Prelude

Prelude

Let's for a moment entertain the idea of a best case scenario:

we see signs of new physics in several final states at the LHC.

What would we do?

Prelude

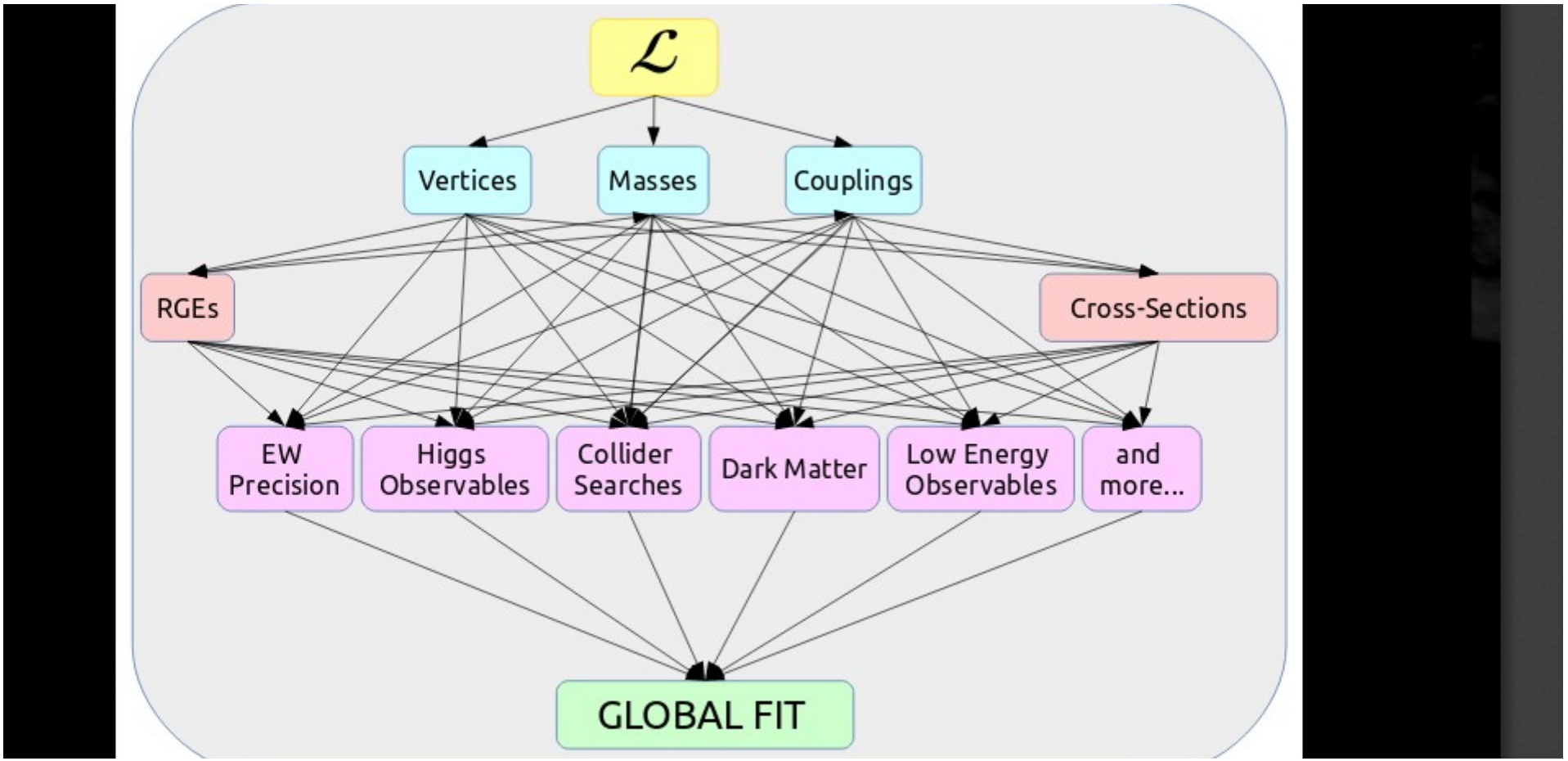
From Jamie's slides [*] (“top-down” approach):

- **We would write our favorite Lagrangian down**
- **We would press “enter”**
- **The computer would fit the model parameters to all relevant observables!**

[*] Would we do the same for positive results?

Prelude

From Jamie's slides:



Prelude

I would want to work on a “bottom-up” approach:

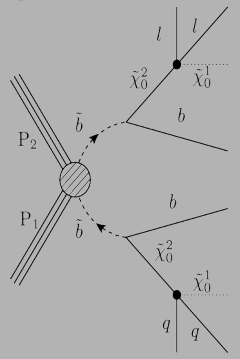
- **I would try to go bottom-up, starting from the data.**
- **I would tackle every excess individually.**
- **First I would try to find out the set of simplified models that best describes the data using Bayes factors.**
- **Then I would estimate the model parameters. The result would be a statistical description of a signal excess in the context of simplified models spectra:**

(From a talk I gave about the description of signal excesses using Bayes factors)

MOCKUP! Not real data, not even real simulation. assuming binned data.

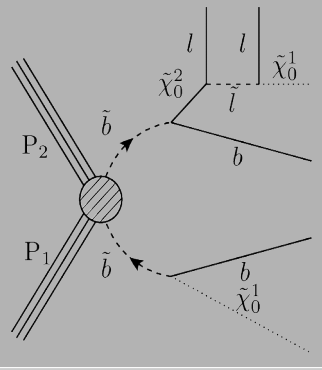
Signal model 1

e.g.

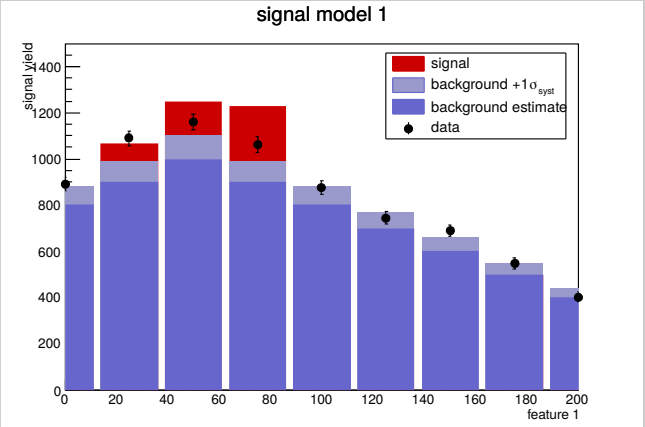


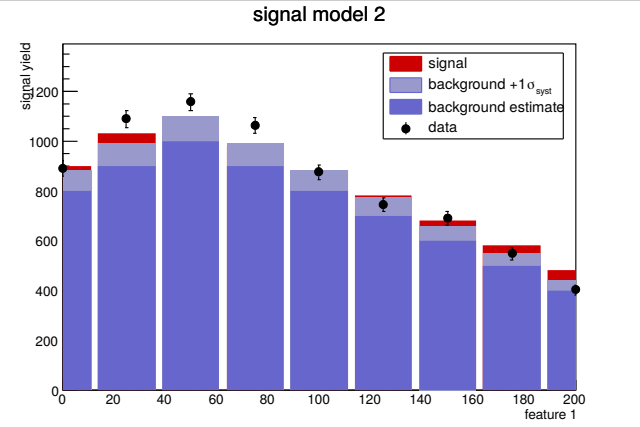
Signal model 2

e.g.

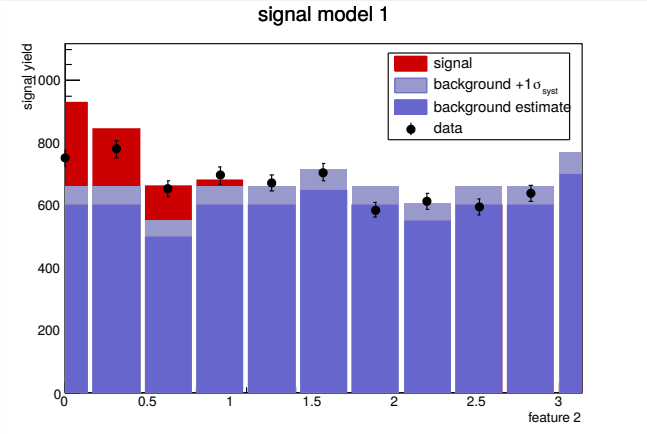


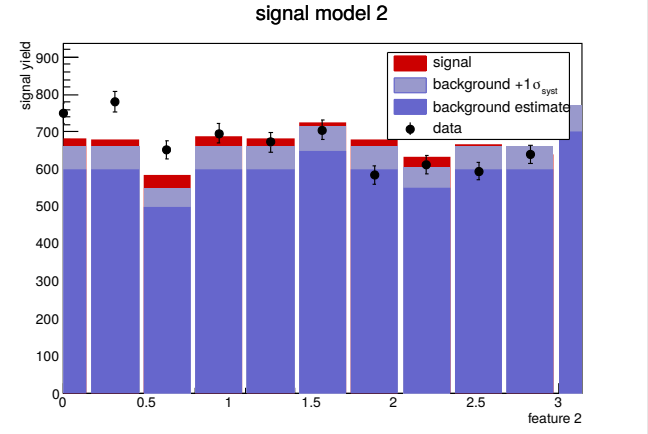
Feature 1 (e.g. „mll“)





Feature 2 (e.g. „φ(ll)“)





(From a talk I gave about the description of signal excesses using Bayes factors)

Bayes factors

- Bayesian alternative to frequentist hypothesis testing

- Given a model selection problem in which we have to choose between two models:

$$K = \frac{P(D|M_1)}{P(D|M_2)} = \frac{\int P(\theta_1|M_1)P(D|\theta_1, M_1)d\theta_1}{\int P(\theta_2|M_2)P(D|\theta_2, M_2)d\theta_2}$$

K = Bayes Factor

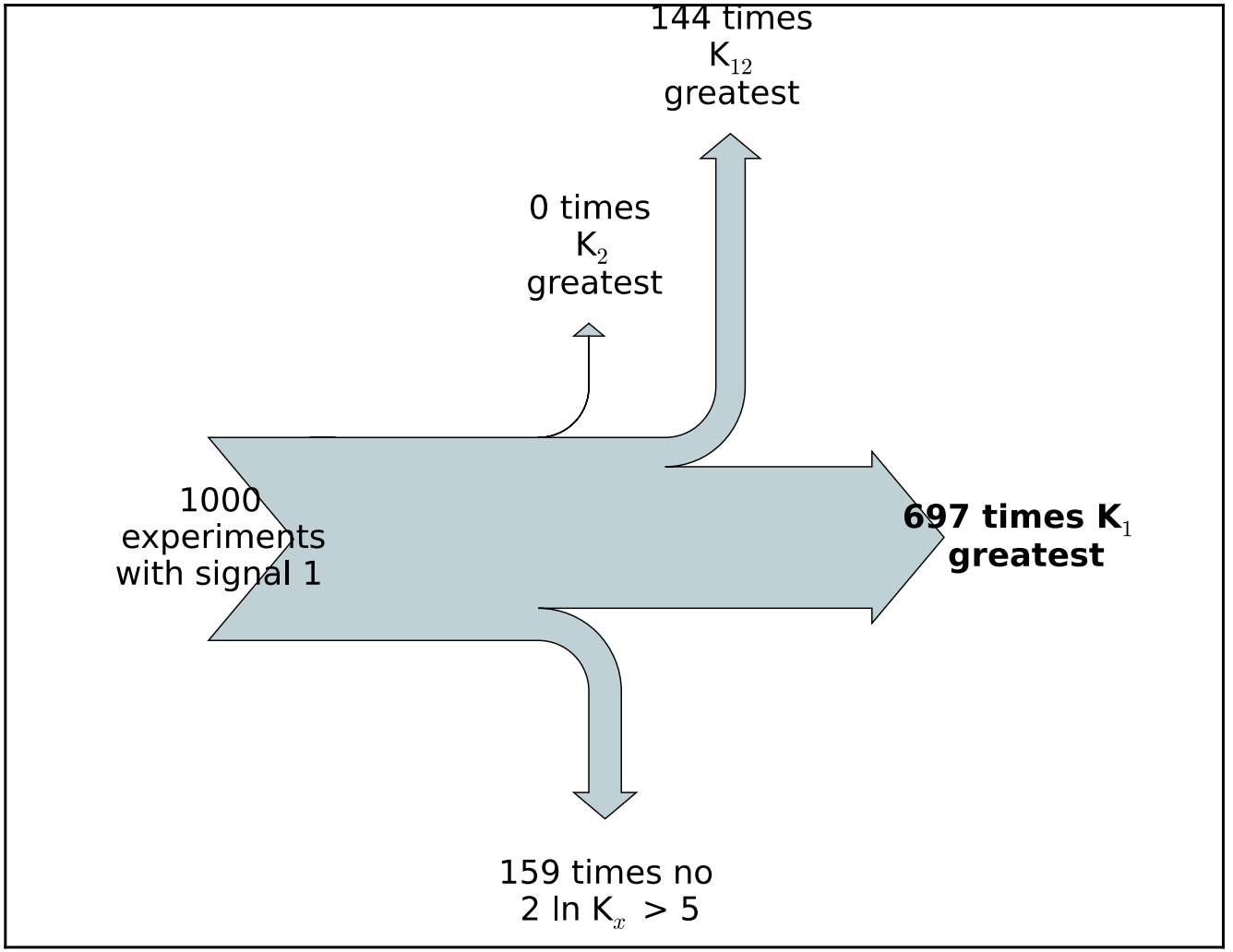
D = Data

θ = Model Parameter Vector: signal strength μ, mixing ratio r, particle masses m_j , assuming that nuisances are already dealt with

M_i = Model, e.g. T2tt or a mixture of models, e.g. T6bbWW + T2tt

(From a talk I gave about the description of signal excesses using Bayes factors)

Which signal model would be used to describe signal 1, feature_1 = 100 GeV?



„Pure“ signal model 1 is recognised in 70% of all cases, or in 83% of all „identifiable cases“. We never identify model 2 as the correct one.

Prelude

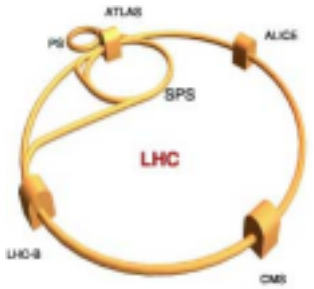
Only in a last step would I worry about “putting it all together” to infer our next, best, fundamental Lagrangian.

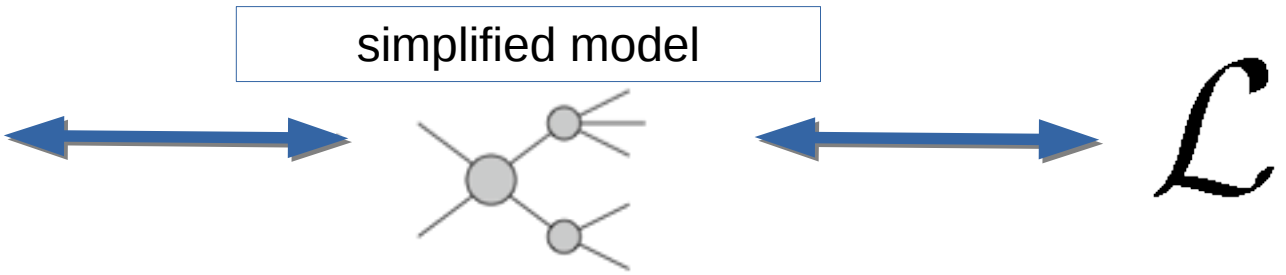
Top-down global fit

vs

Bottom-up incremental inference

(No doubt, as a community we want to see both approaches being followed.)





Prelude

The prelude, off-topic as it was, ends here.

Upgrading to LHC Run-2

Wolfgang Waltenberger
HEPHY Vienna

for the SModelS group

Federico Ambrogi, Sabine Kraml, Suchita Kulkarni, Ursula Laa, Andre Lessa, Veronika Magerl, Wolfgang Magerl, Doris Proschofsky, Michael Traub, Jory Sonneveld, WW <smodels-users@lists.oeaw.ac.at>

Fit(s) for LHC Run-2
Edinburgh, Scotland, 10 – 12 october 2016

Table of content

- The Idea behind SModelS
- Experiment
  - Anatomy of experimental SMS results
  - Formal description of the applicability of an SMS result
  - Construction of the SModelS database
  - Construction of likelihoods from the SMS results
- Theory
  - Decomposition of a model
  - Computation of cross sections
  - Compressing the spectra

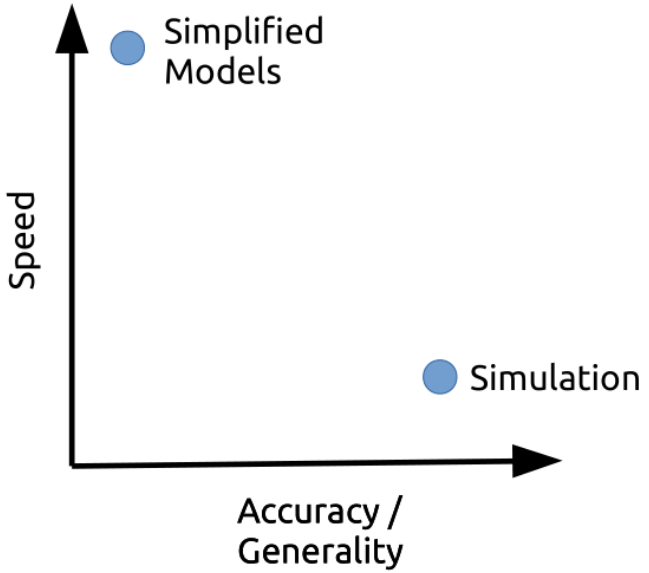
- Applications
- SModelS v1.1
- The Future

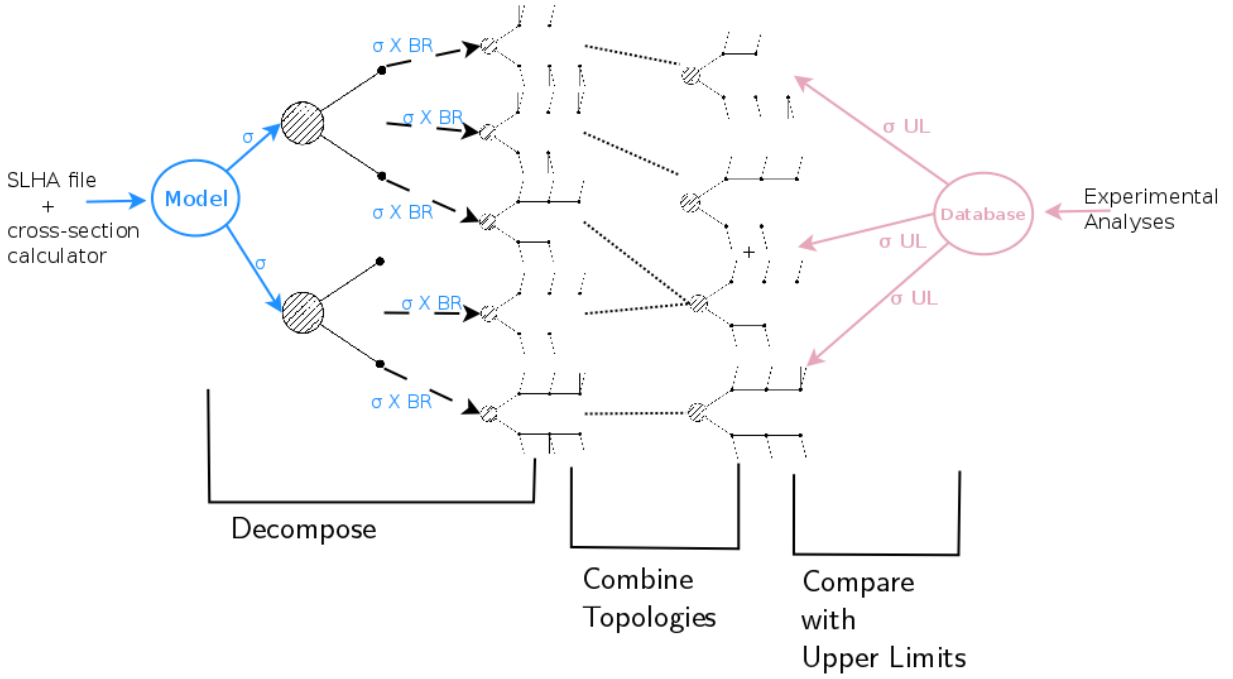
The idea behind SModelS

The Idea behind SModelS

SModelS confronts BSM theory with LHC results by decomposing full models into their simplified models topologies, and comparing

the cross section predictions of these individual topologies with a database of SMS results.





The Idea behind SModelS

SModelS is very soon to release v1.1 (estimate is November, we are in the final, polishing stages of the release).

Novelties in v1.1 are:

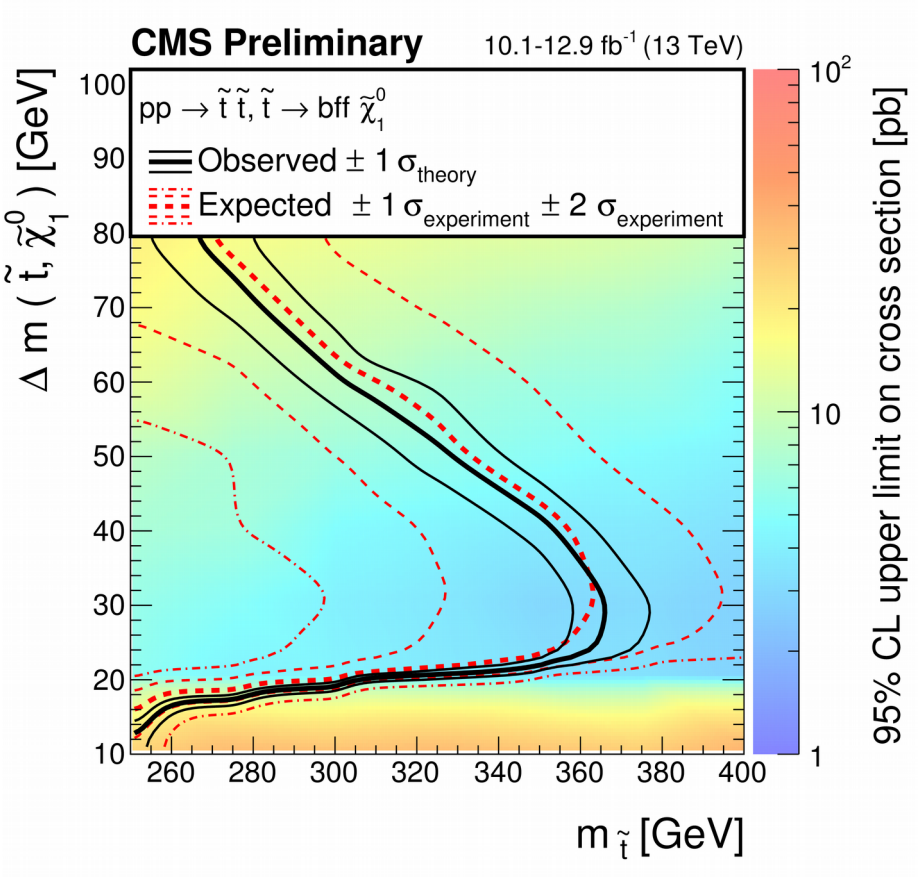
- efficiency maps
- a much larger database
- major speedups
- support for likelihoods

This talk is about SModelS v1.1.

Anatomy of experimental SMS results

Simplified Models results

CMS-SUS-16-025: upper limits



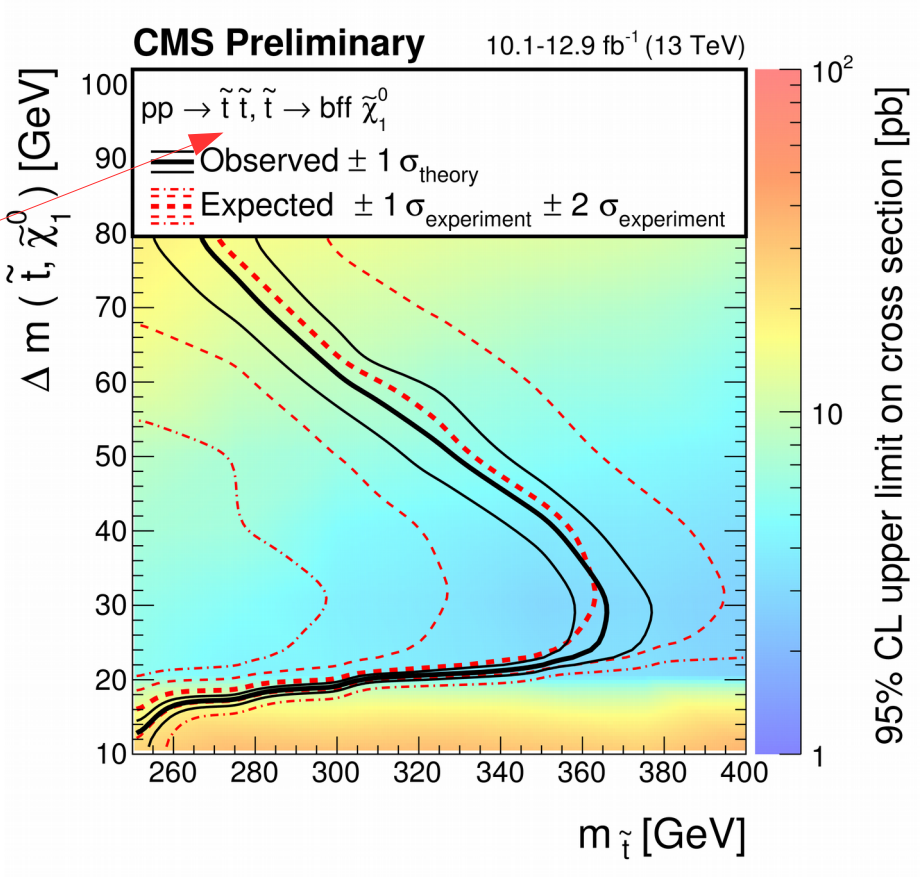
Upper limits, parametrized in the mass space of a simplified model

Simplified Models results

CMS-SUS-16-025: upper limits

$\tilde{t}\,\tilde{t}, \tilde{t} \rightarrow$ bff $\tilde{\chi}$

defines what part of a full theory the result *constrains*



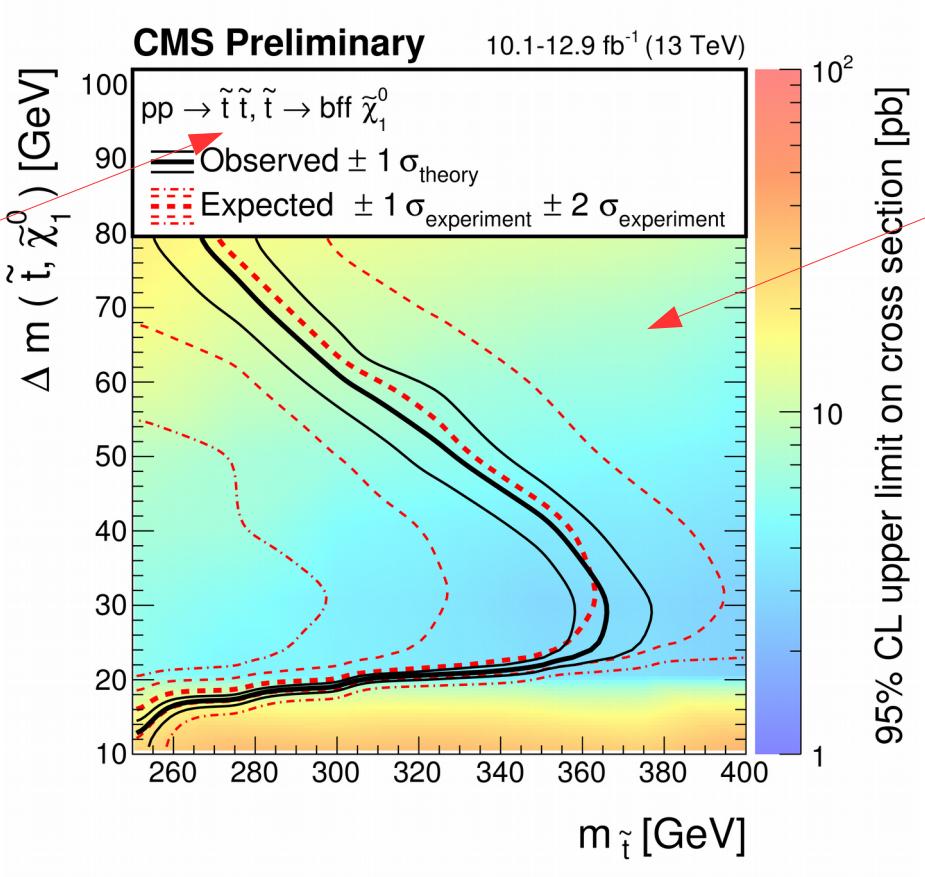
Upper limits, parametrized in the mass space of a simplified model

Simplified Models results

CMS-SUS-16-025: upper limits

$\tilde{t}\,\tilde{t},\ \tilde{t} \rightarrow$ bff $\tilde{\chi}$

defines what part of a full theory the result *constrains*



temperature plot = upper limit as a function of masses

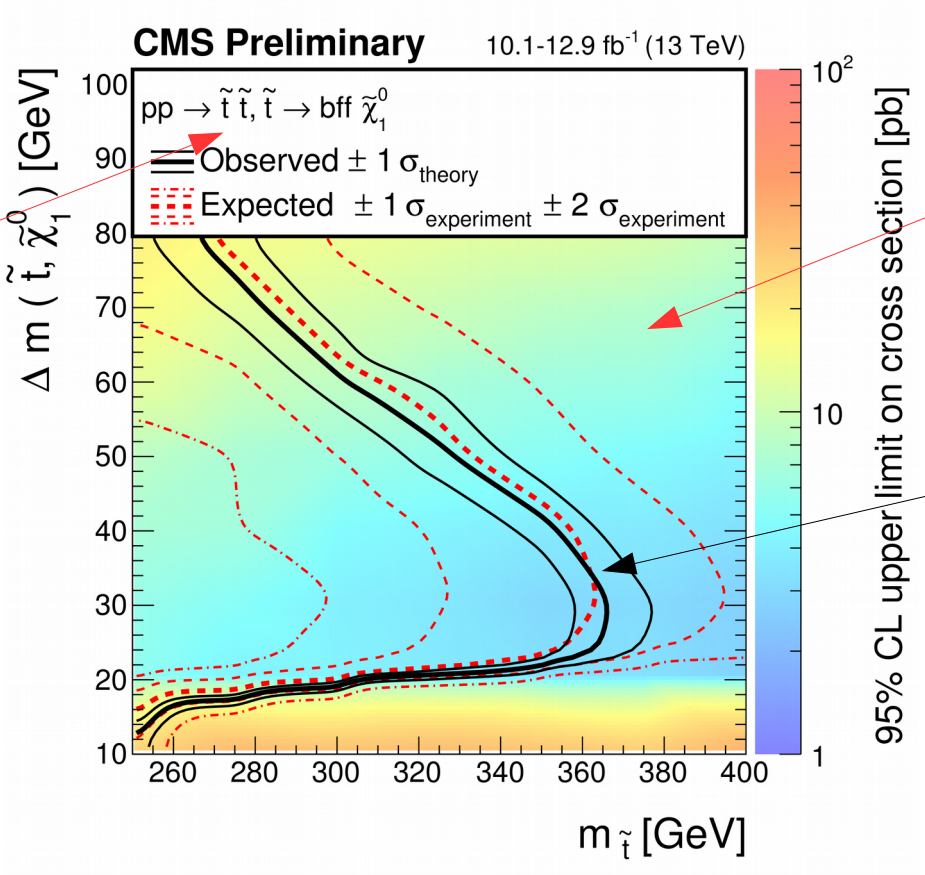
Upper limits, parametrized in the mass space of a simplified model

Simplified Models results

CMS-SUS-16-025: upper limits

$\tilde{t}\,\tilde{t}, \tilde{t} \rightarrow$ bff $\tilde{\chi}$

defines what part of a full theory the result *constrains*



temperature plot = upper limit as a function of masses

The exclusion lines are only here to guide the eye of the reader!
(We kept stressing this in every one of our talks)

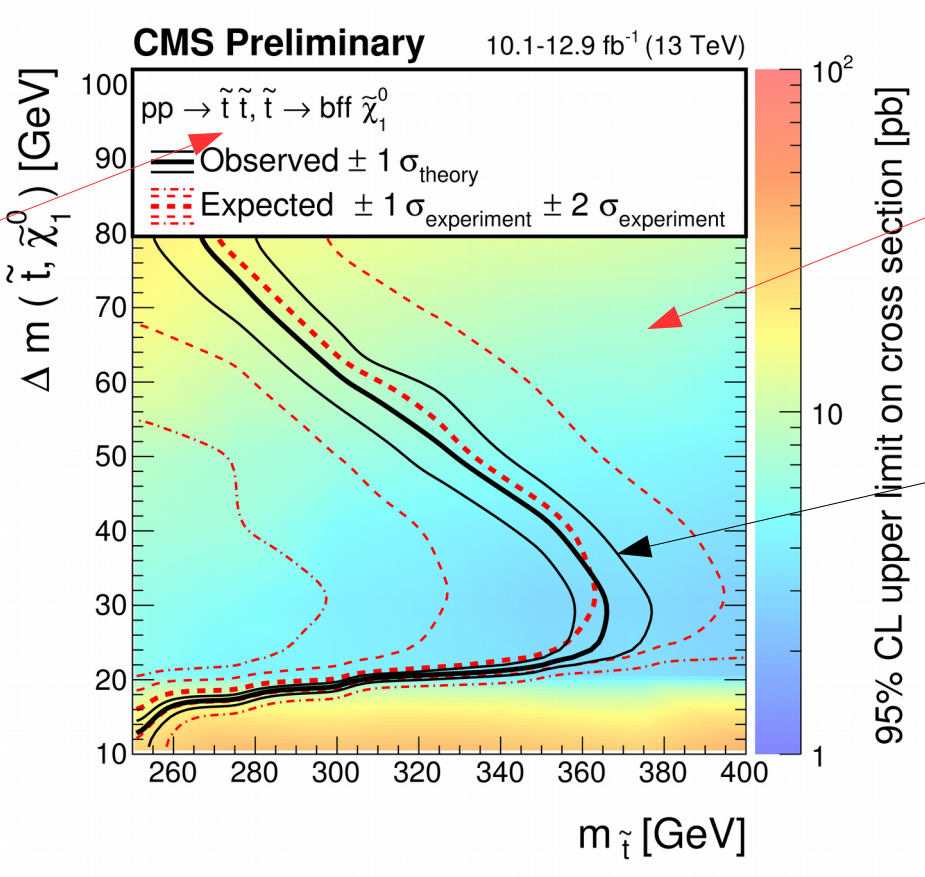
Upper limits, parametrized in the mass space of a simplified model

Simplified Models results

CMS-SUS-16-025: upper limits

$\tilde{t}\,\tilde{t},\ \tilde{t} \rightarrow$ bff $\tilde{\chi}$

defines what part of a full theory the result *constrains*



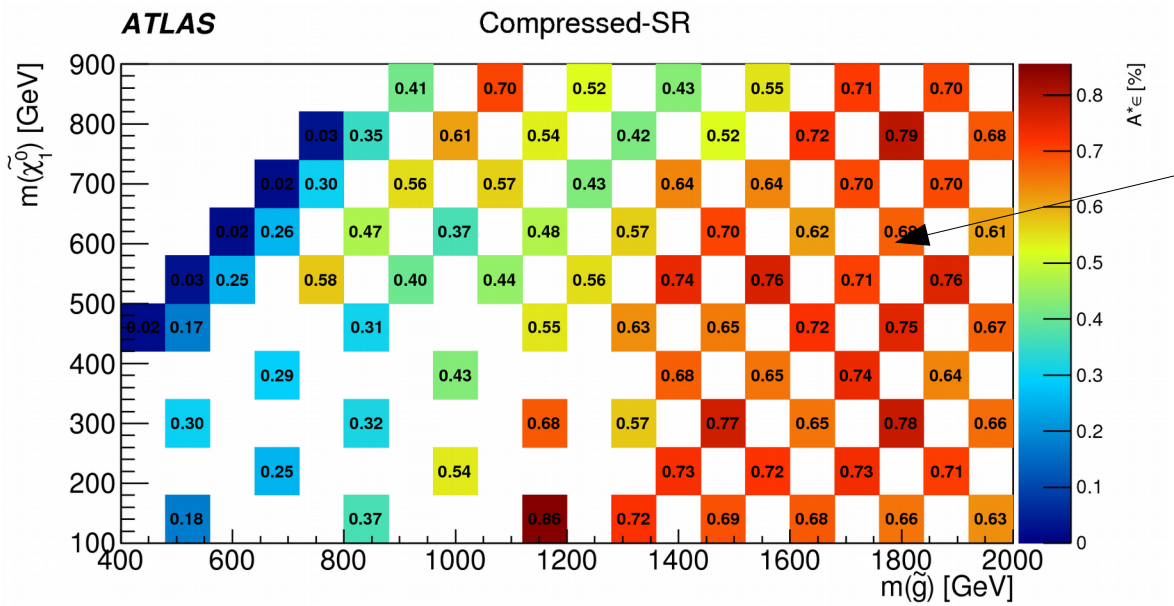
temperature plot = upper limit as a function of masses

We use these exclusion lines to validate our database

Upper limits, parametrized in the mass space of a simplified model

Simplified Models Results

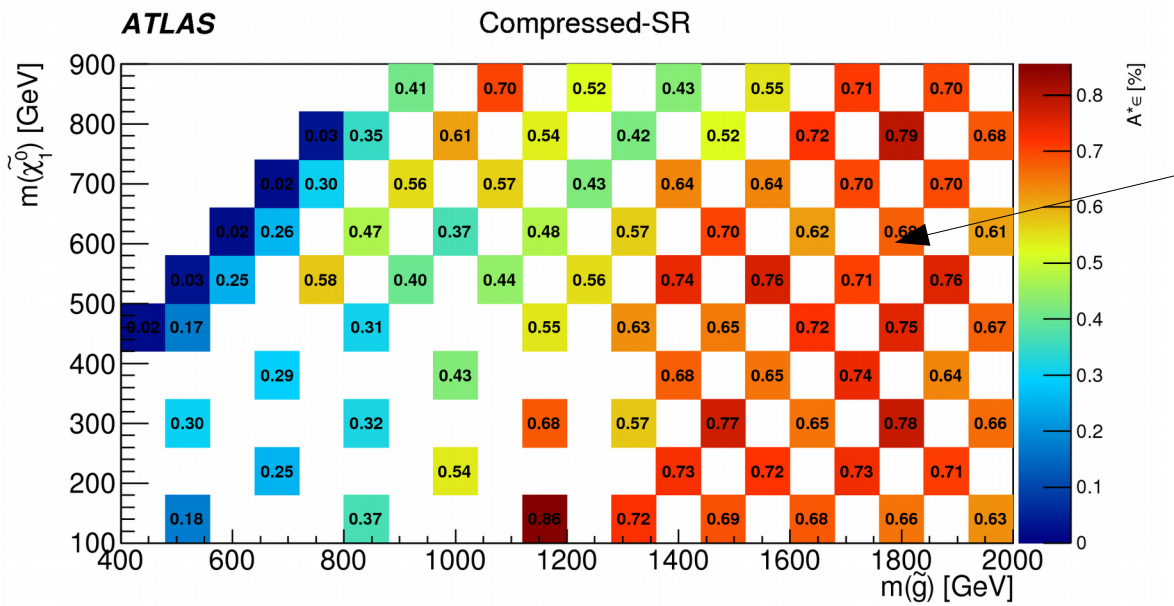
ATLAS-SUSY-2016-01: efficiency maps



Signal efficiencies, parametrized in the mass space of a simplified model

Simplified Models Results

ATLAS-SUSY-2016-01: efficiency maps



signal efficiencies

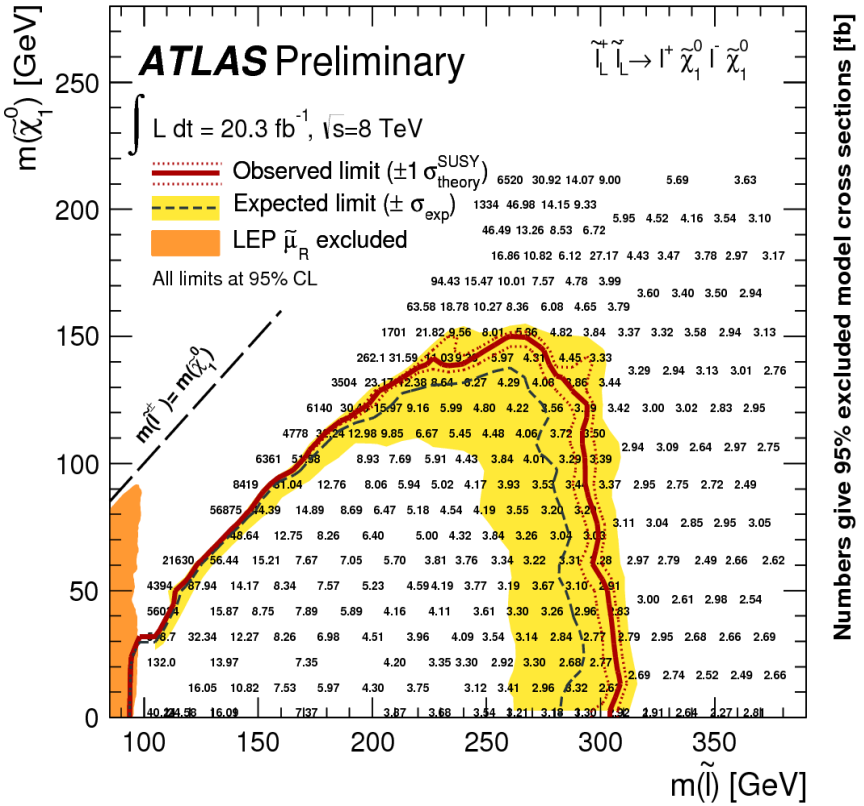
We use the signal efficiencies to compute upper limits on production cross sections. Additional ingredients are: the number of observed events, number of expected (background) events, and the error on the number of expected events!

Signal efficiencies, parametrized in the mass space of a simplified model

Formal language to describe the applicability of an SMS result

A formal language

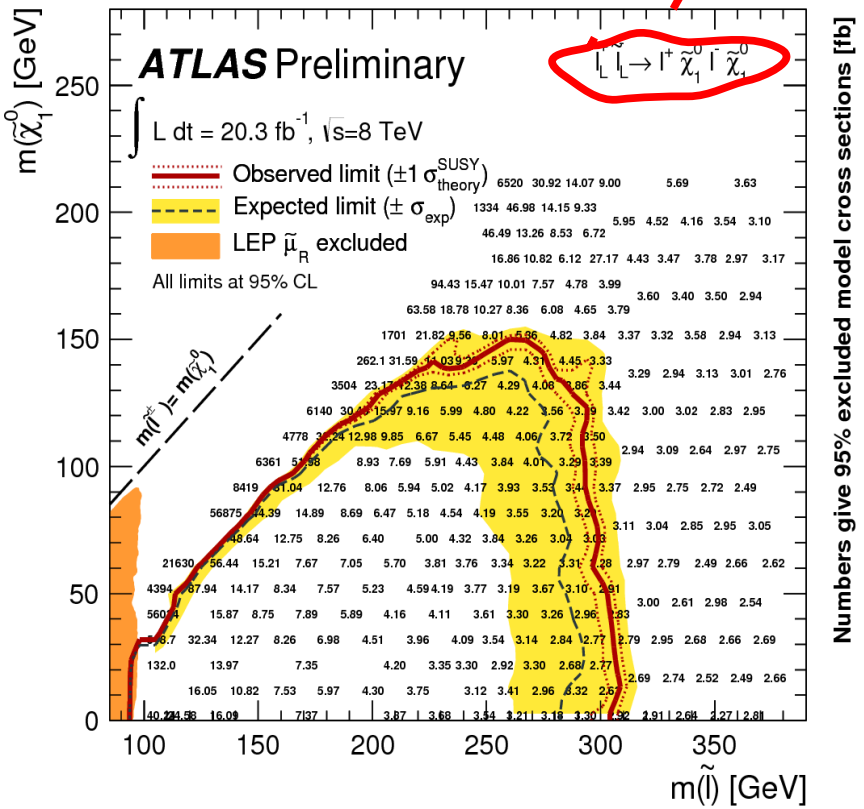
A formalism is needed to describe which part of a fundamental theory is constrained by what model under what conditions.

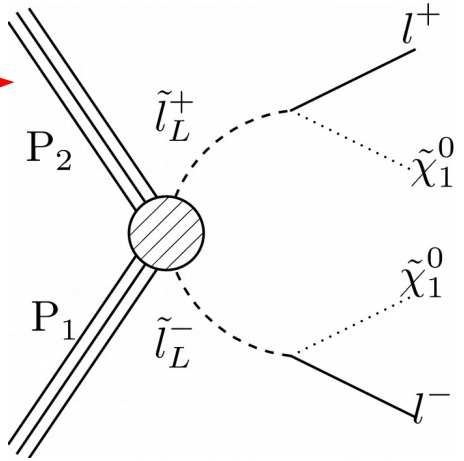


ATLAS-CONF-2013-049

A formal language

A formalism is needed to describe which part of a fundamental theory is constrained by what model under what conditions.

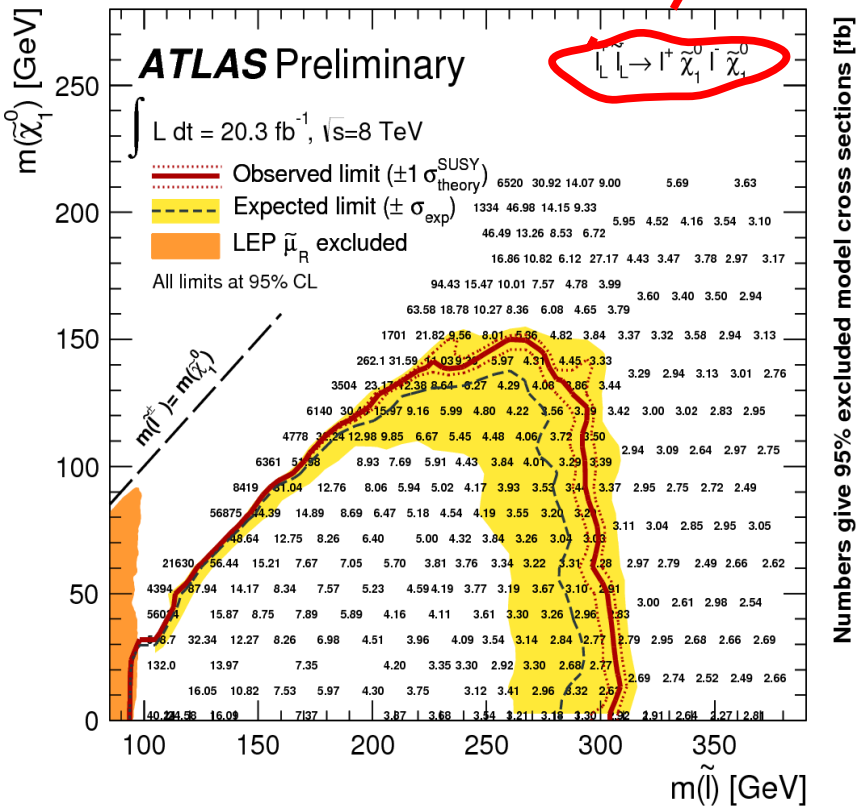


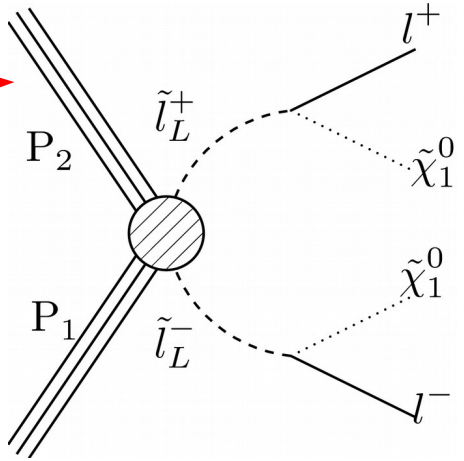


CMS-SUS-16-025

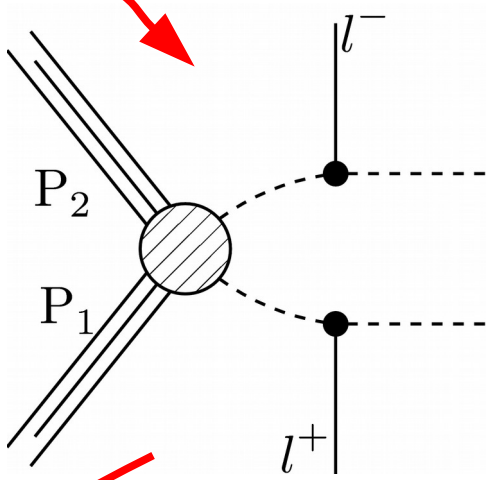
A formal language

A formalism is needed to describe which part of a fundamental theory is constrained by what model under what conditions.





We ignore the nature of the BSM particles



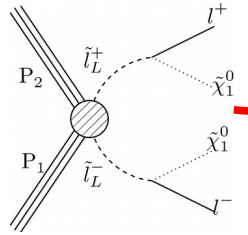
And describe the "constraint" in our formal language:

constraint: [[[e+]],[[e-]] + [[μ+]],[[μ-]]]

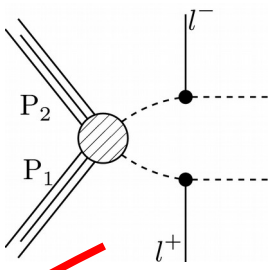
CMS-SUS-16-025

A formal language

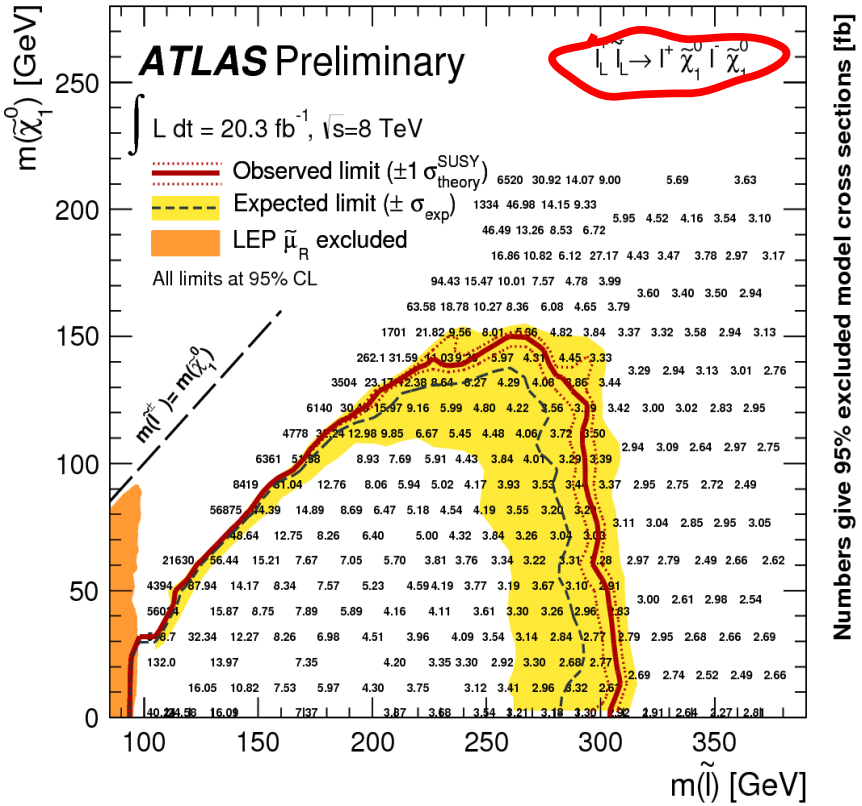
A formalism is needed to describe which part of a fundamental theory is constrained by what model under what conditions.



We ignore the nature of the BSM particles



And describe the "constraint" in our formal language:



CMS-SUS-16-025

constraint: [[[e$^+$]],[[e$^-$]] + [[μ$^+$]],[[μ$^-$]]]

In this case, the experiments make the implicit assumption that:

$$\sigma(\,[\,[[e^+]],[[e^-]]\,]\,)=\sigma(\,[\,[[\mu^+]],[[\mu^-]]\,])$$

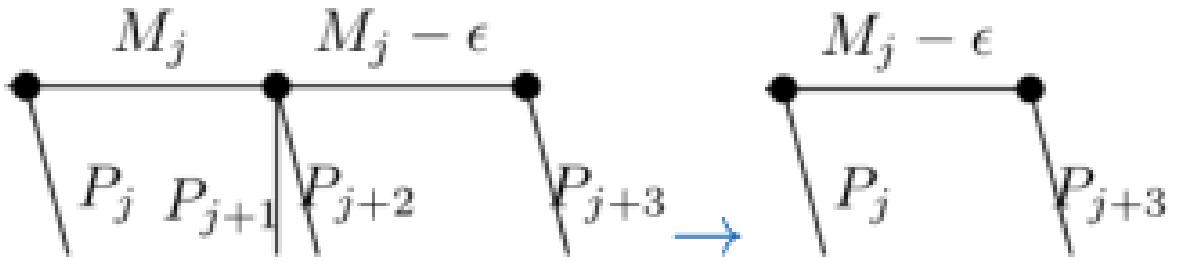
We take this into account by demanding from the theory that:

condition: [[μ$^+$]],[[μ$^-$]] ≥ [[e$^+$]],[[e$^-$]] - ε

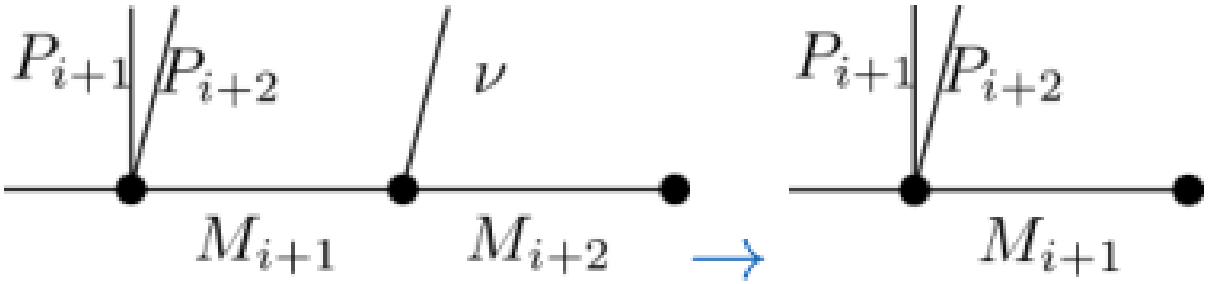
A formal language

Sometimes we can “compress” our spectrum:

“mass” compression



“invisible” compression



Building up a database

Construction of the SModelS database

ID	Topologies	Type	$\mathcal{L}$ [fb^{-1}]	$\sqrt{s}$
CMS-SUS-PAS-15-002	3: T1, T1bbbb, T1tttt[off]	ul	2.2	13
CMS-PAS-SUS-13-015	1: T2tt[off]	ul	19.4	8
CMS-PAS-SUS-13-015	1: T2tt[off]	eff	19.4	8
CMS-PAS-SUS-13-016	1: T1tttt[off]	ul	19.7	8
CMS-PAS-SUS-13-016	1: T1tttt[off]	eff	19.7	8
CMS-PAS-SUS-13-018	1: T2bb	ul	19.4	8
CMS-PAS-SUS-13-023	2: T2tt[off], T6bbWW[off]	ul	18.9	8
CMS-PAS-SUS-14-011	3: T1bbbb, T1tttt[off], T2tt[off]...	ul	19.3	8
CMS-SUS-12-024	2: T1bbbb, T1tttt[off]	ul	19.4	8
CMS-SUS-12-024	2: T1bbbb, T1tttt[off]	eff	19.4	8
CMS-SUS-12-028	6: T1, T1bbbb, T1tttt, T2, T2bb...	ul	11.7	8
CMS-SUS-13-002	1: T1tttt	ul	19.5	8
CMS-SUS-13-004	3: T1bbbb, T1tttt[off], T2tt[off]...	ul	19.3	8
CMS-SUS-13-006	5: TChiChipmSlepL, TChiChipmSlepStau...	ul	19.5	8
CMS-SUS-13-006	1: TChiWH	eff	19.5	8
CMS-SUS-13-007	2: T1tttt[off], T5tttt[off]	ul	19.3	8
CMS-SUS-13-007	1: T1tttt[off]	eff	19.3	8
CMS-SUS-13-011	2: T2tt[off], T6bbWW[off]	ul	19.5	8
CMS-SUS-13-011	1: T2tt[off]	eff	19.5	8
CMS-SUS-13-012	3: T1, T1tttt[off], T2	ul	19.5	8
CMS-SUS-13-013	2: T1tttt[off], T6ttWW[off]	ul	19.5	8
CMS-SUS-13-013	1: T1tttt[off]	eff	19.5	8
CMS-SUS-13-019	6: T1, T1bbbb, T1tttt[off], T2, T2bb...	ul	19.5	8
CMS-SUS-14-010	1: T1tttt[off]	ul	19.5	8
CMS-SUS-14-021	1: T2bbWW[off]	ul	19.7	8
CMS-SUS-14-021	1: T2bbWW[off]	eff	19.7	8

Table 1: SModelS database (CMS)

We collect the results of the experimental collaborations, and augment them with recast analyses (MadAnalysis5, CheckMATE), creating our own efficiency maps.
In addition, fastlim kindly allowed us to also use their efficiency maps (giving them of course the credit)
SModelS v1.1 will ship with ~ 70 analyses, and hundreds of individual results.

ID	Topologies	Type	$\mathcal{L}$ [fb^{-1}]	$\sqrt{s}$
ATLAS-SUSY-2013-09	1: T1tttt	ul	3.2	13
ATLAS-CONF-2013-024	1: T2tt	ul	20.5	8
ATLAS-CONF-2013-024	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	20.5	8
ATLAS-CONF-2013-024	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	1.0	8
ATLAS-CONF-2013-024	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	1.0	8
ATLAS-CONF-2013-035	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	20.7	8
ATLAS-CONF-2013-036	1: TChiChiSlepSlep	ul	20.7	8
ATLAS-CONF-2013-037	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	20.7	8
ATLAS-CONF-2013-047	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	20.3	8
ATLAS-CONF-2013-048	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	20.3	8
ATLAS-CONF-2013-049	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	20.3	8
ATLAS-CONF-2013-053	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	20.1	8
ATLAS-CONF-2013-053	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	1.0	8
ATLAS-CONF-2013-054	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	20.3	8
ATLAS-CONF-2013-061	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	20.1	8
ATLAS-CONF-2013-062	1: T5WW[off]	ul	20.3	8
ATLAS-CONF-2013-062	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	20.3	8
ATLAS-CONF-2013-089	2: T5WW[off], T6WW[off]	ul	20.3	8
ATLAS-CONF-2013-093	1: TChiWH	ul	20.3	8
ATLAS-CONF-2013-093	24: T1, T1bbbb, T1bbbt, T1bbqq, T1bbtt...	eff	20.3	8
ATLAS-SUSY-2013-02	6: T1, T2, T5WW[off], T5tctc, T6WW[off]...	ul	20.3	8
ATLAS-SUSY-2013-02	2: T1, T2	eff	20.3	8
ATLAS-SUSY-2013-04	1: T1tttt	ul	20.3	8
ATLAS-SUSY-2013-04	3: T1tttt, T5WW[off], T5ZZ[off]...	eff	20.3	8
ATLAS-SUSY-2013-05	2: T2bb, T6bbWW[off]	ul	20.1	8
ATLAS-SUSY-2013-05	1: T2bb	eff	20.1	8
ATLAS-SUSY-2013-08	1: T6ZZtt	ul	20.3	8
ATLAS-SUSY-2013-09	1: T1tttt	ul	20.3	8
ATLAS-SUSY-2013-09	1: T1tttt	eff	20.3	8
ATLAS-SUSY-2013-11	4: TChiWW, TChiWZ, TChipChimSlepSnu...	ul	20.3	8
ATLAS-SUSY-2013-11	1: TSlepSlep	eff	20.3	8
ATLAS-SUSY-2013-12	4: TChiChipmSlepL, TChiChipmStauL...	ul	20.3	8
ATLAS-SUSY-2013-14	2: TChiChipmStauL, TChipChimStauSnu...	ul	20.3	8
ATLAS-SUSY-2013-15	3: T2bbWW, T2tt, T6bbWW	ul	20.3	8
ATLAS-SUSY-2013-15	1: T2tt	eff	20.3	8
ATLAS-SUSY-2013-16	1: T2tt	ul	20.1	8
ATLAS-SUSY-2013-16	2: T2tt, T6bbWW[off]	eff	20.1	8
ATLAS-SUSY-2013-18	2: T1bbbb, T1tttt	ul	20.1	8
ATLAS-SUSY-2013-18	2: T1bbbb, T1tttt	eff	20.1	8
ATLAS-SUSY-2013-19	3: T2bbWW, T2tt, T6bbWW[off]	ul	20.3	8
ATLAS-SUSY-2013-21	3: T2bb, T2bbWW[off], T2cc	eff	20.3	8
ATLAS-SUSY-2013-23	1: TChiWH	ul	20.3	8
ATLAS-SUSY-2014-03	1: TScharm	eff	20.3	8

Construction of the SModelS database

Alongside with the actual data (upper limits, efficiency maps), we keep track of
All the meta information that comes with results, at the level of a specific analysis,
at the level of one signal region, and at the level of one specific “result”:

```
sqrts: 8*TeV
lumi: 20.3/fb
id: ATLAS-CONF-2013-048
url: https://atlas.web.cern.ch/Atlas/GR...
implementedBy: WW
contact: fastlim
comment: created from fastlim-1.0
lastUpdate: 2016/08/16
```

analysis level

```
dataType: efficiencyMap
dataId: SR M90
observedN: 260
expectedBG: 300.0
bgError: 40.0
upperLimit: 2.5*fb
expectedUpperLimit: 4.24*fb
```

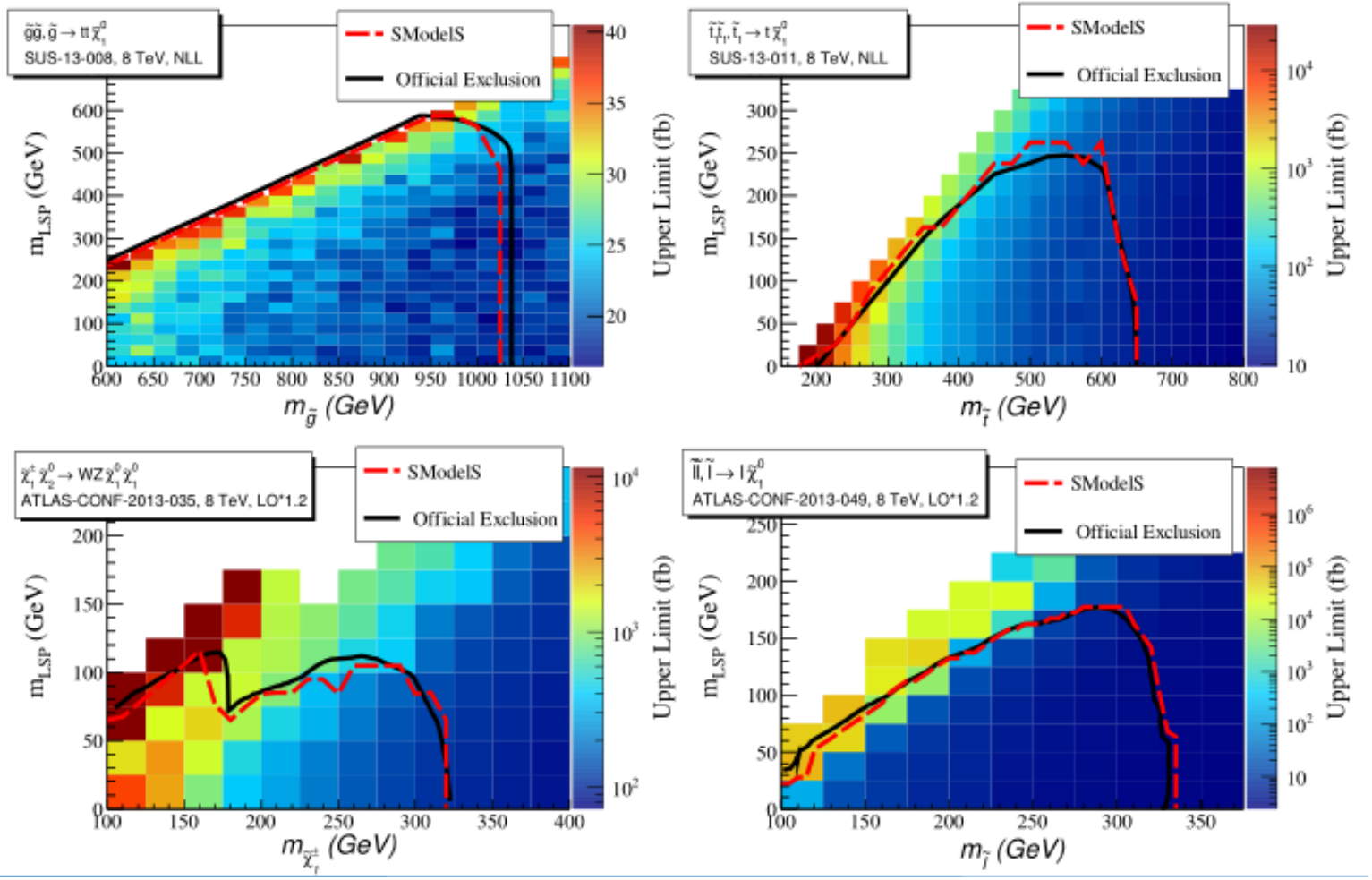
signal region

```
txName: T2
conditionDescription: None
condition: None
constraint: [[['jet']],[['jet']]]
validated: None
axes: 2*Eq(mother,x)_Eq(lsp,y)
publishedData: False
```

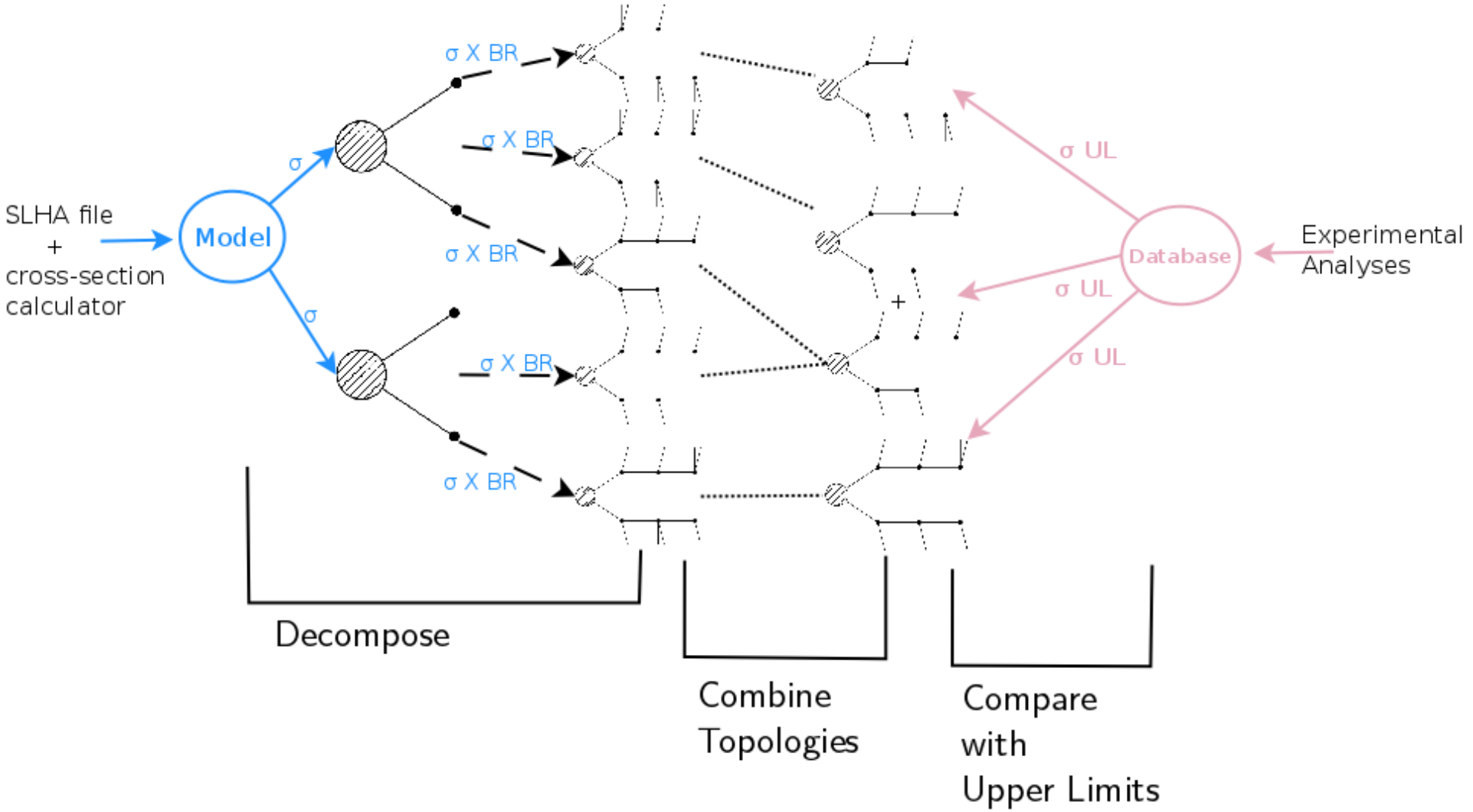
result

Construction of the SModelS database

We validate by defining the full model := simplified model, and check if we can reproduce the exclusion curves

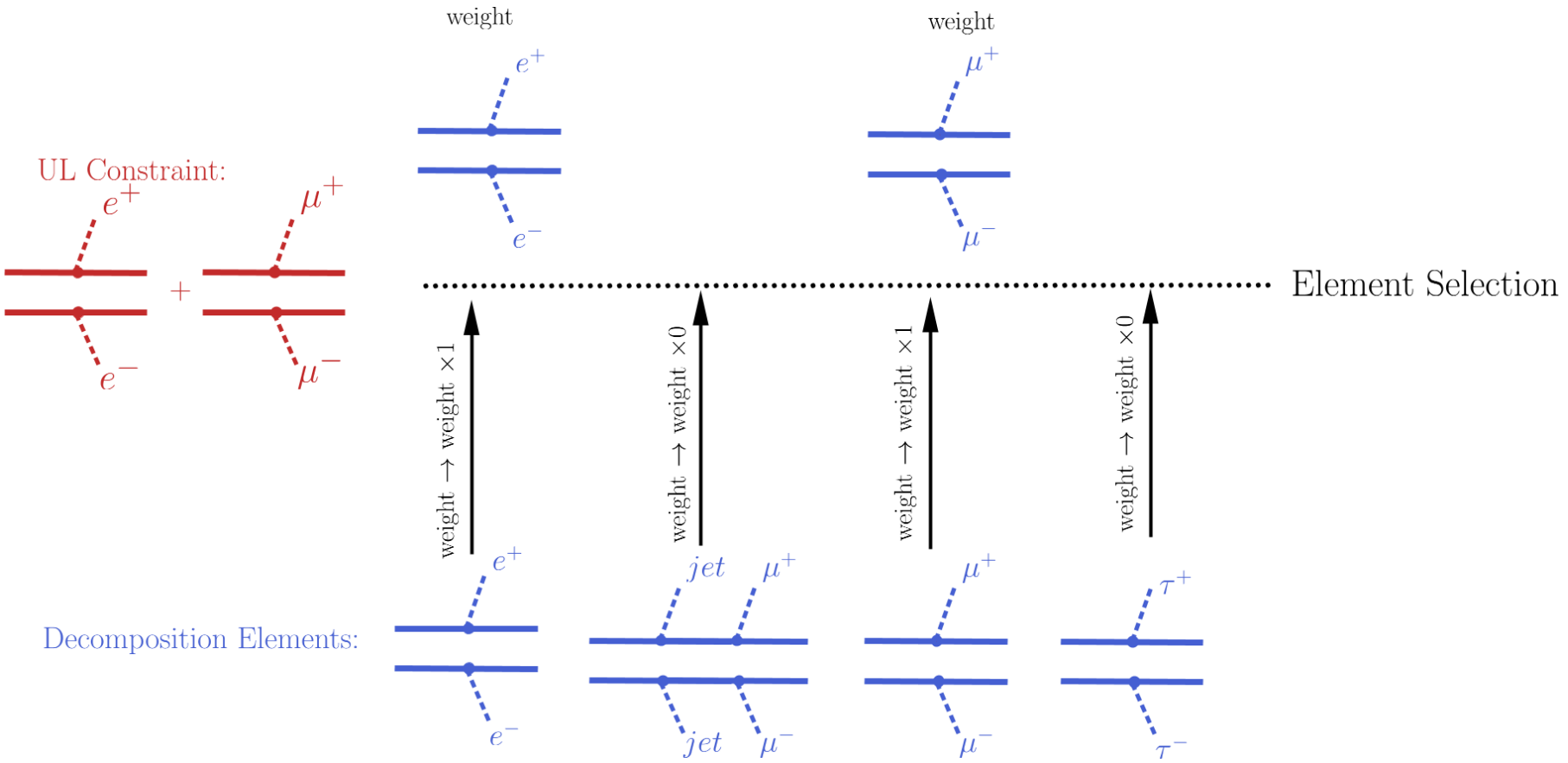


Putting it all together



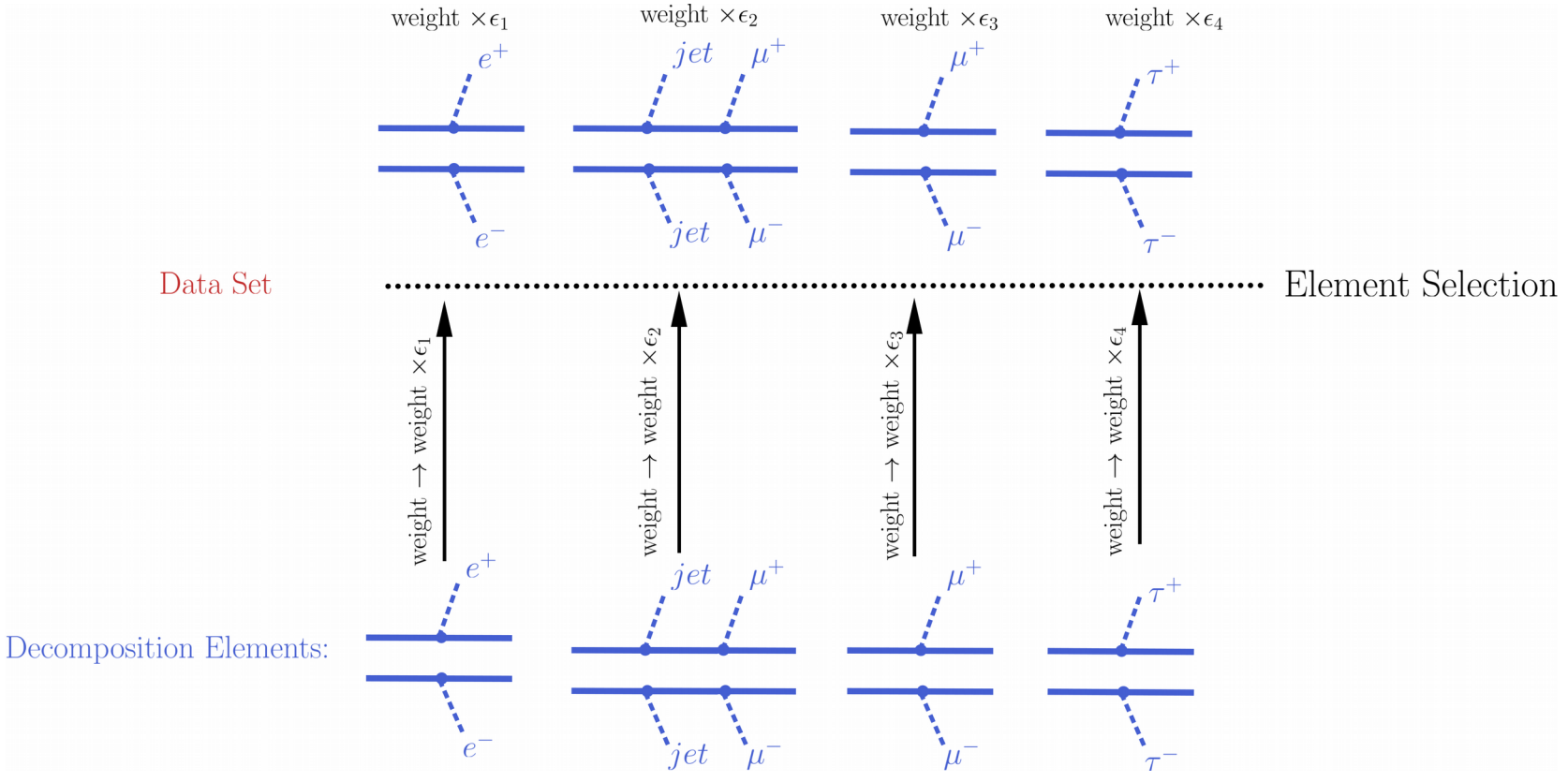
Putting it all together: upper limits

For upper limit results, we cannot combine results given for different topologies. We can only apply them one-by-one. Therefore the final upper limits are very conservative. (On the positive side: it is fairly straightforward for us to add an official SMS result to our database)



Putting it all together: efficiency maps

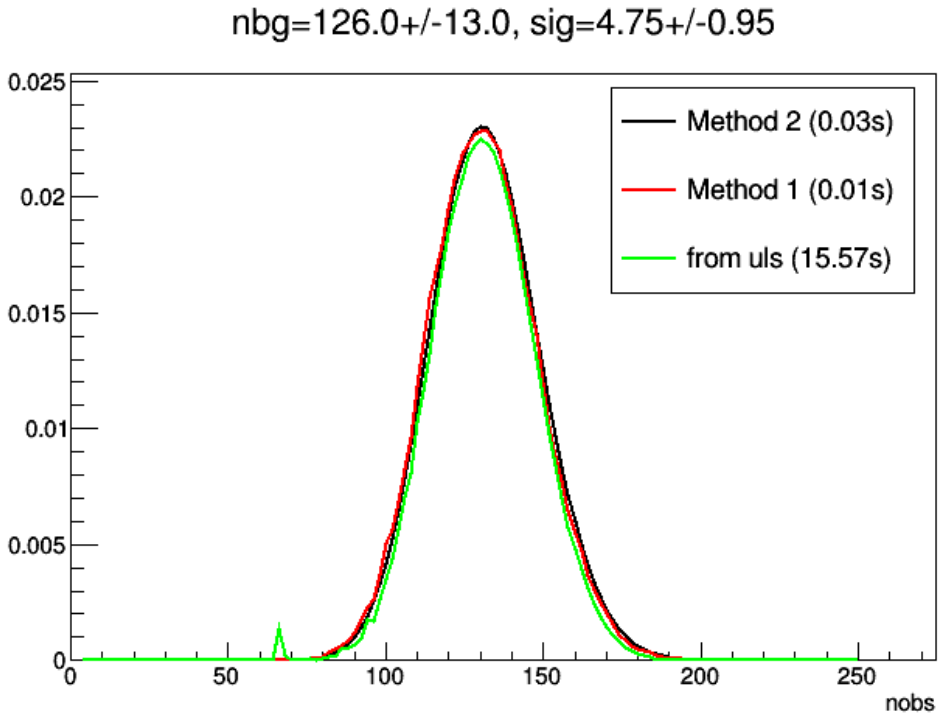
In the case of efficiency maps results we can combine efficiency maps for all topologies that an analysis has results for. (If e.g. an analysis vetoes jets, then the efficiencies for topologies involving jets would ~ 0). As a consequence, the final upper limits tend to be less conservative.



From eff maps and upper limits to likelihoods

For efficiency-maps results we can easily create a full likelihood.
Given for a signal region the number of observed events, the number of expected background events, the error on the background estimate, the efficiency (and its error), we can model a likelihood as a convolution of a Poissonian (for the statistical error) with a Gaussian (for the systematic error).

For upper limits results we can also reconstruct an approximate likelihood, if we:
a) have the information of the expected upper limit
b) assume an approximately Gaussian form of the likelihood



For v1.1 we supply code for likelihoods from efficiency maps, but not for upper limit results (we need a better understanding of when the Gaussian approximation holds)

Applications

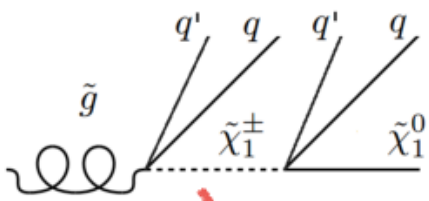
We envisage SModelS to be used for the following goals:

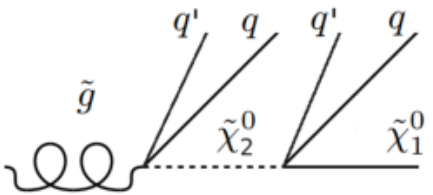
- To quickly eliminate uninteresting points in model scans
- To identify the most constraining analyses, and thus guide recasting efforts
- To identify “missing” topologies and “uncovered” mass ranges and thus give feedback on the design of new analyses and changes in existing analyses
- A quick and cheap method to reinterpret LHC analyses in the context of models other than the pMSSM (UEDs, NMSSM, ...)
- A quick and cheap way to produce conservative likelihoods that may be combined with other data.

Applications: ATLAS pMSSM scan (SModelS v1.0)

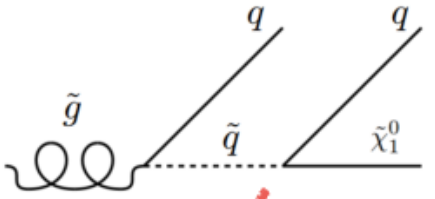
Study Case: Missing Topologies

Examples of Missing Topologies

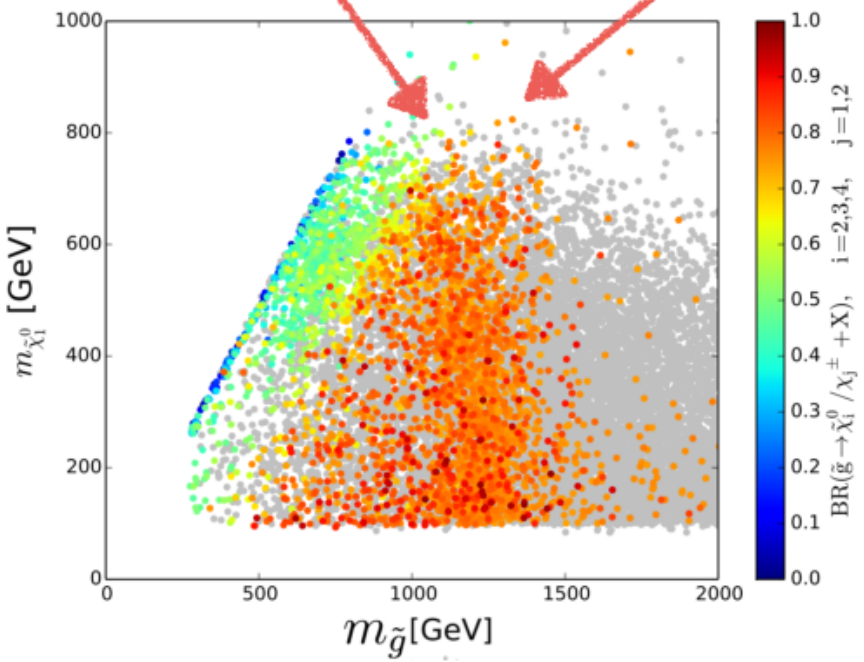




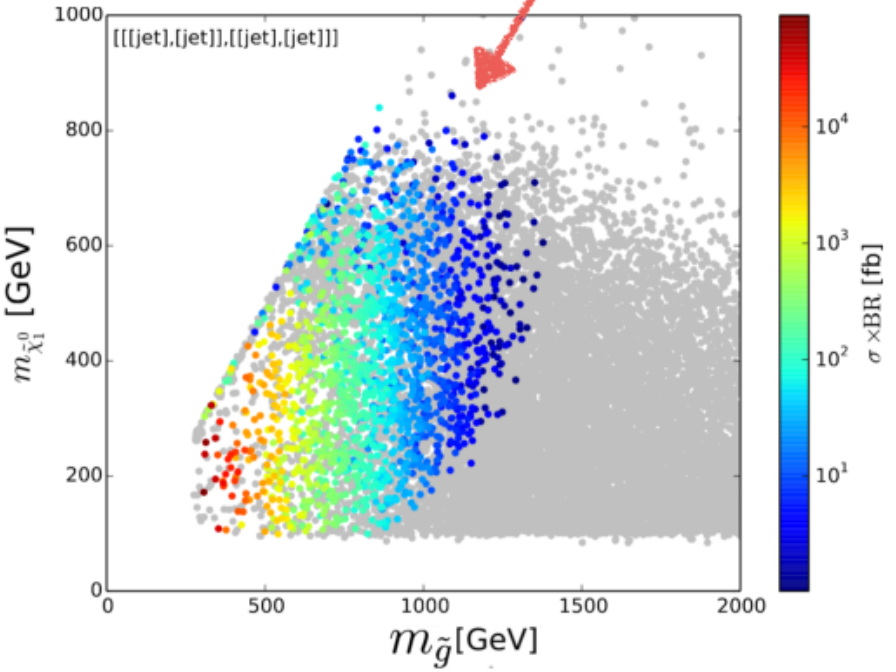
Non-degenerate right and left handed squarks



Branching Ratio Plot $\mathrm{BR}(\tilde{g} \to \tilde{\chi}_i^0 / \chi_j^{\pm} + X)$



Plot (σxBR)

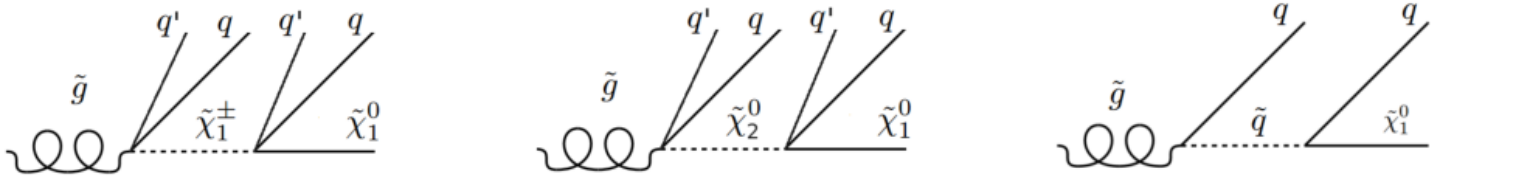


Applications: ATLAS pMSSM scan

Study Case: Missing Topologies

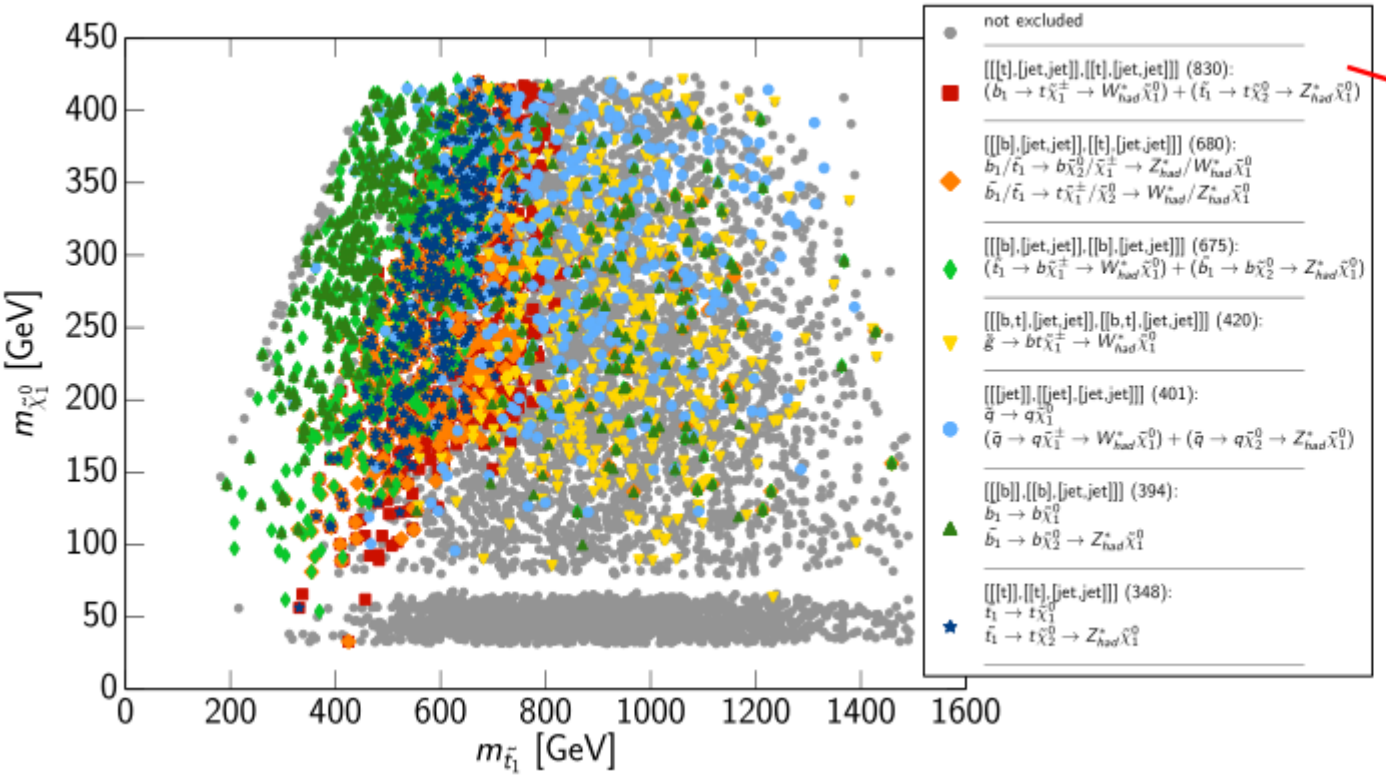
Examples of Missing Topologies

Non-degenerate right and left handed squarks

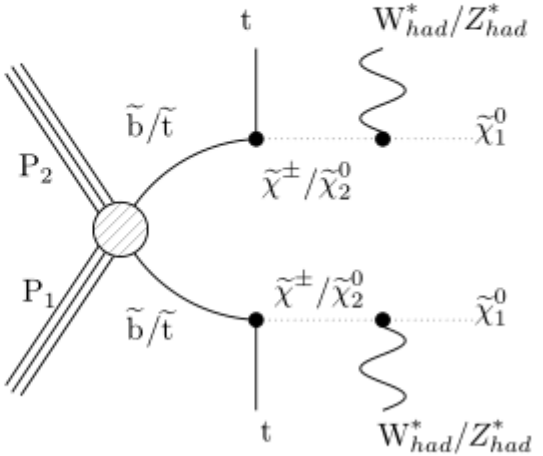


We are producing "Homegrown" efficiency maps in order to cover important missing topologies and improve SModelS constraining power

Applications: Low Fine-Tuning Scan

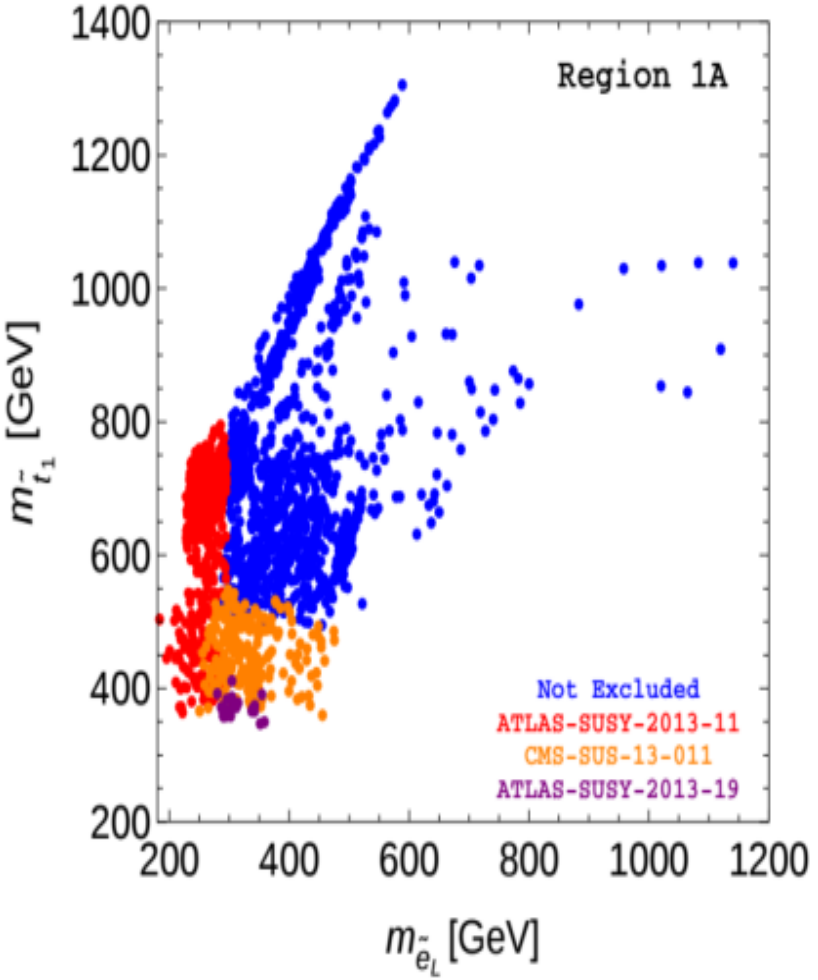


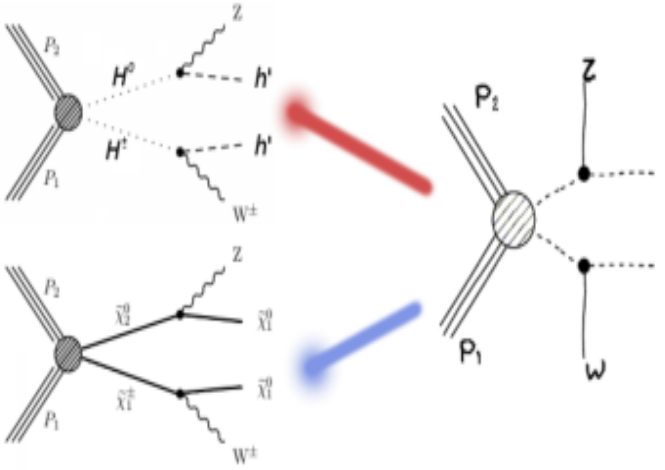
[[[*t*], [*jet*, *jet*]], [[*t*], [*jet*, *jet*]]]

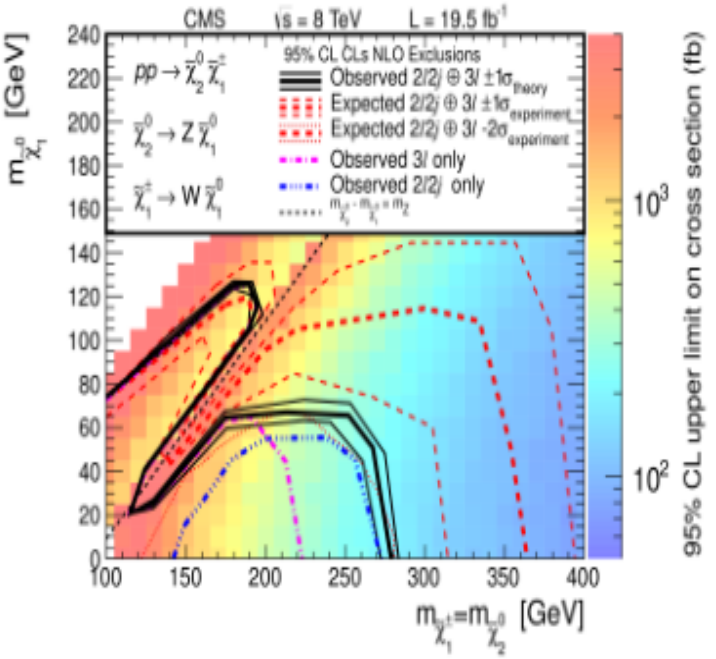


Which of our analyses is sensitive to this guy?

Applications: 2HDM







D. Barducci, G. Bélanger, C. Hugonie and A. Pukhov, JHEP 1601 (2016) 050

⇒ LHC constraints on 2HDM

Future

We intend to extend the functionality of SModelS in several ways:

- Extend to non-Z_2 / non-MET topologies
- Extend to long-lived particles (HCSP scenarios) and other “exotic” signatures
- Create a map of likelihoods that we can safely combine

  (in good approximation: identify analyses with non-overlapping signal regions)
- Support for positive results. Prepare for the next di-photon frenzy
- Create mockup analyses that extrapolate to HL-LHC
  (is quite easy for us)

http://smodels.hephy.at

Thank you!