https://sites.google.com/view/mark-d-goodsell

# Automating the link between strings and BSM


Mark Goodsell

@PitifulRed

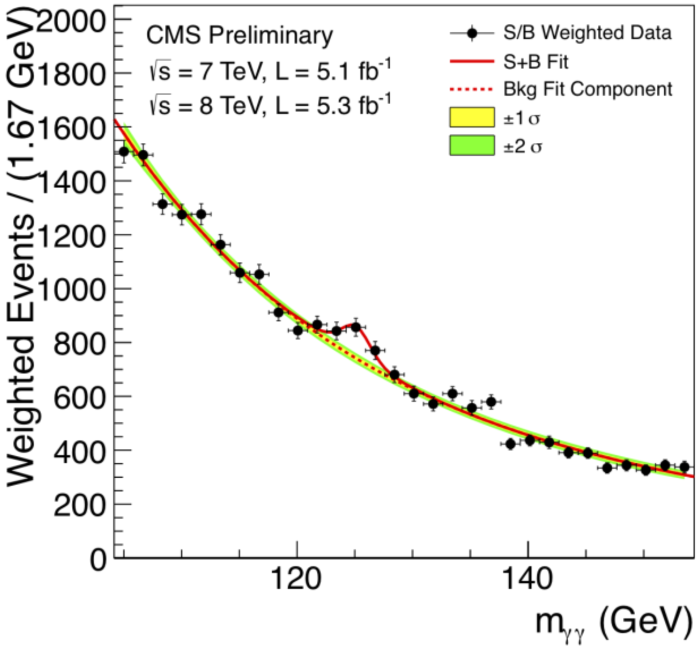
https://realselfenergy.blogspot.com

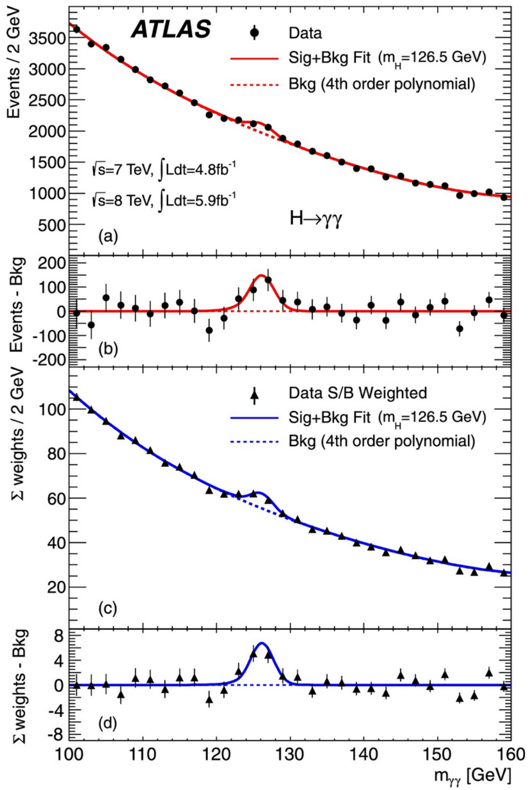
LABORATOIRE DE PHYSIQUE THEORIQUE ET HAUTES ENERGIES

# Overview

- How to connect your favourite model to BSM
- Latest developments in generic models
- Active Learning parameter space exploration

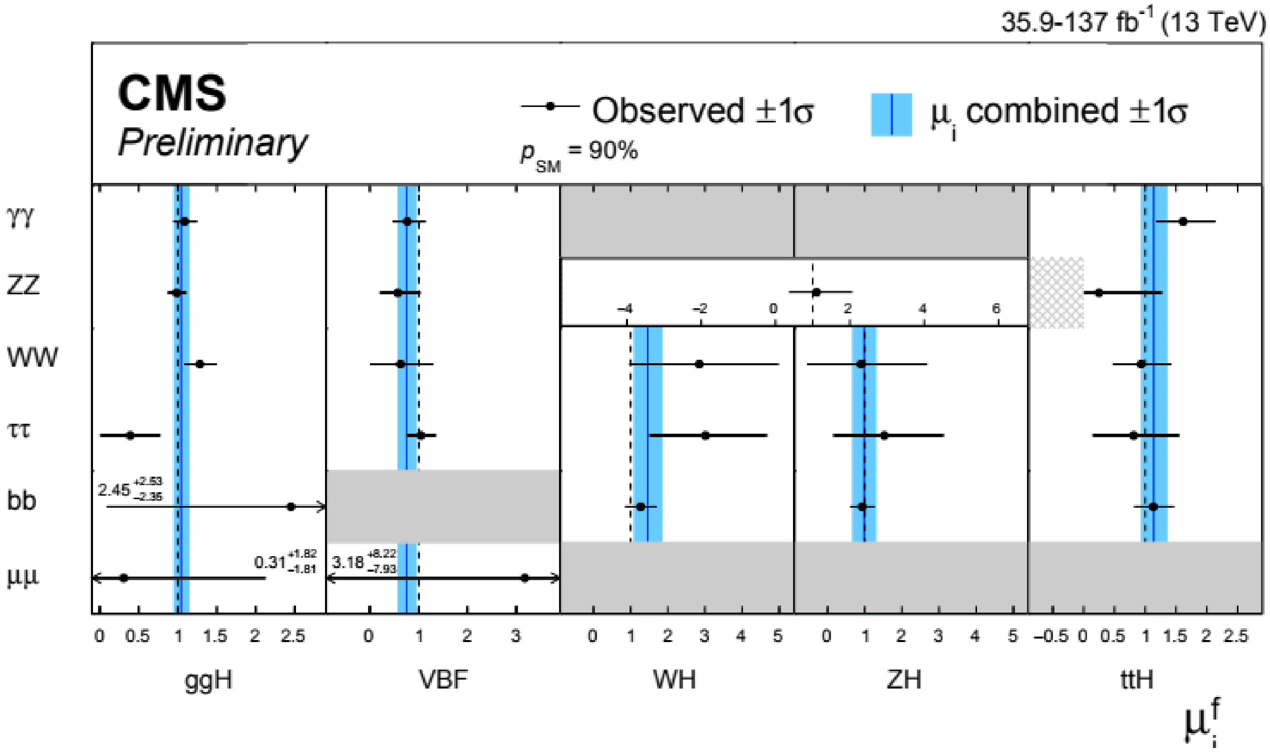
# Happy 10th anniversary of the Higgs discovery!

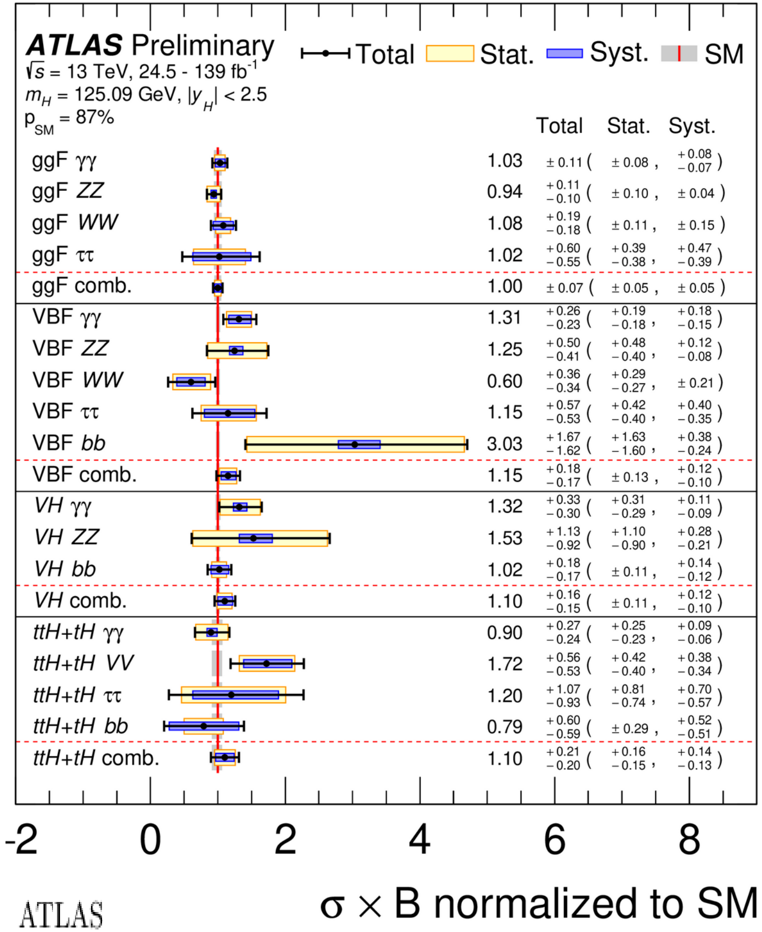



There was a symposium yesterday at CERN

With 10 years of data, we have much more precise measurements of its couplings:




$$\mu = 1.06 \pm 0.07 \text{ (ATLAS)}, \qquad \mu = 1.02^{+0.07}_{-0.06} \text{ (CMS)}.$$

$$\mathrm{BR}(h \to \text{invisible}) < \begin{cases} 0.145 & \text{ATLAS} \\ 0.18 & \text{CMS} \end{cases}$$

**And run 3 starts today!**

# Anomalies to keep looking out for

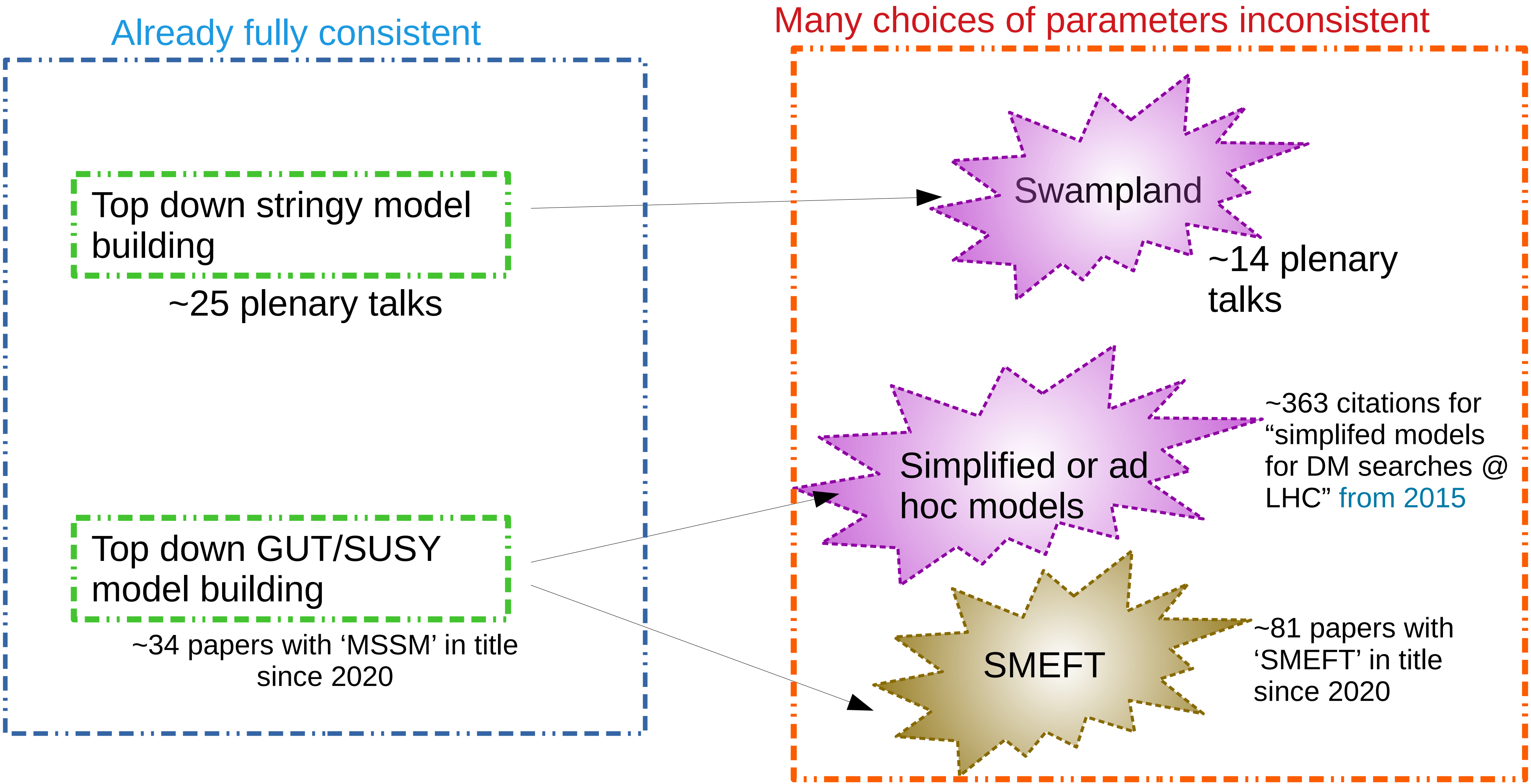
- W boson mass (CDF measurement 7σ above SM prediction!)
- B meson anomalies, may be due to leptoquarks, exotic Z' ...
- Muon g-2: longstanding discrepancy between experiment and theory now looks like discrepancy between lattice (now three independent calculations) and R-ratio. But it may yet be new physics.
- Three 3σ anomalies reported at Moriond 2022.
- Xenon 1T excess
- ...

I will assume you want something more than the SM at low energies!

- If you only predict heavy or decoupled BSM physics → maybe you can predict the Higgs mass? Maybe you have a thermal DM relic that is heavy, or flavour constraints? But otherwise I can’t help.
- LHC limits on colourful particles are O(TeV), but much lighter for EW particles (e.g. higgsinos are notoriously hard to find)
- If you are willing to bet on something at accessible energies: is there a good reason (anymore) to consider only the MSSM from your favourite string construction?

Want (and now have!) *generic* tools for ***any*** model

In recent years, there has been a distinct trend in the both string and BSM community:



From a **top-down** construction:

- Specify some set of fields and pattern of masses at high energy scale
- Potentially have a constrained set of UV parameters (e.g. mSUGRA ~ 5 parameters) but large set of observables we want to check

- EWPT
- Higgs mass
- Higgs couplings
- Rare decays

→ Compute spectrum and decays with SARAH → Check with HiggsBounds, HiggsSignals, Lilleth, flavio, smelli, ...

- Dark matter

→ Produce CalcHEP /UFO from SARAH, compute with MicrOmegas/MadDM/...

- LHC searches

→ Produce UFO from SARAH, events with MadGraph (or maybe Pythia) → MadAnalysis, SModelS, ColliderBit, Hackanalysis...

- ...

From the **bottom up:**

- May have *many* parameters not fixed by observations, e.g.
  - e.g. > 100 in MSSM
  - 2499 in SMEFT
- Don't necessarily have good priors
- Need to apply experimental and ***theoretical*** constraints:

- EWPT
- Higgs mass
- Higgs couplings
- Rare decays
- Dark matter
- LHC searches
- ...

- Vacuum stability
- Landau pole vs cutoff
- Unitarity / Positivity bounds
- ???

Not very well exploited

# (My) recent developments in genericity

- Unitarity for scalars at any centre of mass energy, including colourful scalars
- In paper with S. Passehr in 2019 I presented the full two-loop generic scalar self energies … the next challenge will be to implement these for the Higgs mass.
- Improvements to the W mass calculation
- Improvements to the muon g-2 calculation, EDMs, …

To be made public soon

- New frontier in recasting: Long Lived Particles

# Muon g-2 and EDMs

Only a handful of SMEFT operators are important for lepton g-2:

Nowadays if we can compute the SMEFT coefficients, can include running effects equivalent to the leading logs of 2-loop fixed-order … but more precise because we resum the logs

Implementation of (most of) these in SARAH will be available soon

Dipoles, contribute at leading order (one loop)

$$\begin{aligned}
\Delta a_\ell^{250\text{ GeV}} = \frac{m_\ell}{m_\mu}\mathrm{Re}\Bigg[ & \begin{matrix}2.9_\mu\\2.8_e\end{matrix}\times 10^{-3}\widetilde{C}_{\substack{eB\\\ell\ell}} - \begin{matrix}1.6_\mu\\1.5_e\end{matrix}\times 10^{-3}\widetilde{C}_{\substack{eW\\\ell\ell}} \\
& - \begin{matrix}4.3_\mu\\4.1_e\end{matrix}\times 10^{-5}\widetilde{C}^{(3)}_{\substack{\ell equ\\\ell\ell 33}} - \left(2.6 + 0.37 c_T^{(c)}\right)\times 10^{-6}\widetilde{C}^{(3)}_{\substack{\ell equ\\\ell\ell 22}} \\
& - 7.9\times 10^{-8}\widetilde{C}_{\substack{\ell e\\\ell 33\ell}} + \left(5.7 c_T - \begin{matrix}0.49_\mu\\0.48_e\end{matrix}\right)\times 10^{-8}\widetilde{C}^{(3)}_{\substack{\ell equ\\\ell\ell 11}} + 1.4\times 10^{-8}\widetilde{C}^{(1)}_{\substack{\ell equ\\\ell\ell 33}} \\
& + \left(\begin{matrix}10_\mu\\9.8_e\end{matrix} + 2.5 c_T^{(c)}\right)\times 10^{-9}\widetilde{C}^{(1)}_{\substack{\ell equ\\\ell\ell 22}} - \begin{matrix}4.6_\mu\\4.7_e\end{matrix}\times 10^{-9}\widetilde{C}_{\substack{\ell e\\\ell 22\ell}} \\
& + \frac{m_\ell}{m_\mu}\Bigg\{\begin{matrix}2.5_\mu\\2.4_e\end{matrix}\times 10^{-8}\left(\widetilde{C}_{HWB} + i\widetilde{C}_{H\widetilde{W}B}\right) - \begin{matrix}1.8_\mu\\1.7_e\end{matrix}\times 10^{-8}\left(\widetilde{C}_{HB} + i\widetilde{C}_{H\widetilde{B}}\right) \\
& - \begin{matrix}6.0_\mu\\5.7_e\end{matrix}\times 10^{-9}\left(\widetilde{C}_{HW} + i\widetilde{C}_{H\widetilde{W}}\right) + 3.8\times 10^{-9}\widetilde{C}_{\substack{He\\\ell\ell}} - \begin{matrix}3.7_\mu\\3.6_e\end{matrix}\times 10^{-9}\widetilde{C}^{(1)}_{\substack{Hl\\\ell\ell}} \\
& + \begin{matrix}3.6_\mu\\3.3_e\end{matrix}\times 10^{-9}\widetilde{C}^{(3)}_{\substack{Hl\\\ell\ell}} + \begin{matrix}1.8_\mu\\1.7_e\end{matrix}\times 10^{-9}\widetilde{C}_{HD} + \begin{matrix}2.1_\mu\\2.0_e\end{matrix}\times 10^{-9}\widetilde{C}_{W} \\
& + 1.1\times 10^{-9} i\widetilde{C}_{\widetilde{W}}\Bigg\}\Bigg]
\end{aligned}$$

From Aebischer *et al*, 2102.08954

# W mass

In the SM, the W mass is known to 2-loop order

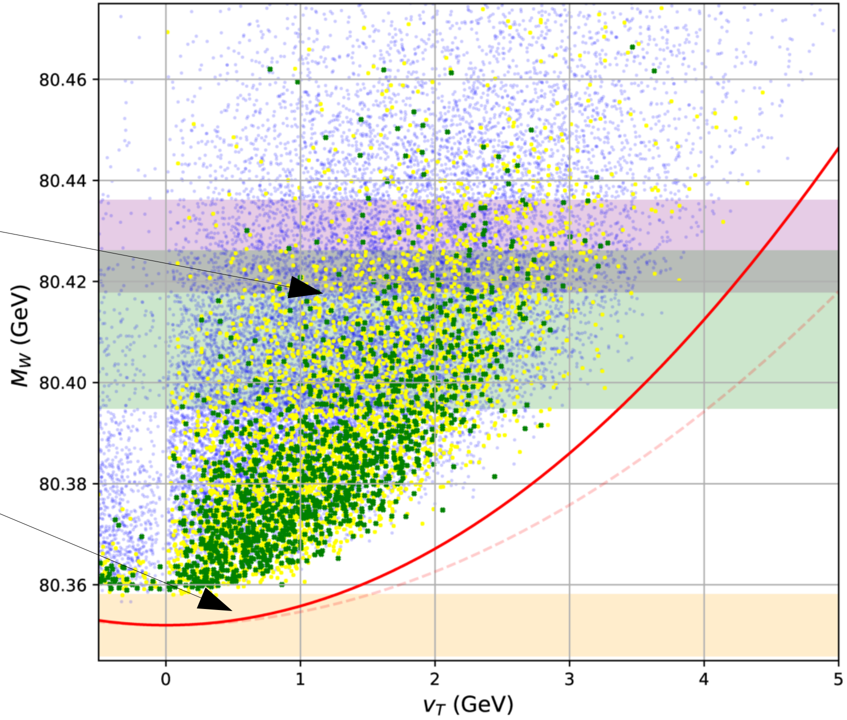
So we can use a pseudo-EFT computation, by subtracting SM contributions from BSM ones and adding to the known full SM computation:

Athron *et al*, 2204.05285

$$M_W^2 = (M_W^2)_{SM}\left(1 + \frac{s_W^2}{c_W^2 - s_W^2}\left[\frac{c_W^2}{s_W^2}(\Delta\rho_{\rm tree} + \Delta\rho) - \Delta r_W - \Delta\alpha\right]\right)$$

Can incorporate tree-level shifts, which occur in models with extra vevs, e.g. triplets in Dirac Gaugino models

$$\rho = 1 + 4\frac{v_T^2}{v^2} \longrightarrow M_W^2 = \rho c_W^2 M_Z^2$$




Or shifts from Z’ mixing with the Z boson … such as from heterotic-inspired Z’ models (work in progress with A. Faraggi)

# Parameter space exploration

Suppose I have some model to confront to data with a few parameters. Simplest (old school) approach is to run the codes on a grid or do a random selection

Usually people have their own codes (reinventing the wheel each time), but there existed SARAH Scan and Plot (Mathematica package).

Not much else AFAIK

- This is massively inefficient if the number of parameters is large! E.g. scan 10 points per variable. So no good for even pMSSM, don’t even think about full MSSM.
- Also there is the problem of how to combine constraints.

But: it’s simple to understand and very useful for making line plots & simple models.

# Likelihood sampling

A solution to the problem of combining constraints is to construct a likelihood function:

$$\mathcal{L} = \mathcal{L}_{\text{collider}} \times \mathcal{L}_{\text{DM}} \times \mathcal{L}_{\text{Higgs}} \times \mathcal{L}_{\text{Flavour}} \cdots$$

This can be a simple gaussian:

$$\mathcal{L}_i = \exp\left[ -\frac{(\mathcal{O}_i - \overline{\mathcal{O}}_i)^2}{2\sigma^2} \right]$$

Or we can take log likelihoods and add.

We can also take correlations into account

Markov-Chain-Monte-Carlo techniques give a sample of points where, after a long enough time $\rightarrow$ $\text{density} \propto \mathcal{L}$

There now exist sophisticated generalisations of this (e.g. multinest, diver, …) developed for other fields, but the aim is always to have points distributed proportional to the likelihood

The advantages claimed for this are that there is a rigorous statistical interpretation in a Bayesian/frequentist approach:

- Can make statements about statistics
- Can find ‘most likely point’
- Can use likelihoods to compare models

Numerous papers using this; e.g. Mastercode collaboration scanning the MSSM

As a result the community has mostly adopted this strategy.

**GAMBIT** now includes **GUM** (Gambit Universal Machine) which generates the model files from SARAH or FeynRules for any model.

Suppose we aren’t interested in having points around the most ‘likely’ regions, e.g.:

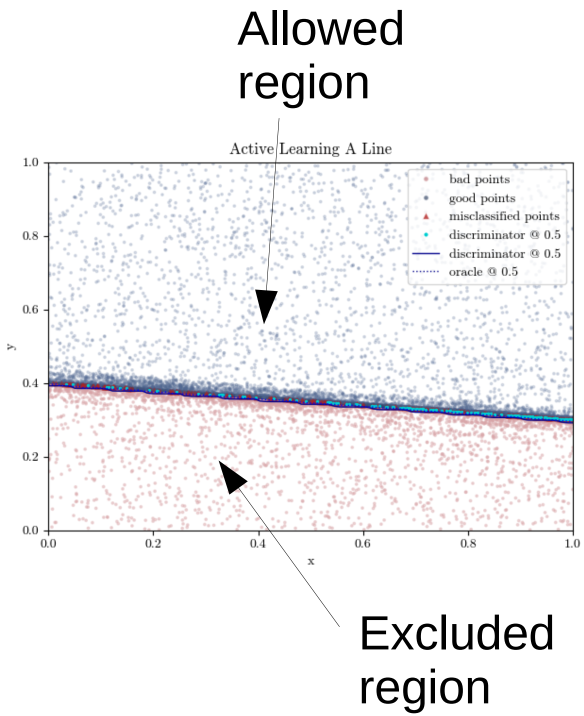
- If we ignore low-energy anomalies, why should new physics be very light? The likelihood function should be flat at high energies.
- What if we are interested instead in the exclusion boundary? By definition these are unlikely regions, all standard algorithms give only very few points there – need a supercomputer/long time to map it out.

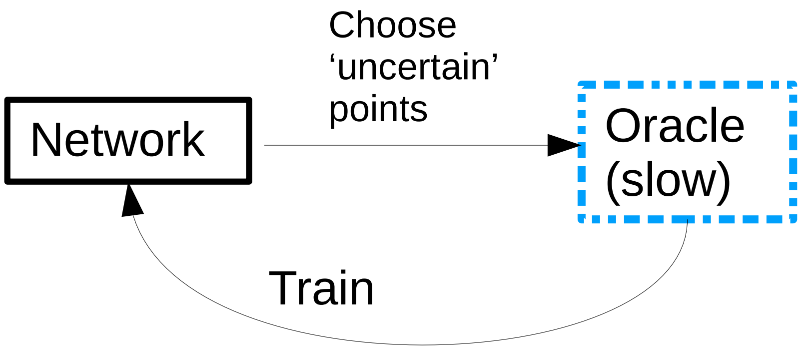
I argue that this is the sort of question many people want to ask instead – especially from string models where the overall mass scale may vary by orders of magnitude.

Want something simple that runs on a laptop that tells us what the allowed values are.

# Active learning

- The simple question: “is this point excluded or not” sounds automatically like a classification problem.
- Why not train a neural net to learn this?
- The question is then how to find the decision boundaries.
- Why not get the network to select points for you!
- Need a measure of the points that the network is most uncertain about.




The case of a slow physics ‘Oracle’ is exactly what we have from our chain of codes: we want to choose only the most interesting points to sample.

# Classic Active Learning

- Used for classification on discrete sets, e.g. language processing, looking up entries in a library.
- Then using a random forest can have an ensemble of models, ‘uncertain’ points are ones where there is no agreement among the models.
- Can refine to include taking batches of size N (can be slow to train the network) by including points that are sufficiently diverse (avoid choosing the same point N times) → KL divergence.
- We start with L random points and select K (<<L) from them with the best score to send to the oracle

# Active Learning with a Neural Network

In our case (2204.13950 with my student Ari Joury) we train a *neural network discriminator:*

- Use a sigmoid final layer → discriminator yields values between 0 and 1
- When the result is near 0.5 it is maximally uncertain → can assign a score for each point: $s_i = y_i(1 - y_i)$
- Choose L points, 10% purely at random, 90% within jumps of 'good' points (like MCMC)
- Need a diversity measure which depends on a *repulsion r* between points with physical distance *d.*

$$r(\underline{x}_i, \underline{x}_j) = \begin{cases} -\frac{a}{a+d^2}, & d < 0.01 \\ 0 & d \geq 0.01 \end{cases}$$

- Then we create a batch of pK points iteratively (p depends on the performance of the discriminator):

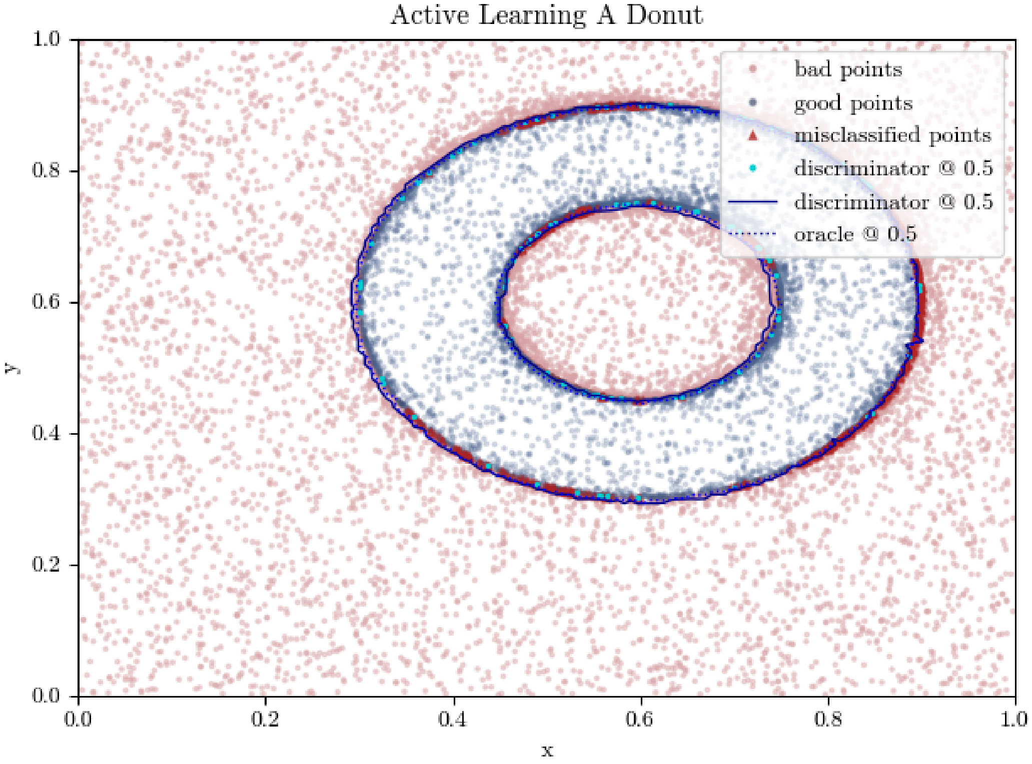
1. Compute $r_j \equiv \sum_i r(\underline{x}_i, \underline{x}_j)$ for $\{\underline{x}_i\} \in \overline{pK}, \{\underline{x}_j\} \in \overline{P}$ for each point $\underline{x}_j$
2. Compute the maximum total repulsion $r_{\max} = \max(\{r_j\})$ and the standard deviation $\sigma$ of the uncertainty scores $\{s_i\}$.
3. Assign to each point a score:

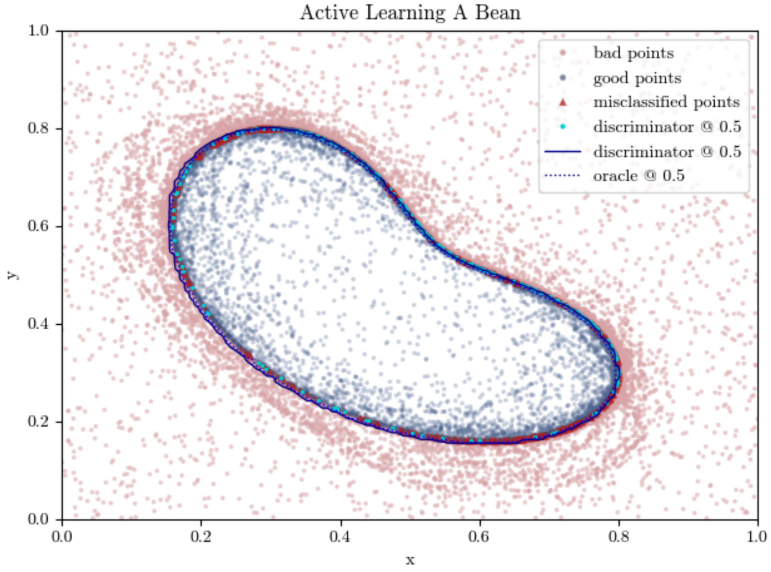
$$S_i = (1-\alpha)s_i + \alpha\ r_i\ \sigma \times \frac{1}{4r_{\max}}$$

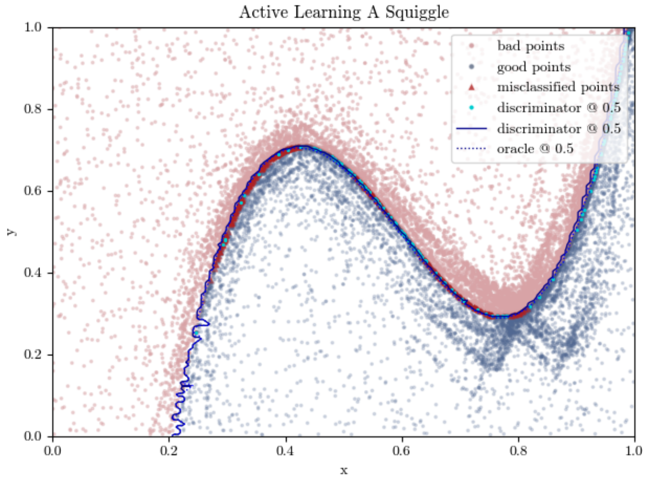
4. Add the point with the highest score $S_i$ to the set $\overline{pK}$ and remove it from the pool $\overline{P}$.

Remaining (1-p)K points are chosen randomly from the batch of L.

# Toy models


Active Learning A Donut
bad points
good points
misclassified points
discriminator @ 0.5
discriminator @ 0.5
oracle @ 0.5
x
y
0.0
0.2
0.4
0.6
0.8
1.0


Active Learning A Bean
bad points
good points
misclassified points
discriminator @ 0.5
discriminator @ 0.5
oracle @ 0.5
x
y
0.0
0.2
0.4
0.6
0.8
1.0


Active Learning A Squiggle
bad points
good points
misclassified points
discriminator @ 0.5
discriminator @ 0.5
oracle @ 0.5
x
y
0.0
0.2
0.4
0.6
0.8
1.0

Take a simple, concrete physics question: **how *heavy* could (thermal) dark matter be?**

There is a classic unitarity bound by Griest and Kaimionkowski

Total cross-section in partial waves is

$$\sigma_{CM}(a \to b) = \frac{4\pi}{(p^a_{CM})^2} 2^{\delta_a} \sum_J (2J+1)|a_J|^2$$

Leads to a bound for each partial wave of

$$\sigma_J \leq \pi \frac{2J+1}{(p^a_{CM})^2} 2^{\delta_a}$$

Gives a maximum mass for thermal DM!

$$\Omega h^2 \simeq 0.11 \times \frac{1.2 \times 10^{-9} (\mathrm{GeV})^{-2}}{\langle \sigma v \rangle}$$

But this is an absolute bound, since we take the maximum possible partial wave at the CM momentum in the annihilation

$$p^a_{CM} = \frac{1}{2} m v_{rel}$$

$$\Omega h^2 \propto m^2 \longrightarrow m_{\mathrm{DM}} \lesssim \mathcal{O}(100\ \mathrm{TeV})$$

We can do better by imposing that the *couplings* are perturbative ***at all scales up to the cutoff***

E.g. consider a toy model of DM scalar S annihilating to X where $m_S \gg m_X$

$$\mathcal{L}_{\text{toy}} \supset -\frac{1}{2}\lambda_{\text{toy}} S^2 |X|^2. \xrightarrow[\textit{Unitarity bound}]{} \lambda_{\text{toy}} < 8\pi\sqrt{2} \textit{ at tree level.}$$

If we calculate the cross-section for this maximum coupling we get

$$\sigma_0 \le 8\pi \frac{|\mathbf{p}_b|}{|\mathbf{p}_a| s}.$$

Momenta of X outgoing $p_b \simeq m_S$

Momenta of incoming S states $p_a \simeq m_S v/2$

For non-relativistic S, relativistic X this gives

$$\boxed{\sigma_0 \le \frac{2\pi}{m_S^2 v}}$$

BUT the absolute Griest-Kamionkowski bound is

$$\boxed{\sigma_0 \le \frac{2\pi}{m_S^2 v^2}}$$

But we are non-relativistic, v is small! → Unitarity of couplings is more stringent, gives lighter mass!

To showcase the DM + momentum-dependent unitarity + colourful unitarity want:

- (Neutral!) scalar DM
- Colourful mediators
- Trilinear couplings

Our example:

DM candidate

$$\mathcal{L} = \mathcal{L}_{SM} - \frac{1}{2}m_S^2 S^2 - m_E^2|Q_E|^2 - m_O^2|Q_O|^2 - \lambda_S S^4 - \frac{1}{2}\lambda_{HS}S^2|H|^2 - \lambda_3|H|^2|Q_E|^2 - \lambda_4|H|^2|Q_O|^2$$
$$- \frac{1}{2}\lambda_1 S^2|Q_O|^2 - \frac{1}{2}\lambda_2 S^2|Q_E|^2 - \lambda_5|Q_E|^4 - \lambda_6|Q_O|^4 - \lambda_7|Q_O|^2|Q_E|^2 - \lambda_8|Q_O Q_E^*|^2$$
$$- \left[\kappa_1 S Q_E Q_O^* + Y_Q^{ij} Q_E q_i q_j + \frac{1}{4}\lambda_C (Q_E Q_O^*)^2 + h.c.\right]$$

Colourful mediators, fundamental reps of SU(3)

Important trilinear coupling

$$m_O > m_S > m_E$$

So colourful odd states can decay to S + even

Have developed a code framework for scanning parameter spaces and linking together all relevant tools in the SARAH family, with plug+play scans (e.g. user written)

Will be made public soon

We compare our AL algorithm against a MCMC with a biased likelihood to find the maximum DM mass allowed by all constraints

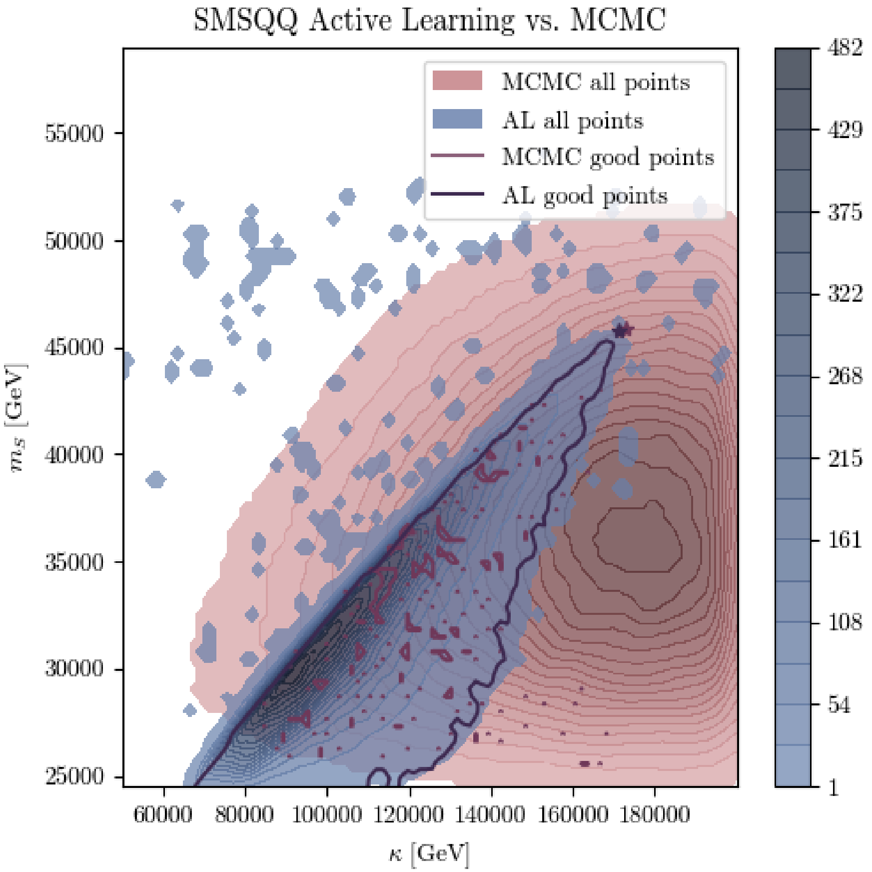
$$\mathcal{L}_{\mathrm{bias}}(m_S, \overline{m}_S, s) = \left(\frac{m_S}{\overline{m}_S}\right)^s$$

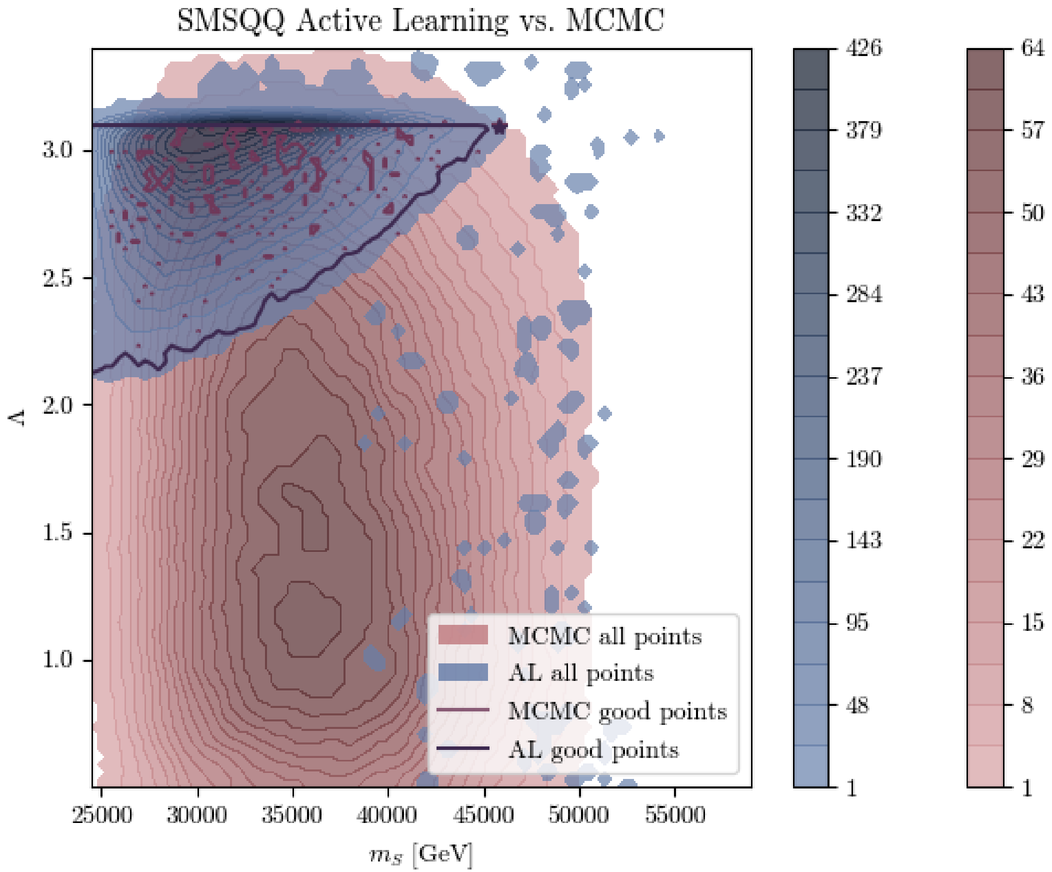
$$\mathcal{L} \equiv \mathcal{L}_{\mathrm{upper}}(\Omega h^2, 0.112, 0.001) \times \mathcal{L}_{\mathrm{upper}}(a_0, 0.5, 0.001) \times \mathcal{L}_{\mathrm{upper}}(\delta_{\mathrm{stability}}, 1, 0.2) \times \mathcal{L}_{\mathrm{bias}}(m_S, \overline{m}_S, 0.2)$$

$$\mathcal{L}_{\mathrm{upper}}(x, \overline{x}, s) \equiv \frac{1}{1 + \exp((x - \overline{x})/s)}$$

We were only interested in the maximum mass allowed (nb since we are using MicrOmegas we are ignoring issues such as Sommerfeld enhancement, bound states etc which should be relevant.

# Active learning vs MCMC




AL automatically searches out the boundaries

MCMC with bias samples most points in uninteresting regions

# Conclusions

- Have several other application of AL: e.g. training to find points with the correct Higgs mass in SUSY models.
- BUT lots more development needed in applying AL/ML to parameter space exploration.
- E.g. want to have automatic tuning of hyperparameters.
- More developments in spectrum generation and EFT approaches are in the pipeline.
- … and much yet to be exploited in purely theoretical constraints!
- Still much more to be done, even with unitarity

# BACKUP

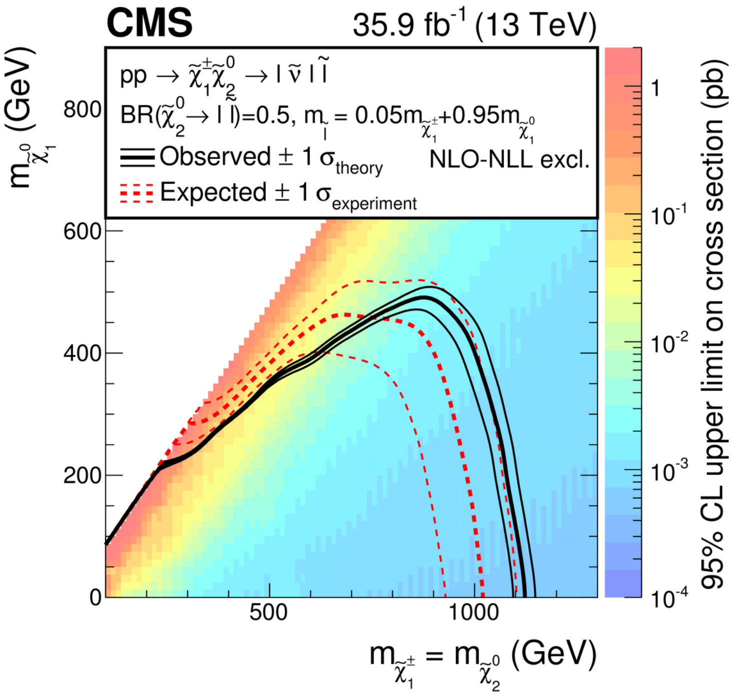
(CODA on LLPs)

- Most LHC searches are for promptly decaying particles
- Huge number of BSM searches from runs 1 and 2, various subsets have been reinterpreted in codes – can then be applied to different models with same signature
- Main frameworks are MadAnalysis (MA5), CheckMATE and ColliderBit (part of GAMBIT)

Limits on colourful particles are strong and well represented, BUT limits on electroweak sector are poor – and have more room for improvement

Heavy SUSY/split SUSY/minimal DM are classic examples which may be hiding under our noses

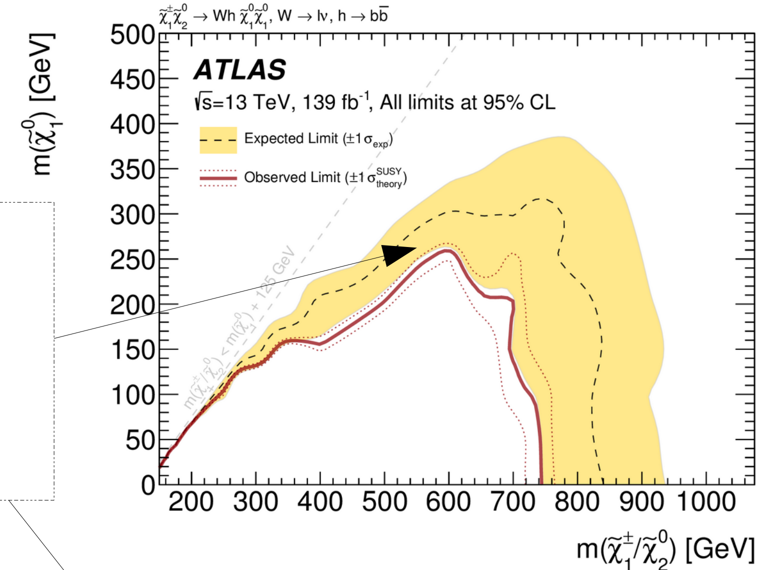
E.g. headline plots in MSSM from ATLAS/CMS:

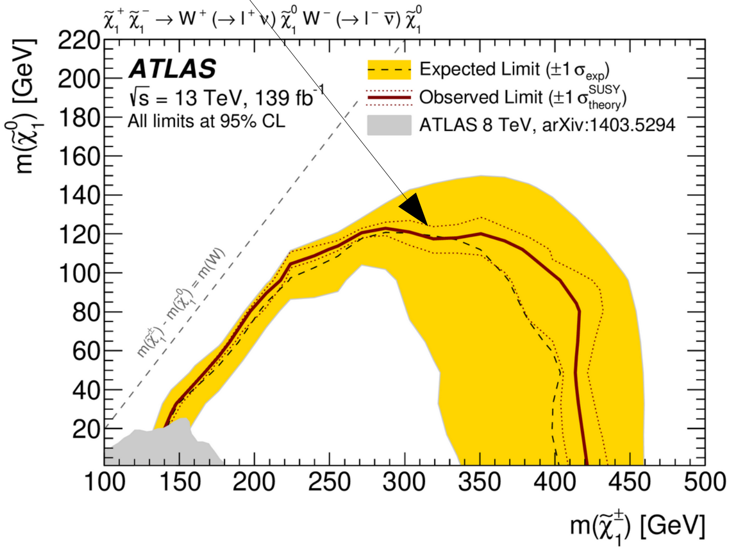



A very optimal case with light sleptons!

Essentially limits on bino/higgsino are very low compared to winos

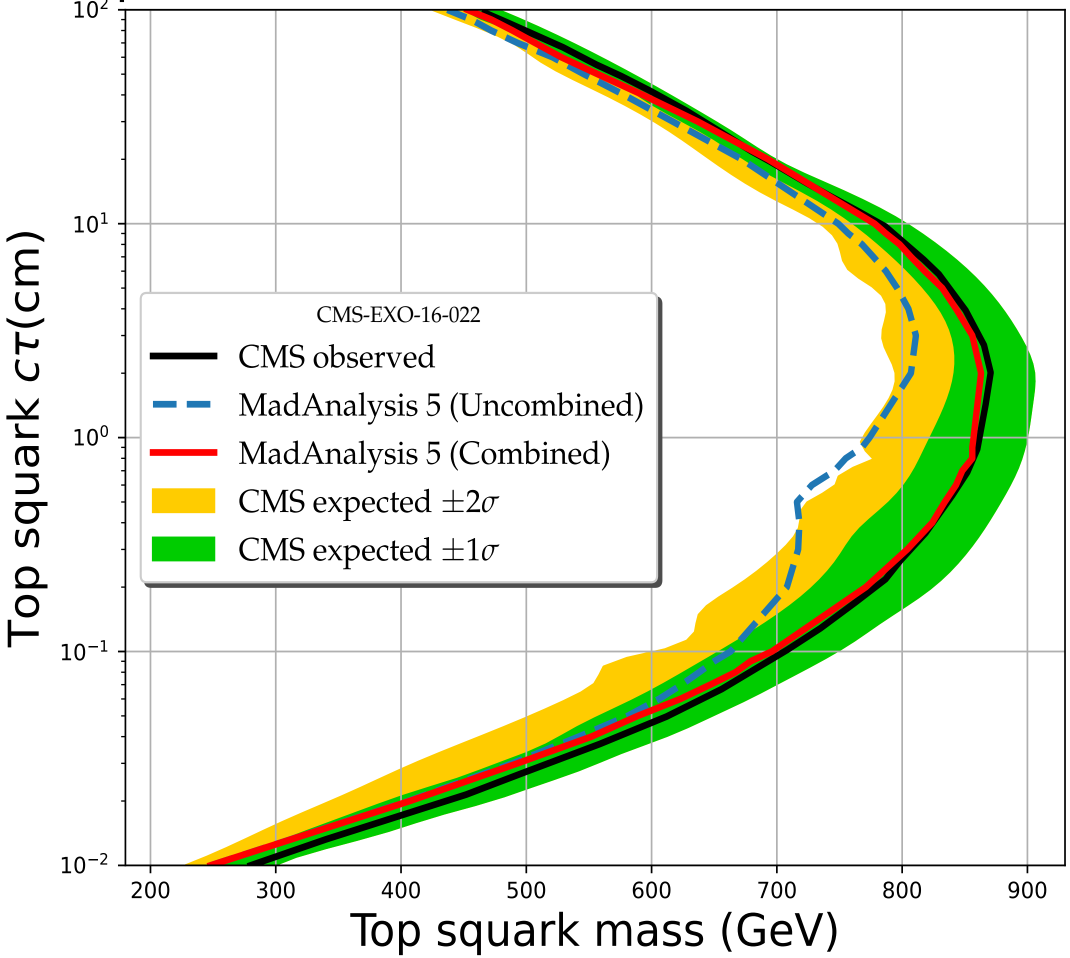
Actually rather weak limits on the LSP, and these are the best cases!



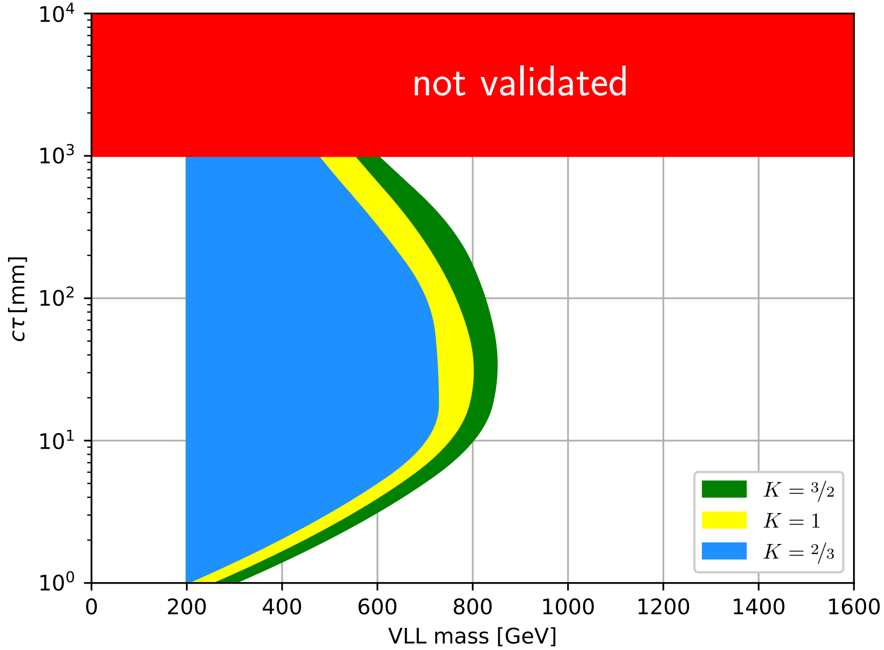



- Such models also may come with LLPs which stick out like a sore thumb!
- They are the current frontier in recasting and developing new LHC searches – large possibility of improvements in future!
- Only CheckMATE has a handful of LLP searches.
- There is an LLP recasting github with 5 older LLP analyses in
- I wrote a code called HackAnalysis (2106.08815) initially for a couple of searches
- … and now we have several LLP searches implemented in MadAnalysis 2112.05163 with J. Araz, B. Fuks and M. Utsch.

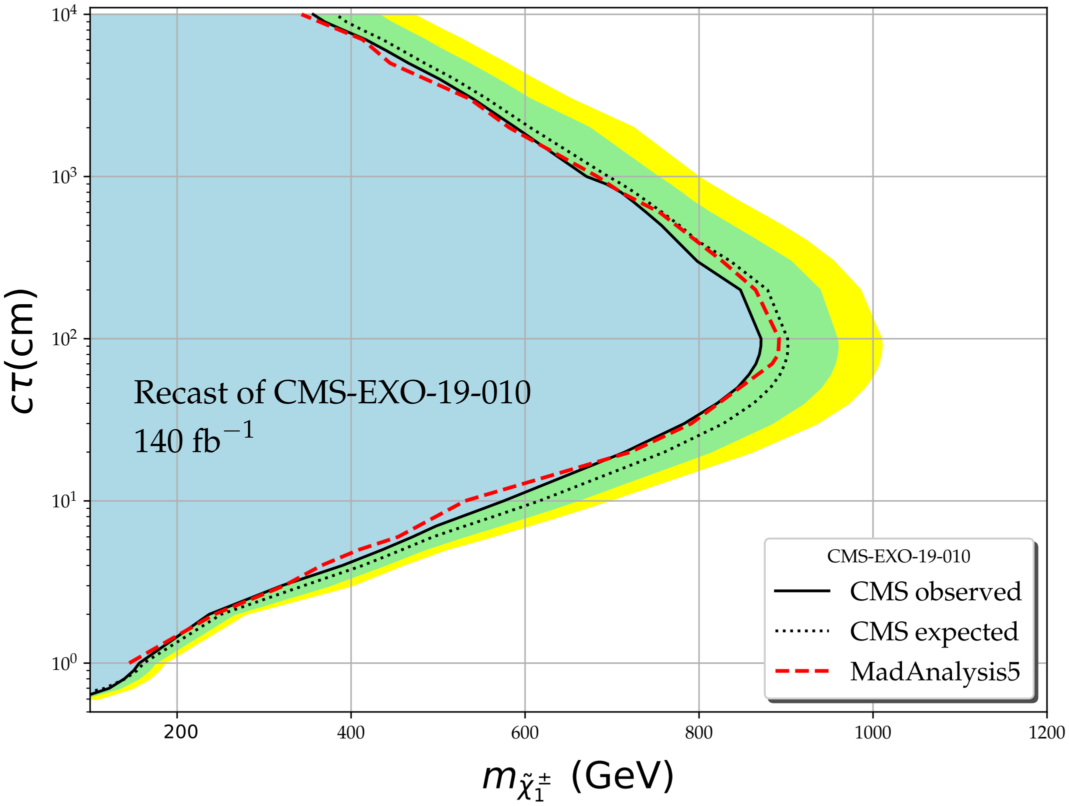
An early run 2 CMS displaced lepton search:



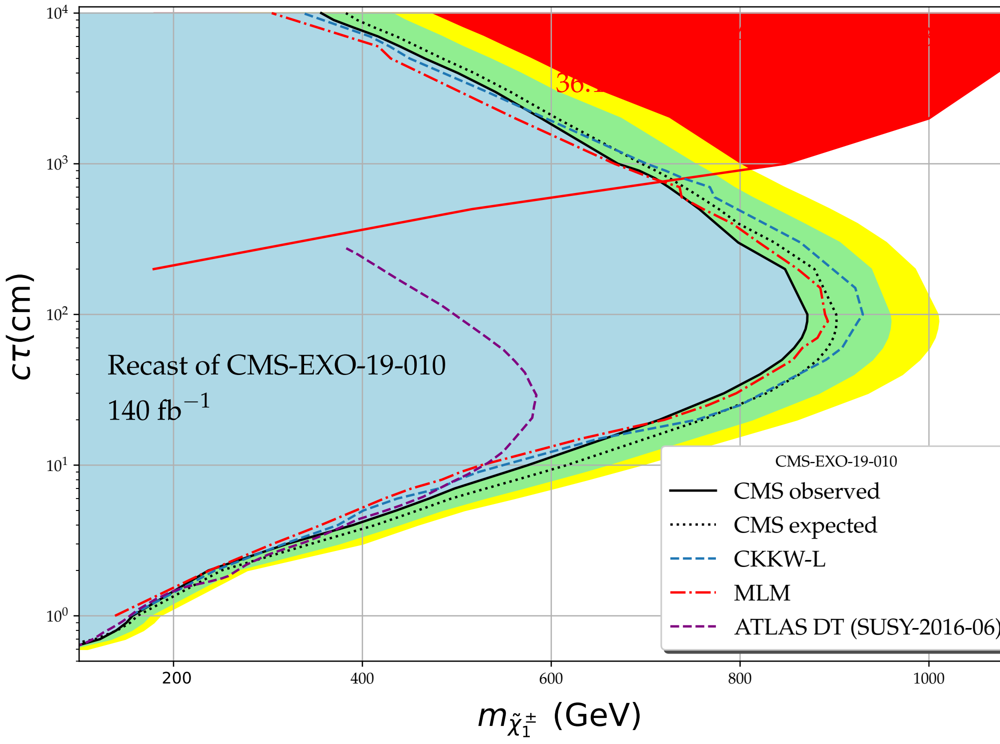

ATLAS displaced displaced leptons (RPV, ATLAS-SUSY-2017-04) search, here reinterpreted as vector-like leptons:



# CMS Disappearing track searches




MadAnalysis




HackAnalysis