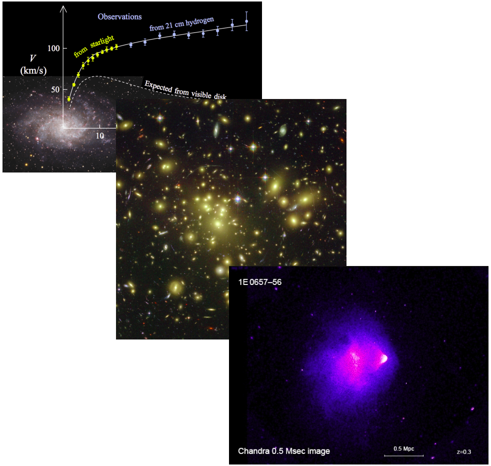
Dark matter, jets, and substructure at the CMS experiment

Siddharth Narayanan

SLAC - 2018/07/19

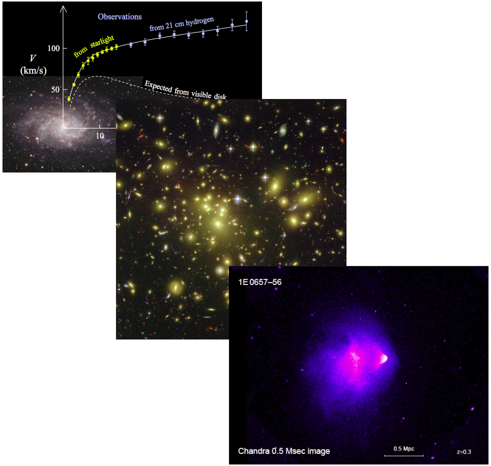
Dark matter at the LHC

- Existence of dark matter strongly suggested by astrophysical observations
  - Galaxy rotation curves
  - Gravitational lensing
  - Bullet Cluster
  - ...



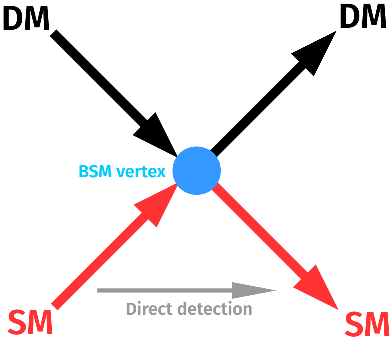
Dark matter at the LHC

- Existence of dark matter strongly suggested by astrophysical observations
  - Galaxy rotation curves
  - Gravitational lensing
  - Bullet Cluster
  - ...
- Most evidence relies on gravitational effects of DM
- Can we find other evidence for interaction between SM and DM?



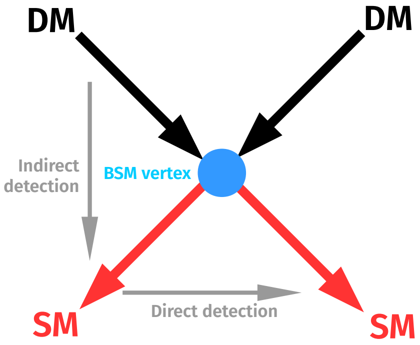
Dark matter at the LHC

- Existence of dark matter strongly suggested by astrophysical observations
  - Galaxy rotation curves
  - Gravitational lensing
  - Bullet Cluster
  - ...
- Most evidence relies on gravitational effects of DM
- Can we find other evidence for interaction between SM and DM?



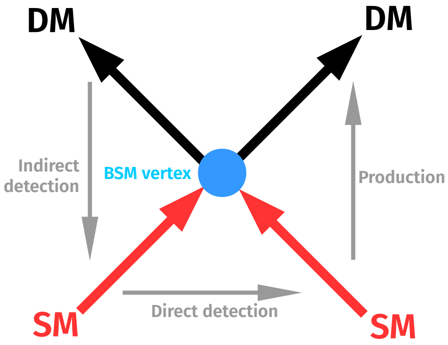
Dark matter at the LHC

- Existence of dark matter strongly suggested by astrophysical observations
  - Galaxy rotation curves
  - Gravitational lensing
  - Bullet Cluster
  - ...
- Most evidence relies on gravitational effects of DM
- Can we find other evidence for interaction between SM and DM?



Dark matter at the LHC

- Existence of dark matter strongly suggested by astrophysical observations
  - Galaxy rotation curves
  - Gravitational lensing
  - Bullet Cluster
  - ...
- Most evidence relies on gravitational effects of DM
- Can we find other evidence for interaction between SM and DM?

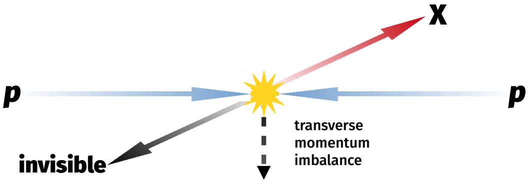


Seeing the invisible

- DM candidates would couple weakly to collider detectors
- Not much point in producing DM if we can't see it!
- The solution: searches for "mono-**X**"
  - DM produced in association with one or more SM particles (**X**)

Seeing the invisible

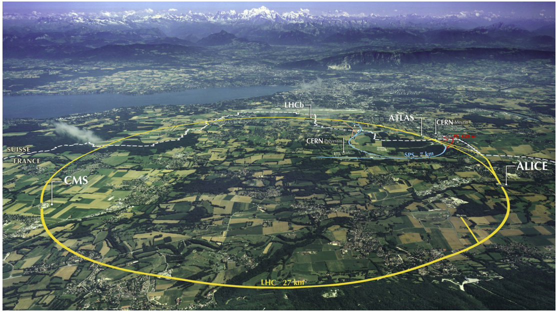
- DM candidates would couple weakly to collider detectors
- Not much point in producing DM if we can't see it!
- The solution: searches for "mono-**X**"
  - DM produced in association with one or more SM particles (**X**)



- **X** creates a transverse momentum imbalance (p_T^{miss})
- Large p_T^{miss} + conservation of momentum $\Rightarrow$ invisible particles!
- In certain cases, can trigger on **X**

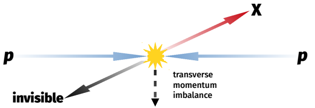
$p_{\mathrm{T}}^{\mathrm{miss}}$ at the LHC

- CMS records proton collisions from the LHC
  - Today: $\sqrt{s} = 13\,\mathrm{TeV}$ results



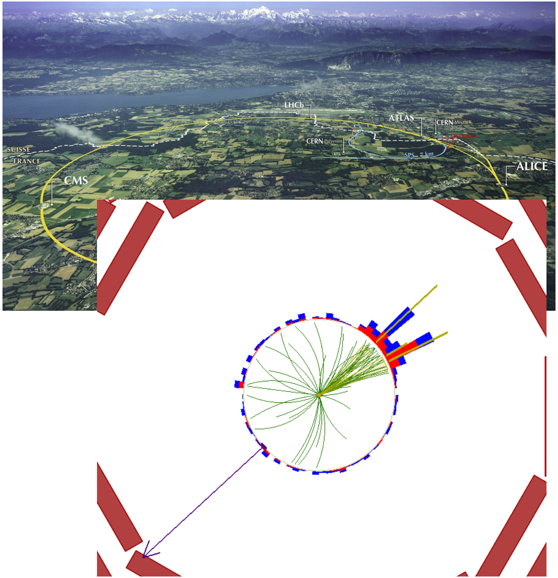
$p_{\mathrm{T}}^{\mathrm{miss}}$ at the LHC

- CMS records proton collisions from the LHC
  - Today: $\sqrt{s} = 13\,\mathrm{TeV}$ results
- pp events are messy, so replace:



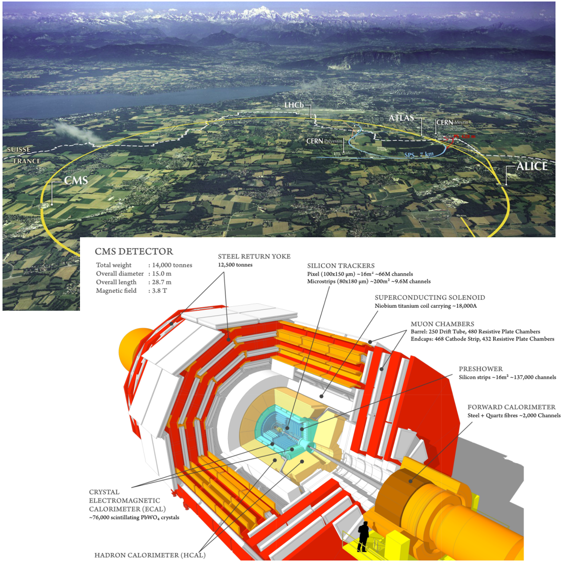
- with:

$$\vec{p}_{\mathrm{T}}^{\mathrm{miss}} = -\left(\sum_{i \in \mathrm{particles}} \vec{p}_i\right)_{\mathrm{T}}$$



Compact Muon Solenoid

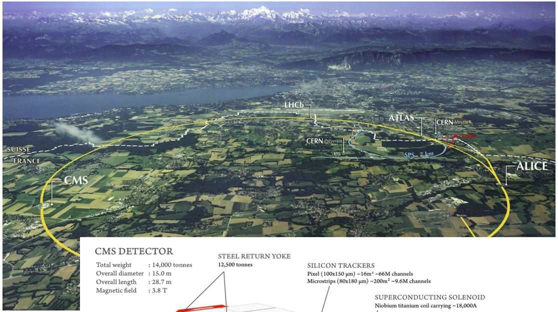
All particles in sum $\Rightarrow$
all subdetectors help measure p_T^{miss}!

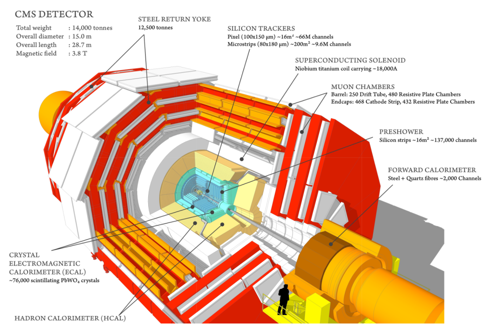


Compact Muon Solenoid

All particles in sum $\Rightarrow$
all subdetectors help measure $p_{\mathrm{T}}^{\mathrm{miss}}$!

- Solenoidal magnet
  - 3.8 T B field
- Silicon tracker
  - Charged particles' $\vec{p}$
  - Track vertices

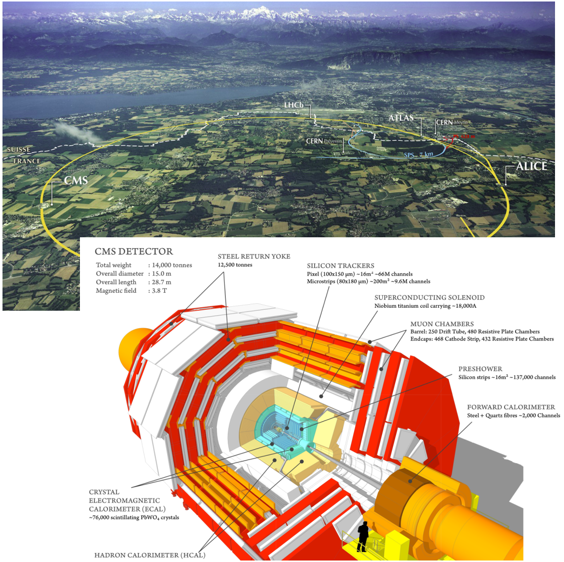




Compact Muon Solenoid

All particles in sum $\Rightarrow$
all subdetectors help measure $p_{\mathrm{T}}^{\mathrm{miss}}$!

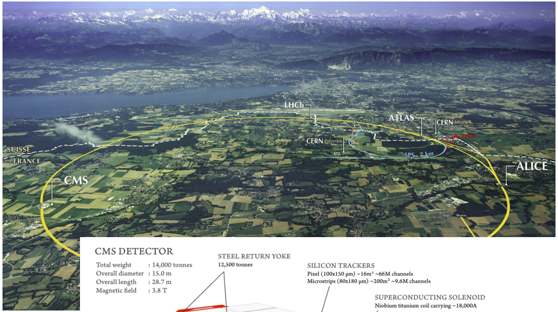
- Solenoidal magnet
  - 3.8 T B field
- Silicon tracker
  - Charged particles' $\vec{p}$
  - Track vertices
- Calorimeters
  - EM and hadronic
  - Good energy resolution
  - Large coverage

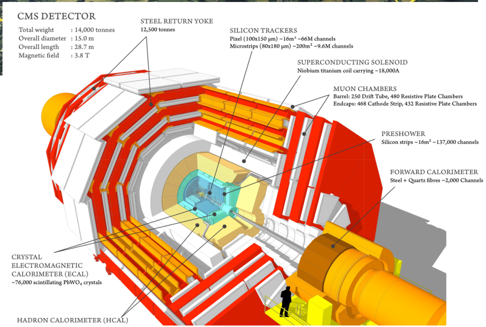


Compact Muon Solenoid

All particles in sum $\Rightarrow$
all subdetectors help measure $p_{\mathrm{T}}^{\mathrm{miss}}$!

- 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 $\vec{p}$





A broad spectrum of DM models and X

	Spin-1 mediated	Spin-0 mediated	Fermion portal DM	Non-thermal DM	ADD	2HDM	Baryonic Z'
X$=q, g$	■	■	■	■	■		
qq'		■					
$V(q\bar{q}')$	■	■					
γ	■				■		
$Z(\ell^+\ell^-)$	■	■			■		
$t\bar{t}$		■					
$b/b\bar{b}$		■					
t	■						
$H(b\bar{b})$						■	■

All signatures characterized by high $p_{\mathrm{T}}^{\mathrm{miss}}$, but choice of X necessitates different reconstruction and background estimation strategies

A broad spectrum of DM models and **X**

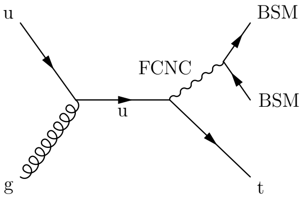
	Spin-1 mediated	Spin-0 mediated	Fermion portal DM	Non-thermal DM	ADD	2HDM	Baryonic Z'
X$=q,g$	■	■	■	■	■		
qq'		■					
$V(q\bar{q}')$	■	■					
γ	■				■		
$Z(\ell^+\ell^-)$	■	■			■		
t	■						
$t\bar{t}$		■					
$b/b\bar{b}$		■					
$H(b\bar{b})$						■	■

This talk: focus on DM+jets

Outline

Signature

Mono-top

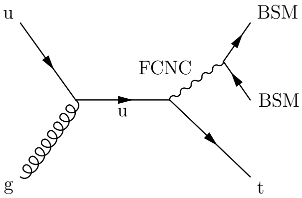


Highlights

Outline

Signature

Mono-top



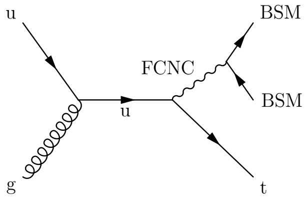
Highlights

Jet substructure
Invisible background estimation

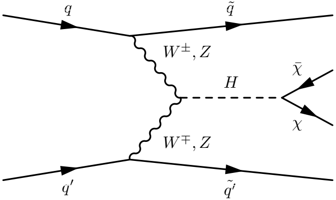
Outline

Signature

Mono-top



VBF Higgs→DM



Highlights

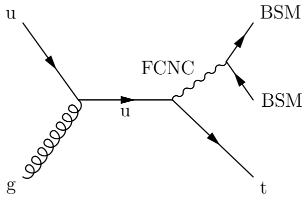
Jet substructure

Invisible background estimation

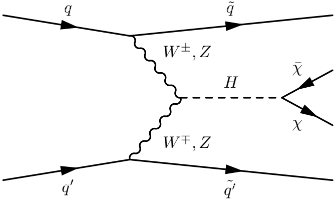
Outline

Signature

Mono-top



VBF Higgs→DM



Highlights

Jet substructure
Invisible background estimation

Electroweak SM backgrounds
Forward jets

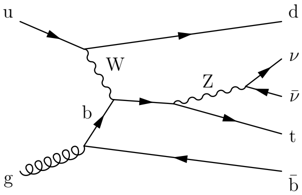
Mono-Top

Hallmarks of top quark+p_T^{miss}

- Let's forget any specific DM model
- This final state must violate flavor conservation
  - SM FC processes will have a b quark in the final state (up to CKM suppression)
- Excess mono-top production $\Rightarrow$ flavor-changing BSM

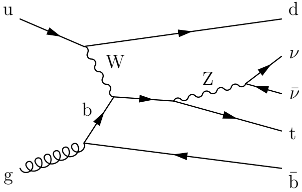
Hallmarks of top quark+$p_{\mathrm{T}}^{\mathrm{miss}}$

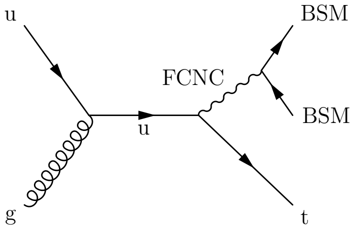
- Let's forget any specific DM model
- This final state must violate flavor conservation
  - SM FC processes will have a b quark in the final state (up to CKM suppression)
- Excess mono-top production ⇒ flavor-changing BSM
- SM process is tiny: 0.14 pb

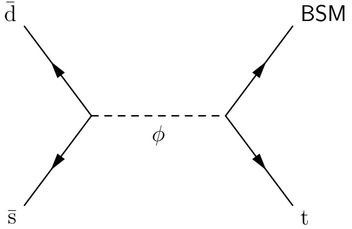


Hallmarks of top quark+p_T^{miss}

- Let's forget any specific DM model
- This final state must violate flavor conservation
  - SM FC processes will have a b quark in the final state (up to CKM suppression)
- Excess mono-top production ⇒ flavor-changing BSM
- SM process is tiny: 0.14 pb

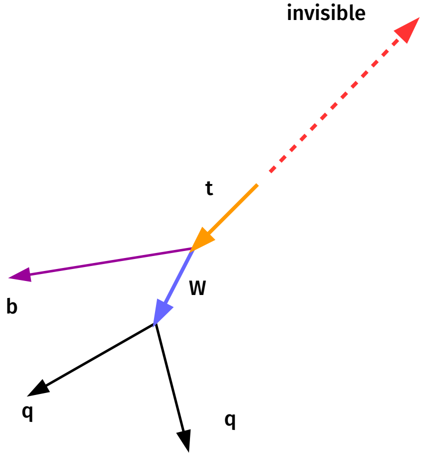






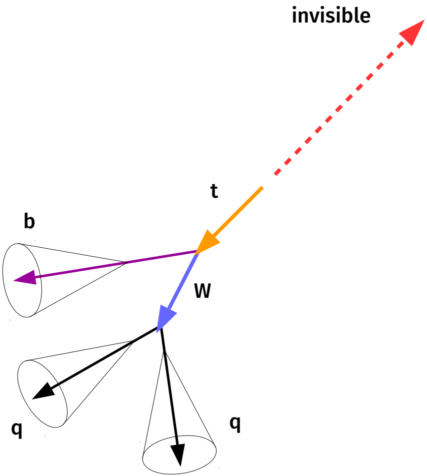
Anatomy of a mono-top event

Hadronic decay ⇒ larger BR, no $p_{\mathrm{T}}^{\mathrm{miss}}$

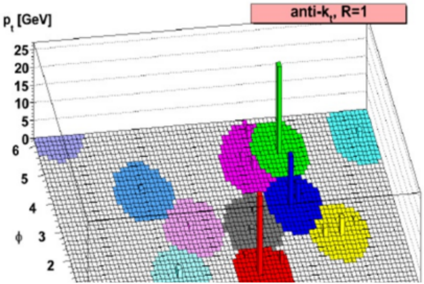


Anatomy of a mono-top event

Bare quarks hadronize into jets

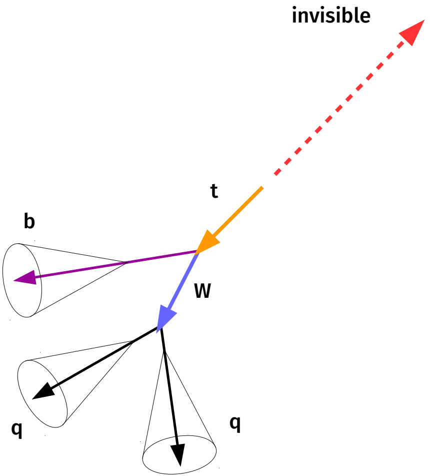


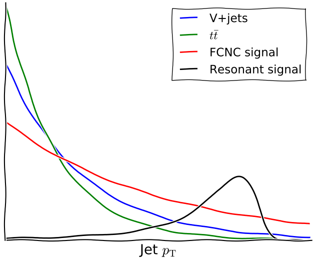
- Bare quarks are not color singlets
- Color confinement $\Rightarrow$ production of color singlets (hadrons)
- Hadrons collimated in direction of parton
- "Jets" are reconstructed using iterative algorithms at LHC



Anatomy of a mono-top event

Bare quarks hadronize into jets

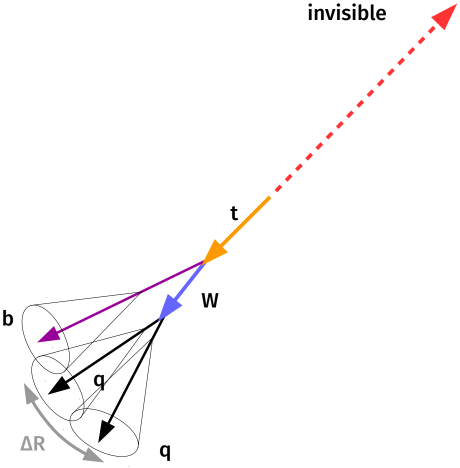


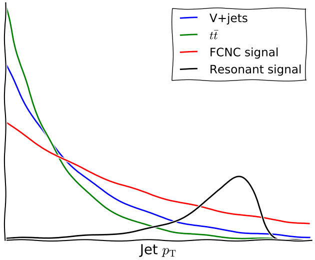


- Signal events generally more energetic than SM events
- Want to increase S/B? Look for very energetic top quarks!

Anatomy of a mono-top event

Decay products collimate

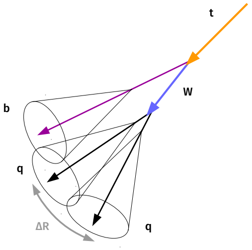




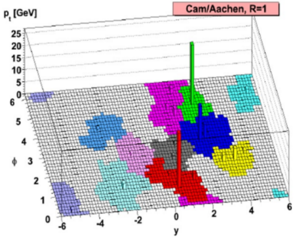
- Signal events generally more energetic than SM events
- Want to increase S/B? Look for very energetic top quarks!
- Separation between jets: $\Delta R \sim 2m_t/p_T$
  - $p_T > 250\,\text{GeV} \Rightarrow$ jets ($R = 0.4$) overlap

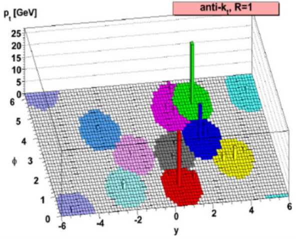
Reconstruction of top quark

- Three $R = 0.4$ jets $\rightarrow$ single $R = 1.5$ jet
- Circular jets $\rightarrow$ oddly-shaped jets
  - Jet is the sum of 3 jets
  - Anti-k_{T} $\rightarrow$ C/A



- These are **big** jets
  - $R = 1.5$ can contain up to half the detector
- Lots of extra radiation in jet
  - PU, ISR, UE/MPI
- Combinatorial fakes (q/g)

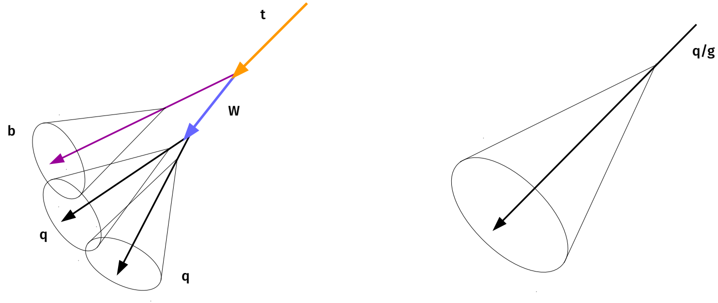




J. Phys.: Conf. Ser. 645 012008

Jet substructure

- Top quark → 3q ⇒ top jet has 3 “prongs”: regions of correlated radiation



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

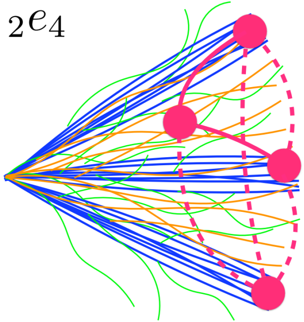
Energy correlation functions

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

$_2e_4$



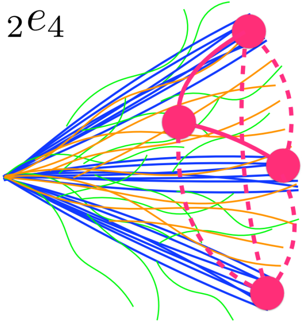
[arXiv:1609.07473]

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



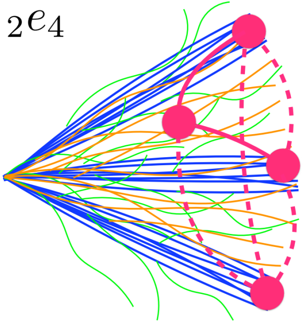
[arXiv:1609.07473]

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

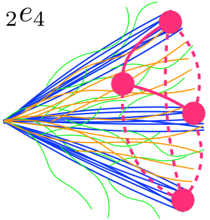


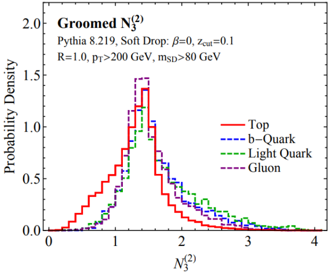
[arXiv:1609.07473]

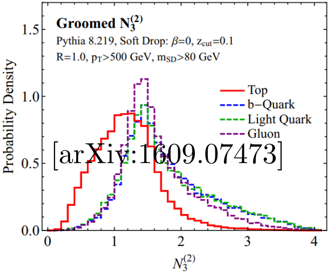
ECF behavior

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

- Top jet: $\mathbf{N}=3$ correlations are strong, $\mathbf{N}=4$ are weak
- q/g jets: $\mathbf{N}=3$ and $\mathbf{N}=4$ are both weak
- $e(\mathbf{N}=4)/e(\mathbf{N}=3)$







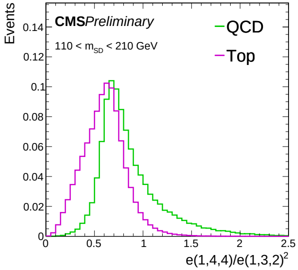
[arXiv:1609.07473]

Space of ECF ratios

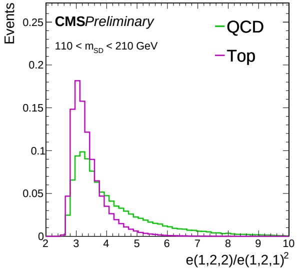
- Can extend argument to infinitely large ratio space

$$\frac{e(a,\mathbf{N},\alpha)}{e(b,\mathbf{M},\beta)^{x}}, \text{ where } M \leq N \text{ and } x = \frac{a\alpha}{b\beta}$$

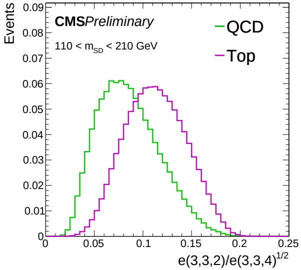
- Turns out many correlation function ratios can separate signal and background



$e(\mathbf{N}=4)/e(\mathbf{N}=3)$



$e(\mathbf{N}=2)/e(\mathbf{N}=2)$

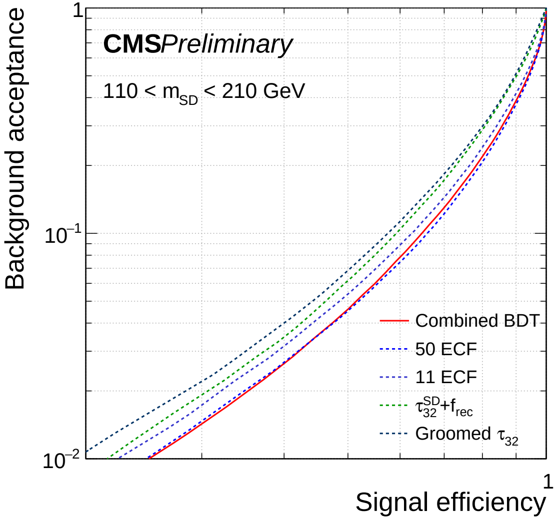


$e(\mathbf{N}=3)/e(\mathbf{N}=3)$

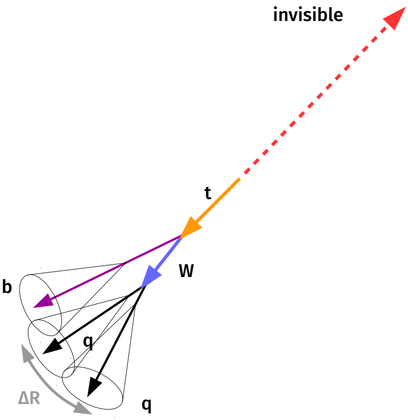
Building a discriminator

- Large number of substructure observables to choose from
- Many are highly correlated or not useful
- Use boosted decision trees to prune space of observables and extract useful information

	$\epsilon_{bkg}(\epsilon_{sig} = 0.5)$
τ_{32}	6.9%
Combined BDT	**4.7%**

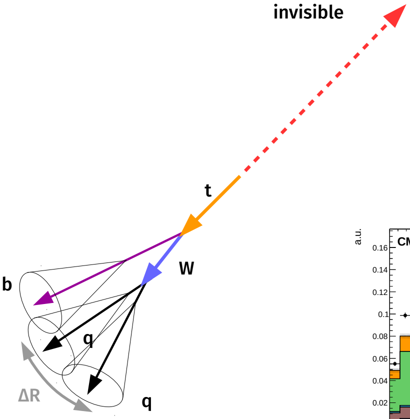


Selecting mono-top events

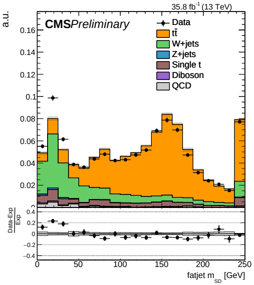


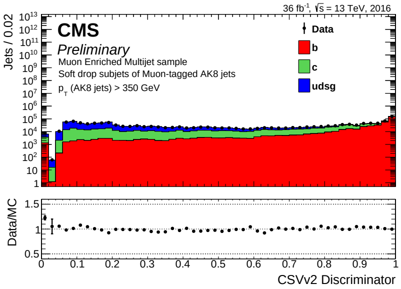
- $p_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$ (trigger threshold)
- CA15 jet, $p_{\mathrm{T}} > 250\,\mathrm{GeV}$
  - Selected by BDT

Selecting mono-top events



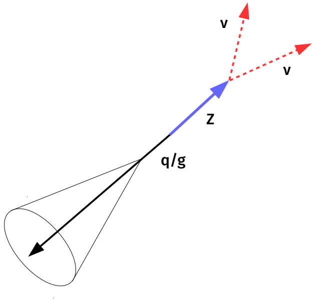
- $p_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$ (trigger threshold)
- CA15 jet, $p_{\mathrm{T}} > 250\,\mathrm{GeV}$
  - Selected by BDT
  - Mass consistent with m_t
  - Signature of B meson decay inside jet
    - Lab frame $c\tau \sim \mathcal{O}(\mathrm{mm})$





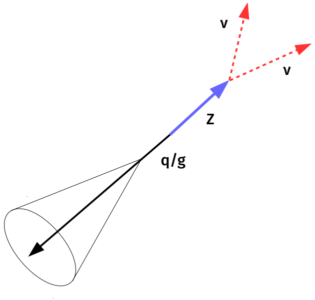
SM backgrounds

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

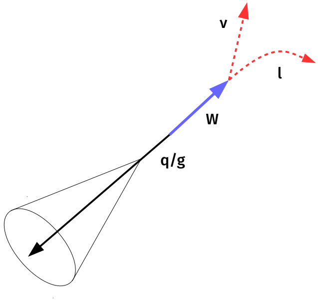


SM backgrounds

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

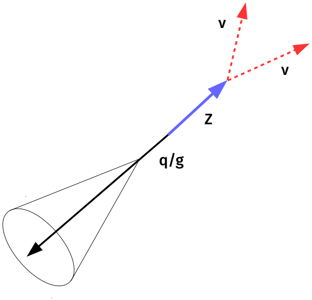


$W \to (\ell)\nu$ (15%)

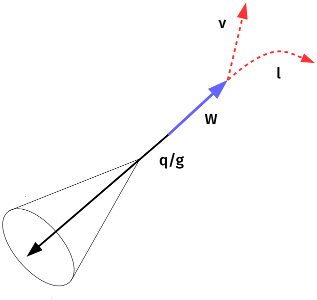


SM backgrounds

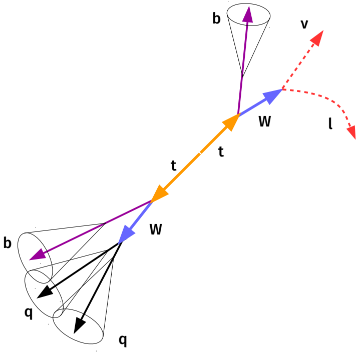
$Z \to \nu\nu$ (30%)



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

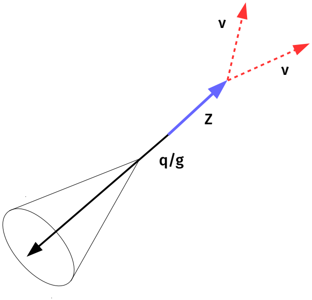


t quark pair (50%)

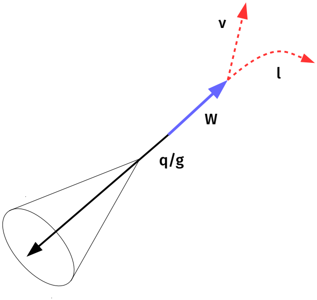


SM backgrounds

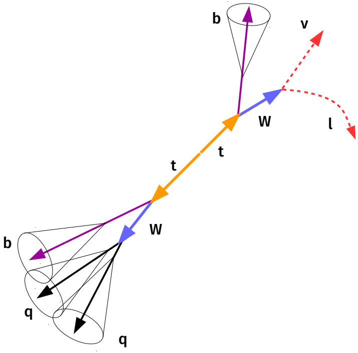
$Z \to \nu\nu$ (30%)



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



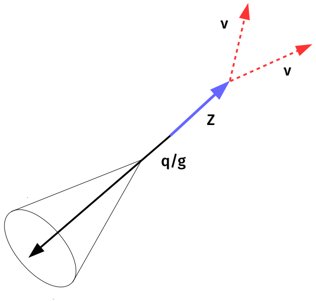
t quark pair (50%)



Note that $p_{\mathrm{T}}^{\mathrm{miss}}$ is the transverse momentum of the **vector boson**

Background estimation

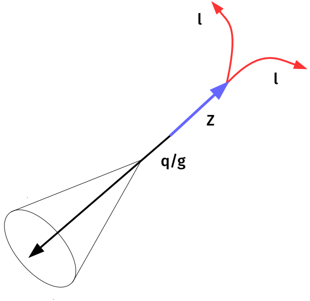
$Z \to \nu\nu$



Ideally:

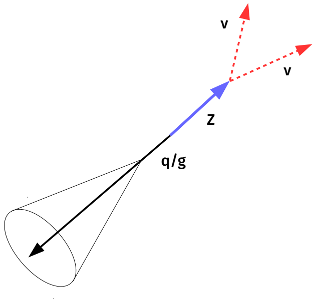
$$p_{\mathrm{T}}^{\mathrm{miss}} = p_{\mathrm{T}}^{Z} = p_{\mathrm{T}}^{\ell\ell}$$

$Z \to \ell\ell$



Background estimation

$Z \to \nu\nu$



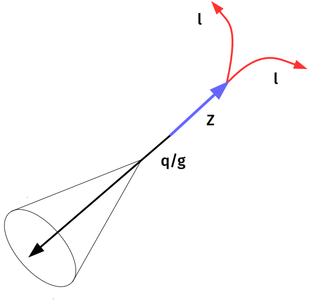
Ideally:

$$p_T^{miss} = p_T^Z = p_T^{\ell\ell}$$

In practice:

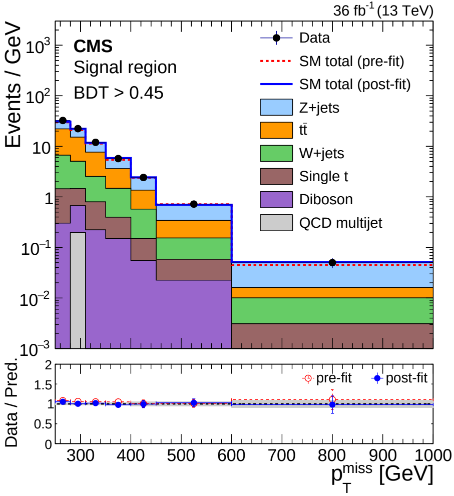
$$p_T^{miss} \approx p_T^Z \approx p_T^{miss}(\text{no } \ell)$$

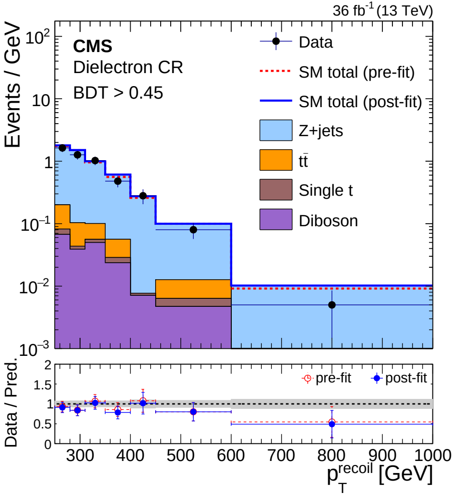
$Z \to \ell\ell$



Hadronic recoil $\equiv$ momentum imbalance if we pretend $\ell^{\pm}$ are invisible.
Only syst. uncertainty on this extrapolation comes from ℓ ID ($\sim$ 1–3%)

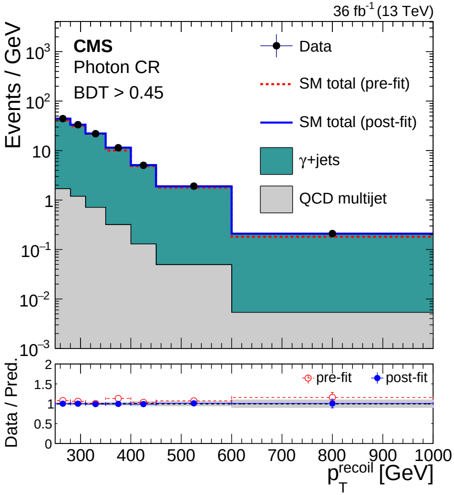
$\mathcal{B}(Z \to \nu\nu) > \mathcal{B}(Z \to \ell\ell)$

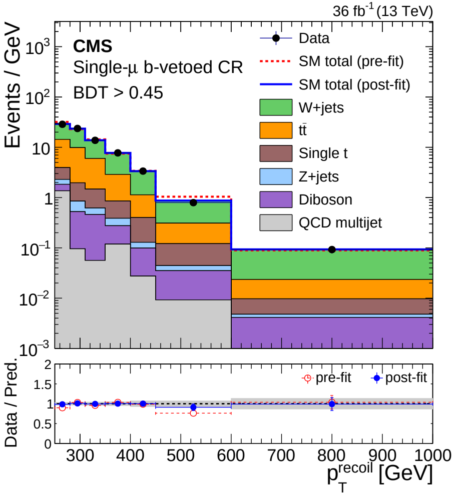




$\sigma(\gamma) \gg \sigma(W) > \sigma(Z)$

Use production of γ and W to estimate production of Z





Fitting ratios

- Free parameters in fit are:

$$\boldsymbol{\mu}_i^{Z\to\nu\nu} \text{ - number of } Z \text{ events in SR}$$

$$\boldsymbol{R}_i^X(\boldsymbol{\theta}) \text{ - ratio of events in SR and CR } X$$

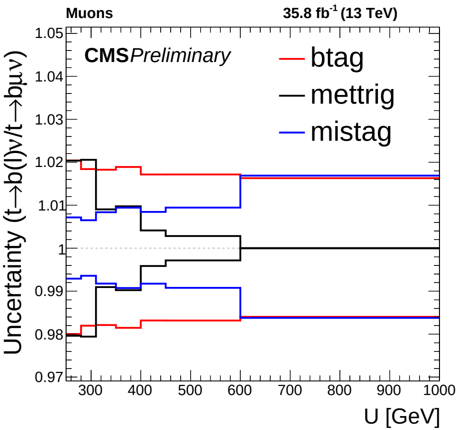
- Each extrapolation $\Rightarrow$ an additional $\boldsymbol{R}$

$$\mathcal{L}(\boldsymbol{\mu}^{Z\to\nu\nu};\boldsymbol{\theta}) = \prod_{i\in\text{bins}} \text{Poisson}\left(d_i^{\text{signal}}\middle| B_i^{\text{signal}}(\boldsymbol{\theta}) + (1+f_i(\boldsymbol{\theta}))\boldsymbol{\mu}_i^{Z\to\nu\nu} + \mu S_i(\boldsymbol{\theta})\right)$$
$$\times \prod_{i\in\text{bins}} \text{Poisson}\left(d_i^{\ell\ell}\middle| B_i^{\ell\ell}(\boldsymbol{\theta}) + \frac{\boldsymbol{\mu}_i^{Z\to\nu\nu}}{\boldsymbol{R}_i^{\ell\ell}(\boldsymbol{\theta})}\right)$$
$$\times \cdots$$

- Physics challenge boils down to predicting and assigning uncertainty on $\boldsymbol{R}$

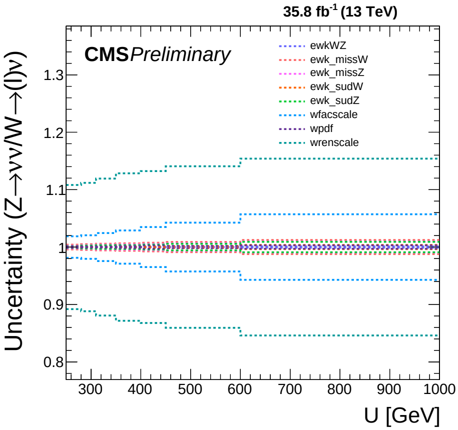
Extrapolation uncertainties

$t\bar{t} \to b\mu\nu$+jets



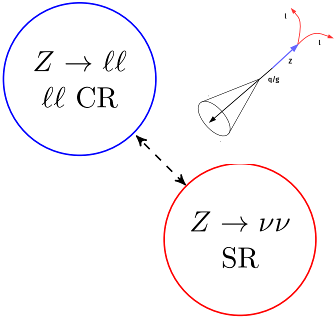
Small experimental uncertainties

$W \to \ell\nu$

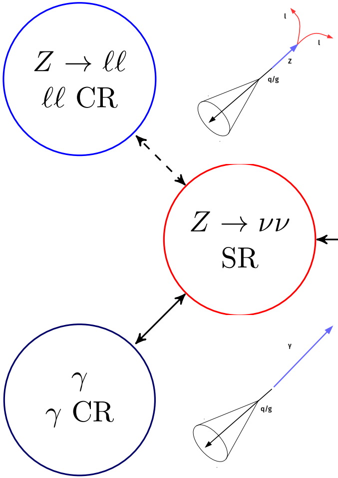


Large theoretical uncertainties

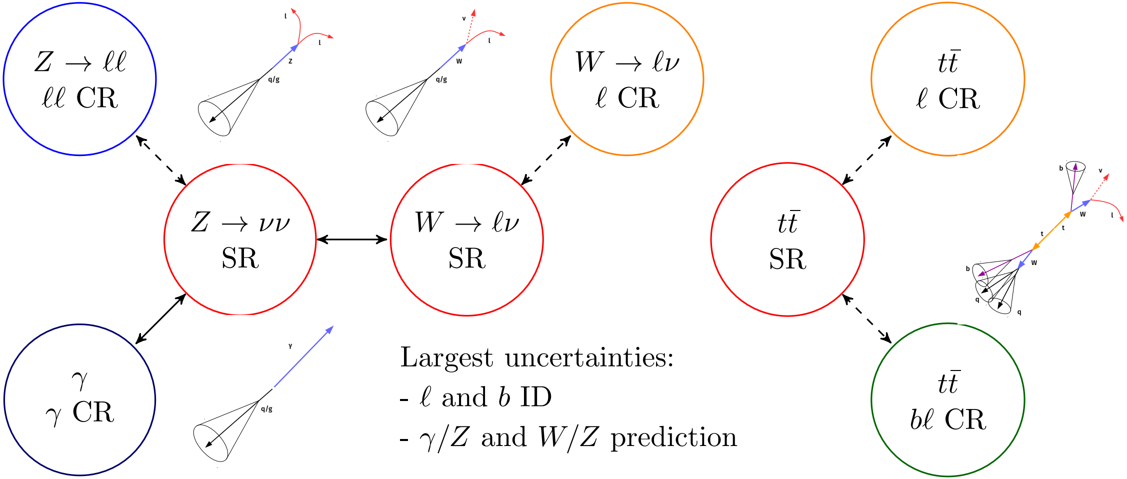
Background estimation summary



Background estimation summary



Background estimation summary

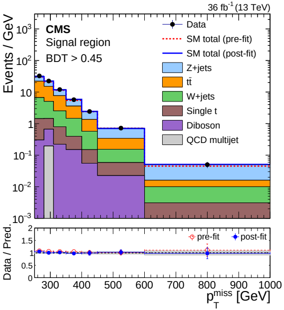


Largest uncertainties:

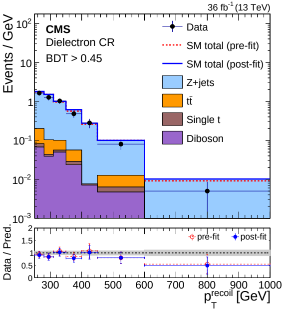
- ℓ and b ID
- γ/Z and W/Z prediction

Unblinding the data

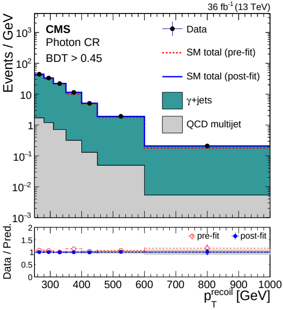
Signal



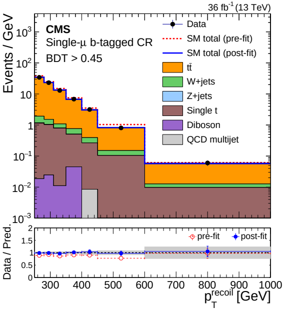
$Z \to ee$



γ



$t\bar{t} \to \mu\nu$+jets



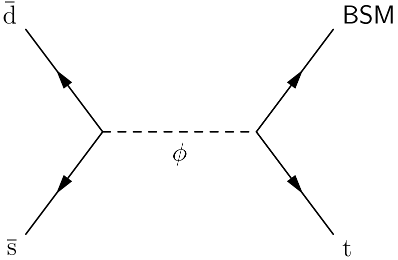
- Too many regions to show all here
- SM processes are able to describe data quite well in all regions, including the SRs
- No observation of an excess

No observation? Set limits

Benchmark models probe different mono-top kinematics.

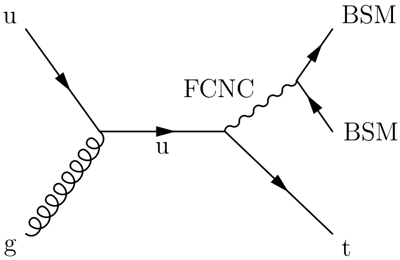
Resonant scalar

- p_{T} of top quark increases with m_ϕ
- Therefore, efficiency of signal selection improves at high m_ϕ



FCNC

- Falling $p_{\mathrm{T}}^{\mathrm{miss}}$ spectra ⇒ worse signal eff.
- Interesting parameters to constrain are m_V and couplings g_χ, g_q

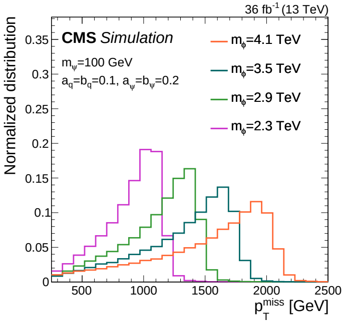


No observation? Set limits

Benchmark models probe different mono-top kinematics.

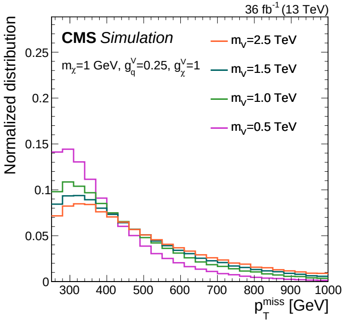
Resonant scalar

- p_T of top quark increases with m_ϕ
- Therefore, efficiency of signal selection improves at high m_ϕ



FCNC

- Falling p_T^{miss} spectra ⇒ worse signal eff.
- Interesting parameters to constrain are m_V and couplings g_χ, g_q

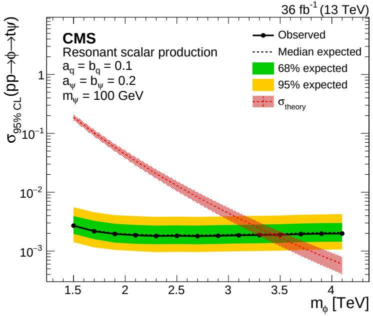


No observation? Set limits

Benchmark models probe different mono-top kinematics.

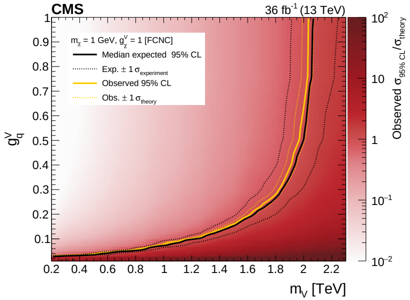
Resonant scalar

- p_T of top quark increases with m_ϕ
- Therefore, efficiency of signal selection improves at high m_ϕ



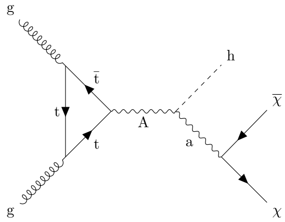
FCNC

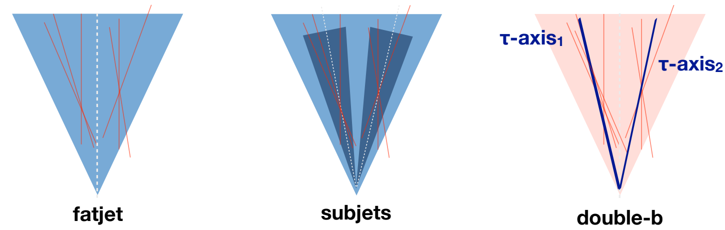
- Falling p_T^{miss} spectra $\Rightarrow$ worse signal eff.
- Interesting parameters to constrain are m_V and couplings g_χ, g_q

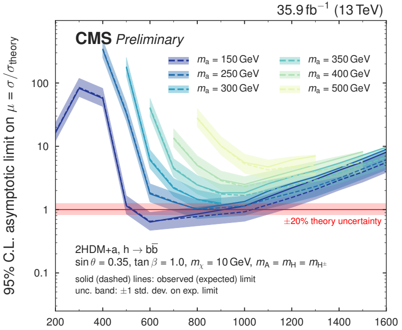


Another DM+substructure search: mono-Higgs(bb)

- Backgrounds and estimation technique very similar to mono-top
- Sensitive to extended Higgs sectors (2HDM+a, baryonic Z',...)
- Replace 3-prong/1-b with 2-prong/2-b large-radius jet







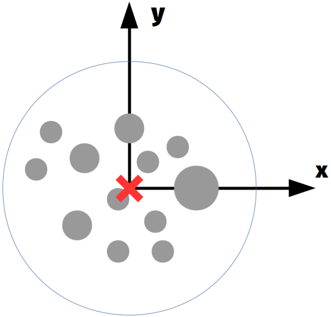
Detour: ML for substructure

- Top-tagging using QCD-motivated observables works very well
- Are we reaching a “maximum” performance threshold?
- One approach is to brute-force the problem using deep learning
- Factorize the question: physics effects vs. detector effects
- Following studies are done using hadron-level simulation
  - Madgraph5 at LO for hard scattering
  - Pythia8 for hadronization
  - No detector simulation
- Training is done on a desktop computer
  - NVIDIA GTX 1080 GPU
  - Keras[1] with tensorflow[2] backend

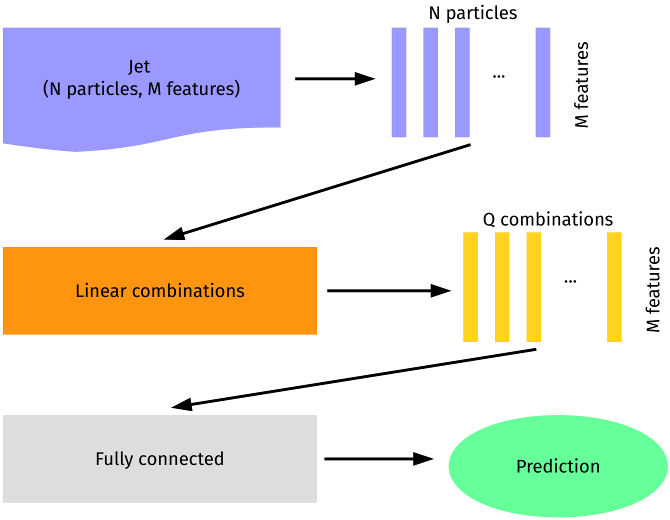
[1] `https://github.com/keras-team/keras`
[2] `https://github.com/tensorflow/tensorflow`

Observables

- For each particle in the jet, 7 features:
  - p^μ
  - ΔR(particle,jet)
  - Soft drop survival
  - Particle type ($e^\pm$, $\mu^\pm$, γ, charged hadron$^\pm$, neutral hadron)
- Rotate the jet so:
  - Jet axis coincides with z-axis
  - Hardest particle away from jet axis lies in x-z plane

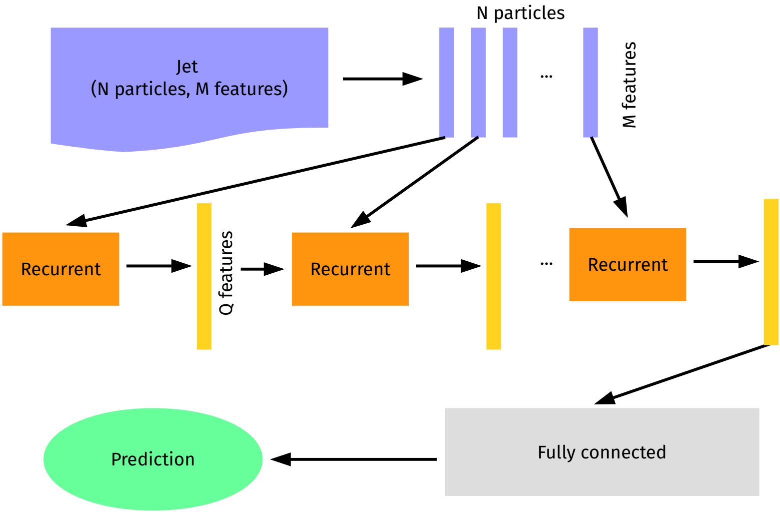


Network architectures



- Fully connected: brute force approach

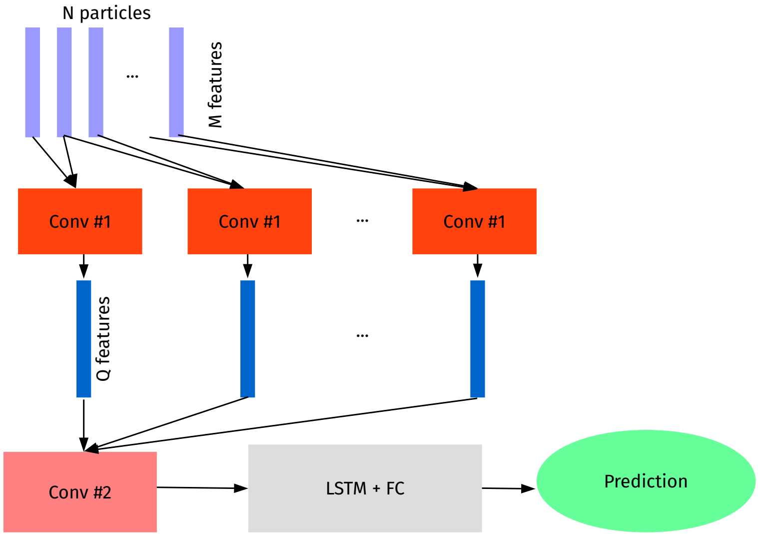
Network architectures



- Fully connected: brute force approach

- Recurrent NN: read the jet as a “sentence”, where a particle is a “word”

Network architectures



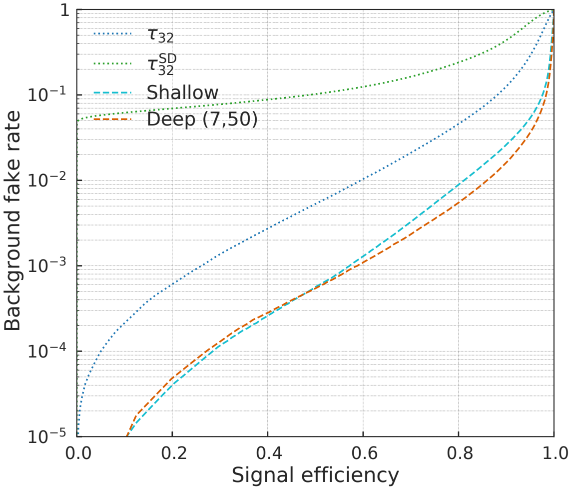
- Fully connected: brute force approach

- Recurrent NN: read the jet as a "sentence", where a particle is a "word"

- 1D convolutions: allows some invariance to incorrect ordering

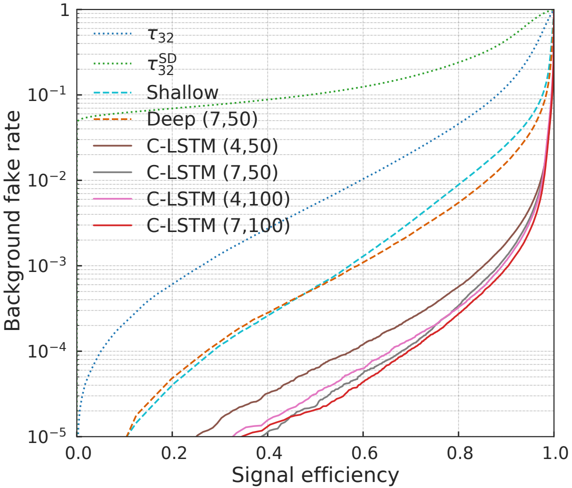
DNN performance

- Compare fully-connected network to "shallow" network using ECFs
- $\mathcal{O}(10^6)$ parameters
- Positive: performant classifier without thinking about physics
- Negative: that's it?



DNN performance

- Compare fully-connected network to "shallow" network using ECFs
- $\mathcal{O}(10^6)$ parameters
- Positive: performant classifier without thinking about physics
- Negative: that's it?
- Dramatic improvement from giving structure to the network
- Adding more information ($4 \to 7$) or more particles ($50 \to 100$) helps
- C-LSTMs have $\mathcal{O}(10^5)$ parameters



Next steps/WIP

- Quantifying how realistic this improvement is
  - What are we learning that QCD observables don't capture?
  - Is it IRC unsafe things?
  - How does detector smearing hurt?
    - Hint: it's painful

Next steps/WIP

- Quantifying how realistic this improvement is
  - What are we learning that QCD observables don't capture?
  - Is it IRC unsafe things?
  - How does detector smearing hurt?
    - Hint: it's painful
- Removing correlation with various nuisances
  - Kinematics of jet: mass, p_{T}
  - Pile-up
  - QCD uncertainties
    - Again, IRC unsafety plays a role

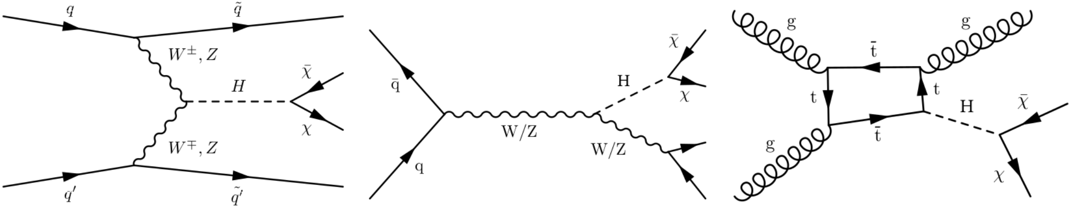
Next steps/WIP

- Quantifying how realistic this improvement is
  - What are we learning that QCD observables don't capture?
  - Is it IRC unsafe things?
  - How does detector smearing hurt?
    - Hint: it's painful
- Removing correlation with various nuisances
  - Kinematics of jet: mass, p_{T}
  - Pile-up
  - QCD uncertainties
    - Again, IRC unsafety plays a role
- There is a lot of promise in these approaches!

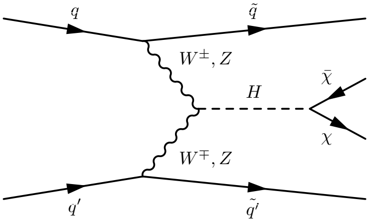
VBF $H \to$ invisible

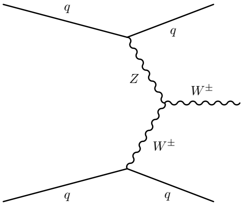
Invisible Higgs

- DM fermion could be given mass through Higgs mechanism
- If $2m_\chi < m_H$, should observe $H \to \chi\bar{\chi}$
- Production mode $\Rightarrow$ mono-**X** channels
  - $gg \to H$ + ISR $\Rightarrow$ mono-jet
  - $VH \Rightarrow$ mono-$V(qq')$ and mono-$Z(\ell\ell)$
  - VBF $\Rightarrow$ VBF+H $\to$inv



VBF production of bosons



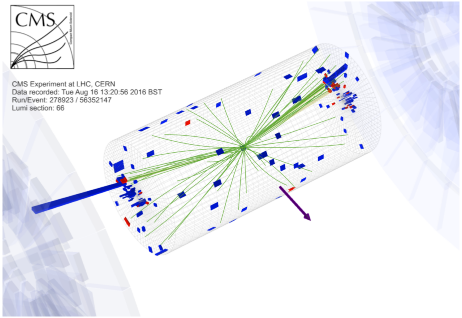


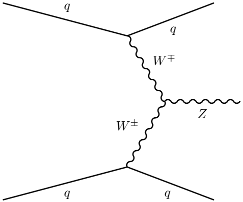
Characterized by:

- Two forward jets
- Large p_{T}^{H}

Can replace H with Z or W

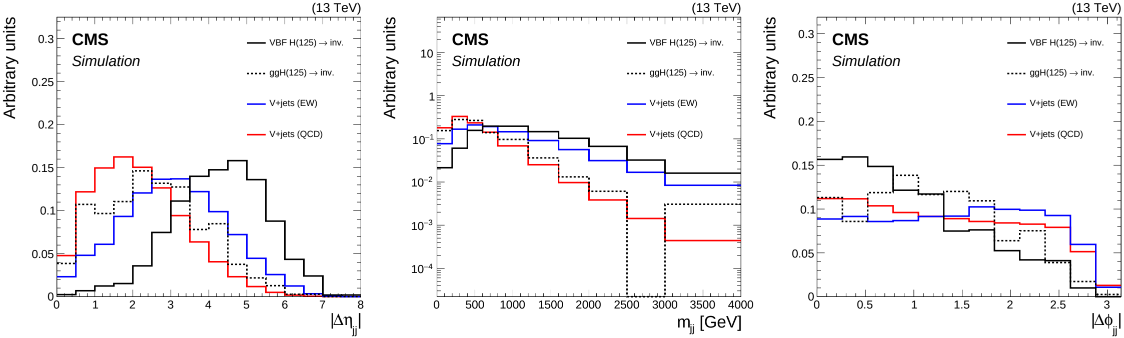
- Irreducible background





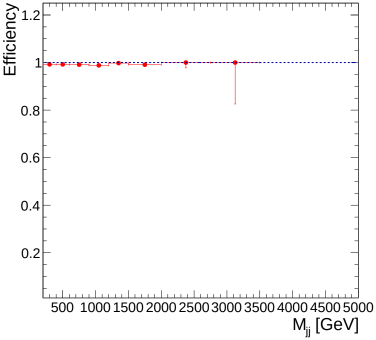
Forward jets are important

- As with mono-top and mono-Higgs, we use the jets to mitigate backgrounds
- In this case, the jets can be resolved distinctly

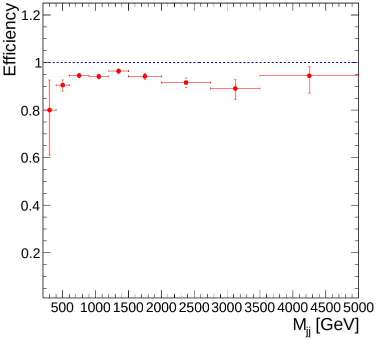


Forward jets are challenging

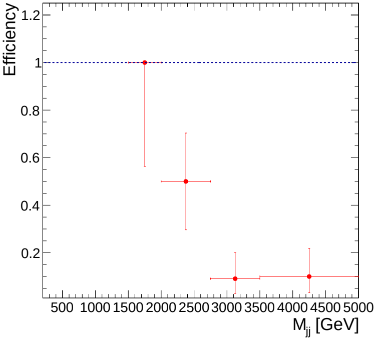
- "Forward" typically refers to jets outside of the tracker's acceptance
- Rely entirely on calorimeters
- Energy resolution and trigger efficiency degrade in this region



Two central jets



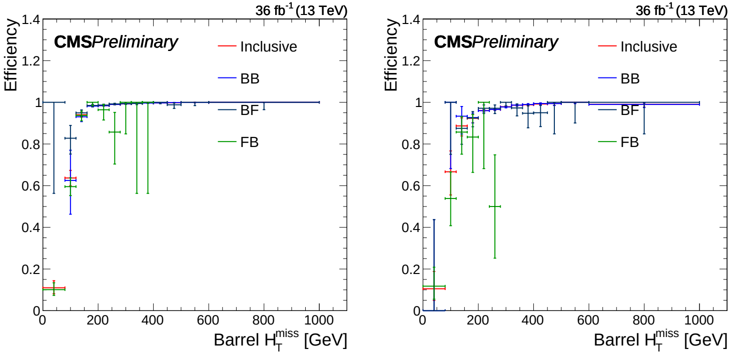
One central, one forward



Two forward jets

Forward jets are challenging

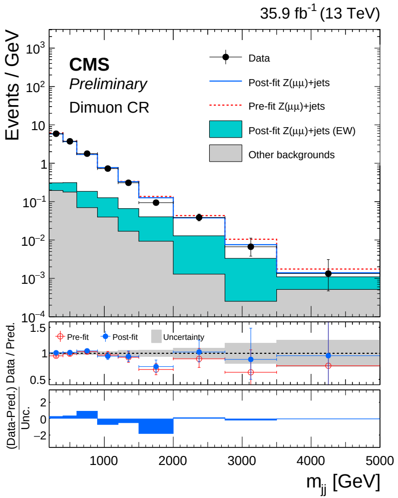
- "Forward" typically refers to jets outside of the tracker's acceptance
- Rely entirely on calorimeters
- Energy resolution and trigger efficiency degrade in this region
- Characterize events using quality within tracker acceptance



V+jet estimation

- Need to precisely estimate EW+QCD components of V+jets
- Prediction is made to NLO in QCD and EW
- As with mono-top, correlate Z and W production

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



Validation of ***R***

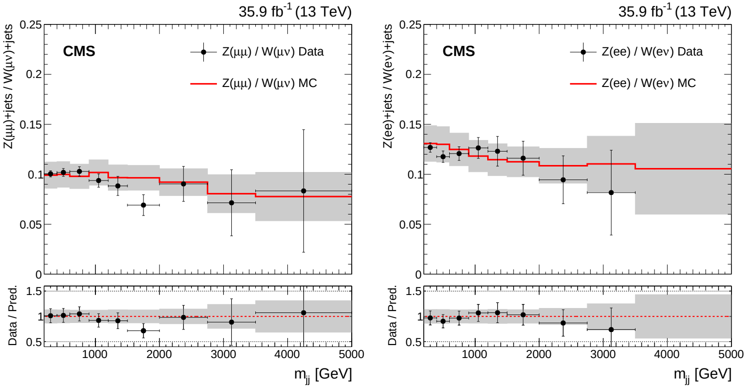
- How do we know our prediction and uncertainties make sense?
- Cannot check in data whether we correctly predict $R = \frac{Z\to\nu\nu}{W\to\ell\nu}$
- However, we can check:

$$\frac{Z \to \ell\ell}{W \to \ell\nu} \approx \frac{Z \to \nu\nu}{W \to \ell\nu}$$

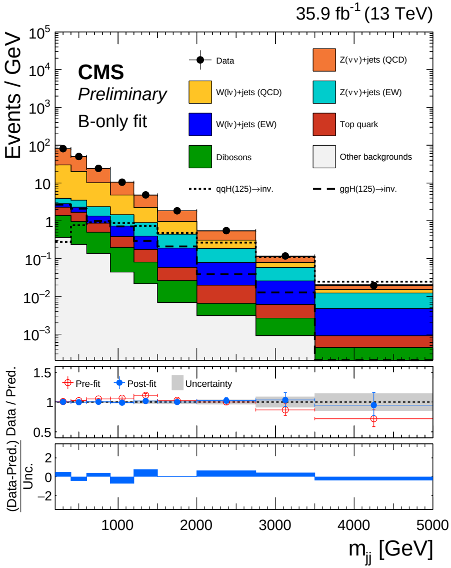
Validation of R

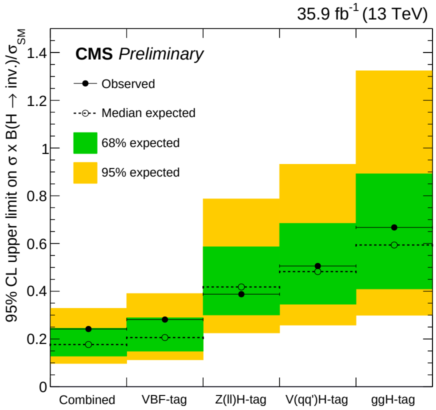
- How do we know our prediction and uncertainties make sense?
- Cannot check in data whether we correctly predict $R = \frac{Z\to\nu\nu}{W\to\ell\nu}$
- However, we can check:

$$\frac{Z \to \ell\ell}{W \to \ell\nu} \approx \frac{Z \to \nu\nu}{W \to \ell\nu}$$



Results

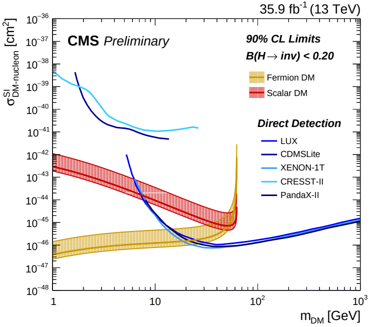




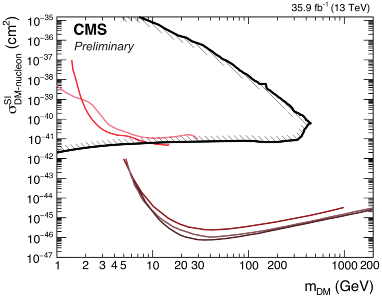
- Combine with other production modes to directly constrain $\mathcal{B}(H \to \text{inv})$
- VBF drives the combination

Comparison of LHC and direct detection constraints

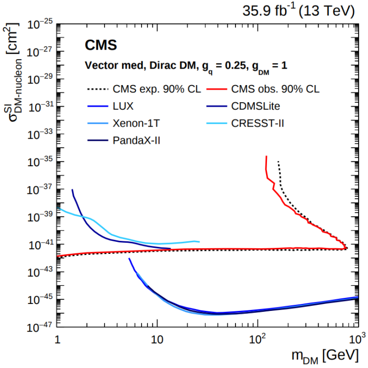
$H \to$ inv



Mono-Higgs



Mono-jet

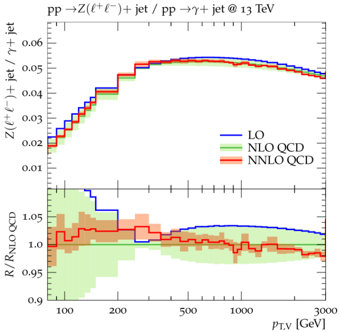


- LHC constraints strongest at low DM mass
- Constraints depend strongly on choice of model

Theoretical prediction of V+jets

Uncertainty	Size	Impact on sensitivity
W/Z EW	15%	50%
W/Z QCD	15%	25%
Trigger	2%	20%
Lepton ID	2-3%	15%

- Theoretical uncertainties dominate VBF (and most mono-X searches)
- Inclusive predictions were dramatically improved in 2016
  - 15% $\rightarrow$ 5%
  - [arXiv:1705.04664]
- Strong relationship with theory community on this effort
- Expect VBF predictions at similar level by Run 3



Conclusions

Mono-top

- Jet substructure is critical
  - Resolved case not feasible in Run 2
  - ECF-based tagger came out of interactions with theory community
- Strong constraints on flavor-changing DM models
  - Search designed to be model independent $\Rightarrow$ further re-interpretation
- ECFs and other substructure tools not limited to mono-top
  - Mono-Higgs
  - Visible mediator searches
  - SM, Higgs, etc.

Conclusions

Mono-top

- Jet substructure is critical
  - Resolved case not feasible in Run 2
  - ECF-based tagger came out of interactions with theory community
- Strong constraints on flavor-changing DM models
  - Search designed to be model independent $\Rightarrow$ further re-interpretation
- ECFs and other substructure tools not limited to mono-top
  - Mono-Higgs
  - Visible mediator searches
  - SM, Higgs, etc.

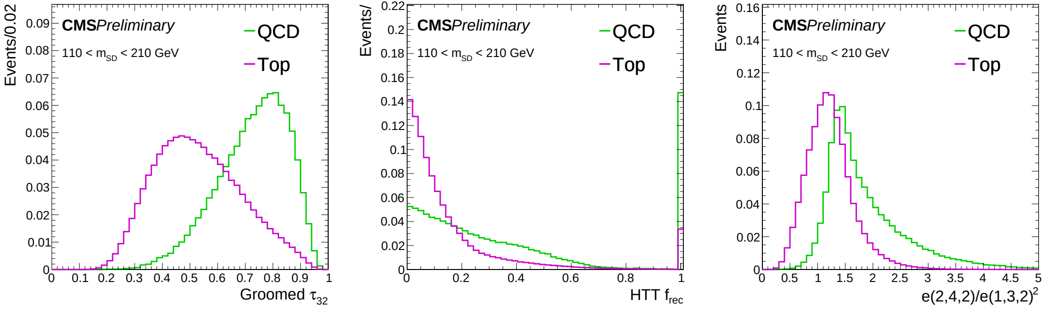
VBF $H \rightarrow$ invisible

- Very different set of jet challenges
  - Difficulty in energy measurement, triggering
  - Simpler reconstruction, but huge combinatoric background
- Key here is accurate measurement of SM backgrounds
  - EW and QCD components (at LO)
- Reducing theoretical uncertainties
  - Better prediction of W and Z spectra
  - Understanding correlation between W and Z

BACKUP

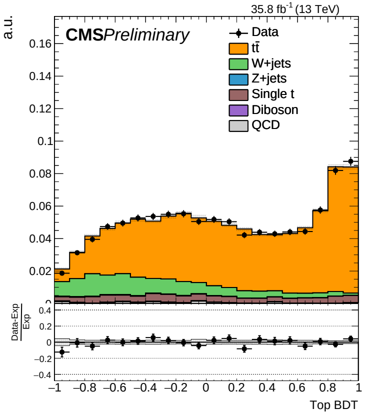
Some substructure observables

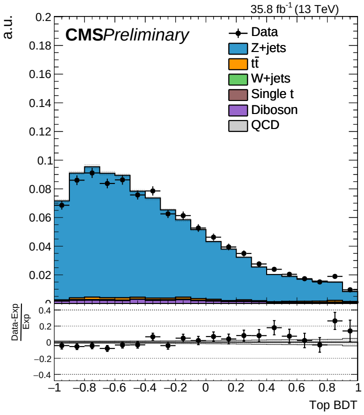
- N-subjettiness [Thaler *et al*, arXiv:1011.2268]
  - $\tau_{\mathbf{N}}$: compatibility of jet with **N**-axis hypothesis
- HEPTopTagger [Anders *et al*, arXiv:1312.1504]
  - Reconstruct W and t decay products inside jet
- Energy correlation functions [Moult *et al*, arXiv:1609.07473]
  - $e(\alpha, \mathbf{N}, a)$ sensitive to **N**-point correlations in the jet

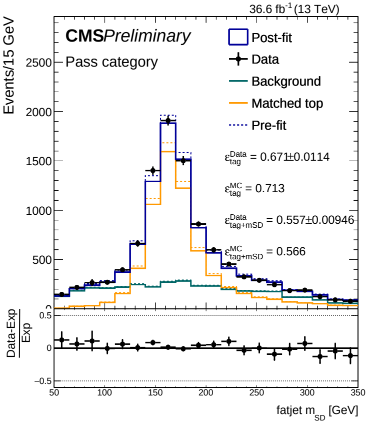


Comparison to data

Substructure relies on physics that may not be well-simulated by hadronization models.
Comparison to data shows that the BDT classifier is well-described.



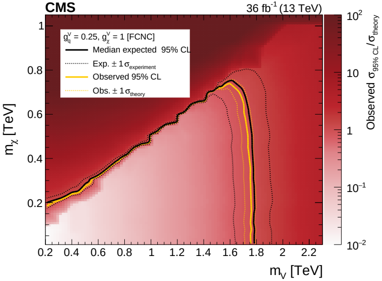




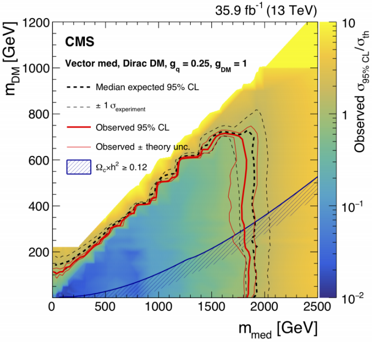
How does mono-top compare?

- Sensitivity of mono-top and mono-jet similar (with same assumptions on g_χ, g_q, m_V)
- If FCNC is embedded in DM model, sensitivity similar to mono-jet
  - No DD limits for 3rd gen FCNC because $\sigma_{\mathrm{DM,N}}$ re-interpretation is tricky

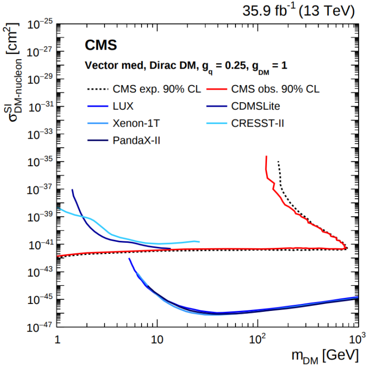
Mono-top spin-1



Mono-jet spin-1



DM exclusion



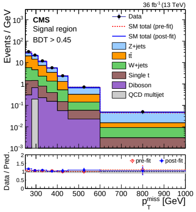
Re-interpretation using simplified likelihoods

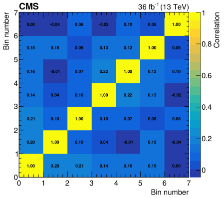
- Searches are designed in a semi-model-independent way
- Need a way for new models to be constrained using these results
- We cannot release all of our data and expect theory community to redo analysis
- Even a complete likelihood is tricky - 100s of parameters and constraints

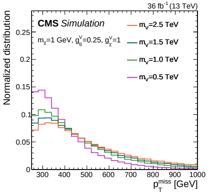
Re-interpretation using simplified likelihoods

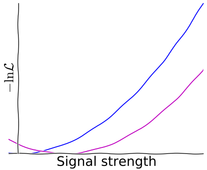
- Searches are designed in a semi-model-independent way
- Need a way for new models to be constrained using these results
- We cannot release all of our data and expect theory community to redo analysis
- Even a complete likelihood is tricky - 100s of parameters and constraints

Solution: simplified likelihood









p_T^{miss} distribution (CMS) + p_T^{miss} correlations (CMS) + BSM hypothesis (anyone) ≈ Likelihood (anyone)

Summary and outlook

- Run 2 has seen a significant improvement in mono-**X** searches at CMS
  - Increased luminosity and cross-sections
  - New techniques to estimate backgrounds and identify **X**

Summary and outlook

- Run 2 has seen a significant improvement in mono-**X** searches at CMS
  - Increased luminosity and cross-sections
  - New techniques to estimate backgrounds and identify **X**
- What are the key challenges for the remainder of Run 2 and beyond?

Summary and outlook

- Run 2 has seen a significant improvement in mono-**X** searches at CMS
  - Increased luminosity and cross-sections
  - New techniques to estimate backgrounds and identify **X**
- What are the key challenges for the remainder of Run 2 and beyond?
- Triggering becomes harder as instantaneous luminosity and pile-up increase
  - Mono-**X** signature has relatively few trigger handles
  - $p_{\mathrm{T}}^{\mathrm{miss}}$ depends on online jet resolution

Summary and outlook

- Run 2 has seen a significant improvement in mono-**X** searches at CMS
  - Increased luminosity and cross-sections
  - New techniques to estimate backgrounds and identify **X**
- What are the key challenges for the remainder of Run 2 and beyond?
- Triggering becomes harder as instantaneous luminosity and pile-up increase
  - Mono-**X** signature has relatively few trigger handles
  - $p_{\mathrm{T}}^{\mathrm{miss}}$ depends on online jet resolution
- Many searches rely on accurate theoretical predictions of backgrounds
  - $\mathcal{O}(1\%)$ uncertainties on V/Z ratios key for mono-**jet** sensitivity

Summary and outlook

- Run 2 has seen a significant improvement in mono-**X** searches at CMS
  - Increased luminosity and cross-sections
  - New techniques to estimate backgrounds and identify **X**
- What are the key challenges for the remainder of Run 2 and beyond?
- Triggering becomes harder as instantaneous luminosity and pile-up increase
  - Mono-**X** signature has relatively few trigger handles
  - $p_{\mathrm{T}}^{\mathrm{miss}}$ depends on online jet resolution
- Many searches rely on accurate theoretical predictions of backgrounds
  - $\mathcal{O}(1\%)$ uncertainties on V/Z ratios key for mono-**jet** sensitivity
  - Uncertainties are larger for other V+jet topologies
    - Among limiting factors for **VBF**+$p_{\mathrm{T}}^{\mathrm{miss}}$, mono-**top**, mono-**Higgs**

Summary and outlook

- Run 2 has seen a significant improvement in mono-**X** searches at CMS
  - Increased luminosity and cross-sections
  - New techniques to estimate backgrounds and identify **X**
- What are the key challenges for the remainder of Run 2 and beyond?
- Triggering becomes harder as instantaneous luminosity and pile-up increase
  - Mono-**X** signature has relatively few trigger handles
  - $p_{\mathrm{T}}^{\mathrm{miss}}$ depends on online jet resolution
- Many searches rely on accurate theoretical predictions of backgrounds
  - $\mathcal{O}(1\%)$ uncertainties on V/Z ratios key for mono-**jet** sensitivity
  - Uncertainties are larger for other V+jet topologies
    - Among limiting factors for **VBF**+$p_{\mathrm{T}}^{\mathrm{miss}}$, mono-**top**, mono-**Higgs**
  - VV ratios $\Rightarrow$ mono-$\boldsymbol{Z(\ell\ell)}$ and mono-γ
  - $t\bar{t}\,V$ prediction $\Rightarrow$ dileptonic $\boldsymbol{t\bar{t}}$ +DM

Generalized ECFs

- Extension of original ECFs to allow for different angular orders:

$$e(o,N,\beta) \equiv {}_oe_N^\beta = \sum_{i_1<i_2<\dots<i_N\in J}\left[\prod_{1\leq k\leq j} z_{i_k}\right] \times \min\left\{\prod_{k,l\in \text{pairs}\{i_1,\dots,i_N\}}^{o} \Delta R_{kl}^\beta\right\}$$

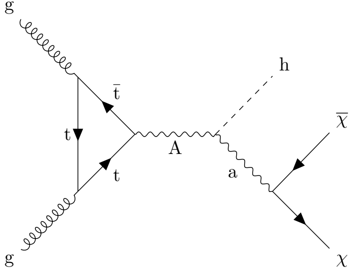
- e.g.

$$ {}_2e_3^1 = \sum_{a<b<c\in J} z_a z_b z_c \times \min\{\Delta R_{ab}\Delta R_{ac}, \Delta R_{ab}\Delta R_{bc}, \Delta R_{bc}\Delta R_{ac}\}$$

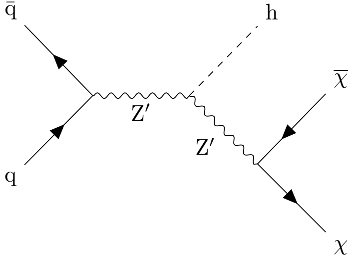
- Summary of parameters:
  - N = order of the correlation function. An N-pronged jet should have $e_N \gg e_M$, for $N < M$
  - o = order of the angular factor.
  - β = angular power

Mono-Higgs

2HDM+a



Baryonic Z'

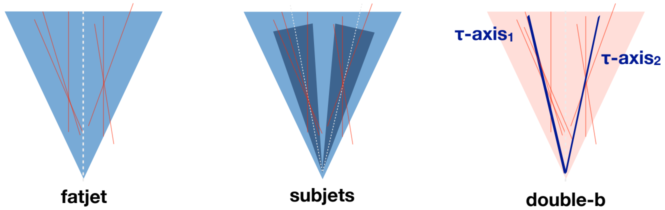


- 5 additional Higgs bosons, including heavy (A) and light (a) pseudoscalars

- Quantize baryon number with gauge field Z'
- "SM" h mixes with baryonic h_B, providing effective coupling to Z'

Identifying $H \to b\bar{b}$

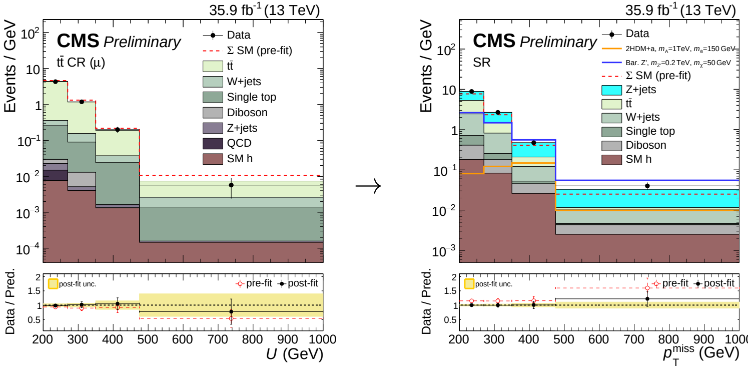
- As with mono-top, we focus on highly-boosted decays
- Two-prong substructure tagging is done using ECFs
- Identifying flavor content of $H \to b\bar{b}$ is more important
  - Two B mesons $\Rightarrow$ difficult to fake signature
- Subjet tagging becomes less efficient at high p_{T}



- Use "double-b" tagger to see if entire jet is consistent with 2 bs [CMS-BTV-15-002]

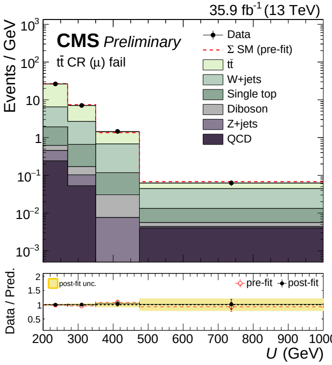
Background estimation

- As with mono-top, use visible $Z/W/t\bar{t}$ processes to constrain invisible analogs

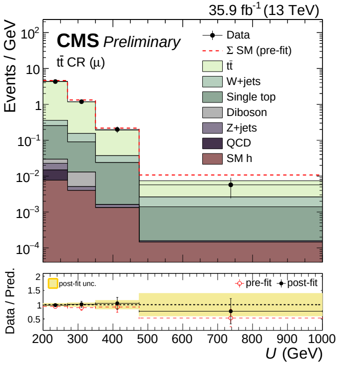


Background estimation

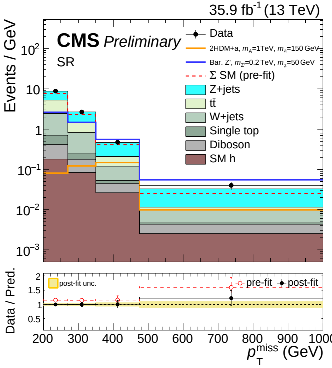
- As with mono-top, use visible $Z/W/t\bar{t}$ processes to constrain invisible analogs
- Control data includes events that both pass and fail the double-b selection
- Use this ratio to correct the efficiency of backgrounds in the signal region



/



→

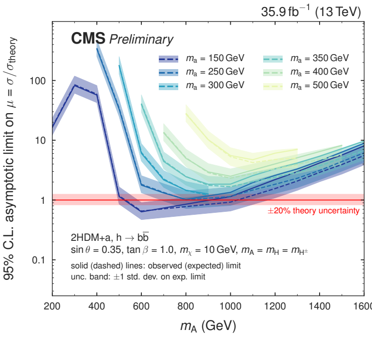


Constraints on 2HDM+a

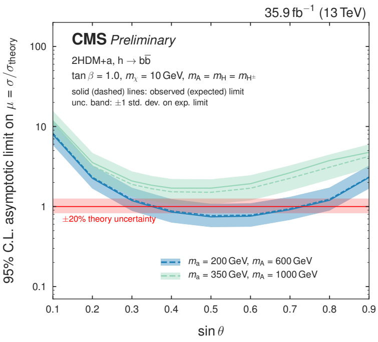
- 2HDM+a is a rich model ⇒ many free parameters
- Some couplings constrained by unitarity and perturbativity
- Assume that heavy Higgses all have same mass m_A

Constraints on 2HDM+a

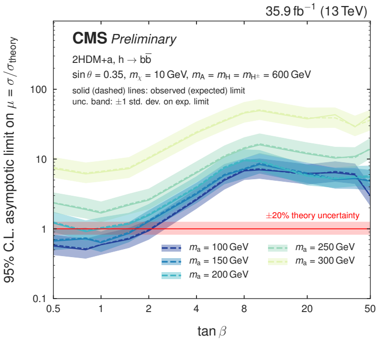
- 2HDM+a is a rich model ⇒ many free parameters
- Some couplings constrained by unitarity and perturbativity
- Assume that heavy Higgses all have same mass m_A



Heavy Higgs mass m_A



a-A mixing angle θ



$\beta = \langle h \rangle / \langle H \rangle$

Constraints on baryonic Z'

- Only free parameters are masses $m_{Z'}$, m_χ and couplings g_q, g_χ
- Can re-cast constraints as a function of $\sigma_{\mathrm{DM-N}}$ for comparison to direct detection