LHC recasting in a nutshell

Fuks Benjamin

LPTHE / UPMC

1st MadAnalysis 5 workshop on LHC recasting
@ High 1, Gangwon Province, Korea

August 20 - 28, 2017

Outline

Introduction

New physics at the LHC

- The quest for physics beyond the Standard Model has started!
  - How to get hints of new physics?
    - Confront data to the Standard Model expectation in search channels
    - Observe unexplained deviations at a good confidence level
  - Ingredient 1: predictions for the Standard Model background
  - Ingredient 2: predictions for the new physics signals

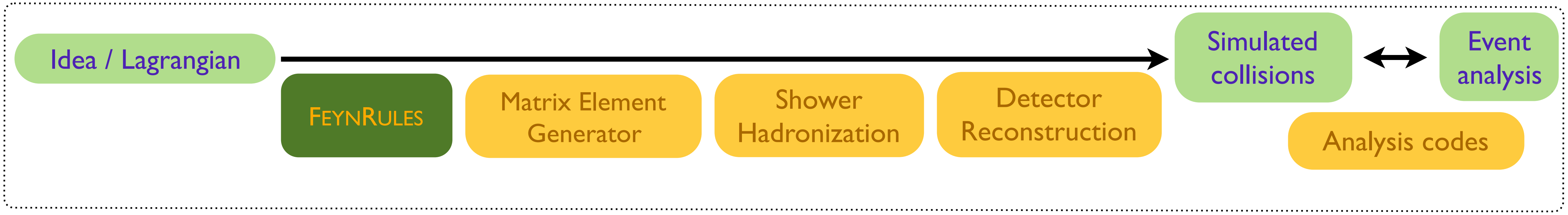
- More on the new physics nature
  - Fitting deviations by new physics signals
    - Designing new analyses to probe new ideas
    - Reinterpretation of data in possibly different theoretical frameworks
    - Leading order MC tools do a proper job

Predictions; signal and background analysis

Confronting models to data

New physics simulations so far

✦ Streamlining the links between models and simulations



✦ Implementation of any theory in MC tools is straightforward (LO and NLO)

Let’s reverse the chain…

Reinterpreting LHC physics analyses

- ✦ Exploit the full potential of the LHC (for new physics)
  - ✤ Priority #1 of the European strategy for particle physics
  - ✤ Designing new analyses to probe new ideas — Prospectives (based on MC simulations)
  - ✤ Recasting LHC analyses to study models not considered — The LHC legacy

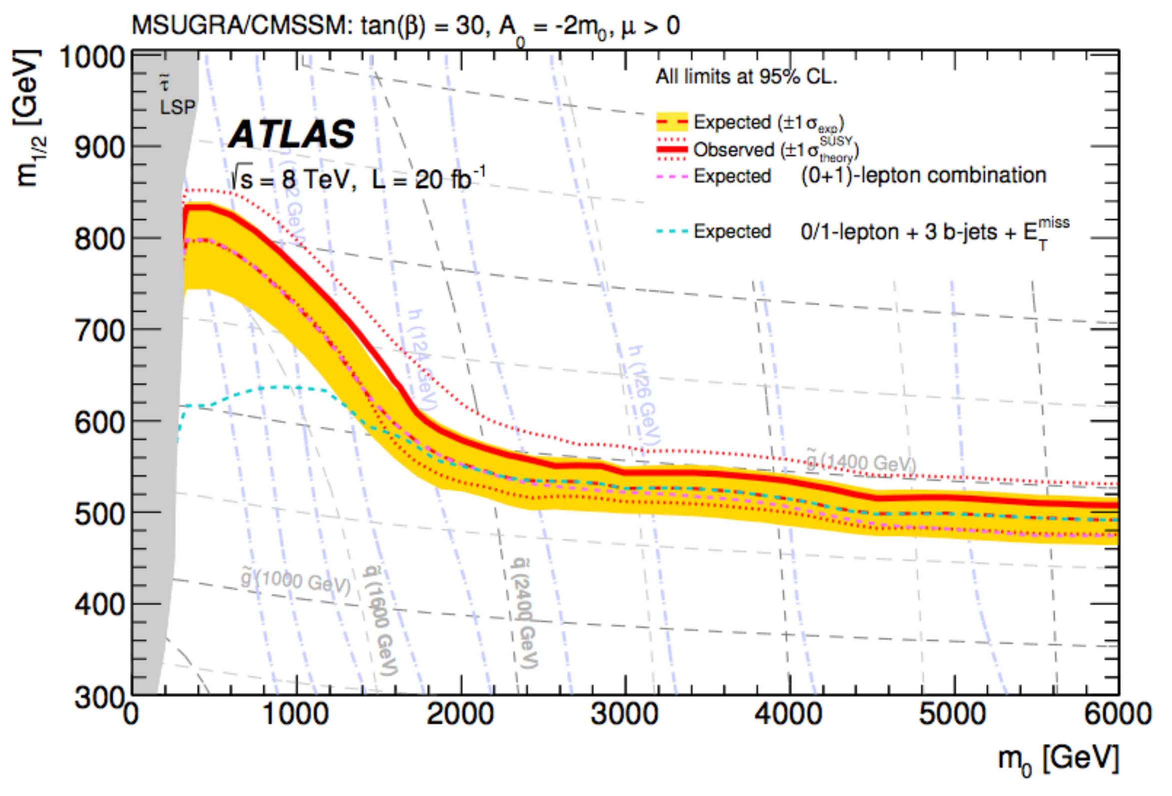
- ✦ LHC data has been collected with significant human and financial efforts
  - ✤ Important for on-going analyses (within popular theoretical contexts of today)
  - ✤ Important for future opportunities (within future scientific contexts)
- ✦ Data preservation in high-energy physics is mandatory [Kogler, South & Steder (JPCS'12)]
  - ✤ Studies are on-going and go beyond raw data (ICFA DPHEP Study Group)

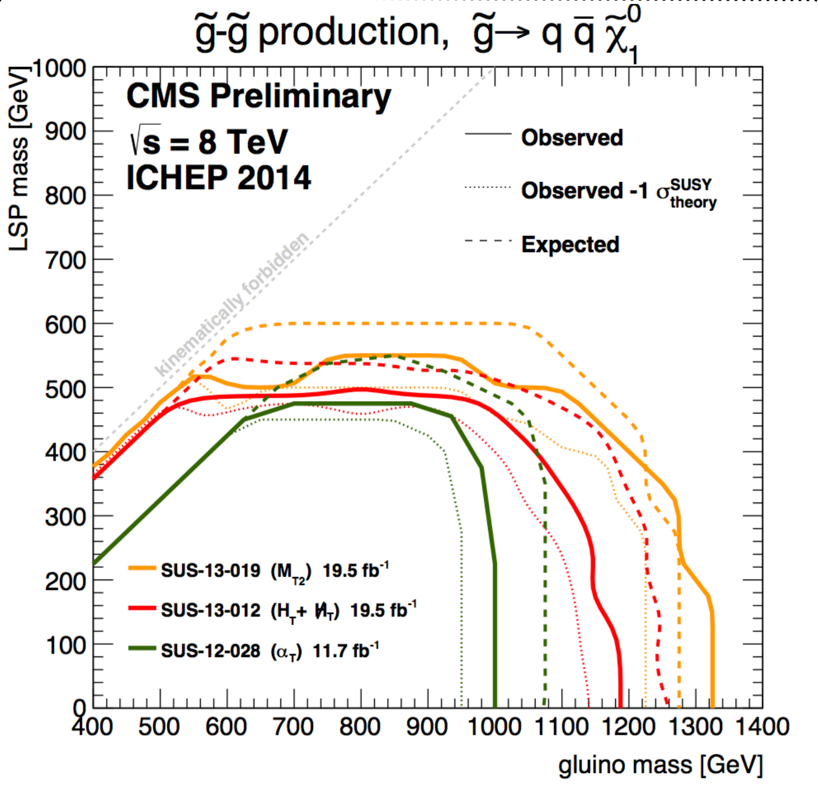
- ✦ Related tools need to be supported by the entire community [Kraml *et al.* (EPJC'12)]
  - ✤ Both theorists and experimentalists
  - ✤ Allowing for the reinterpretation of the LHC analysis results

The Simplified Model Spectra (SMS) approach

New physics results at the LHC

- ✦ The LHC has been built as a discovery machine
  - ✤ There are many ATLAS and CMS searches for new physics
  - ✤ Interpretation within popular frameworks and simplified models (SMS)





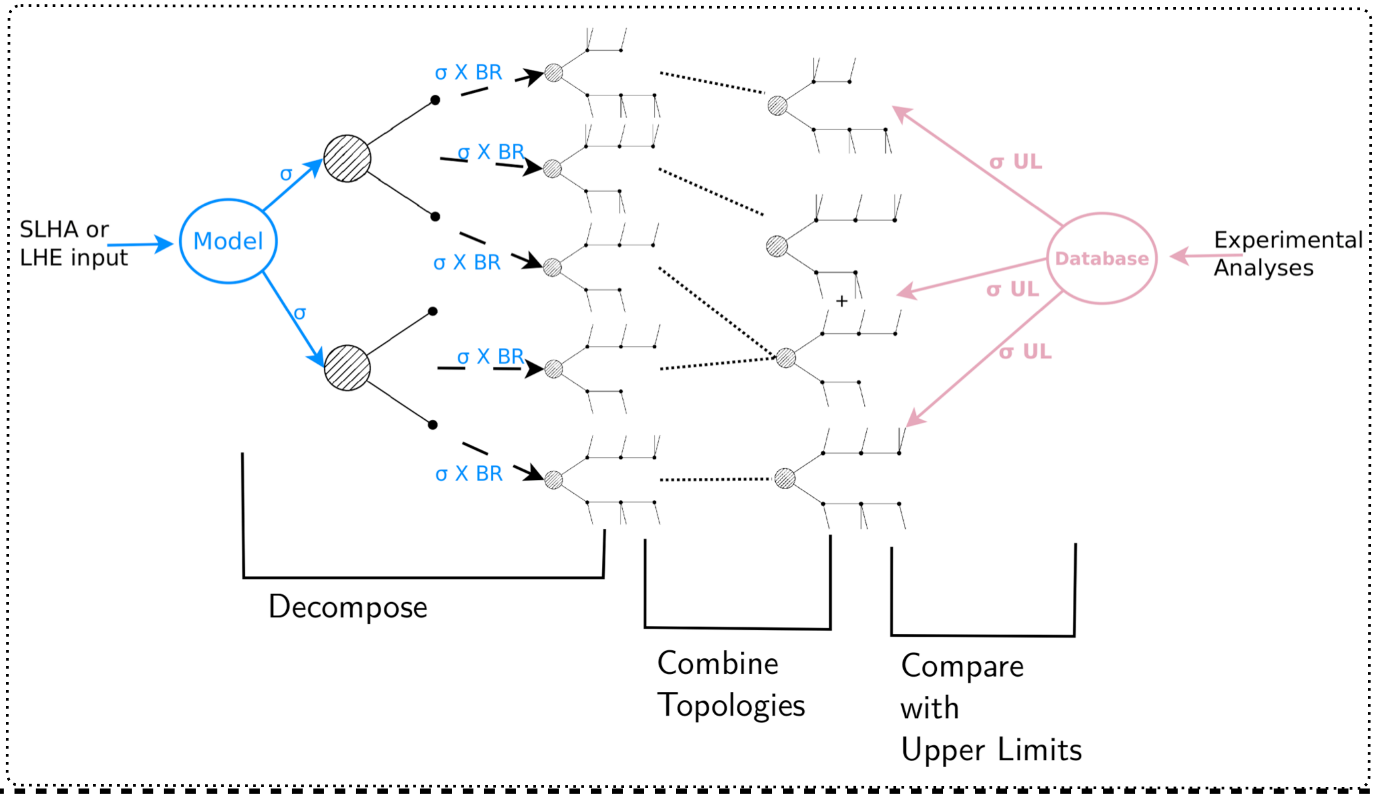
- ✦ We need to reinterpret the results for all kinds of models

The SMS approach for reinterpretations

✦ The SMS-based reinterpretation framework

- ✤ All signatures of a theory are decomposed according to those of the SMS searches
- ✤ Fiducial cross sections are calculated on the basis of public efficiency maps
- ✤ Comparisons to published upper bounds are made

✦ Main features



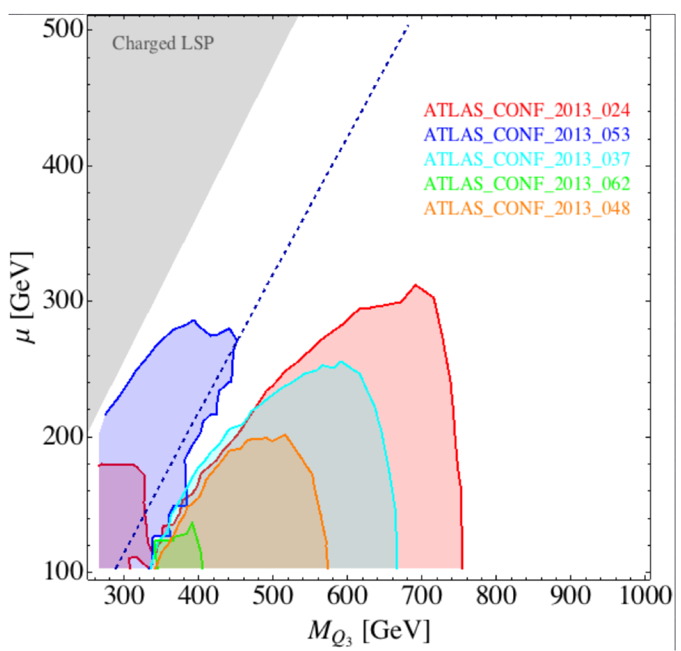
- ✤ Extremely fast
- ✤ Moderately accurate and general
  - ★ Kinematical configurations often not close to the SMS ones
  - ★ Multistep or asymmetric decays

SMS reinterpretation tools

✦ Existing tools: FastLim, SModelS, XQCat

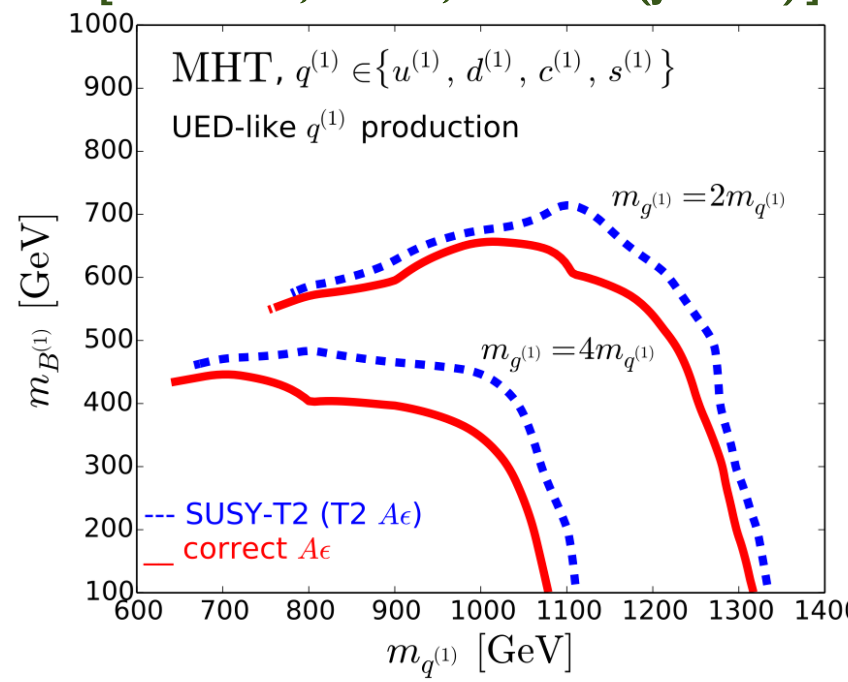
[Papucci, Sakurai, Weiler & Zeune (EPJC'14)] [Kraml *et al.* (EPJC'14)] [Barducci *et al.* (CPC'15)]

✦ Examples



MSSM reinterpretations with FastLim

[Edelhäuser, Krämer, Sonneveld (JHEP'15)]



Limitations (using SModelS): SUSY *versus* UED

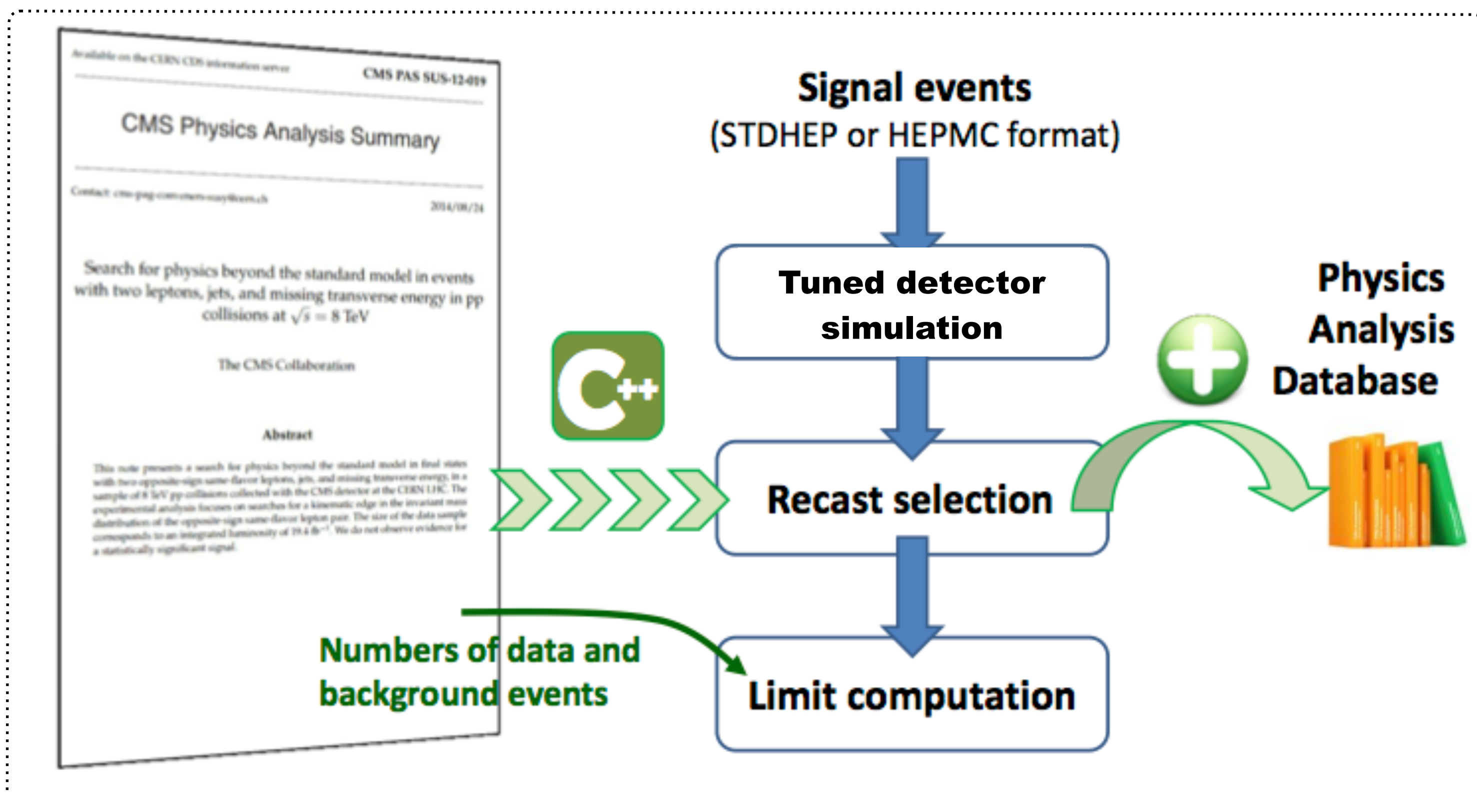
The ‘fastsim’-based approach

Beyond the SMS approach

✦ There are plethora of new physics realizations that deserve to be studied

- ✤ Experimentalists cannot study all the options
- ✤ The simplified model approach is often not sufficient (*e.g.*, different topologies)
- ✤ Our choice: rely on a public detector simulator mimicking ATLAS and CMS
- ✤ Need for a (public) framework where LHC analyses can be easily implemented

✦ Another recasting strategy (as used in MADANALYSIS 5)



✤ 2 options for detector effects

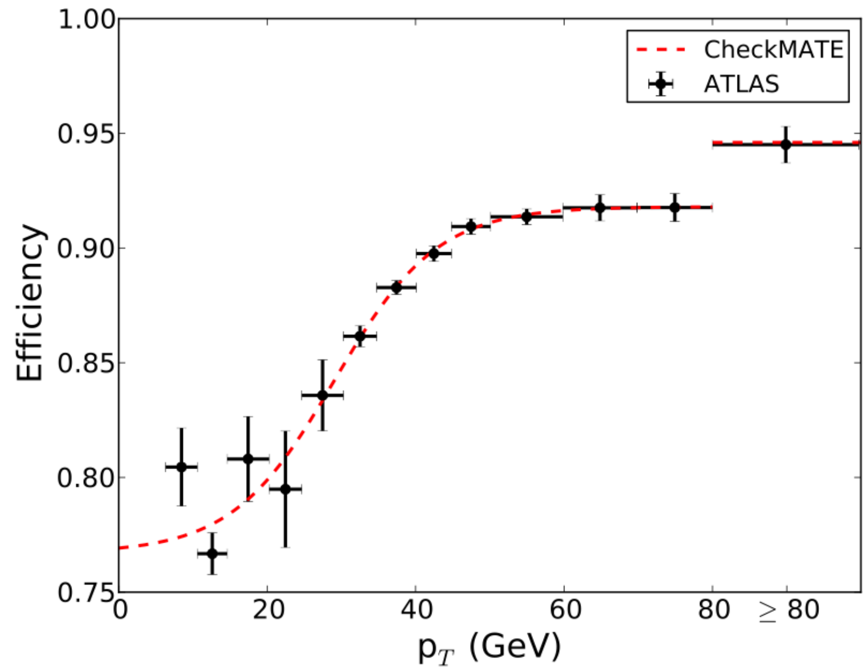
- ★ DELPHES/PGS-like (resolutions, efficiencies, *etc.*)
- ★ RIVET-like (transfer functions)

Detector modeling with DELPHES

[de Favereau *et al.* (JHEP'14)]

✦ Detector simulation with DELPHES 3

- ✤ Starts from hadron-level MC information
- ✤ Derive calorimetric and track information; object reconstruction is then necessary
  - ★ Close to what actually happens in a real experiment
- ✤ DELPHES is modular ➢ extra modules and tuning can be added / included
  - ★ Information on lepton isolation or track information; skimming of the output files, *etc*.



Medium electron efficiency in CHECKMATE

```
module Efficiency ElectronEfficiency {
  set InputArray ElectronEnergySmearing/electrons
  set OutputArray electrons

  # efficiency formula for electrons
  # medium efficiency from a fit to ATLAS medium electron efficiencies
  set EfficiencyFormula {
      (pt < 90.) * ((1.65892e-11)*pt^6 + \
                    (-5.71108e-09)*pt^5 + \
                    (8.08921e-07)*pt^4 + \
                    (-5.88213e-05)*pt^3 + \
                    (0.00219812)*pt^2 + \
                    (-0.0345875)*pt + 0.968282) + \
      (pt >= 90.) * 0.945514}
}
```

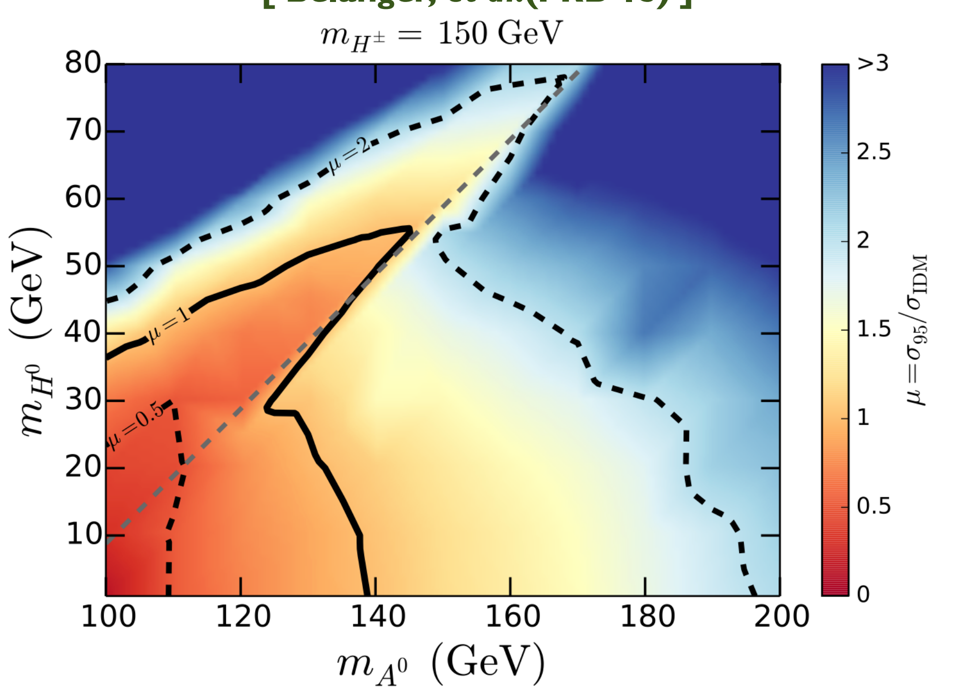
Corresponding implementation in MADANALYSIS 5

Current existing programs

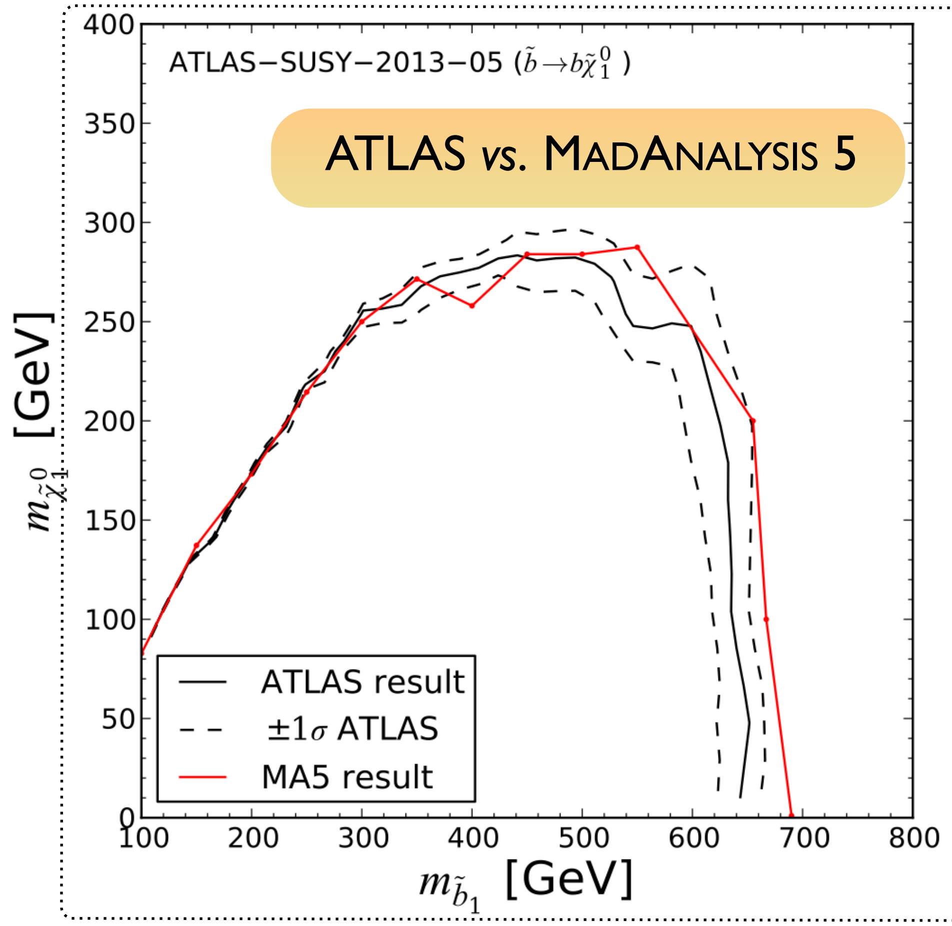
- Two public programs using DELPHES: CHECKMATE and MADANALYSIS 5
  [Drees *et al.* (CPC'14; 2016)] [Conte, BF & Serret (CPC'12); Conte, Dumont, BF & Wymant (EPJC'14)]
- One private program based on RIVET: ATOM [Kim, Papucci, Sakurai & Weiler]
- RIVET can now be used for new physics [Buckley *et al.* (CPC`13)]
- CONTUR uses Standard Model searches [Butterworth *et al.*]

- Examples

[Belanger, *et al.*(PRD'15)]



Constraining the inert Higgs doublet model with SUSY searches and MADANALYSIS 5



Recasting made easy with MADANALYSIS 5 (I)

[Conte, Dumont, BF, Wymant (EPJC '14); Dumont, BF, Kraml *et al.* (EPJC '15)]

✦ Confronting a BSM signal to LHC analyses is straightforward

- Starting point: a showered/hadronized event file
- Installation of the detector simulators: 'install DELPHES'
- Installation of the analysis libraries: 'install PAD' (more analyses with the MA5tune)

✦ In practice:

```
ma5>set main.recast = on
MA5-WARNING: DelphesMA5tune and/or the PADForMA5tune are not installed (or deactivated): the corresponding analyses will be unavailable
ma5>import samples/stops.hep.gz
MA5:    -> Storing the file 'stops.hep.gz' in the dataset 'defaultset'.
ma5>submit
MA5:    Creating folder 'ANALYSIS_0'...
MA5: Would you like to edit the recasting Card ? (Y/N)
Answer: Y
```

✦ Snippet of the recasting card (only on/off switches to be set by the user)

- O(20) 8 TeV ATLAS and CMS analyses; O(5) 13 TeV ATLAS+CMS analyses

```
atlas_1605_03814     v1.2   on   delphes_card_ATLAS_1604_07773.tcl        # ATLAS - 13 TeV - multijet (2-6 jets) + met
ATLAS_1604_07773     v1.2   on   delphes_card_ATLAS_1604_07773.tcl        # ATLAS - 13 TeV - monojet
ATLAS_EXOT_2014_06   v1.2   on   delphes_card_atlas_sus_2013_05_pad.tcl   # ATLAS - 8 TeV - monophoton
cms_exo_12_047       v1.2   on   delphes_card_cms_b2g_12_012.tcl          # CMS   - 8 TeV - monophoton
cms_exo_12_048       v1.2   on   delphes_card_cms_b2g_12_012.tcl          # CMS   - 8 TeV - monojet
cms_b2g_14_004       v1.2   on   delphes_card_cms_b2g_14_004.tcl          # CMS   - 8 TeV - Dark matter production with a ttbar pair
cms_b2g_12_022       v1.2   on   delphes_card_cms_b2g_14_004.tcl          # CMS   - 8 TeV - Monotop search
CMS_B2G_12_012       v1.2   on   delphes_card_cms_b2g_12_012.tcl          # CMS   - 8 TeV - T5/3 partners in the SSDL channel
```

Recasting made easy with MADANALYSIS 5 (2)

[Conte, Dumont, BF, Wymant (EPJC '14); Dumont, BF, Kraml *et al.* (EPJC '15)]

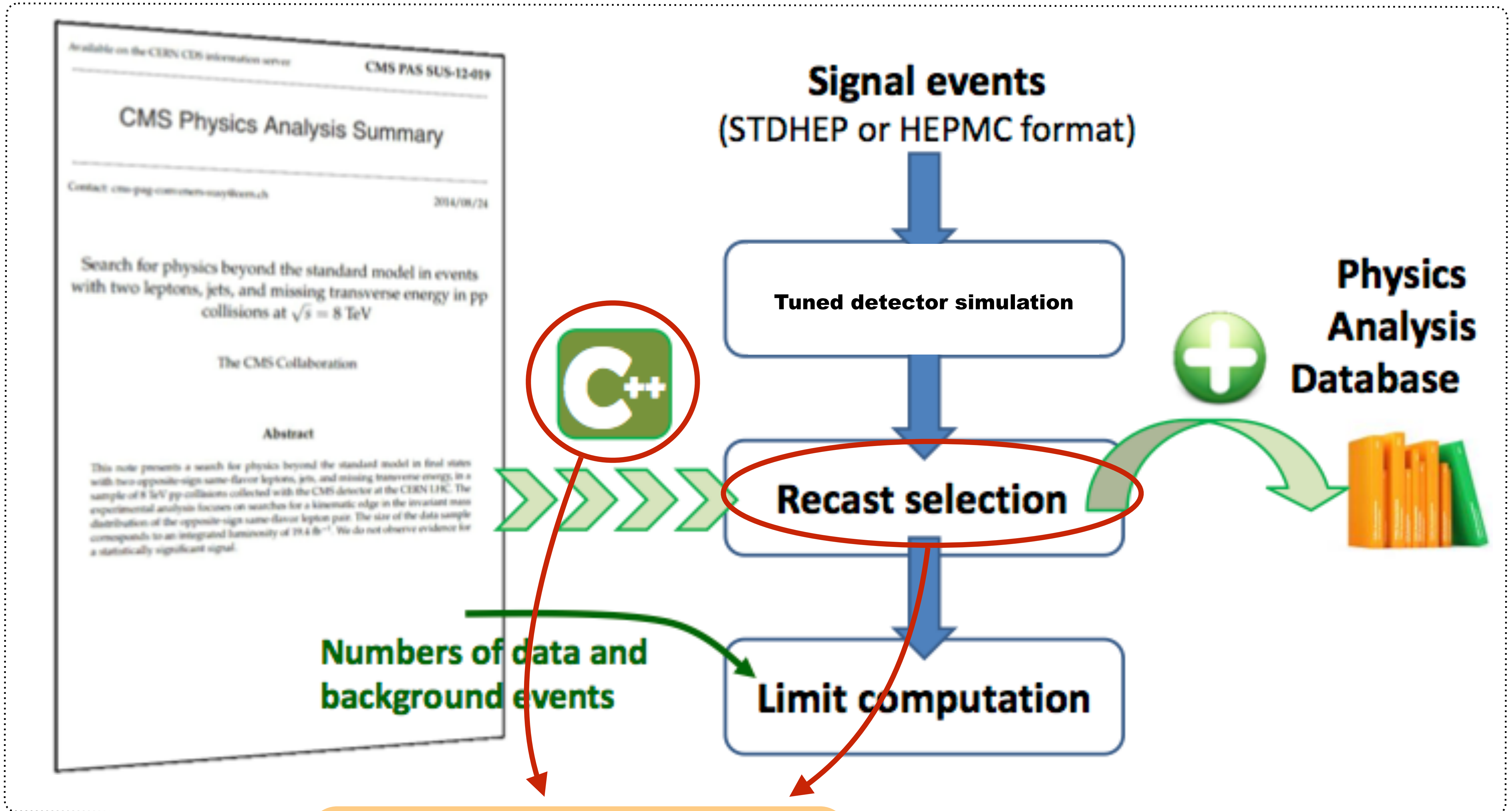
- Snippet of the output file (example: low statistics ➢ lots of '-1' in the example)
  - CLs if a signal cross section is provided
  - Cross sections excluded at the 95% CL

```
ATLAS_1604_07773    EM1    25.8538538    27.4980471    ||    0.0100000    0.0099499    0.0000000    0.0099499
ATLAS_1604_07773    EM2    -1            -1            ||    0.0000000    0.0000000    0.0000000    0.0000000
ATLAS_1604_07773    EM3    -1            -1            ||    0.0000000    0.0000000    0.0000000    0.0000000
ATLAS_1604_07773    EM4    -1            -1            ||    0.0000000    0.0000000    0.0000000    0.0000000
ATLAS_1604_07773    EM5    -1            -1            ||    0.0000000    0.0000000    0.0000000    0.0000000
ATLAS_1604_07773    EM6    -1            -1            ||    0.0000000    0.0000000    0.0000000    0.0000000
ATLAS_1604_07773    EM7    -1            -1            ||    0.0000000    0.0000000    0.0000000    0.0000000
ATLAS_1604_07773    IM1    58.3118133    52.7020233    ||    0.0100000    0.0099499    0.0000000    0.0099499
ATLAS_1604_07773    IM2    -1            -1            ||    0.0000000    0.0000000    0.0000000    0.0000000
ATLAS_1604_07773    IM3    -1            -1            ||    0.0000000    0.0000000    0.0000000    0.0000000
ATLAS_1604_07773    IM4    -1            -1            ||    0.0000000    0.0000000    0.0000000    0.0000000
ATLAS_1604_07773    IM5    -1            -1            ||    0.0000000    0.0000000    0.0000000    0.0000000
ATLAS_1604_07773    IM6    -1            -1            ||    0.0000000    0.0000000    0.0000000    0.0000000
```

Reimplemenation challenges

Reimplementing an analysis: the challenges

✦ Recasting strategy (as in MADANALYSIS 5)



How easy is it to implement (and validate) an LHC analysis?

Implementing a new analysis in a recasting tool

✦ Picking up an experimental publication
- ✤ Reading
- ✤ Understanding

✅ **Relatively easy**

✦ Writing the analysis code in the tool internal language ✅ **Relatively easy**

✦ Getting the information missing from the publication for a proper validation
- ✤ Efficiencies (trigger, electrons, muons, b-tagging, JES, *etc*.)
  - ★ Including p_T and/or η dependence
  - ★ Accurate information
- ✤ Detailed cutflows for some well-defined benchmark scenarios
  - ★ Exact definition of the benchmarks (spectra)
  - ★ Event generation information (cards, tunes, *etc*.)
- ✤ Expected number of events in each region and cross sections
- ✤ Digitized histograms (*e.g.*, on HEPDATA)

⚠️ **Essential**
❌ **Often difficult!**

✦ Comparing theory tools and real life

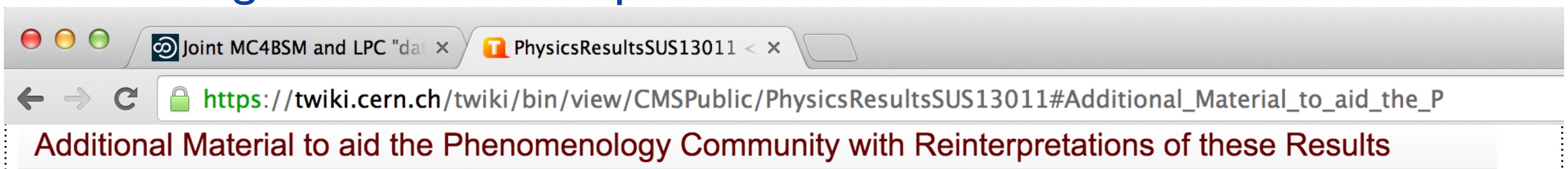
Ex. 1: CMS-SUS-13-11 (stops with one lepton)

[Conte, Dumont, BF, Wymant ('14)]

- Missing information for the validation
  - Efficiencies
  - Cutflows and Monte Carlo information for given benchmarks

✅ **Discussions with CMS needed**

- All missing information was provided



✅ **Update of the analysis wiki page**
Shared LHE files and PYTHIA cards

- Validation

Cut	MadAnalysis 5	CMS
At least one lepton, four jets and 100 GeV of missing transverse energy	31.4	29.7
At least one b-tagged jet	27.1	25.2
No extra loosely-isolated lepton or track	22.5	21.0
No hadronic tau	22.0	20.6
Angular separation between the missing momentum and the two hardest jets	18.9	17.8
Hadronic top quark reconstruction	12.7	11.9
The transverse mass M_T (defined in the text) is larger than 120 GeV	10.4	9.6
At least 300 GeV of missing transverse energy and $M_{T2}^W > 200$ GeV	5.1	4.2

*Ex.*2: ATLAS-EXO-2014-04 (monophoton)

[Barducci ('15)]

- Missing information
  - Crack in the detector: no photons in the [1.37-1.52] η-range
  - Tight photon requirements

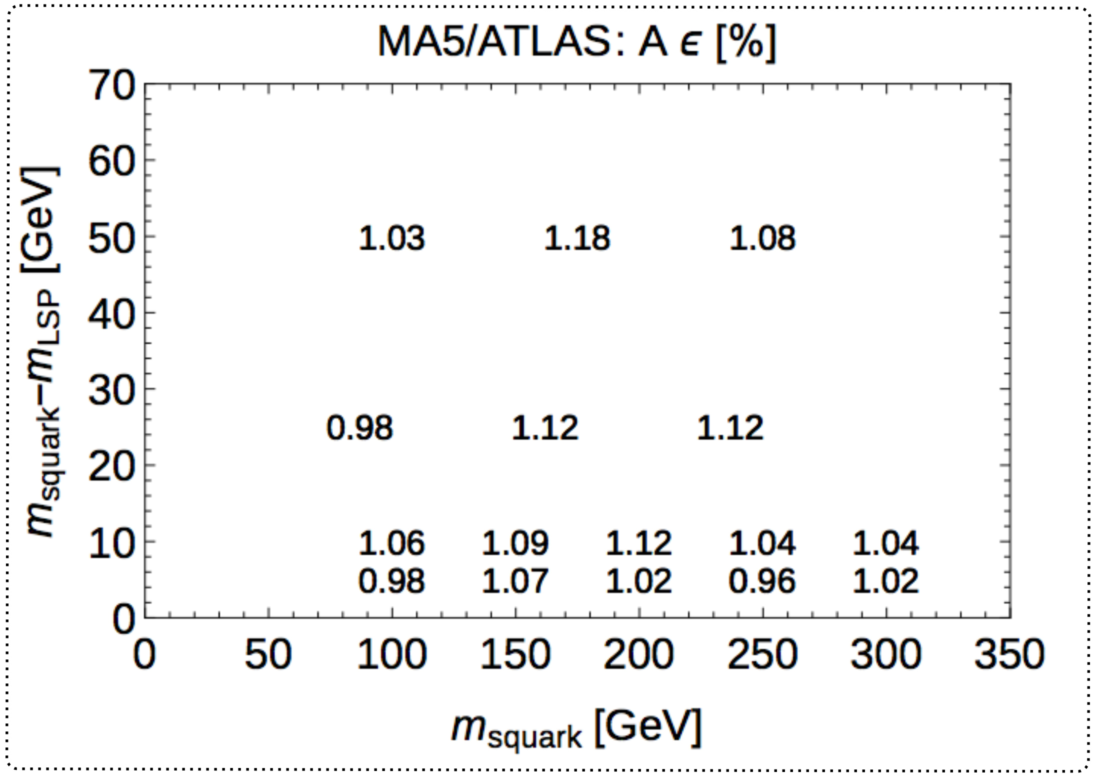
✅ Discussions with ATLAS needed

✅ In ATLAS-COM-PHYS-2014-542

- Event generation for the test benchmarks
  - Monte Carlo information (cards, tunes, *etc.*)

✅ Kindly provided by ATLAS

Very good results
(ratio of efficiencies)



*Ex.*3: Recasting CMS-EXO-12-048

[Conte, BF, Guo ('16)]

✦ Missing information for the validation

- ✤ Discussion with CMS to get validation benchmarks
- ✤ Cutflows and Monte Carlo information for given benchmarks

✅ **Discussions with CMS needed**

✦ Validation:

	Selection step	CMS	$\epsilon_i^{\rm CMS}$	MA5	$\epsilon_i^{\rm MA5}$	$\delta_i^{\rm rel}$
0	Nominal	84653.7		84653.7		
1	One hard jet	50817.2	0.6	53431.28	0.631	5.2%
2	At most two jets	36061	0.7096	38547.75	0.721	1.61%
3	Requirements if two jets	31878.1	0.884	34436.35	0.893	1.02%
4	Muon veto	31878.1	1	34436.35	1.000	0
5	Electron veto	31865.1	1	34436.35	1.000	0
6	Tau veto	31695.1	0.995	34397.54	0.998	0.3%
	$\not\!\!E_T > 250$ GeV	8687.22	0.274	7563.04	0.219	20.00%
	$\not\!\!E_T > 300$ GeV	5400.51	0.621	4477.67	0.592	4.66%
	$\not\!\!E_T > 350$ GeV	3394.09	0.628	2813.70	0.628	0.00%
	$\not\!\!E_T > 400$ GeV	2224.15	0.6553	1753.71	0.623	4.93%
	$\not\!\!E_T > 450$ GeV	1456.02	0.654	1110.92	0.633	3.21%
	$\not\!\!E_T > 500$ GeV	989.806	0.679	722.83	0.650	4.27%
	$\not\!\!E_T > 550$ GeV	671.442	0.678	487.54	0.674	0.59%

✅ **Validated at the 20% level**

Issue with the low-MET modelling in DELPHES

Ex.4 : When things are borderline...

- ✦ ATLAS-EXOT-2014-04 (monophotons)
  - ✤ Effects non-reproducible with DELPHES (cleaning cuts, triggers, good vertexing)
- ✦ ATLAS-SUS-2013-09 (stops in the dilepton channel)
  - ✤ Information on effects non-reproducible with DELPHES lost (student quitted)

⚠ **Efficiencies computed by hand**
Maybe model-dependent

Very good results
(for a SUSY benchmark)

[Barducci ('15)]

Cut	ATLAS	Rel. decr.	MA5 (ul ul~)	Rel. decr.
Nominal	9989		9989	
a. Trigger	8582		?	
b. Good Vertex	8574			
c. Cleaning cuts	8213			
0. $E_T^{\rm miss} > 150$ GeV	4131		4384	
1. 1 loose γ, $p_T > 125$ GeV, $\lvert\eta\rvert < 2.37$	2645	−36.0	2637	−39.8
2. Tight leading γ with $\lvert\eta\rvert < 1.37$	2068	−21.8	2052	−22.2
3. Isolated leading γ	1898	−8.2	1856	−9.6
4. $\Delta\phi(\gamma^{\rm leading}, E_T^{\rm miss}) > 0.4$	1887	−0.6	1840	−0.8
5. $N_{\rm jet} \leq 1$ and $\Delta\phi(jet, E_T^{\rm miss}) > 0.4$	1219	−35.4	1234	−33.0
6. Lepton veto	1188	−2.5	1233	−0.1

Ex.5: And the darkness came…

✦ Missing or incomplete validation information

- ✤ CMS-SUS-12-028 (α_T)
  - ★ No cutflows; no answers from CMS to requests

❌ **Dead end!**

- ✤ CMS-SUS-13-007 (1 lepton+b-jets+met)
  - ★ Semi-official validation material (that cannot be used for a public validation)
  - ★ No cutflows
  - ★ Messy definition of the benchmark points

⚠ **We had to forget about it**

✦ Missing or incomplete analysis information

⚠ **Too vague!**

- ✤ ATLAS-EXOT-2013-10 (monolepton)
  - ★ The average trigger efficiency is 80%-90% in the muon channel
  - ★ 80% of the muons are reconstructed with most of the loss coming from…
  - ★ No precise information on signal event generation
  - ★ No signal distributions on HEPDATA

Unfortunately: many more examples!

A wishlist for experimentalists - analysis

✦ Analysis description

- Clear description of the selections, including their sequence
  - ★ Tabulated form appreciable
- Efficiencies for objects (electrons, muons, jets, taus, b-tagging, *etc.*)
  - ★ Including p_T and η dependence
- Efficiencies for triggers, event cleaning, *etc.*
  - ★ Effects that cannot be modeled in our fast simulation
- Digitized figures (ROOT format, text format, *etc.*)
- Special variables (*e.g.*, the CMS razor, asymmetric M_{T2})
  - ★ Snippets of code highly appreciated

A wishlist to experimentalists - validation

✦ Validation material ➢ quality of the reinterpretation

✤ Public information on benchmark scenarios
  ★ Spectra and decay tables (under an SLHA-form)
  ★ Several scenarios are appreciable

✤ Public information on the Monte Carlo tools configuration
  ★ Cards, tunes, merging information, *etc*.

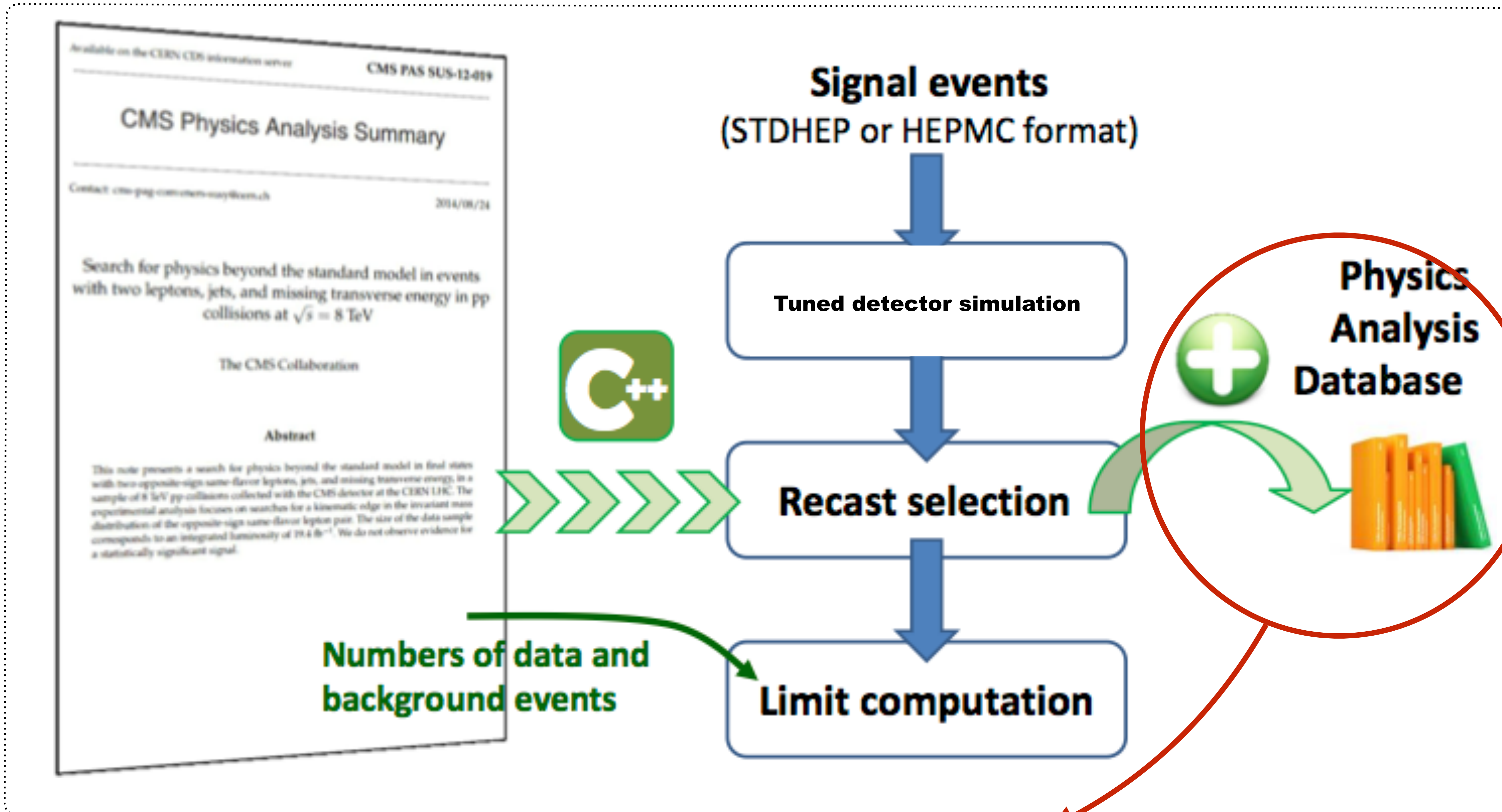
✤ Detailed cutflows for the benchmarks, with the correct cut ordering
  ★ Including each step of the (pre)selection
  ★ For several benchmarks
  ★ The more steps available, the better (preselection, cleaning, *etc*.)
  (pin down the differences of our machinery with CMS-ATLAS simulations)

✤ Kinematical distributions at different steps of the selection
  ★ Extra cross-checks

Preservation

The LHC legacy

✦ Recasting strategy (as in MADANALYSIS 5)



How to store recasted analyses?
Part of the LHC legacy

MADANALYSIS 5 analyses on INSPIRE

[BF, Martini ('16)]

✦ Implementation of LHC analyses can be uploaded on INSPIRE

✤ DOI are assigned: can be cited, searched for, *etc.*

Information | Citations (1) | Files

Files are versioned, can be downloaded

MadAnalysis5 implementation of the CMS search for dark matter production with top quark pairs in the single lepton channel (CMS-B2G-14-004)

DOI and citations

Fuks, Benjamin; Martini, Antony

Description: This is the MadAnalysis5 implementation of the CMS search for dark matter in a channel where a pair of dark matter particles is produced in association with a top-antitop system. This search targets events featuring a single lepton originating from the top decays and a large amount of missing transverse energy.

Information how to use this code and a detailed validation summary are available at http://madanalysis.irmp.ucl.ac.be/wiki/PhysicsAnalysisDatabase. The CMS analysis is documented at https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsB2G14004.

Cite as: Fuks, B., Martiny, A. (2016). MadAnalysis5 implementation of the CMS search for dark matter production with top quark pairs in the single lepton channel (CMS-B2G-14-004). doi: 10.7484/INSPIREHEP.DATA.MIHA.JR4G

Automatic installation of all implemented analyses from MADANALYSIS 5

Record added 2016-05-09, last modified 2016-05-09

The Public Analysis Database of MADANALYSIS

[Dumont, BF, Kraml *et al.* (EPJC '15)]

- A database with MADANALYSIS 5 implementations of LHC analyses exists
  - http://madanalysis.irmp.ucl.ac.be/wiki/PublicAnalysisDatabase
- Snippet of the webpage

ATLAS analyses, 13 TeV

Analysis	Short Description	Implemented by	Code	Validation note	Version
ATLAS-EXOT-2015-03	monojet + missing transverse energy	D. Sengupta	Inspire	PDF	v1.3/Delphes3
ATLAS-SUSY-2015-06	jets + missing transverse momentum	S. Banerjee, B. Fuks, B. Zaldivar	Inspire	PDF	v1.3/Delphes3

Delphes card for ATLAS-EXOT-2015-03

Dedicated DELPHES cards

Code from INSPIRE

Validation information (cutflows, distributions, *etc.*)

- Can be automatically installed within MADANALYSIS 5

Physics

NLO effects on a CLs: top-philic dark matter

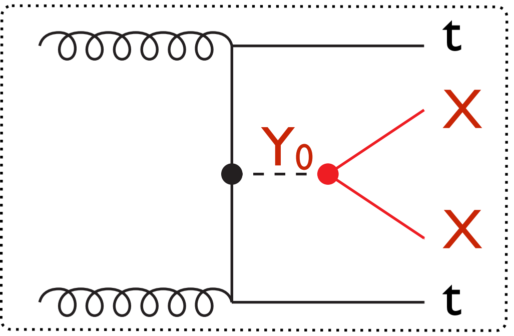
[Arina, Backovic, Conte, BF, Guo *et al.* (JHEP'16)]

✦ A simplified model for top-philic dark matter

[BF & Martini (2016)]

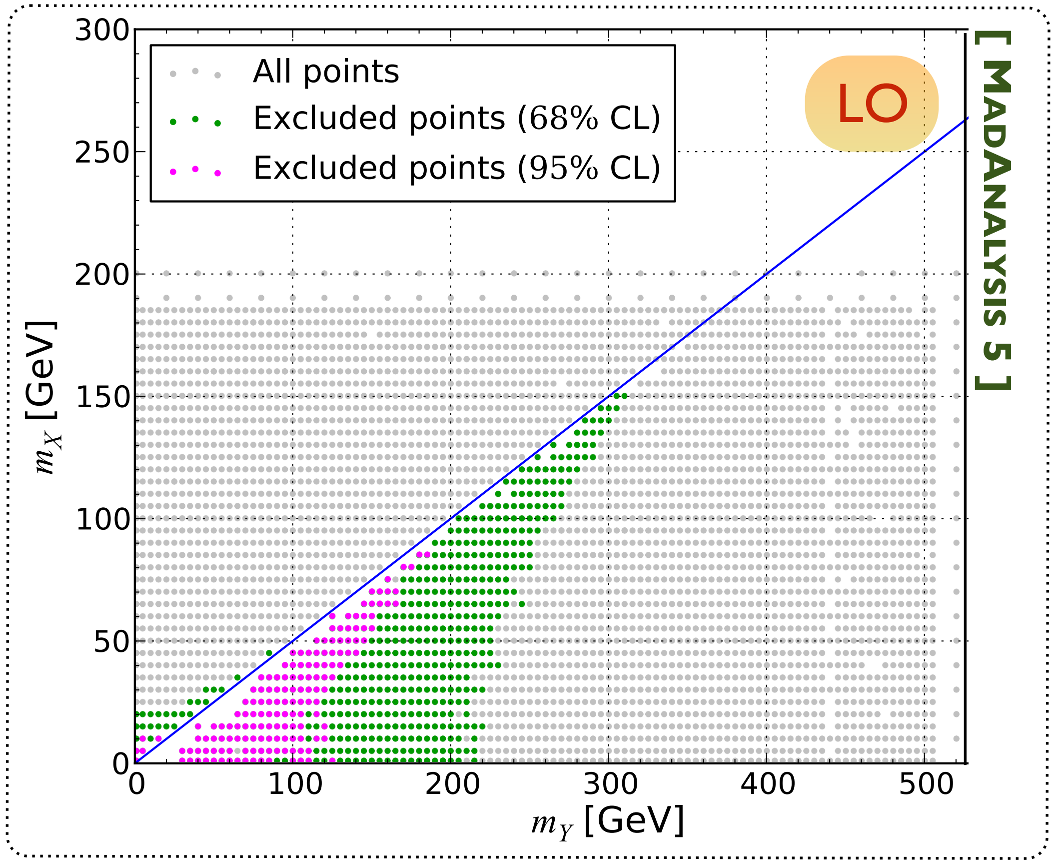
- A dark sector with a fermionic dark matter candidate X
- A (scalar) mediator Y_0 linking the dark sector and the top

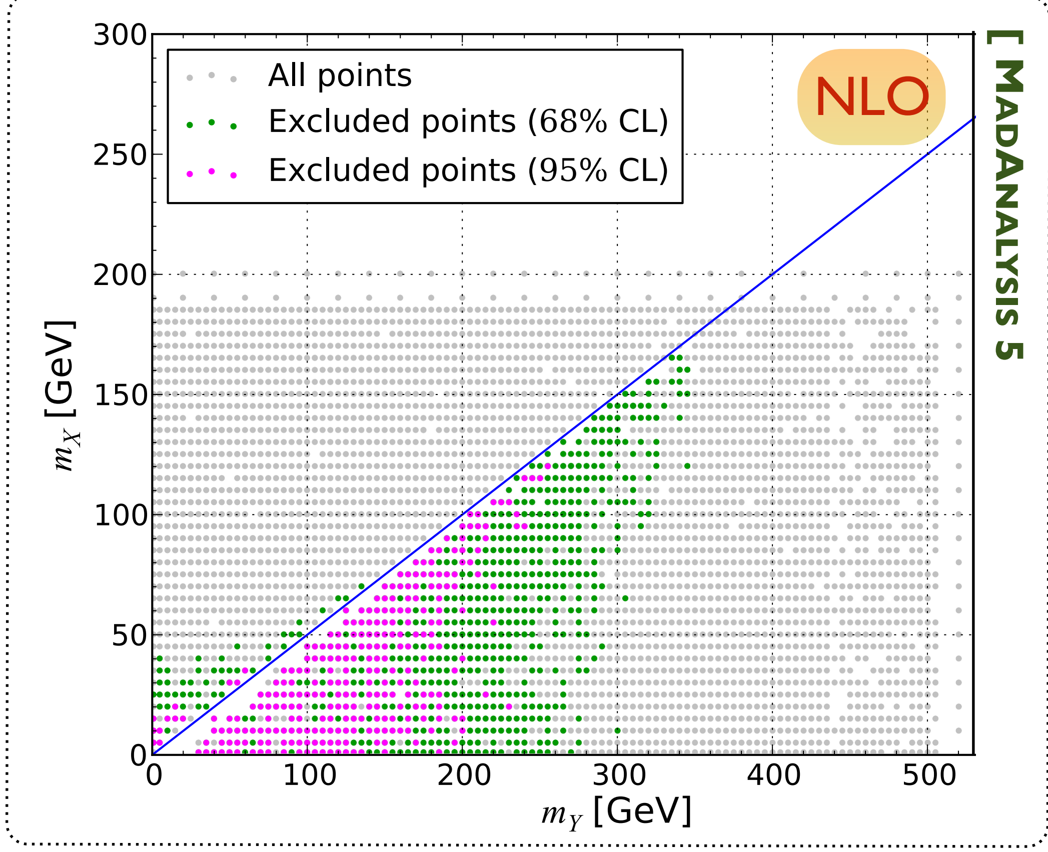
$$\mathcal{L}_{t,X}^{Y_0} = -\Big(g_t \frac{y_t}{\sqrt{2}} \bar{t}t + g_X \bar{X}X\Big) Y_0$$



- Could be probed with tt+MET events (CMS-B2G-14-004)

✦ For central scales: mild (but visible) NLO effects on the exclusions





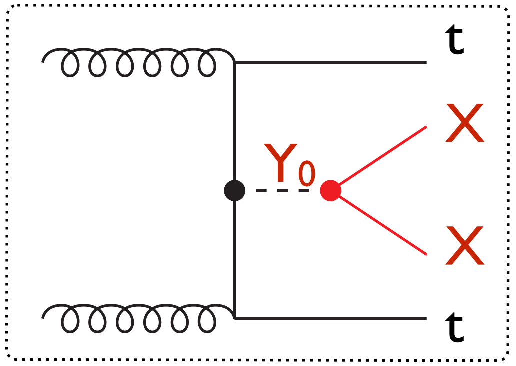
- How is the picture changing when including scale variations?

NLO effects on a CLs: top-philic dark matter

[Arina, Backovic, Conte, BF, Guo *et al.* (JHEP'16)]

✦ There are theoretical uncertainties on a CLs number

	(m_Y, m_X)	$\sigma_{\rm LO}$ [pb]	$\rm CL_{LO}$ [%]	$\sigma_{\rm NLO}$ [pb]	$\rm CL_{NLO}$ [%]
I	(150, 25) GeV	$0.658^{+34.9\%}_{-24.0\%}$	$98.7^{+0.8\%}_{-13.0\%}$	$0.773^{+6.1\%}_{-10.1\%}$	$95.0^{+2.7\%}_{-0.4\%}$
II	(40, 30) GeV	$0.776^{+34.2\%}_{-24.1\%}$	$74.7^{+19.7\%}_{-17.7\%}$	$0.926^{+5.7\%}_{-10.4\%}$	$84.2^{+0.4\%}_{-14.4\%}$
III	(240, 100) GeV	$0.187^{+37.1\%}_{-24.4\%}$	$91.6^{+6.4\%}_{-18.1\%}$	$0.216^{+6.7\%}_{-11.4\%}$	$86.5^{+8.6\%}_{-5.5\%}$



✤ An excluded point (95% CL) may not be excluded when accounting for errors

✤ The CLs number can increase / decrease at NLO

✤ The error band is reduced

Summary

✦ The LHC legacy

- It is crucial to be able to reinterpret the LHC results in any theoretical context
- This is a very active field of the last few years: several tools are available
- **Reproducibility** is the ability of an entire experiment to be reproduced, (possibly by an independent theoretical study)

✦ Two approaches

- The simplified model spectrum approach (based on efficiencies and cross sections)
- The fastsim strategy (simulating the detector in some ways)

✦ Recasting in MADANALYSIS 5

- MADANALYSIS 5 has been actively developed along the 'fastsim strategy' lines
  - ★ User-friendly way to confront any MC-simulated BSM signal to LHC results