# Presenting LHC data in a way that is useful for modelling and reinterpretation

Andrew Pilkington – University of Manchester

*Presented at 'Dark Matter at the Dawn of Discovery', Heidelberg, 9th April 2018*

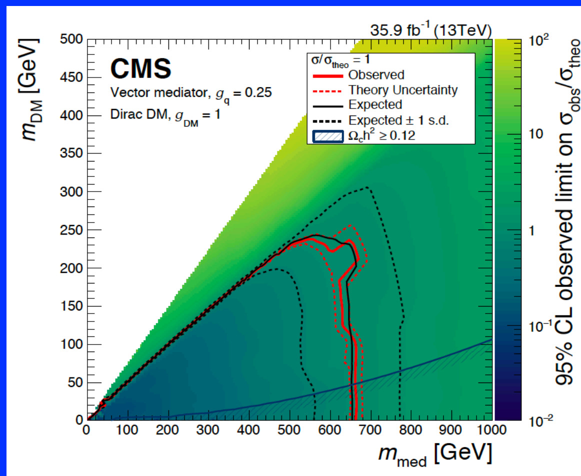
**Outline**

1) The motivation for a new approach
2) Proof-of-principle for dark matter searches
3) The ease of reinterpretations
4) An example from the Higgs sector

# Model-specific searches for dark matter at the LHC

CMS Collaboration, arXiv:1711.00431

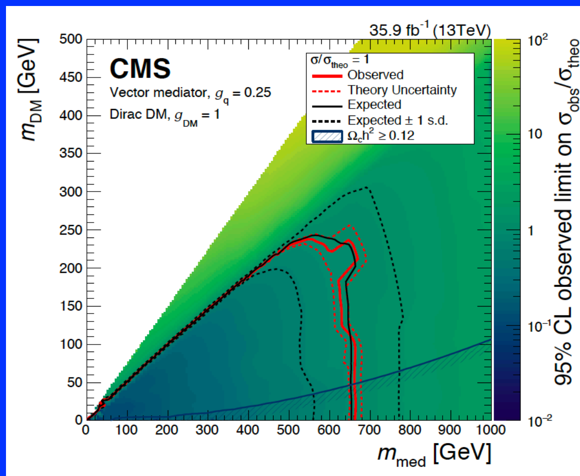
**Limit on specific theoretical model**



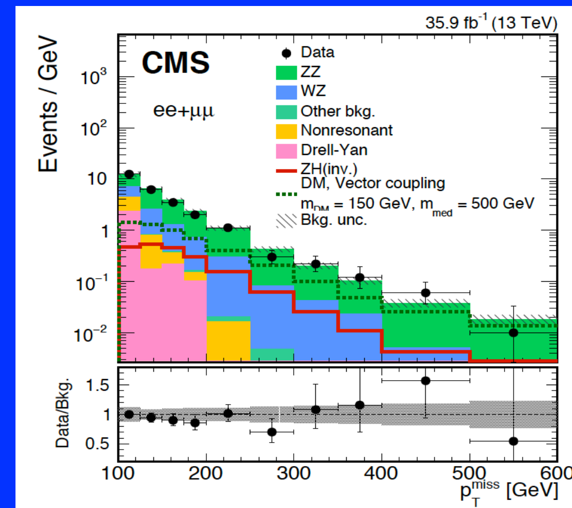
# Model-specific searches for dark matter at the LHC

CMS Collaboration, arXiv:1711.00431

## Limit on specific theoretical model



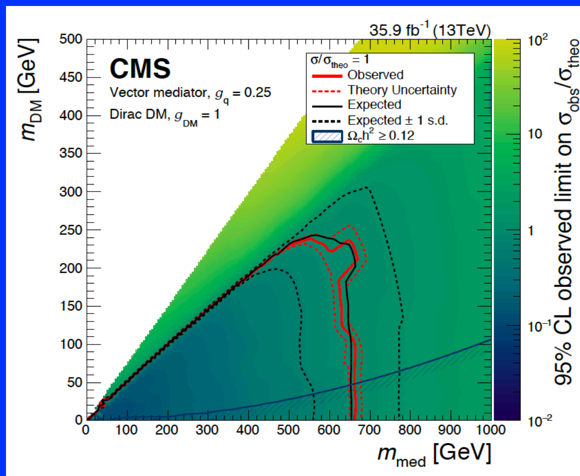

## Signal-enhanced control region



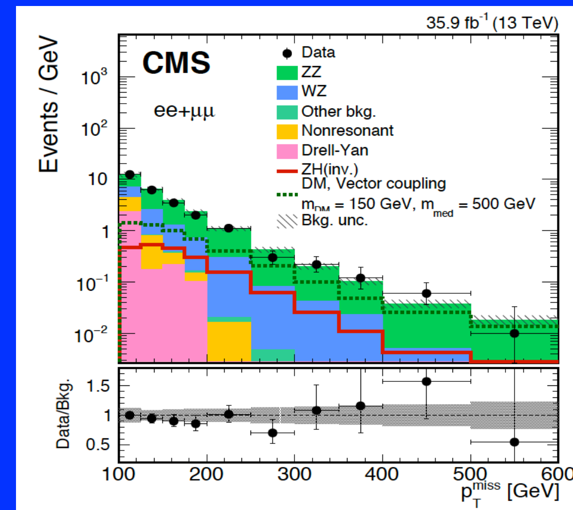
# Model-specific searches for dark matter at the LHC

CMS Collaboration, arXiv:1711.00431

## Limit on specific theoretical model



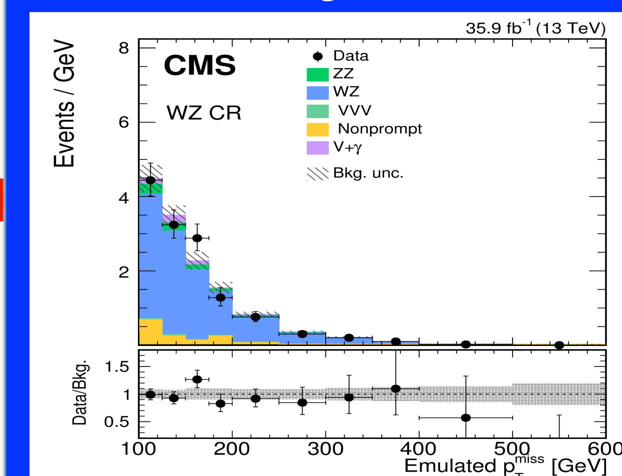

## Signal-enhanced control region




## Systematics

| Source of uncertainty | Effect (%) Signal | ZZ | WZ | NRB | DY | Impact on the exp. limit (%) |
|---|---|---|---|---|---|---|
| * VV EW corrections | — | 10 | −4 | — | — | 14 (12) |
| * Renorm./fact. scales, VV | — | 9 | 4 | — | — | |
| * Renorm./fact. scales, ZH | 3.5 | — | — | — | — | |
| * Renorm./fact. scales, DM | 5 | — | — | — | — | |
| * PDF, WZ background | — | — | 1.5 | — | — | 2 (1) |
| * PDF, ZZ background | — | 1.5 | — | — | — | |
| * PDF, Higgs boson signal | 1.5 | — | — | — | — | |
| * PDF, DM signal | 1–2 | — | — | — | — | |
| NRB extrapolation to the SR | — | — | — | 20 | — | <1 |
| DY extrapolation to the SR | — | — | — | — | 100 | <1 |
| Lepton efficiency (WZ CR) | — | — | 3 | — | — | <1 |
| Nonprompt bkg. (WZ CR) | — | — | — | — | 30 | <1 |
| Integrated luminosity | | | 2.5 | | | <1 |
| * Electron efficiency | | | 1.5 | | | |
| * Muon efficiency | | | 1 | | | |
| * Electron energy scale | | | 1–2 | | | |
| * Muon energy scale | | | 1–2 | | | |
| * Jet energy scale | | | 1–3 (typically anticorrelated w/ yield) | | | 1 (<1) |
| * Jet energy resolution | | | 1 (typically anticorr.) | | | |
| * Unclustered energy ($p_T^{miss}$) | | | 1–4 (typically anticorr.), strong in DY | | | |
| * Pileup | | | 1 (typically anticorrelated) | | | |
| * b tagging eff. & mistag rate | | | 1 | | | |

## Data-driven background constraint



# Circumventing model-dependence?...additional output

- Obvious issues with model-specific limits:
  - results valid for narrow class of theories
  - reinterpretations difficult for models with different event topology
  - how to combine with other measurements? (i.e. different channels/experiments)

# Circumventing model-dependence?...additional output

- Obvious issues with <u>model-specific</u> limits:
  - results valid for narrow class of theories
  - reinterpretations difficult for models with different event topology
  - how to combine with other measurements? (i.e. different channels/experiments)

- To avoid these issues, collaborations publish information about the data:
  - observed and expected event yields in signal region
  - 'model-independent limits' on $\sigma \times A \times \varepsilon$
  - but still need to know $A \times \varepsilon$ to utilise the data

| Selection | $\langle\sigma\rangle_{\text{obs}}^{95}$ [fb] |
|---|---|
| IM1 | 531 |
| IM2 | 330 |
| IM3 | 188 |
| IM4 | 93 |
| IM5 | 43 |
| IM6 | 19 |
| IM7 | 7.7 |
| IM8 | 4.9 |
| IM9 | 2.2 |
| IM10 | 1.6 |

| | **Inclusive Signal Region** | |
|---|---|---|
| Region | Predicted | Observed |
| IM1 | $245900 \pm 5800$ | 255486 |
| IM2 | $138000 \pm 3400$ | 144283 |
| IM3 | $73000 \pm 1900$ | 76808 |
| IM4 | $39900 \pm 1000$ | 41523 |
| IM5 | $12720 \pm 340$ | 13680 |
| IM6 | $4680 \pm 160$ | 5097 |
| IM7 | $2017 \pm 90$ | 2122 |
| IM8 | $908 \pm 55$ | 980 |
| IM9 | $464 \pm 34$ | 468 |
| IM10 | $238 \pm 23$ | 245 |

ATLAS Collaboration, JHEP 01 (2018) 126

# Circumventing A × ε if outside of the collaboration

- Approach 1: use detector smearing functions
  - Example: DELPHES or BuckFast
  - Simple approximation of the detector response for each reconstructed object.
  - Requires validation for each analysis, i.e. reproducing event yields for BSM benchmarks (cumbersome).

- Approach 2: use recasting framework: https://arxiv.org/pdf/1010.2506.pdf
  - Archive of each data analysis and simulation framework from each collaboration
  - Cloud-based service to compare new physics model to data
  - Likely a large overhead in service development and service maintenance

# A different approach: detector-corrected observables

- Idea: construct dark-matter-sensitive observables that are free from detector inefficiency and resolution

ATLAS Collaboration, EPJ C77 (2017) 765

Number of background-subtracted events in MET+jets signal region

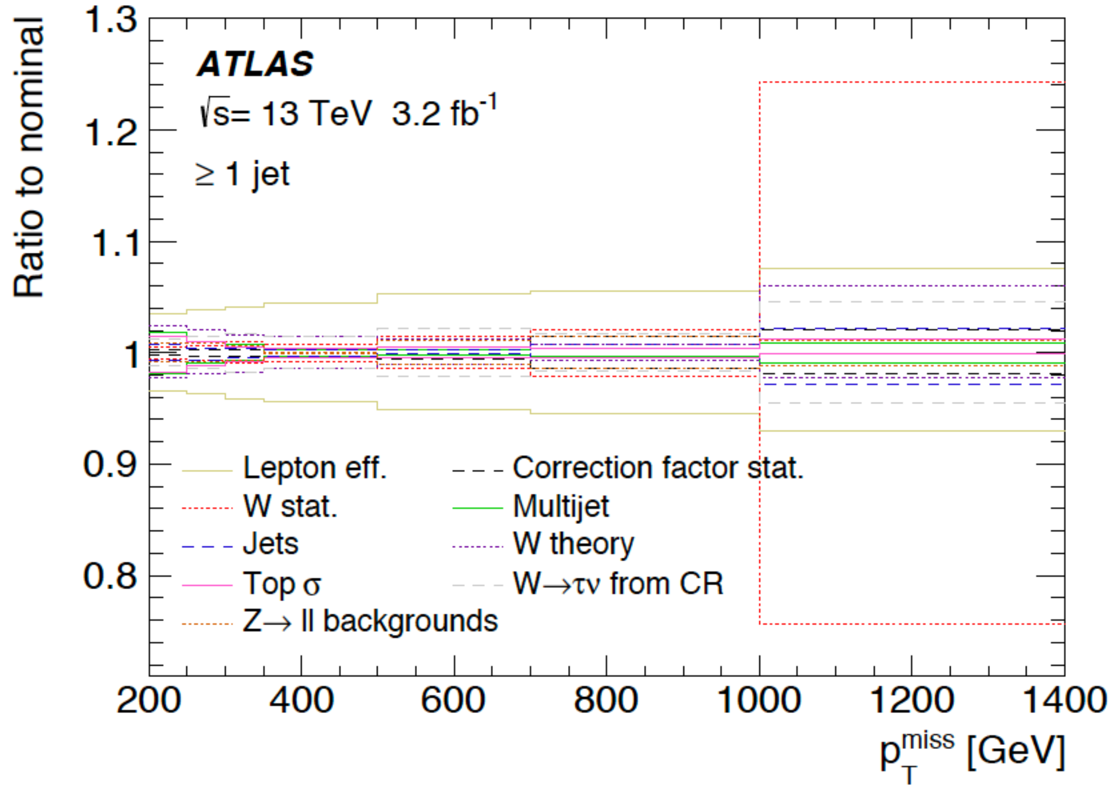
$$R_{\text{miss}} = \frac{\sigma(\not{p}_{\text{T}} + \text{jets})}{\sigma(Z \to \ell^{+}\ell^{-} + \text{jets})} = \frac{1}{C_Z} \frac{N(\not{p}_{\text{T}} + \text{jets})}{N(Z \to \ell^{+}\ell^{-} + \text{jets})}$$

Correction for detector inefficiency and resolution

Number of background-subtracted events in $l^{+}l^{-}$ + jets signal region

- In the SM, the numerator consists of only Z-bosons decaying to neutrinos.

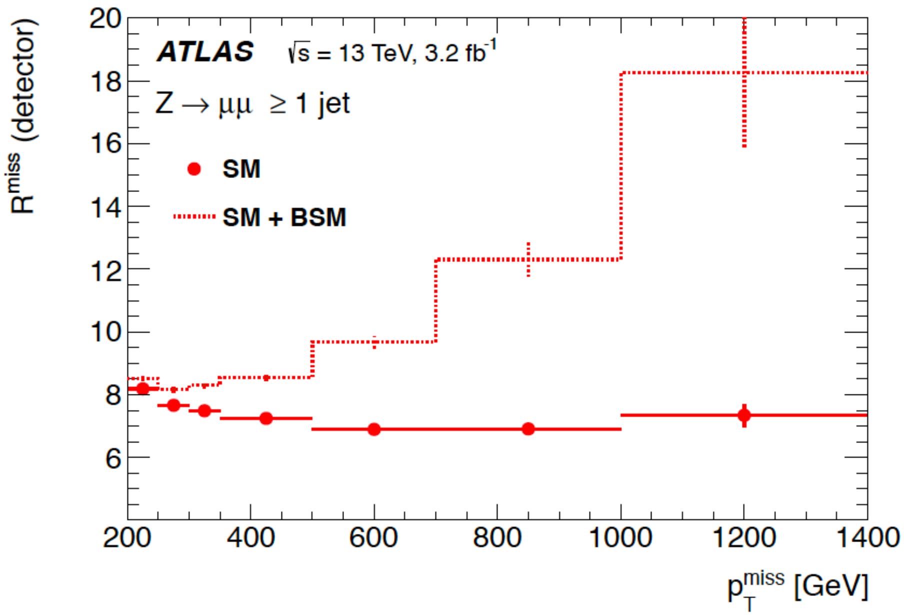
# Cancellation of systematic uncertainties

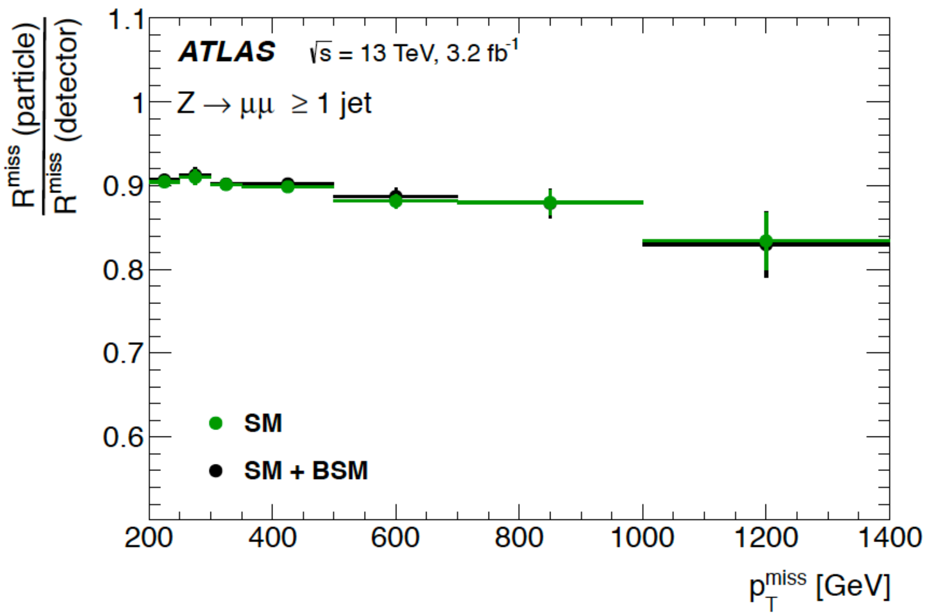



- $R^{miss}$ designed for cancellation of dominant experimental/theoretical uncertainties:
  - Jet energy scale and resolution
  - QCD uncertainties in shape of backgrounds (Z+jets)

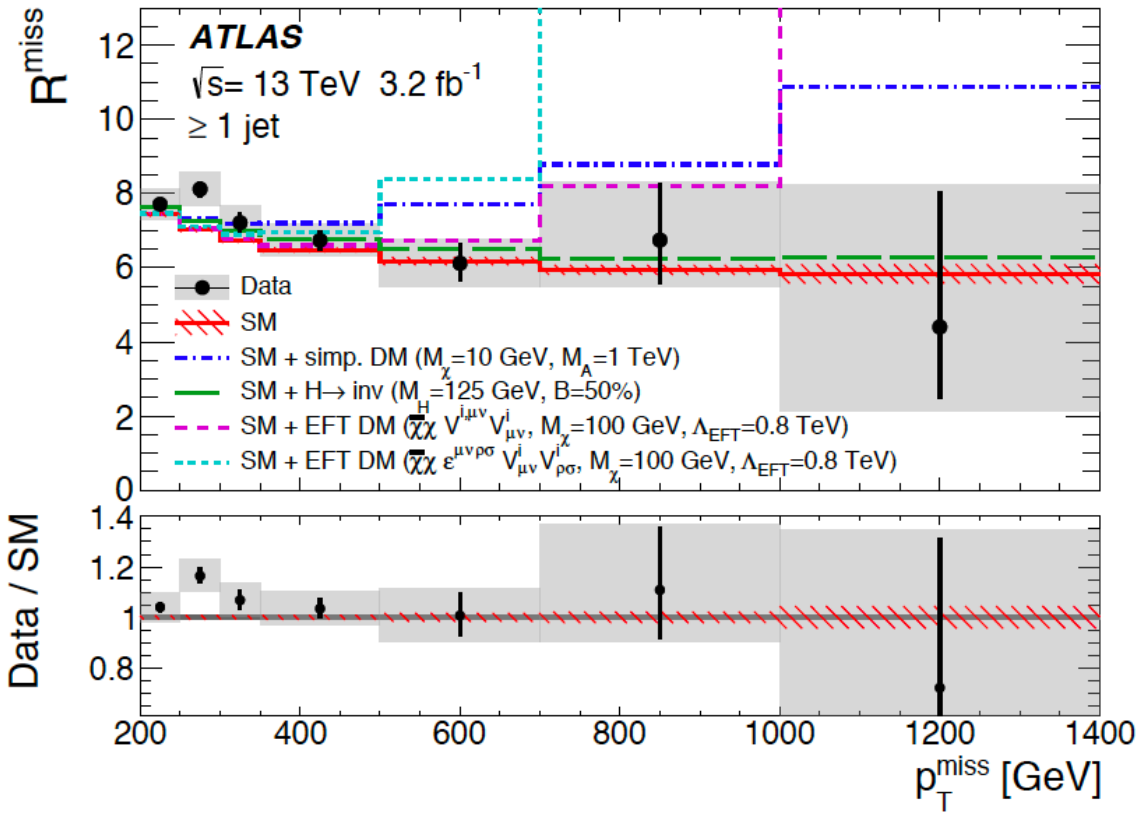
# Impact of BSM physics on the correction for detector effects

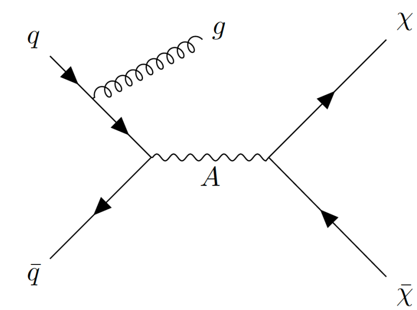
- Correction factor determined from simulation of $R^{miss}$
  - QCD scale variations in Z+jets production cancel in the $R^{miss}$ ratio
  - Presence of BSM physics potentially spoils the cancellation
  - Very small effect for new physics models that produce only $p_T^{miss}$ + jets



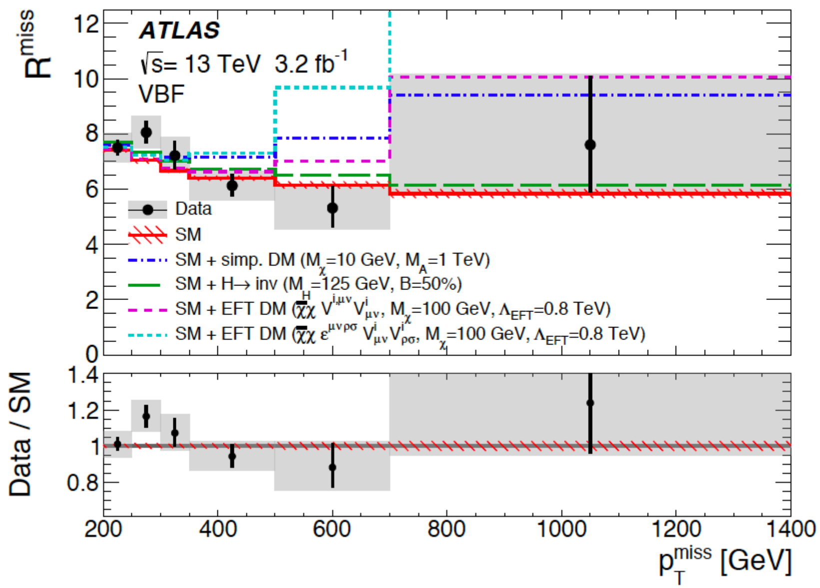
# Measurement of $R^{miss}$ in a 'mono-jet' phase space



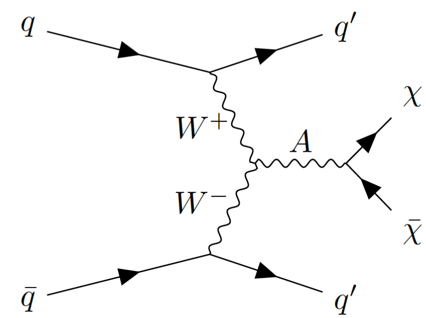



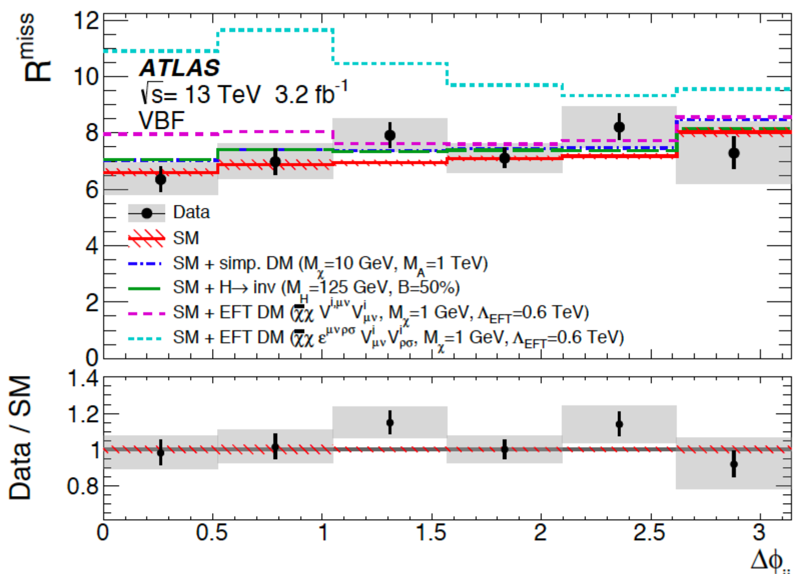

- $p_T^{miss} > 200$ GeV
- zero leptons within fiducial volume
- leading jet $p_T > 120$ GeV

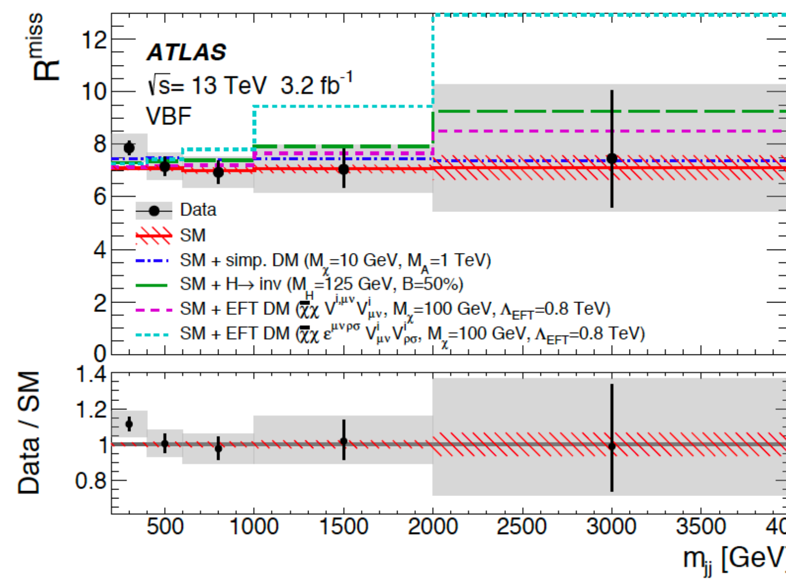
# Measurement of $R^{miss}$ in a VBF phase spaces



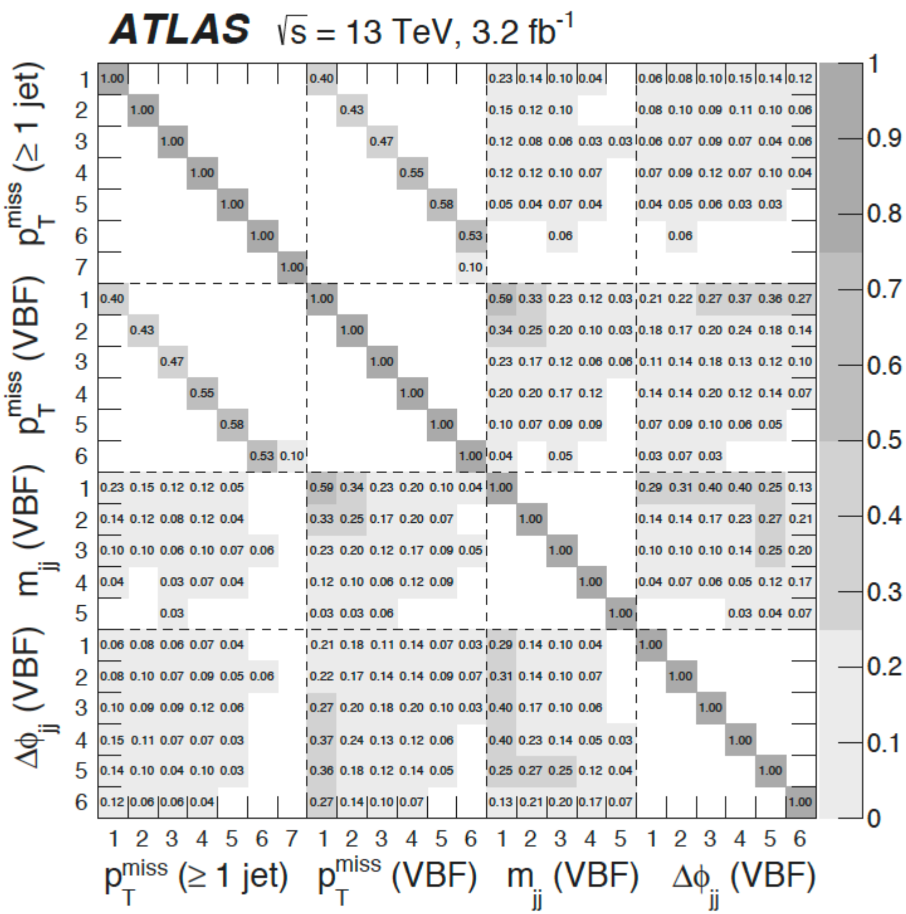

- Two jets: $p_{T,1} > 80$ GeV and $p_{T,2} > 50$ GeV
- No third jet between tagging jets
- Dijet invariant mass > 250 GeV



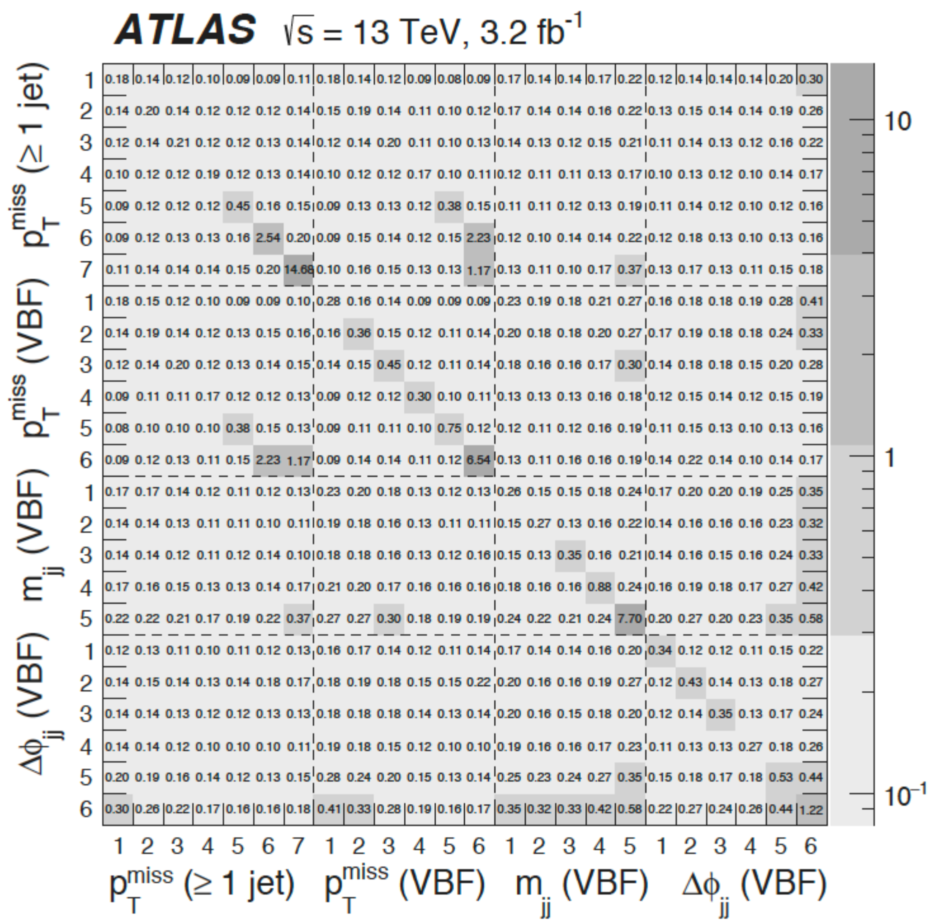



# Determination of statistical and systematic correlations



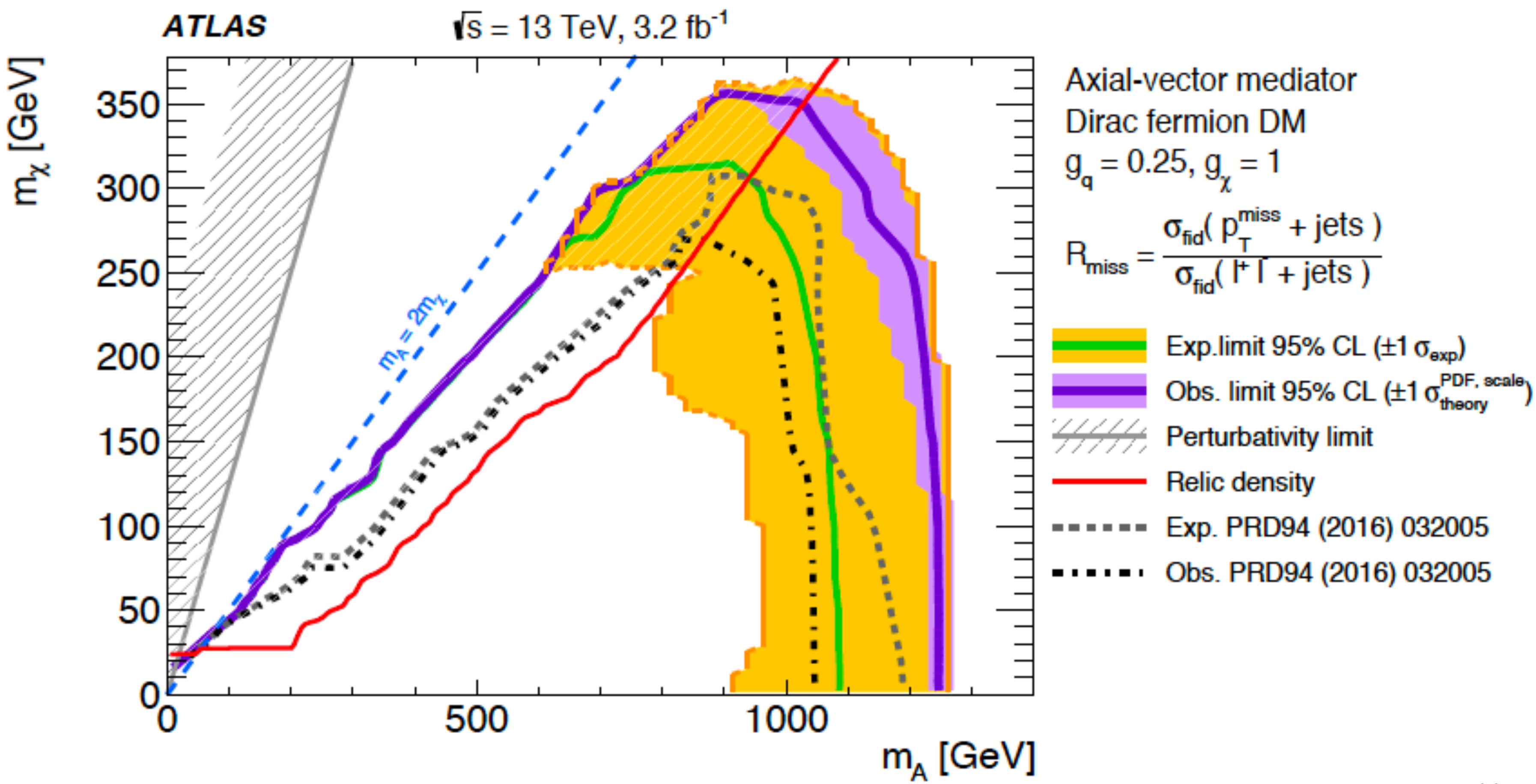

Statistical correlations




Total covariance matrix (C)

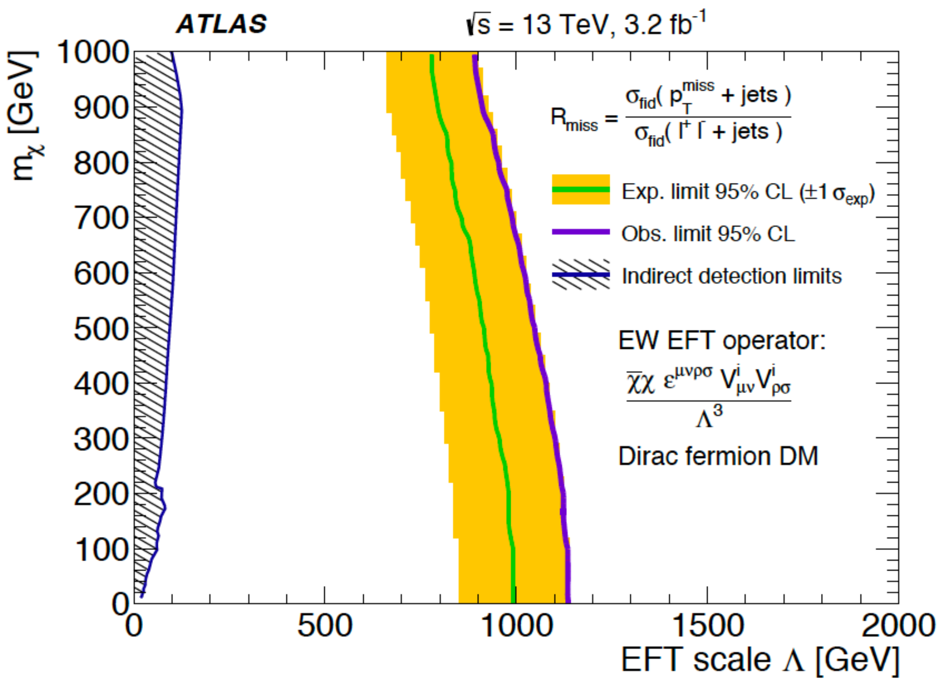
# Constraints on simplified dark matter models

- Limits set using CLs after constructing $\chi^2 = (\mathbf{y}_{\text{data}} - \mathbf{y}_{\text{pred}})^T C^{-1} (\mathbf{y}_{\text{data}} - \mathbf{y}_{\text{pred}})$



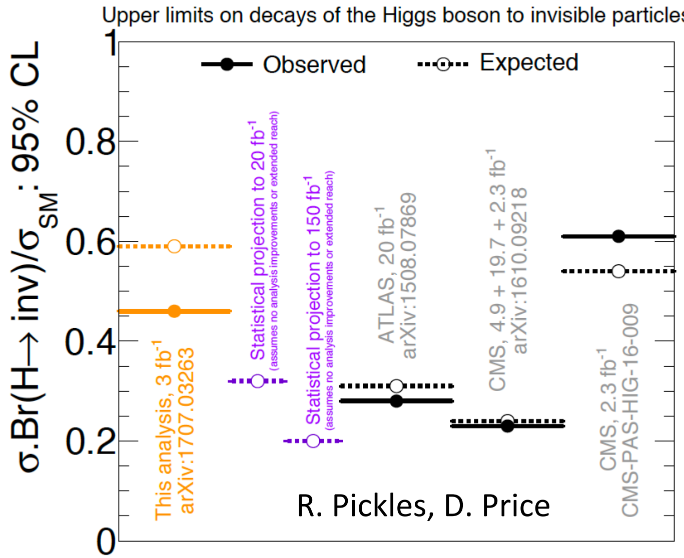
# Easy reinterpretation of data for any model of interest…..

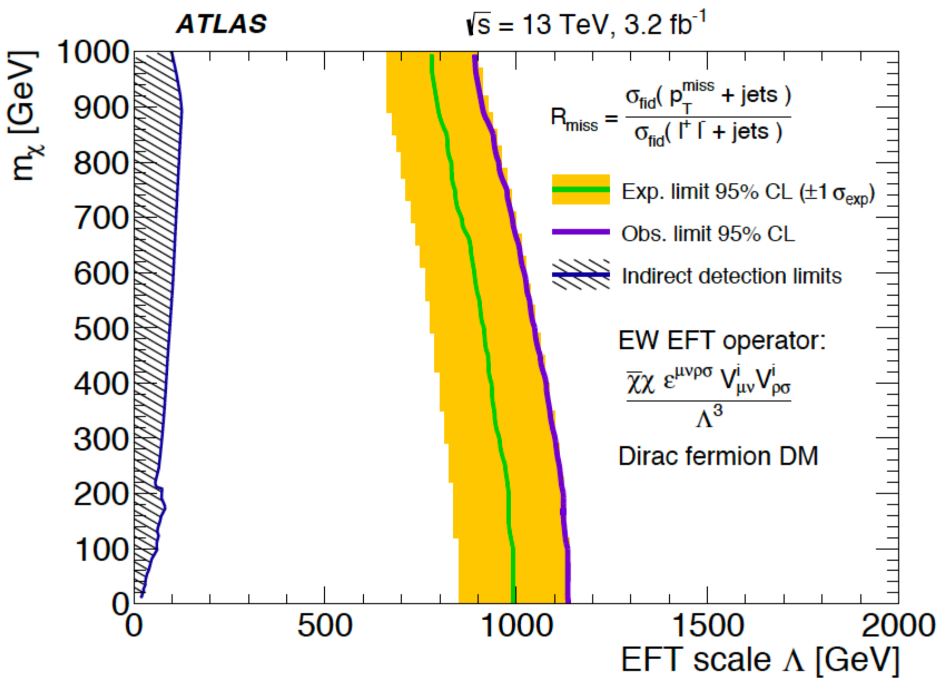
- Observables and covariance matrices published on HEPDATA.
- Rivet routine provided for fast comparison of any model to the data.



# Easy reinterpretation of data for any model of interest…..

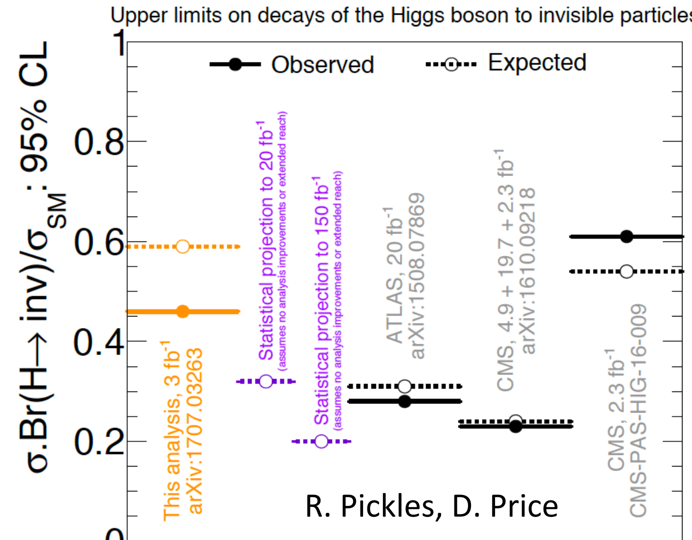
- Observables and covariance matrices published on HEPDATA.
- Rivet routine provided for fast comparison of any model to the data.

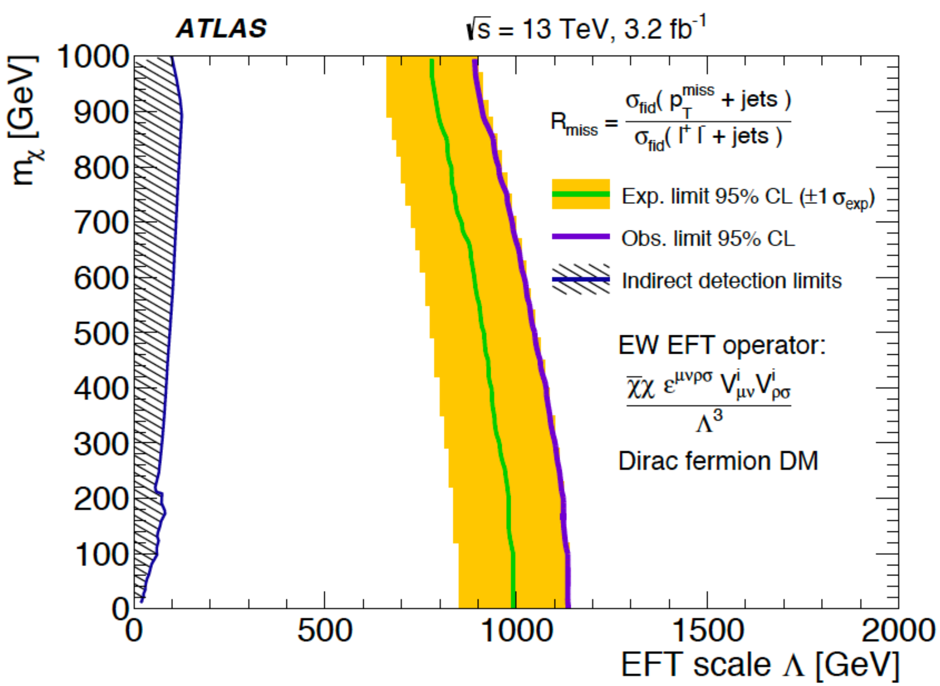


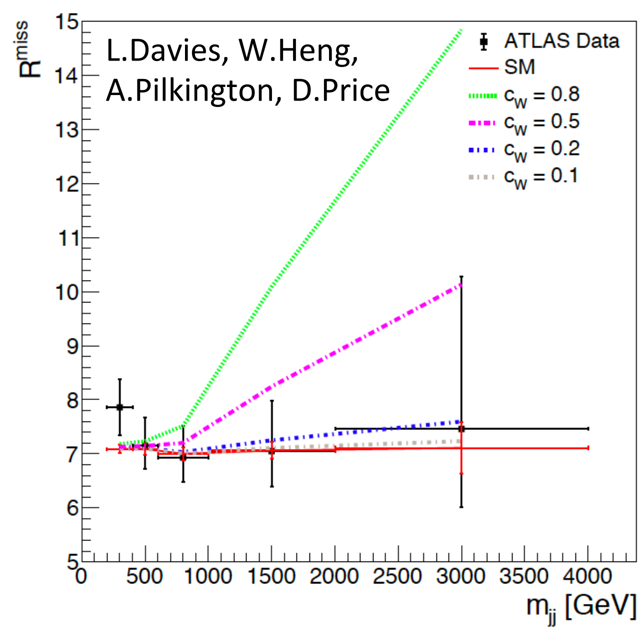



# Easy reinterpretation of data for any model of interest…..

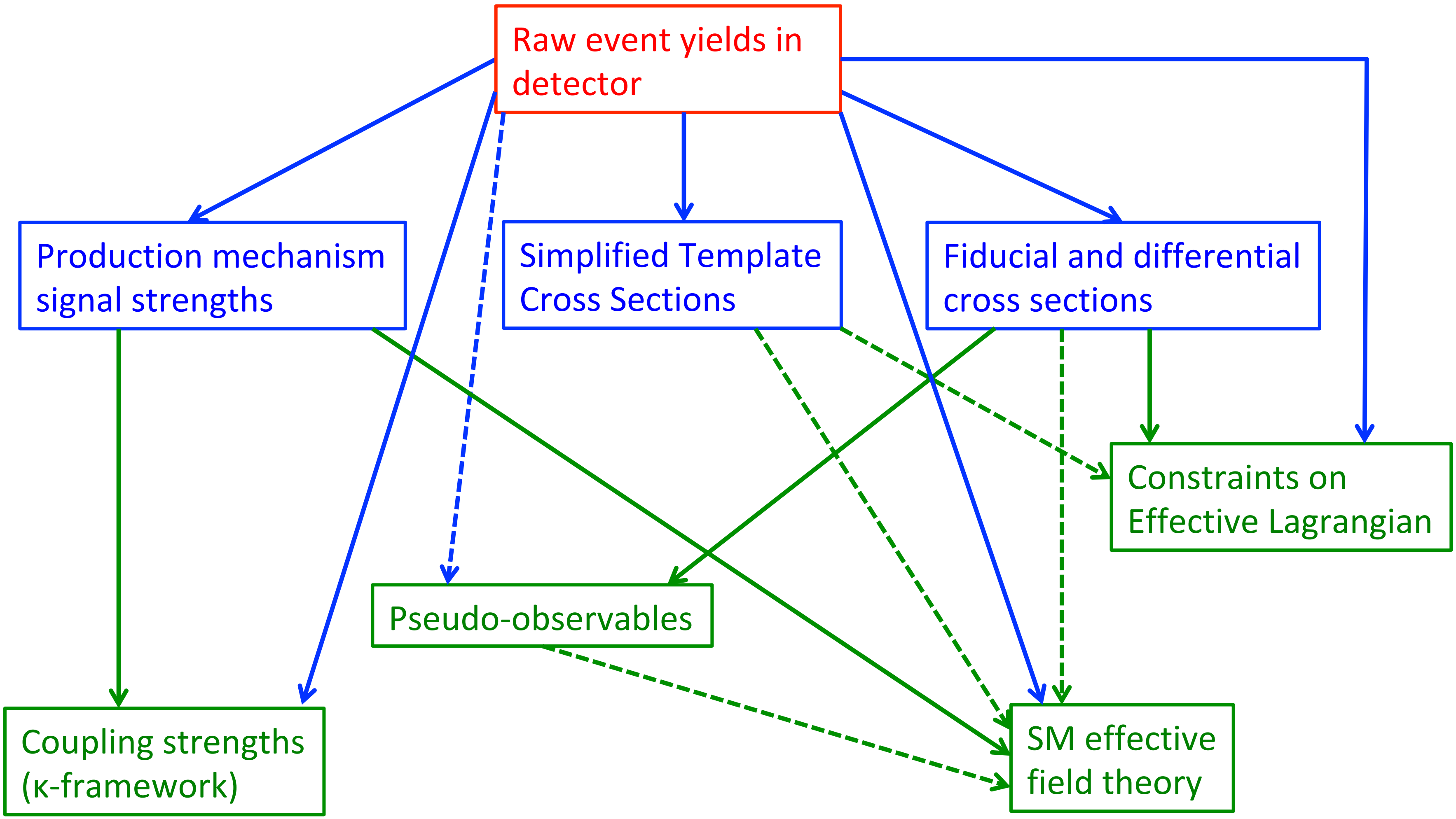
- Observables and covariance matrices published on HEPDATA.
- Rivet routine provided for fast comparison of any model to the data.



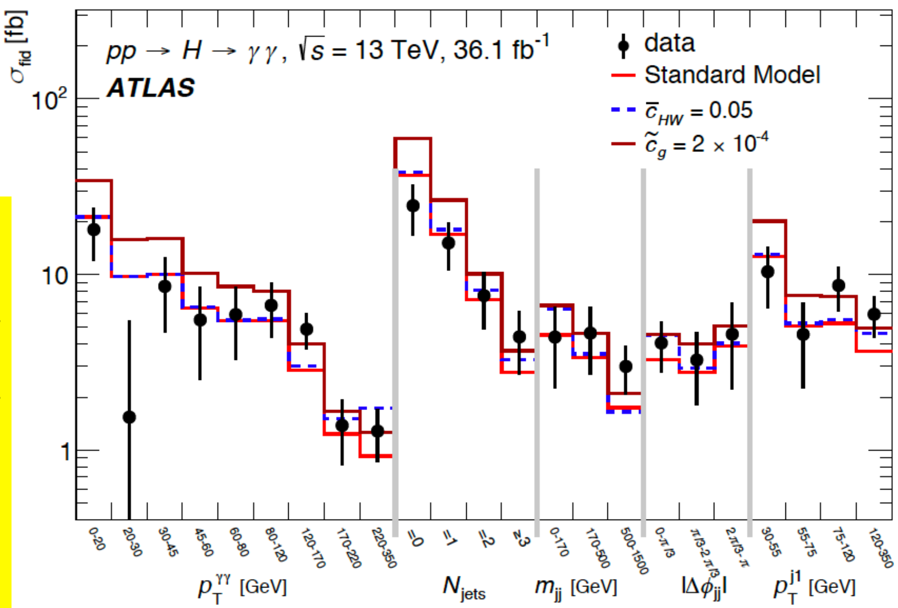



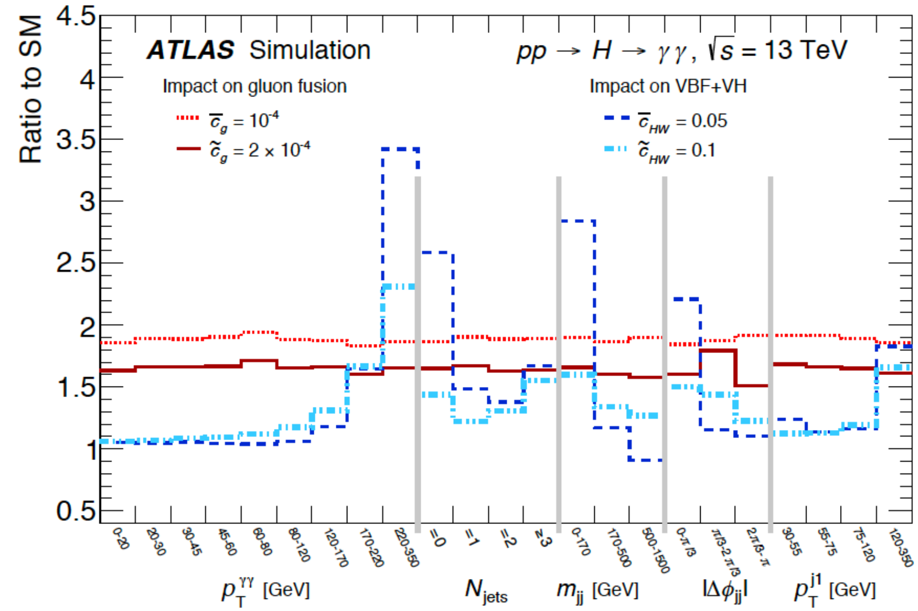
# Different experimental approaches to probing the Higgs couplings




- Experimental measurements
- Theoretical interpretations
- Dashed means not yet demonstrated

# Constraints on effective Lagrangian operators

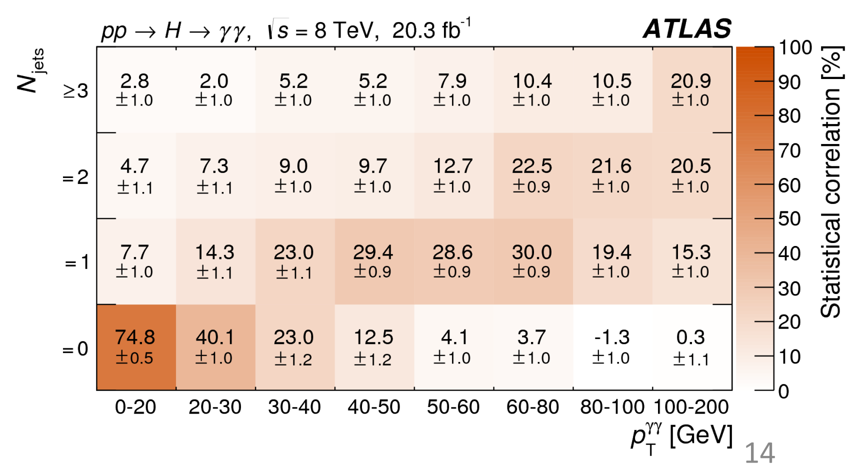



Constraint on specific operators in effective Lagrangian via simultaneous fit to 5 observables

$$\mathcal{L}_{\rm eff} = \bar{c}_\gamma O_\gamma + \bar{c}_g O_g + \bar{c}_{HW} O_{HW} + \bar{c}_{HB} O_{HB} + \tilde{c}_\gamma \tilde{O}_\gamma + \tilde{c}_g \tilde{O}_g + \tilde{c}_{HW} \tilde{O}_{HW} + \tilde{c}_{HB} \tilde{O}_{HB}$$

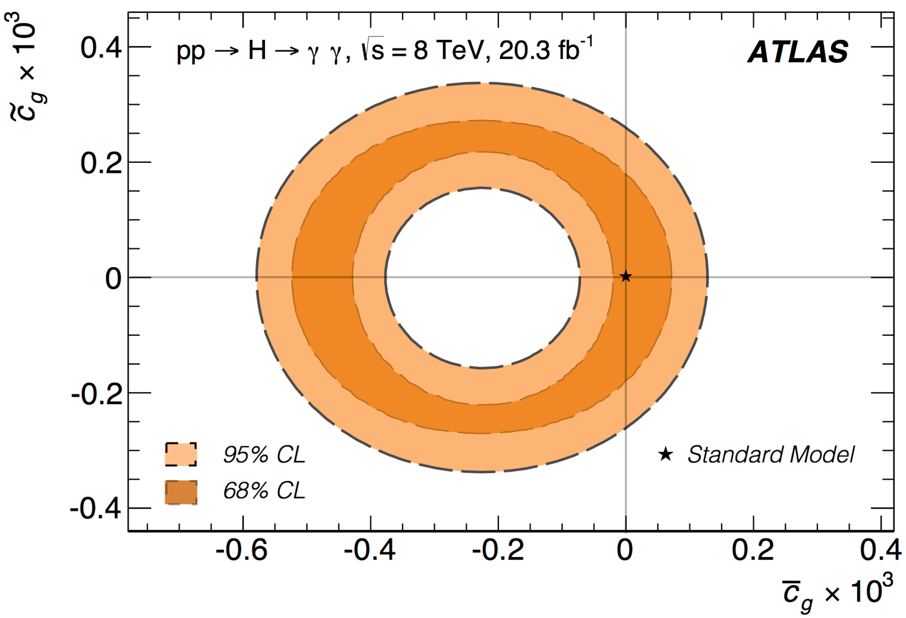
Hγγ  Hgg  HVV

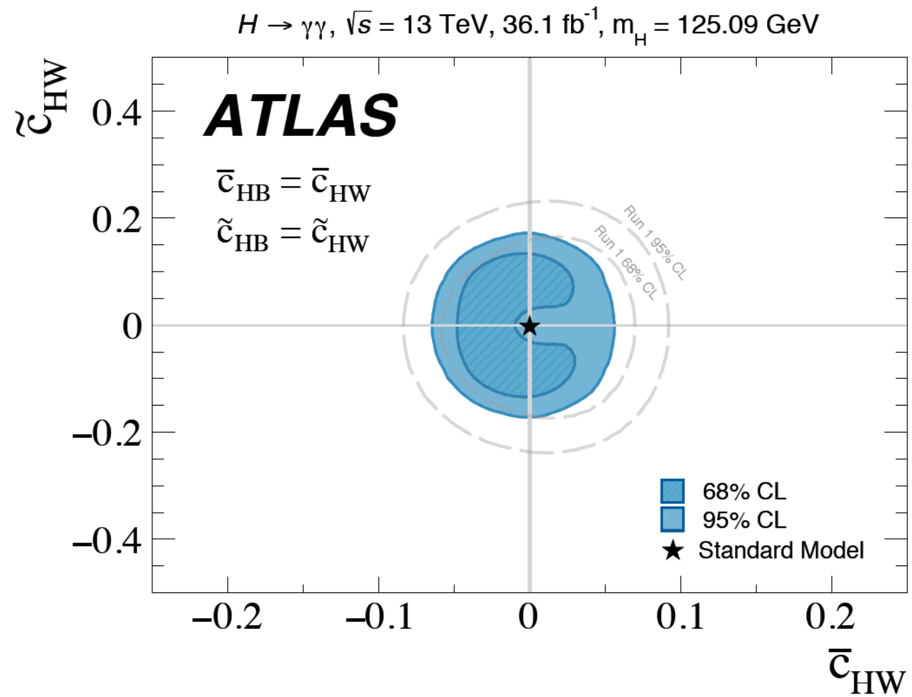



| $N_{jets}$ | 0-20 | 20-30 | 30-40 | 40-50 | 50-60 | 60-80 | 80-100 | 100-200 |
|---|---|---|---|---|---|---|---|---|
| ≥3 | 2.8 ±1.0 | 2.0 ±1.0 | 5.2 ±1.0 | 5.2 ±1.0 | 7.9 ±1.0 | 10.4 ±1.0 | 10.5 ±1.0 | 20.9 ±1.0 |
| =2 | 4.7 ±1.1 | 7.3 ±1.1 | 9.0 ±1.0 | 9.7 ±1.0 | 12.7 ±1.0 | 22.5 ±0.9 | 21.6 ±1.0 | 20.5 ±1.0 |
| =1 | 7.7 ±1.0 | 14.3 ±1.1 | 23.0 ±1.1 | 29.4 ±0.9 | 28.6 ±0.9 | 30.0 ±0.9 | 19.4 ±1.0 | 15.3 ±1.0 |
| =0 | 74.8 ±0.5 | 40.1 ±1.0 | 23.0 ±1.2 | 12.5 ±1.2 | 4.1 ±1.0 | 3.7 ±1.0 | -1.3 ±1.0 | 0.3 ±1.1 |

ATLAS Collaboration, arXiv:1802.04146 and PLB 753 (2016) 69-85

# Constraints on effective Lagrangian terms in H->yy




- Proof of principle: set 2D constraints on CP-even and CP-odd couplings to gluons (left) and weak bosons (right)
- All data/correlations public: can repeat with favourite EFT basis or BSM physics model

# Using Higgs data to search for resonantly-enhanced dark matter production

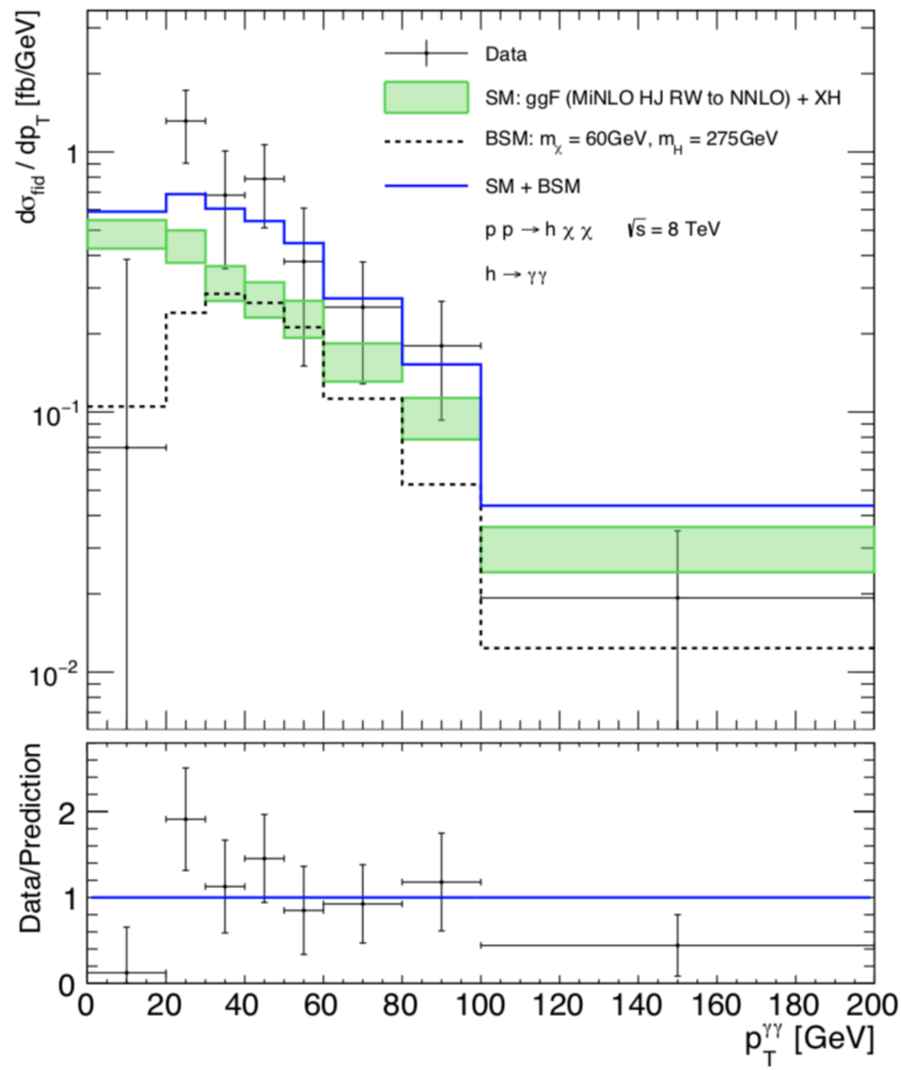
Von Buddenbrock et. al. https://arxiv.org/abs/1506.00612v1

**CERNpress** @CERNpress

Sorry guys, but there is no evidence so far in the #LHC data to support the existence of a hypothetical #Madala #boson

3:14 PM - Sep 7, 2016

199 235 people are talking about this



# Summary

- LHC data cost billions to produce, must make sure it is as useful as possible:
  - Ensure longevity of the data
  - Allow future reinterpretations and combinations of multiple analyses
  - Facilitate model building

- Producing detector-corrected observables achieves all of those goals
  - Proof-of-principle for dark-matter searches published in EPJ C77 (2017) 765
    - HEPDATA: https://www.hepdata.net/record/78366
    - Rivet: http://rivet.hepforge.org/analyses/ATLAS_2017_I1609448.html
  - Existing measurements of Standard Model processes can be used as well