-fireable events

