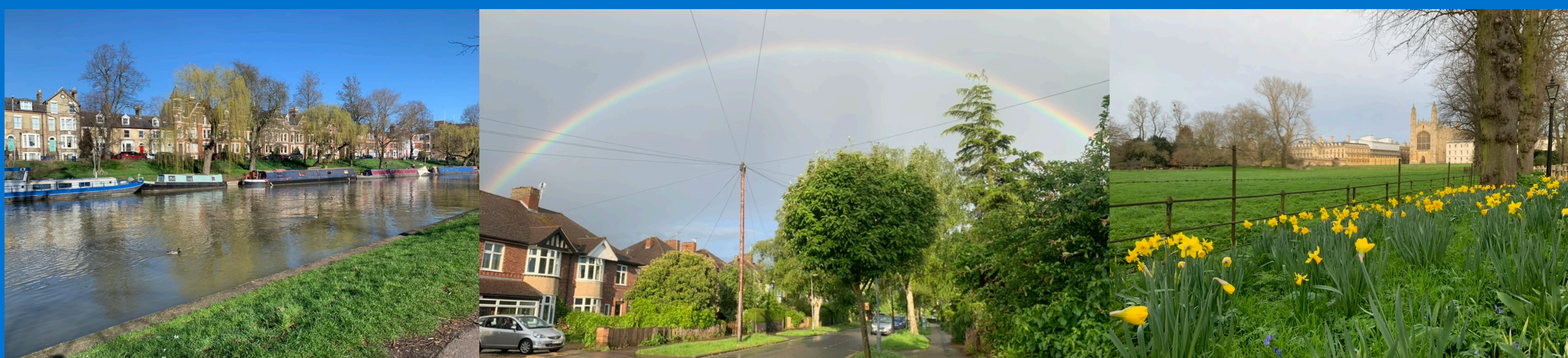
Unfolding in BSM search-sensitive regions of phase space



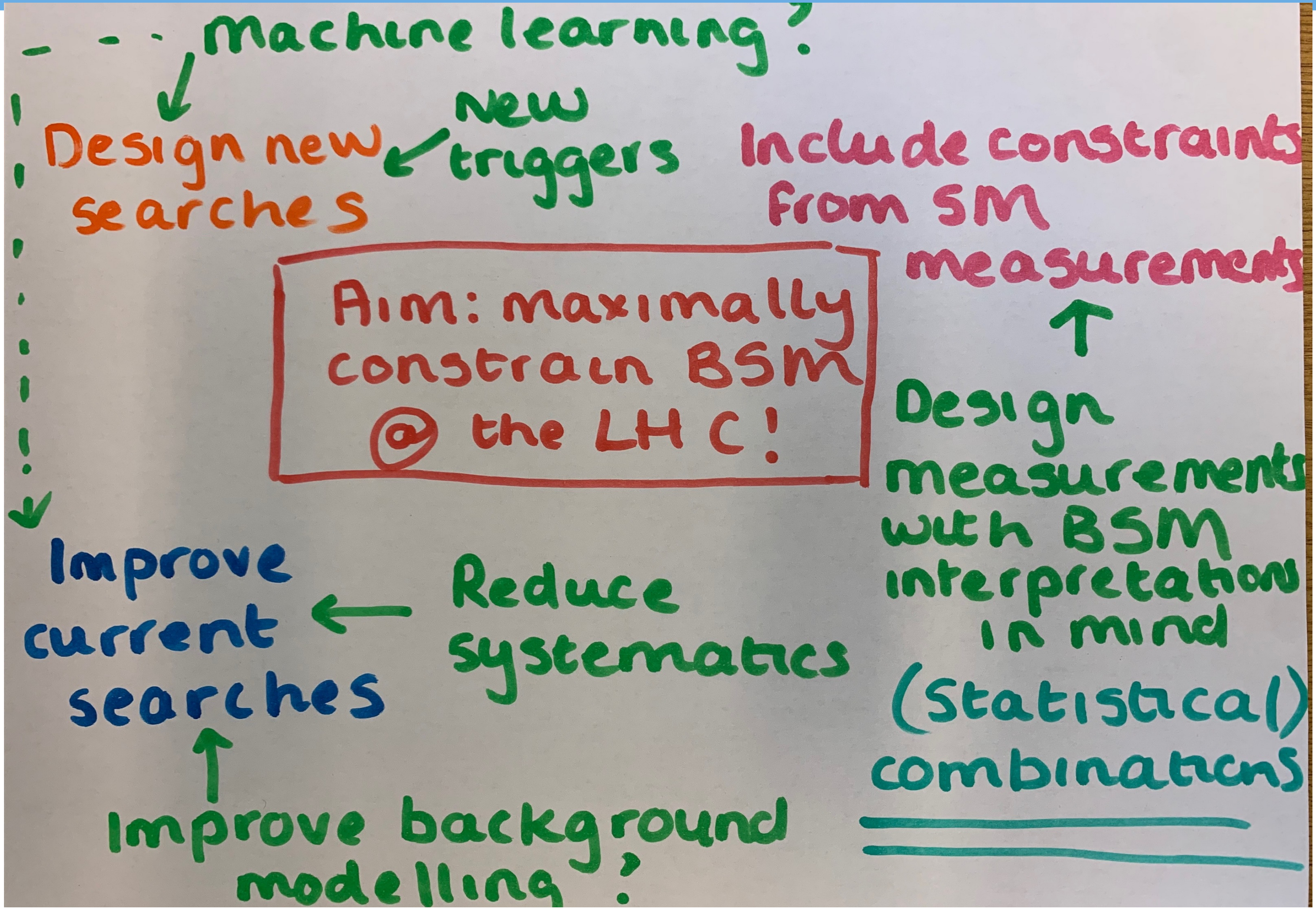
Dr Sarah Williams, University of Cambridge

Thanks to Jon Butterworth and Louis Corpe for input and discussion

Introduction

- I was originally asked to discuss this paper from 2022 which focused on differential measurements of SM WW production in a “SUSY-inspired” phase space (SUSY=supersymmetry).
- Will focus on this paper, but also mention a few other results that fall under the umbrella of “how to better use precision measurements to maximise BSM sensitivity”.
- Aim: promote discussion, and encourage new ideas for the future?

If this talk were a mind-map

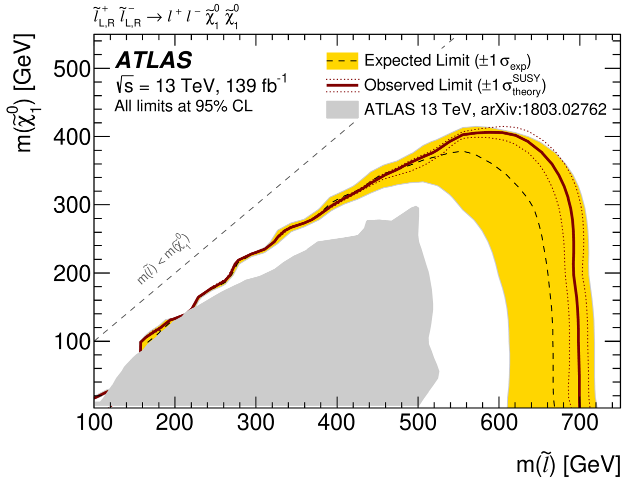


Testing the SM in colliders

Two complementary ways to test the SM in colliders

1. Direct searches for new particles predicted in well-motivated extensions (including SUSY).

2. Indirect searches through precision measurements of SM parameters (and the Higgs) where deviations would be hints of new physics coming into plan.



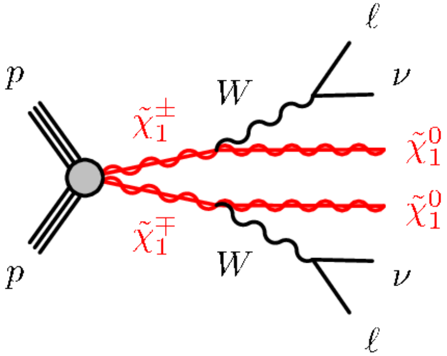
Exclusion limits on SUSY masses for slepton pair production using "simplified model"

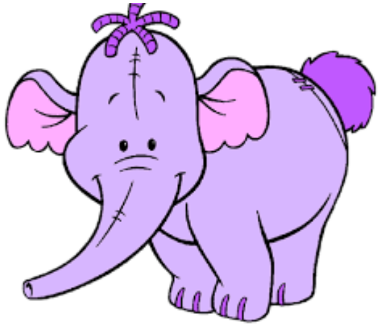
Bridging the gap between these (often disjoint) activities represents an exciting direction for the future…

Reminder: direct searches for new physics

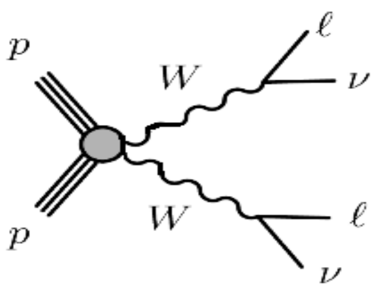
Key point: these are performed using "detector-level" quantities

Signal

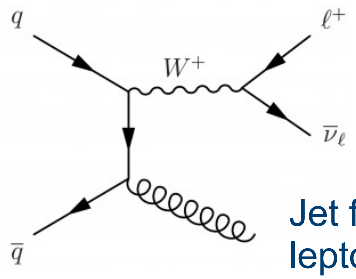




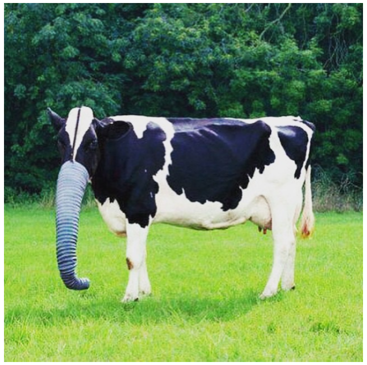
Irreducible background



Reducible backgrounds



Jet faking a lepton

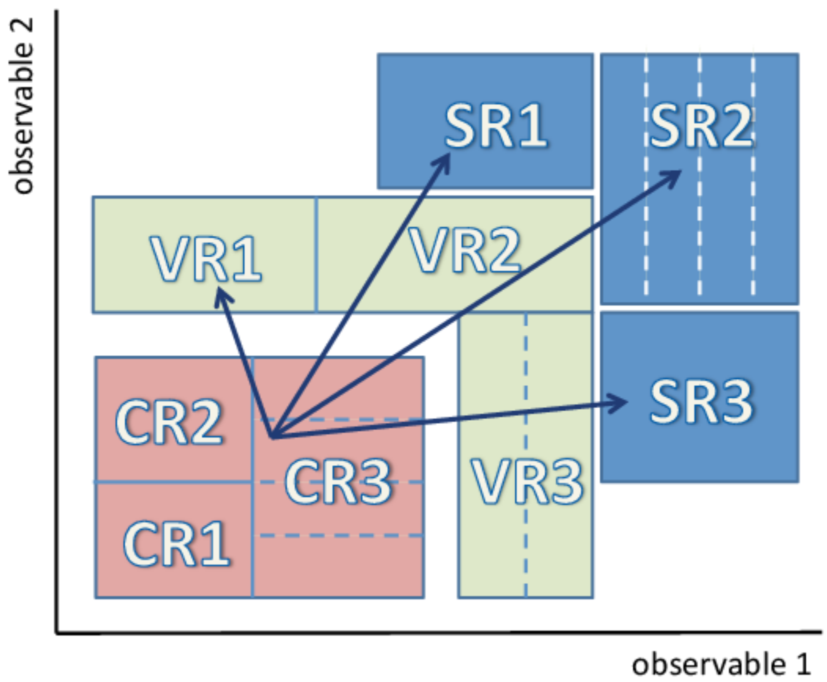


Cow faking an elephant (not very well)

=> Direct searches for new physics at the LHC involve searching for statistically significant deviations from the SM in particular decay channels/event topologies.

Typical SUSY search strategy

Diagram taken from the Histfitter paper



- Identify (binned) **”signal region(s)”** where we would expect to see an excess over the SM prediction if the signal were present.
- Use **“control region(s)”** to extract data-driven normalisation factor(s) for dominant background component(s) using simultaneous likelihood fit.
- Before unblinding, use **“validation regions”** to check that the background-estimates provide accurate normalisation and shapes of kinematic distributions.

We typically produce ”post-fit” yields tables and kinematic distributions, and quote the normalisation factors from the likelihood fit (which depends on the MC generator being used).

Search-inspired SM measurements in ATLAS

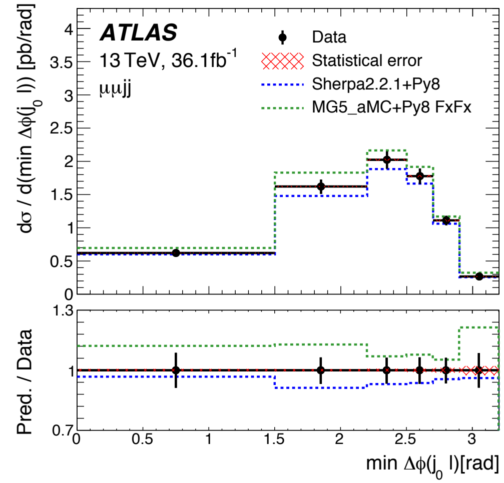
Eur. Phys. J. C 79 (2019) 733

Searches for scalar leptoquarks and differential cross-section measurements in dilepton–dijet events in proton–proton collisions at a centre-of-mass energy of $\sqrt{s}$ = 13 TeV with the ATLAS experiment

⇒ First result of its kind from ATLAS.

⇒ Used bin-by-bin unfolding (not possible here)

$$\frac{d\sigma_i^p}{dX} = \frac{(N_i - \sum_{q \neq p} R_i^q) \cdot \frac{T_i^p}{R_i^p}}{w_i \cdot L},$$



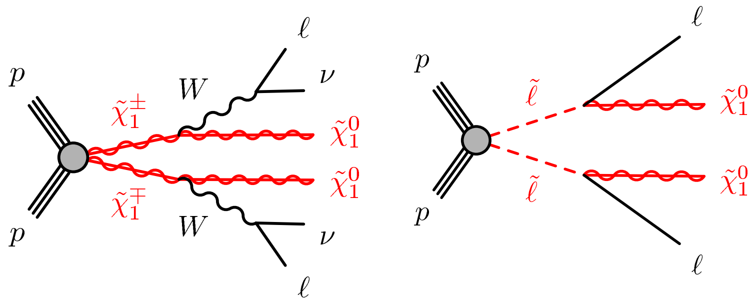
The result I will now discuss was the first effort to do a search inspired unfolding calculation in a SUSY-inspired phase space:

- Follows an early run 2 search for supersymmetric charginos and sleptons decaying to 2-lepton final states.
- Future aim would be to better incorporate measurements into search efforts

Overview EWK 2l+0jets search

Eur. Phys. J. C 80 (2020) 123

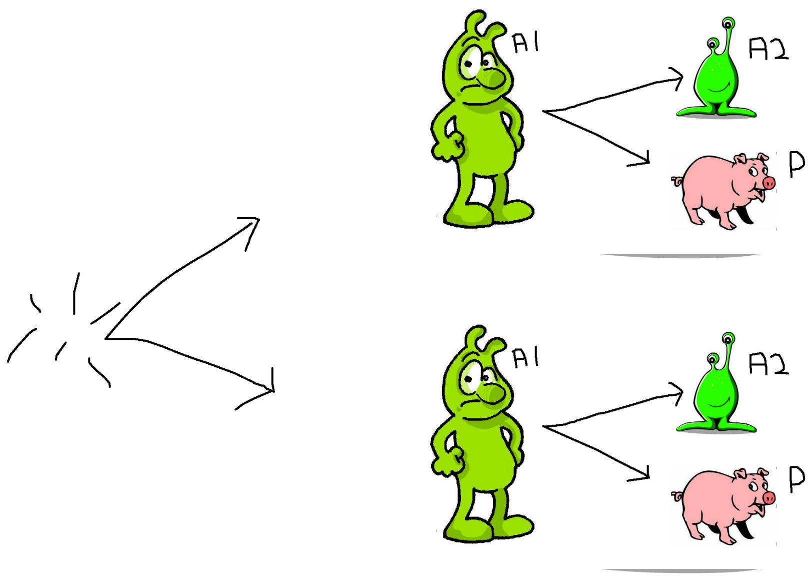
Signal region (SR)	SR-DF-0J	SR-DF-1J	SR-SF-0J	SR-SF-1J
$n_{\text{non-}b\text{-tagged jets}}$	= 0	= 1	= 0	= 1
$m_{\ell_1\ell_2}$ [GeV]	>100		>121.2	
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	>110			
$E_{\mathrm{T}}^{\mathrm{miss}}$ significance	>10			
$n_{b\text{-tagged jets}}$	= 0			
Binned SRs				
m_{T2} [GeV]	∈[100,105) ∈[105,110) ∈[110,120) ∈[120,140) ∈[140,160) ∈[160,180) ∈[180,220) ∈[220,260) ∈[260,∞)			
Inclusive SRs				
m_{T2} [GeV]	∈[100,∞) ∈[160,∞) ∈[100,120) ∈[120,160)			



Set of binned SRs in the 'stransverse mass variable' (m_{T2}) separated into same-flavour SF ($e^\pm e^\mp$ or $\mu^\pm\mu^\mp$) and different flavour DF ($e^\pm\mu^\mp$) categories and by the light (i.e. non-b-tagged) jet multiplicity, at high values of the missing transverse momentum and object-based missing transverse momentum significance.

Dominant background is SM WW background. Aim of this work was to provide a (differential) SM measurement in a region of phase space closer to this search.

The ”stransverse mass” variable m_{T2}



- Originally designed to measure SUSY masses at the LHC.
- Powerful discriminating variable in searches for semi-invisibly decaying pair-produced particles.
- For massless invisible particles, in the absence of reconstruction/misidentification effects expect a kinematic endpoint for $t\bar{t}$ and WW production at the W-boson mass.

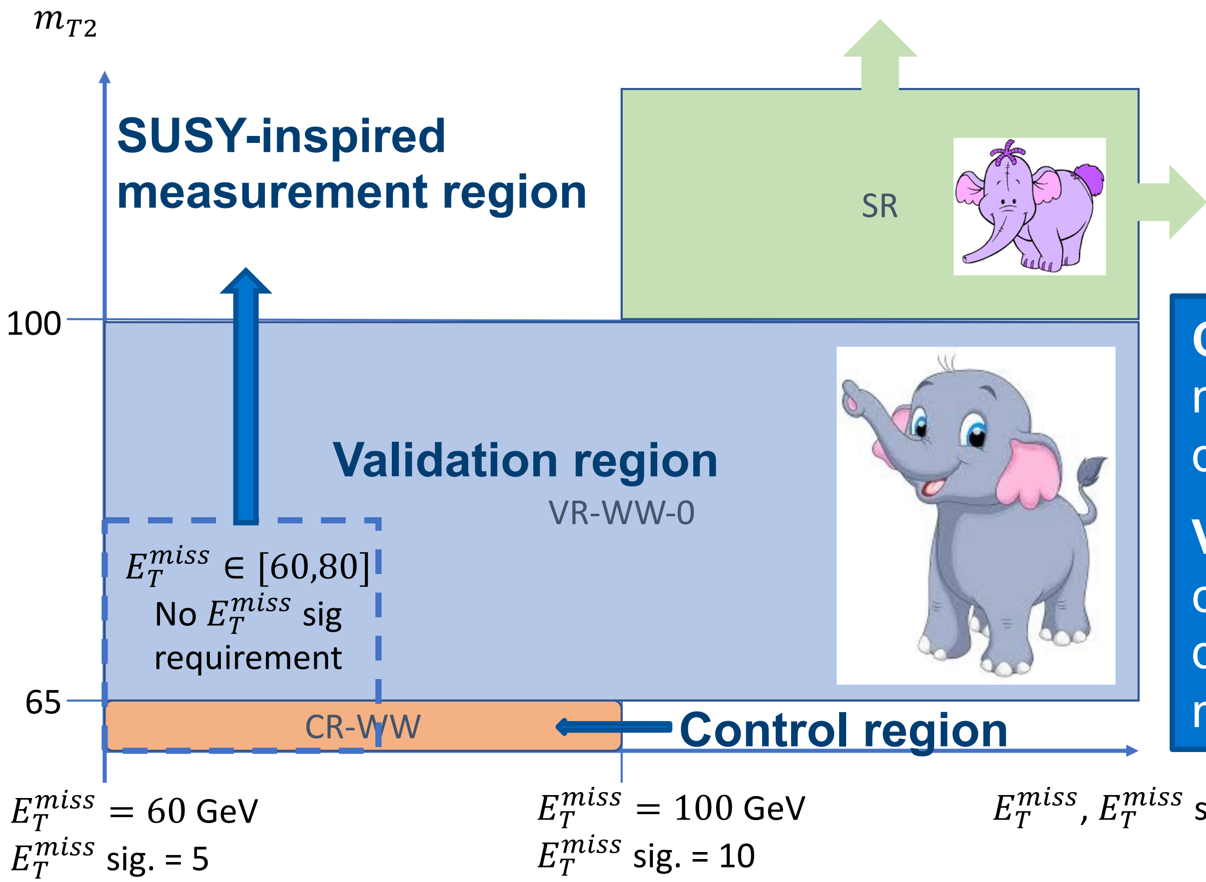
$$m_{A2}^2 \geq m_{T2}^2 = \min_{\mathbf{p}_T^{miss} = \mathbf{p}_T^{A1,a} + \mathbf{p}_T^{A1,b}} [\max\{m_T^2(\mathbf{p}_T^{P,a}, \mathbf{p}_T^{A1,a}), m_T^2(\mathbf{p}_T^{P,b}, \mathbf{p}_T^{A1,b})\}]$$

This meant that the search required high m_{T2} , with events with $m_{T2} \in [60, 100]$ GeV being used for estimation and validation of the WW background

Background estimation

[Eur. Phys. J. C 80 (2020) 123]

2D plane below considers all electron-muon events with no hadronic jets…



Perform search in **signal region** (SR)

Control region (CR)- normalise MC to observed data *

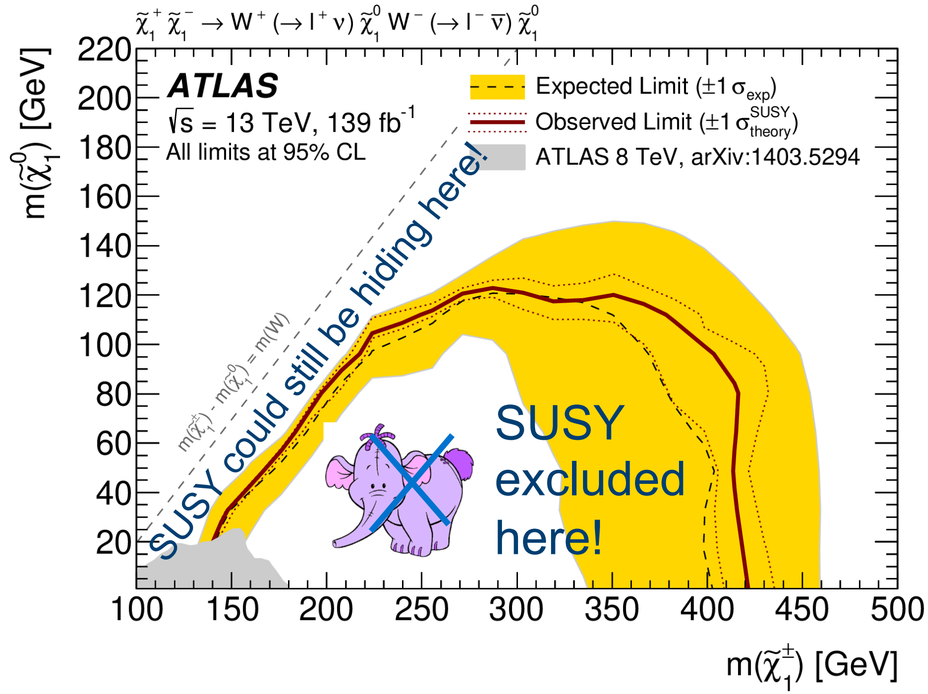
Validation region (VR)- compare yields/ kinematic distributions to check modelling

* In reality this is performed using a simultaneous likelihood fit to data across several CRs.

Headlines results of the search

Eur. Phys. J. C 80 (2020) 123

- **No SUSY!**
- WW scaled by $\mu_{WW} = 1.25 \pm 0.11$
  - This suggests a 25% underestimation of WW- really?
  - Difficult to compare to previous SM measurements.
- Sensitivity limited by theoretical uncertainties- room for improvement?

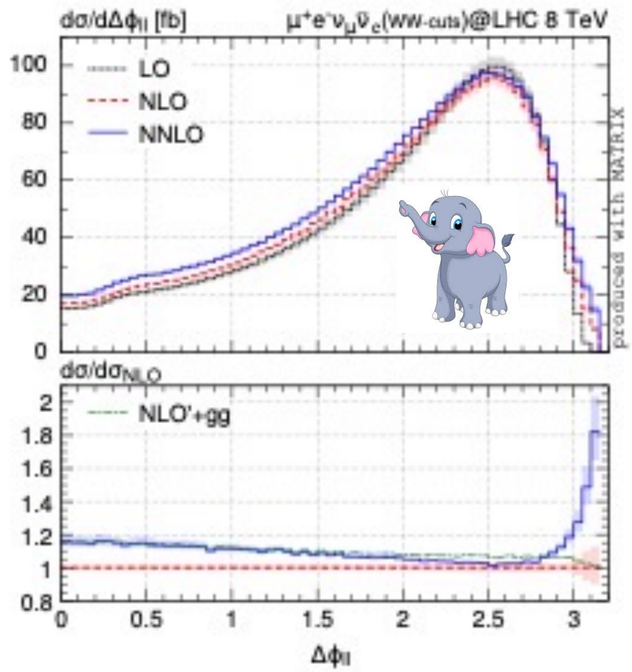


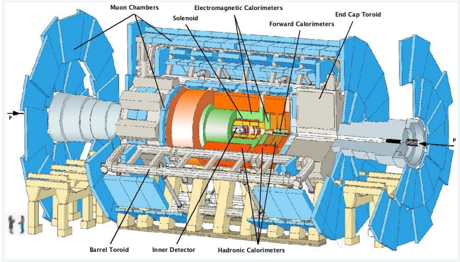
Precision measurements in topologies inspired by searches → improve understanding of backgrounds → increase future search sensitivity!

Precision measurements at the LHC

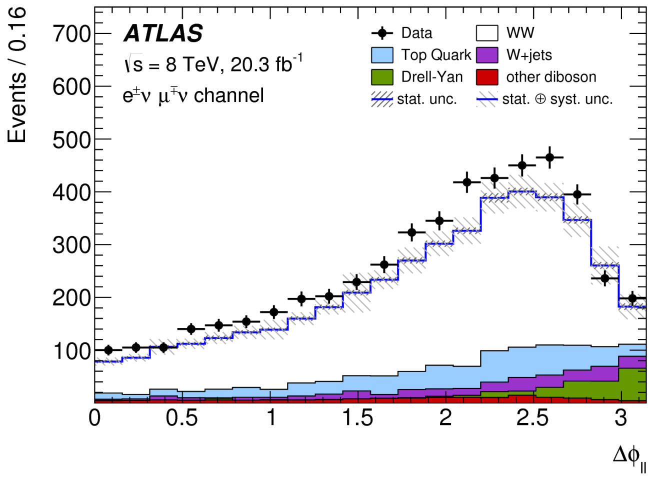
JHEP 09 (2016) 029

Precision fiducial and differential measurements of SM processes at the LHC are presented as "unfolded" particle-level quantities





Folding
(Resulting distributions impacted by detector effects, statistical fluctuations)



Reconstructed detector level distribution

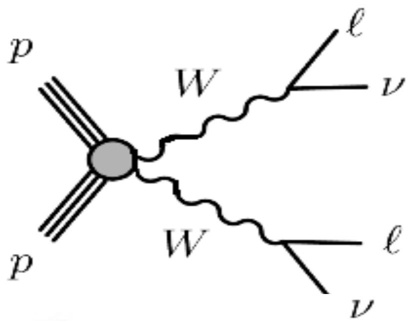
Particle-level theory prediction(s)

Unfolding calculation

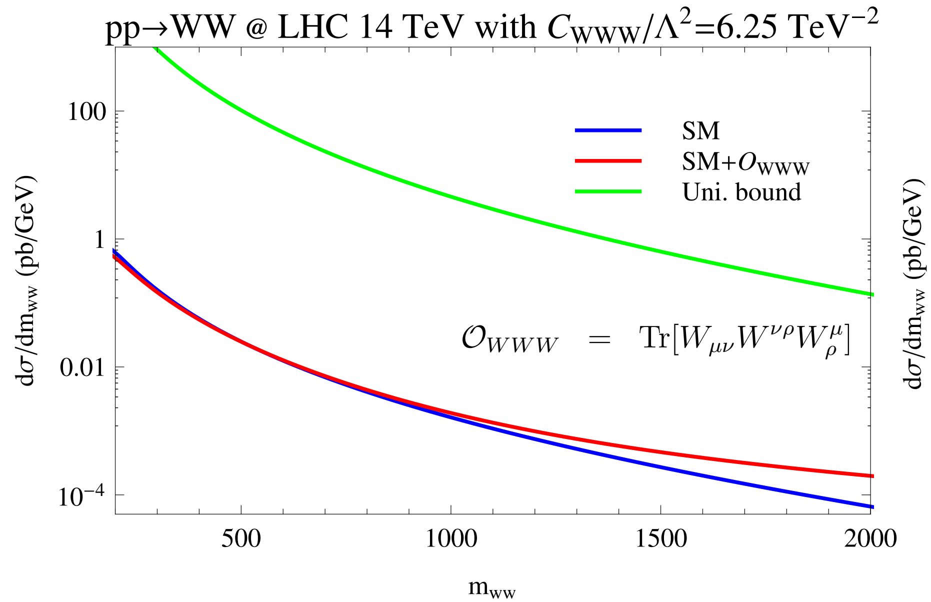
(This entire workshop is devoted to this step)

Precision measurements as indirect probes of new physics

arXiv:1309.7890



Elephant with a "tail"



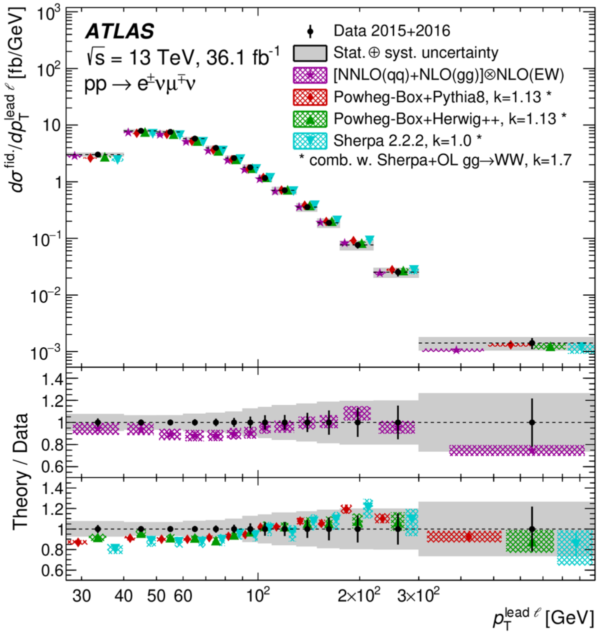
Low energy effects of some high scale new physics can alter the tails of distributions → precision measurements can have sensitivity to mass scales far above those probed in direct searches.

EFTs- alter SM lagrangian using effective operators:

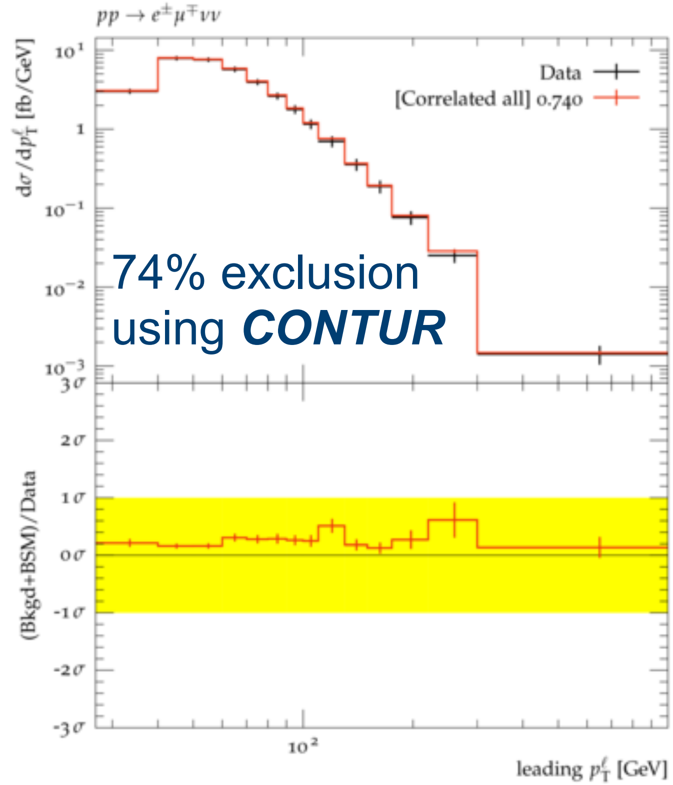
$\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{c_i}{\Lambda^2}\mathcal{O}_i + $ ... to approximate new physics at low scales $\ll \Lambda$

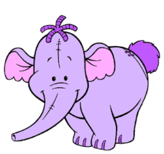
Constraining BSM physics using measurements

Eur. Phys. J. C 79 (2019) 884

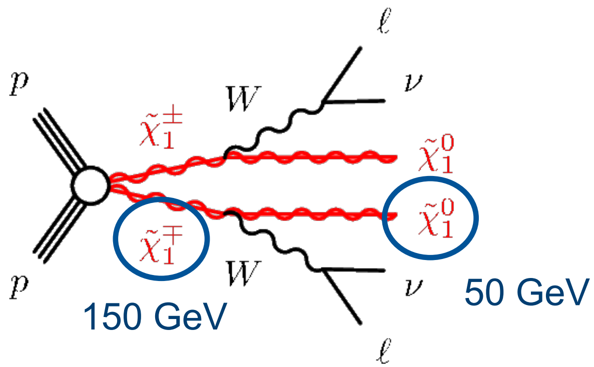


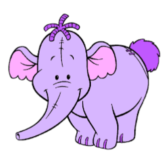
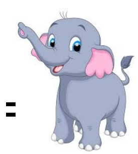
SM measurements in more extreme event topologies would provide exciting new constraints on BSM scenarios!





=



Red=  +  Black =  only

CONTUR and rivet

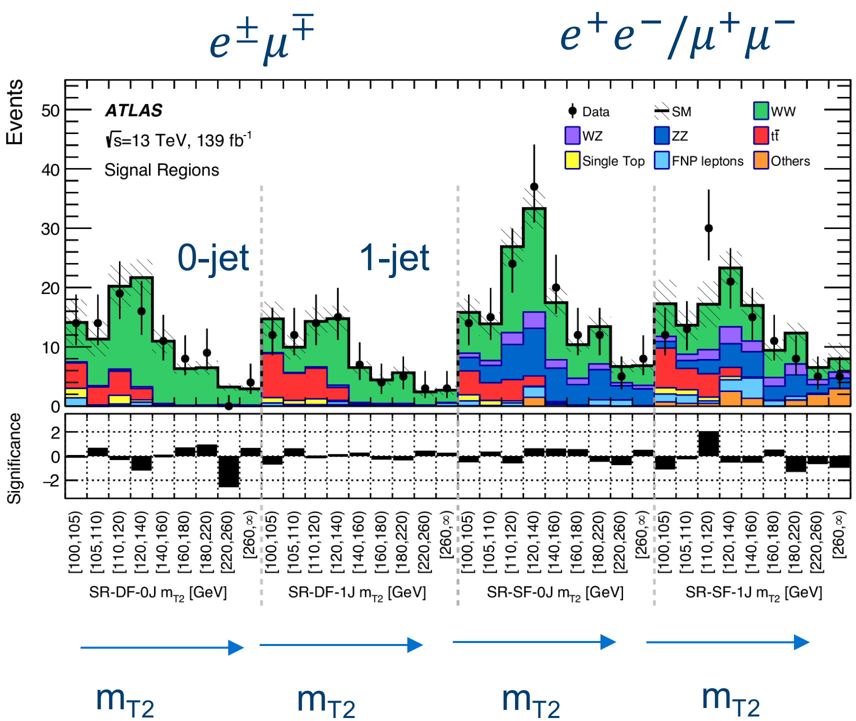
https://rivet.hepforge.org

https://hepcedar.gitlab.io/contur-webpage/

… established frameworks that can be used to for MC tuning and to establish sensitivity of precision measurements to BSM (check out recent results!)

- Rivet= “Robust Independent Validation of Experiment and Theory”- preserves SM measurements for use in MC tuning,.
- CONTUR= “Constraints on new theories using Rivet” => use precision measurements preserved in Rivet to calculate constraints on new physics.
- Divide measurements into groups that have no overlap (and hence no statistical correlations).
- Take the most sensitive measurement from each group and combine them into a single likelihood.
- Calculate likelihood ratio corresponding to S+B vs B-only (where B-only is the unfolded data) and perform statistical test using CL_s technique.

Motivation for SUSY-inspired measurement



(WW scaled by $\mu_{WW} = 1.25 \pm 0.11$)

Region m_{T2} [GeV]	SR-DF-0J $\in$[100,∞)
Total background expectation	96
MC statistical uncertainties	3%
WW normalisation	7%
VZ normalisation	< 1%
$t\bar{t}$ normalisation	1%
Diboson theoretical uncertainties	7%
Top theoretical uncertainties	7%
$E_{\mathrm{T}}^{\mathrm{miss}}$ modelling	1%
Jet energy scale	2%
Jet energy resolution	1%
Pile-up reweighting	< 1%
b-tagging	< 1%
Lepton modelling	1%
FNP leptons	1%
Total systematic uncertainties	15%

No significant excesses across highly “binned” set of signal regions.

Sensitivity limited by theoretical uncertainties… room for improvement!

Fiducial region for unfolding

Selection requirement	Criteria
Lepton flavour	$e^{\pm}\mu^{\mp}$
Lepton p_{T}	> 25 GeV
Lepton $\|\eta\|$	$< 2.47\,(e^{\pm}), < 2.6\,(\mu^{\mp})$
Lepton veto	No additional electrons with $p_{\mathrm{T}} > 10$ GeV, $\|\eta\| < 2.47$
	No additional muons with $p_{\mathrm{T}} > 10$ GeV, $\|\eta\| < 2.6$
$m_{e\mu}$	> 100 GeV
Jet veto	No jets with $p_{\mathrm{T}} > 20$ GeV, $\|\eta\| < 2.4$
m_{T2}	$\in [60, 80]$ GeV
$E_{\mathrm{T}}^{\mathrm{miss}}$	$\in [60, 80]$ GeV

- m_{T2} and E_T^{miss} window adjusted to avoid too narrow m_{T2} window (correlations with angular variables) whilst retaining WW purity (top contamination increases with increasing m_{T2} and E_T^{miss}).
- Top contamination ~ 33%. Use same background estimation as search but peform dedicated check in new top validation region

Comparison to ATLAS 36 fb^{-1} SM WW measurement

https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/STDM-2017-24/

SUSY-inspired fiducial region at higher E_T^{miss}, and $m_{e\mu}$, as well as having an m_{T2} requirement. There's also a different jet veto.

Selection requirement	Selection value
p_T^ℓ	> 27 GeV
η^ℓ	$\lvert\eta^e\rvert < 2.47$ (excluding $1.37 < \lvert\eta^e\rvert < 1.52$), $\lvert\eta^\mu\rvert < 2.5$
Lepton identification	*TightLH* (electron), *Medium* (muon)
Lepton isolation	*Gradient* working point
Number of additional leptons ($p_T > 10$ GeV)	0
Number of jets ($p_T > 35$ GeV, $\lvert\eta\rvert < 4.5$)	0
Number of b-tagged jets ($p_T > 20$ GeV, $\lvert\eta\rvert < 2.5$)	0
$E_T^{\text{miss,track}}$	> 20 GeV
$p_T^{e\mu}$	> 30 GeV
$m_{e\mu}$	> 55 GeV

For SM measurement top background was ~ 25% in SR. Lower than for this effort. Use a similar unfolding strategy (IBU)

By considering 0-jet events, naturally orthogonal to the WW+1jet measurement

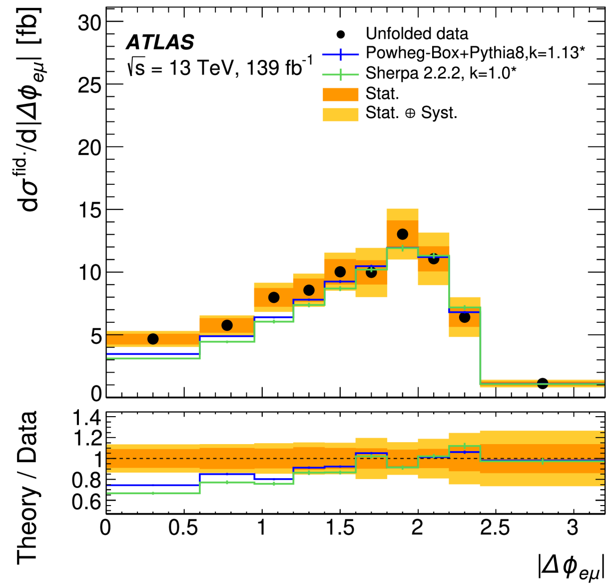
Both efforts produced differential measurements of the same 6 variables!

$$p_T^{e\mu}, m_{e\mu}, p_T^{lep1}, \lvert y_{e\mu}\rvert, \Delta\phi_{e\mu}, \lvert\cos\theta^*\rvert = \left\lvert \tanh\left(\frac{\Delta\eta_{e\mu}}{2}\right)\right\rvert$$

Consistency with SM measurements

$$\sigma_{WW\rightarrow e^{\pm}\nu\mu^{\mp}\nu} = 19.2 \pm 0.3\ (\text{stat})\ \pm 2.5\ (\text{syst})\ \pm 0.4\ (\text{lumi})\ \text{fb} = 19.2 \pm 2.6\ (\text{total})\ \text{fb}.$$

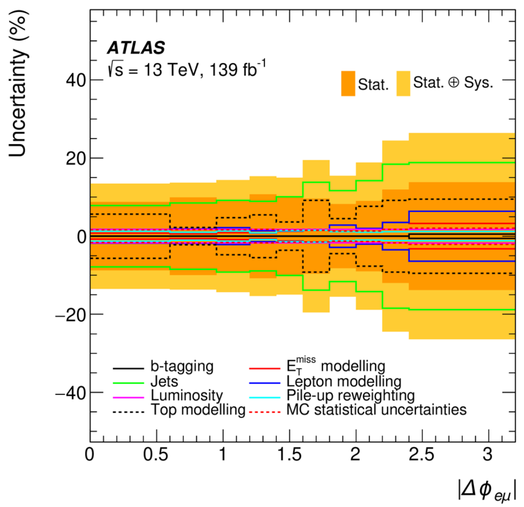
- Powheg+Pythia qq->WW + Sherpa gg->WW predicts 17.8 fb when the NLO k-factors are applied (1.13 for the powheg WW sample).
- Normalisation factor from the EWK 2l+0jets search was 1.25 +- 0.11.
- 1.13* 19.2/17.2 =1.22 -> very consistent.



When the higher order cross-section calculations are used, level of disagreement in the particle-level measurements is consistent with that seen in previous SM measurements

Systematic uncertainties on the measurement

SUSY-inspired measurement



Jet uncertainties ~ 12% impact on fiducial cross-section

2l+0jets search

Region m_{T2} [GeV]	SR-DF-0J $\in[100,\infty)$
Total background expectation	96
MC statistical uncertainties	3%
WW normalisation	7%
VZ normalisation	$< 1\%$
$t\bar{t}$ normalisation	1%
Diboson theoretical uncertainties	7%
Top theoretical uncertainties	7%
E_T^{miss} modelling	1%
Jet energy scale	2%
Jet energy resolution	1%
Pile-up reweighting	$< 1\%$
b-tagging	$< 1\%$
Lepton modelling	1%
FNP leptons	1%
Total systematic uncertainties	15%

WW+0jet 36 /fb measurement

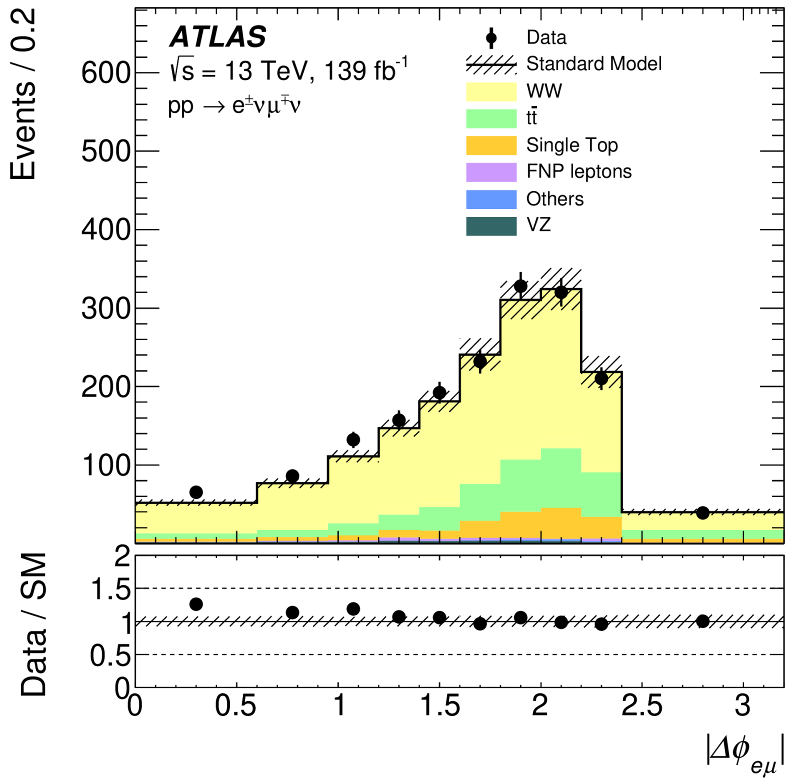
Uncertainty source	Uncertainty [%]
Electron	0.7
Muon	0.9
Jets	3.0
b-tagging	3.4
$E_T^{miss,track}$	0.4
Pile-up	1.6
W+jets background modelling	3.1
Top-quark background modelling	2.6
Other background modelling	1.3
Unfolding, incl. signal MC stat. uncertainty	1.4
PDF+scale	0.1
Systematic uncertainty	6.7
Statistical uncertainty	1.3
Luminosity uncertainty	2.1
Total uncertainty	7.1

(With jet veto optimized to reduce JES/JER uncertainties)

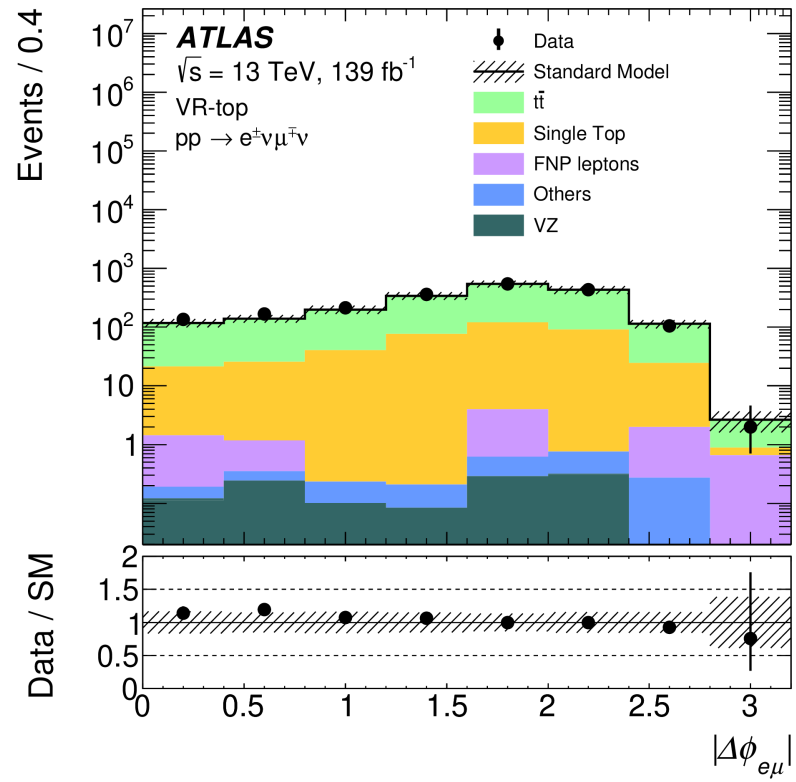
- Plans for the future: should we be optimizing further to reduce specific systematics in our searches!
- Think about possible SM measurements when designing searches.

Top background validation

Check region with same selection as measurement but with one b-tagged jet



Unfolding region

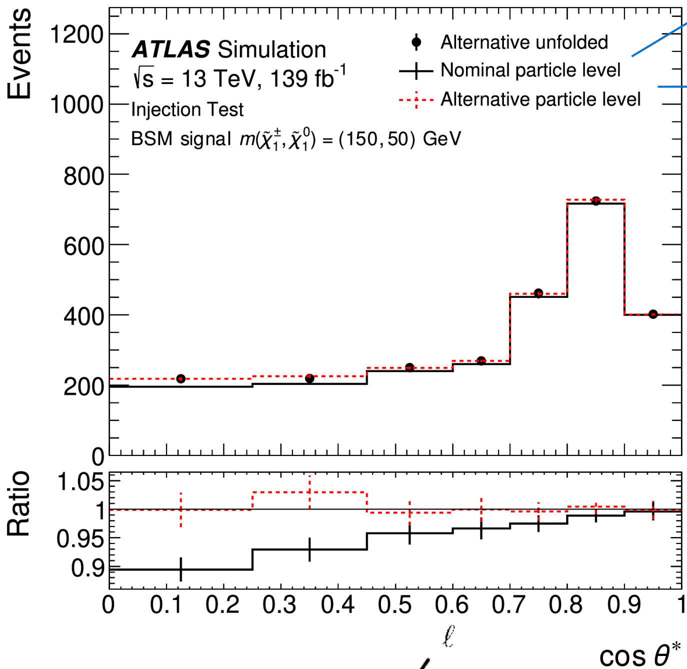


VR-top

Good agreement seen in VR-top for all distributions considered!

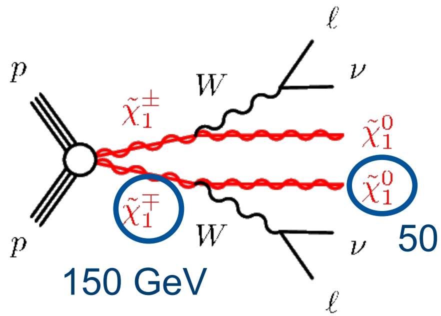
BSM injection tests

Check that unfolding calculation could correctly recover injected detector-level BSM signal

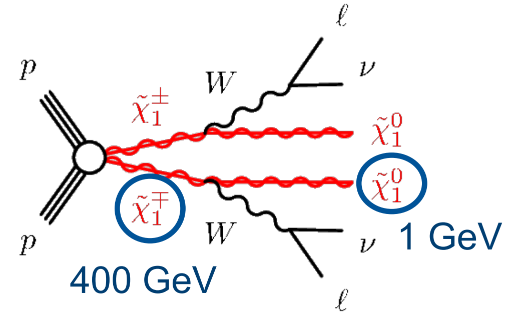


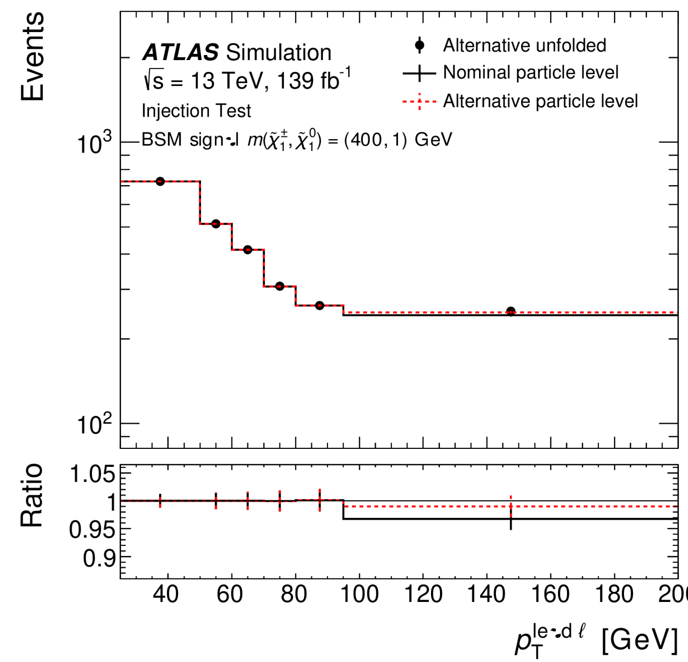
Particle-level WW

Particle-level WW + BSM



Try two signal points on the edge of exclusion from EWK search…





Important checks of the validity of using results for reinterpretation!

Stress tests

- Aim to estimate any bias introduced using information from nominal signal MC sample in the unfolding calculation.

- Data-driven stress-test: reweight simulated signal MC at generator level to obtain better improvement between signal and background-subtracted data at detector level. Check that reweighted detector-level signal unfolds to the reweighted particle-level distribution.

    - Perform using unfolded variables and ”hidden” variables (in this case MET and m_{T2}) => important for BSM interpretations.

Comments and points for discussion

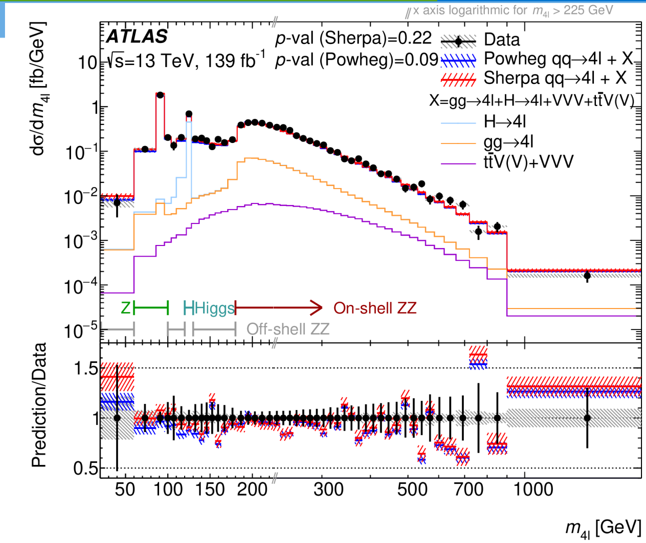
https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/STDM-2018-30/

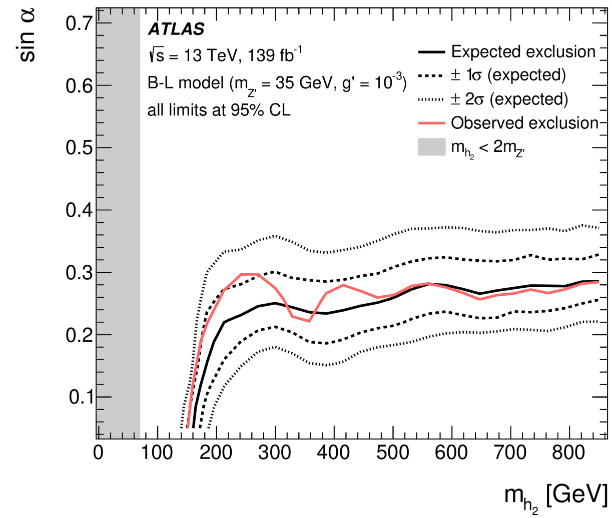
- Unfolding in search CRs/VRs can improve background modelling and provide constraints on new physics – could we extend the unfolding calculation into search regions?

- Such techniques could very naturally use PL unfolding- opportunities/challenges?

- ML classifiers are becoming widespread in search design- how will that impact prospects for unfolding?

- Right level of model-dependence/independence when designing searches/measurements? (see ATLAS 4l measurement on right).

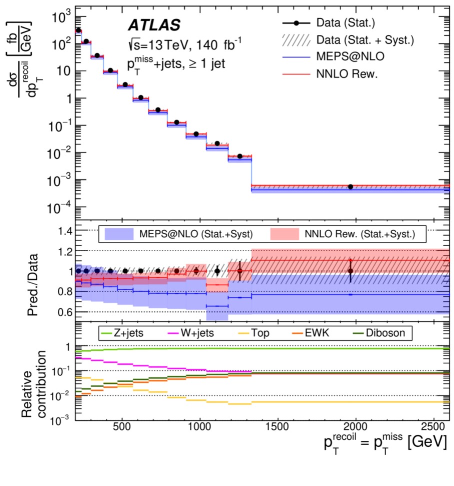


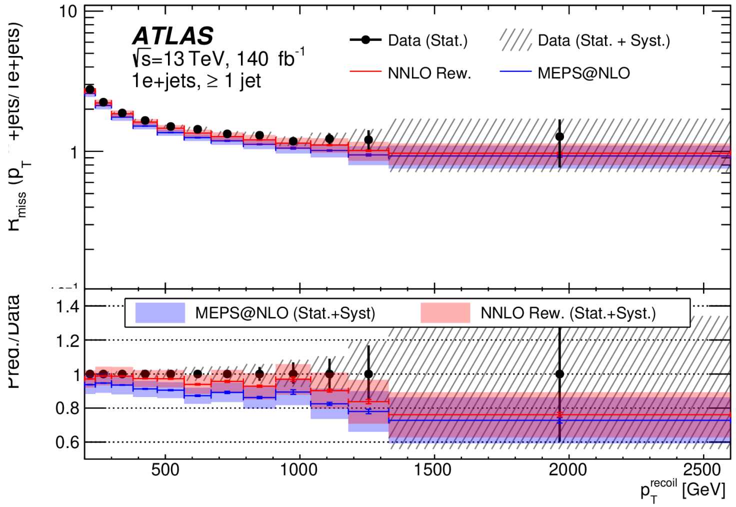


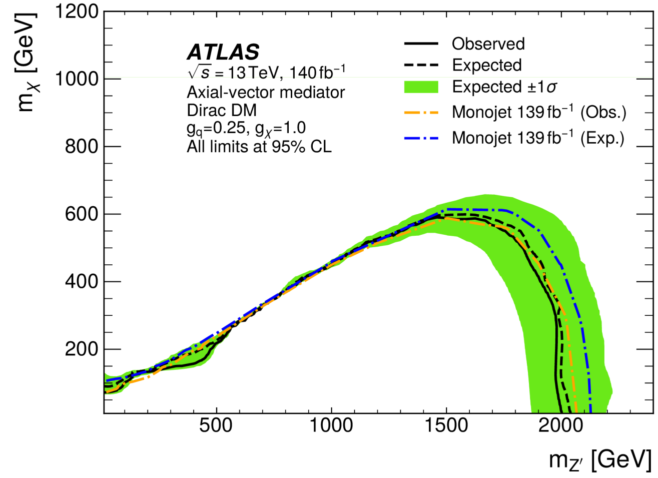
New ATLAS result(s): E_T^{miss} + jets differential measurement(s)

https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/STDM-2018-55/

- Differential measurement of E_T^{miss} and jets, with auxiliary measurements of jets recoiling against isolated leptons and photons in same phase space.
- Ratios between distributions benefit from cancellations in modelling and major systematic uncertainties and provide competitive constraints on new physics to direct searches despite remaining model independent.

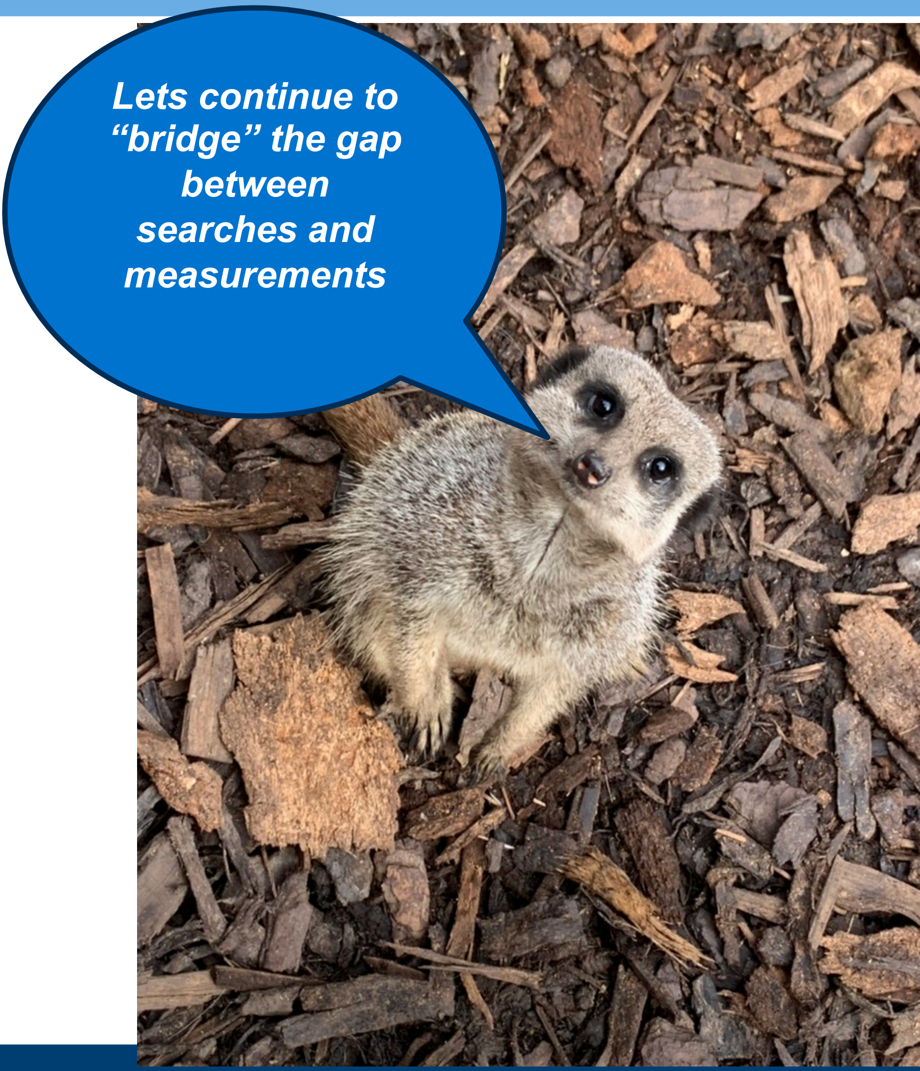






Conclusion

- Have presented a detailed example of an ATLAS result performing unfolding in a SUSY-inspired phase space.
- As well as the points for discussion on slide 24- would be very interested in thoughts/feedback from the statistics community on the unfolding methodology/optimization process.
- Feel free to reach out (sarah.louise.williams@cern.ch) with further thoughts/questions.

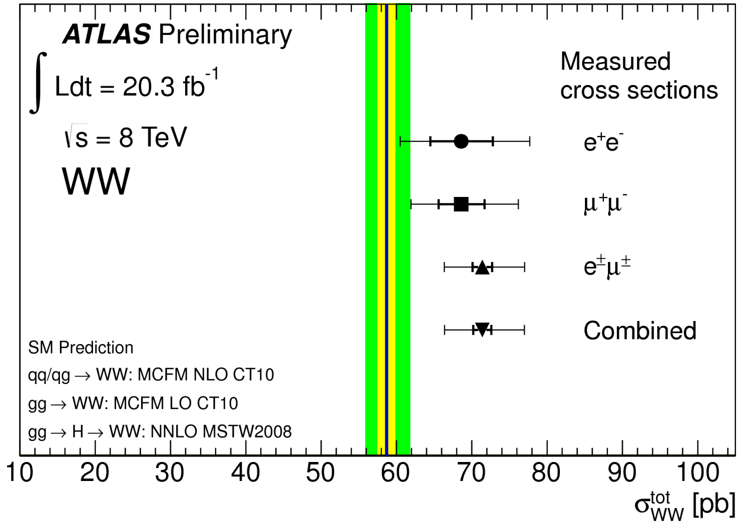


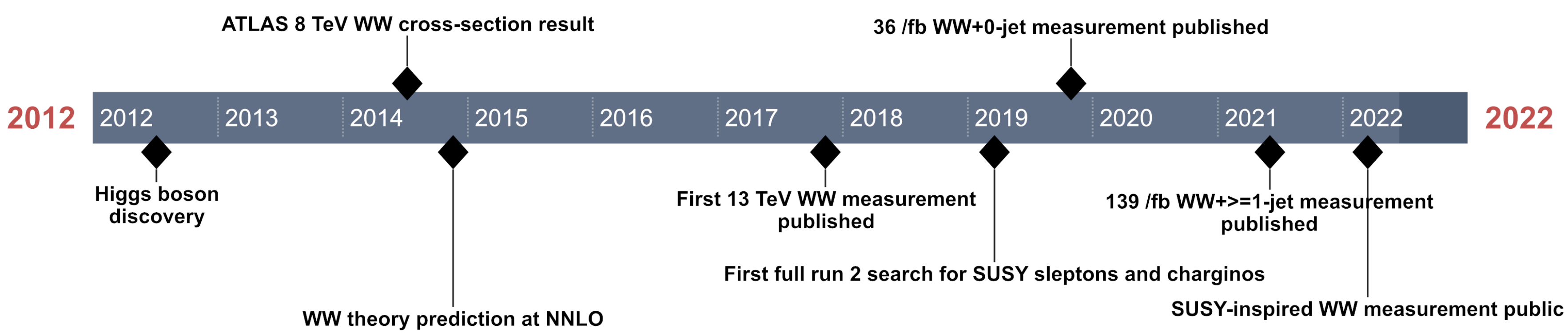
Some history up to 2022

WW precision measurements a first hint of new physics at the LHC?

Quote from physics briefing

"And yes, we should also mention that the WW cross section result comes out a bit high compared to its Standard Model expectation. Not statistically significant, but enough to intrigue theorists and experimentalists to study this tricky channel in more detail."





Some history up to 2022

SUSY to the rescue?

https://arxiv.org/abs/1303.5696

High Energy Physics – Phenomenology

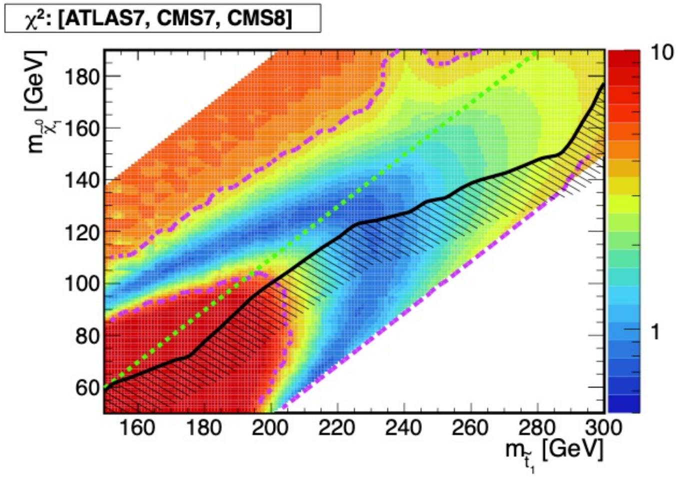
[Submitted on 3 Jun 2014 (v1), last revised 1 Dec 2014 (this version, v3)]

`Stop' that ambulance! New physics at the LHC?

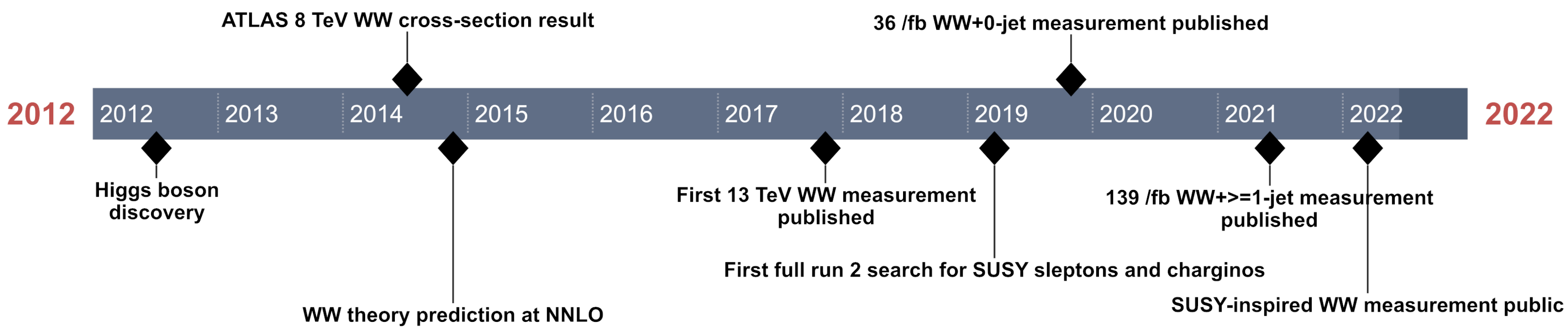
Jong Soo Kim, Krzysztof Rolbiecki, Kazuki Sakurai, Jamie Tattersall

A number of LHC searches now display intriguing excesses. Most prominently, the measurement of the W^+W^- cross-section has been consistently $\sim 20\%$ higher than the theoretical prediction across both ATLAS and CMS for both 7 and 8 TeV runs. More recently, supersymmetric searches for final states containing two or three leptons have also seen more events than predicted in certain signal regions. We show that a supersymmetric model containing a light stop, winos and binos can consistently match the data. We perform a fit to all measurements and searches that may be sensitive to our model and find a reduction in the log-likelihood of 15.4 compared to the Standard Model which corresponds to 3.5-σ once the extra degrees of freedom in the fit are considered.

https://arxiv.org/abs/1406.0858

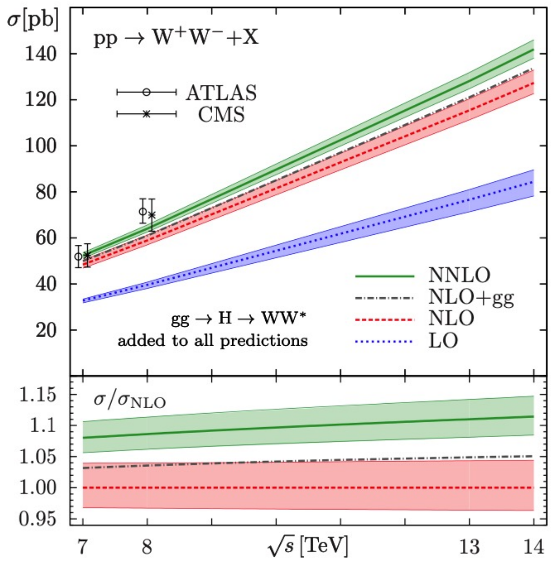


Alternative explanations included charginos and/or sleptons…



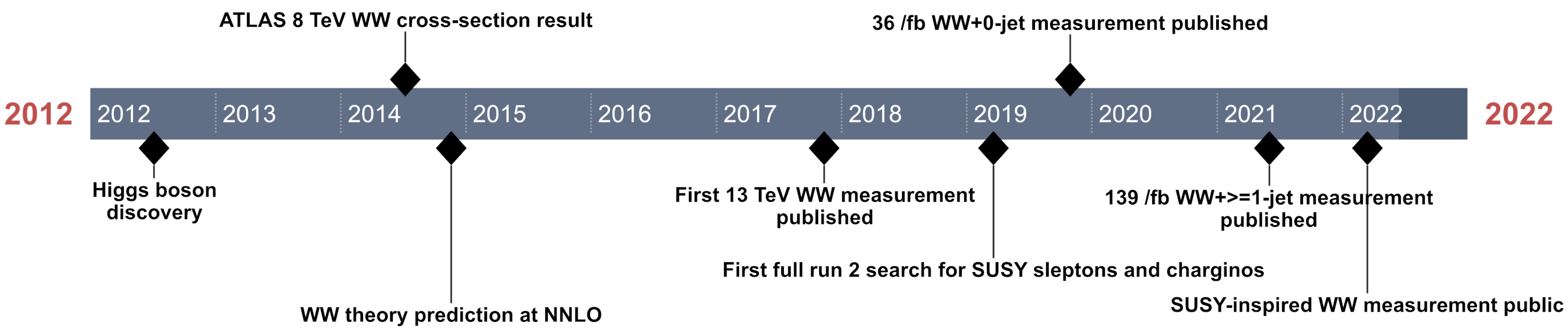
Some history up to 2022

https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/STDM-2013-07/



- Jet veto adds introduces an additional scale into the theoretical calculation -> complicates NNLO calculations/approximations.
- Full calculations of WW to NNLO in QCD reduced tensions with SM measurements

https://arxiv.org/pdf/1408.5243.pdf



Some history up to 2022

Plus exciting developments including $\gamma\gamma \rightarrow WW$ production

Recent run-2 measurements in di-leptonic final states

- 36 fb^{-1} WW+0-jet measurement: ***Eur. Phys. J. C 79 (2019) 884***
- 139 fb^{-1} WW+≥1-jet measurement: ***JHEP 06 (2021) 003***

… and the result I will discuss on WW measurements in a SUSY-inspired topology…