Collider Phenomenology & LHC Recasting with MadAnalysis 5

Jack Y. Araz
Institute For Particle Physics Phenomenology
Durham University

On behalf of Ma5-Dev Team:
JYA, Benjamin Fuks & Eric Conte

Oklahoma State University HEP Group
July 28th 2021

Outline

- Introduction
    - Briefly Normal Mode & Expert Mode
- LHC recasting with MadAnalysis 5
    - NEW Improving exclusion limits statistics with likelihood
    - NEW Particle propagation module for long-lived particles
- Conclusion
- Hands-on session

Introduction

Why designing & recasting is important?

- Exploiting the full potential of LHC (for new physics)
  - *Designing* new analyses (based on MC simulations)
  - *Recasting* LHC analyses (The LHC legacy)
- Data preservation in HEP is mandatory
  - Going beyond raw data via analyses
- Related tools need to be supported by the entire community
  - Both theorists & experimentalists
- Universal recasting tool

Les Houches Recommendations (EPJC '12)

Reinterpretation Forum Report (SciPost '20)

MadAnalysis 5

What is MadAnalysis 5?

- A framework for phenomenological analyses
- Any level of sophistication: partonic, hadronic, detector, reconstructed
- Several input formats: STDHEP, HEPMC, LHE, LHCO, ROOT (from Delphes)
- User-friendly, flexible & Fast!!!
- Interfaces several HEP packages: MadGraph, FastJet, Delphes, pyhf

Normal Mode

- Intuitive commands typed in the Python interface
- Analysis performed behind the scenes (black box)
- Human readable output: HTML and LaTeX

Expert Mode

- C++ programming with the SampleAnalyzer framework
- Support for multiple sub-analyses, an efficient way for handling cuts and histograms, etc.

LHC recasting with MadAnalysis 5

Durham University

IP3

Reimplementing an analysis in MadAnalysis 5

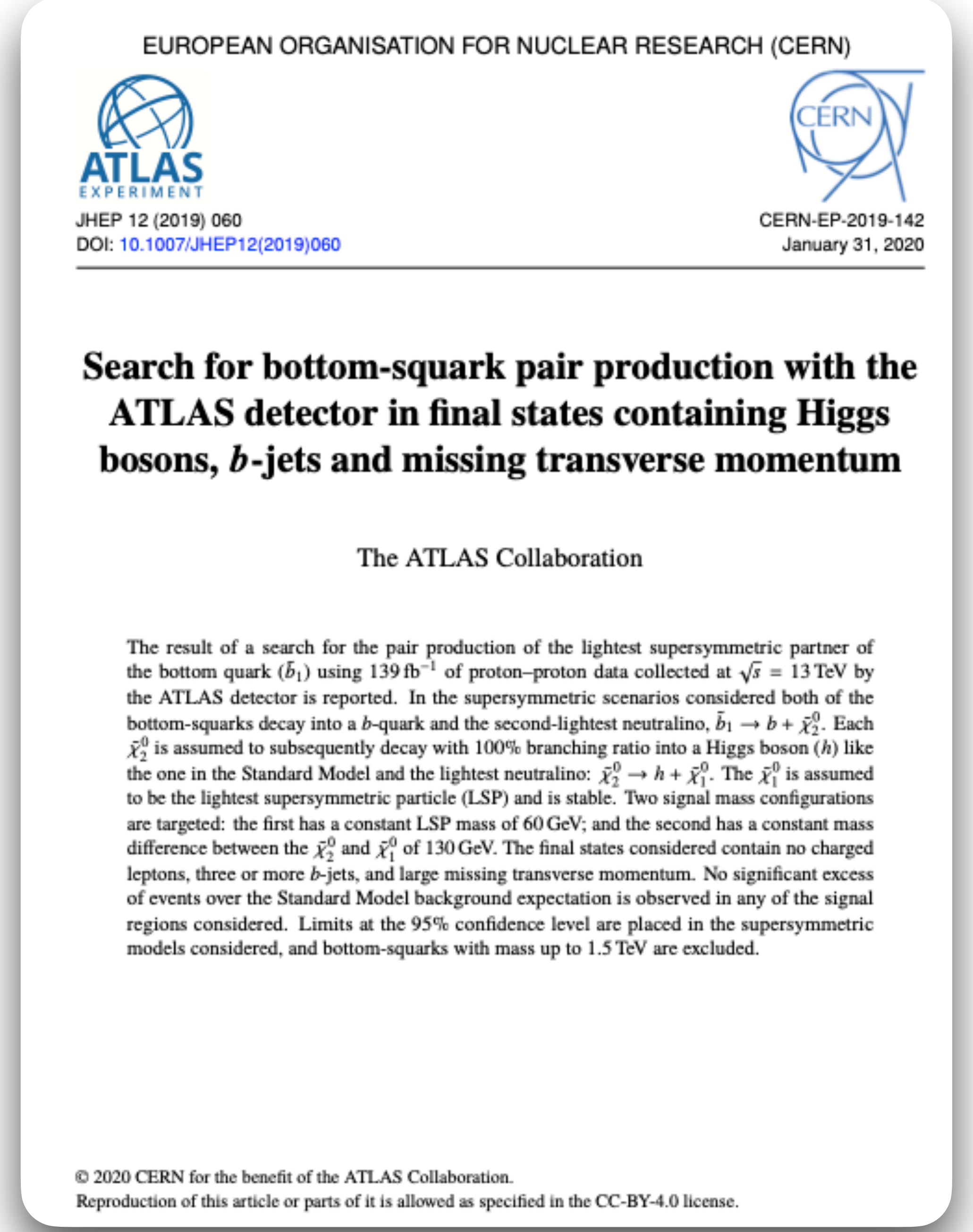
Signal Events
(STDHEP or HEPMC format)

JYA, Fuks, Polykratis EPJC '21

De Favereau et al; JHEP '14

SFS@Ma5 or Delphes

Detector Simulation


EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

JHEP 12 (2019) 060
DOI: 10.1007/JHEP12(2019)060

CERN-EP-2019-142
January 31, 2020

Search for bottom-squark pair production with the ATLAS detector in final states containing Higgs bosons, *b*-jets and missing transverse momentum

The ATLAS Collaboration

The result of a search for the pair production of the lightest supersymmetric partner of the bottom quark ($\tilde{b}_1$) using 139 fb^{-1} of proton–proton data collected at $\sqrt{s} = 13$ TeV by the ATLAS detector is reported. In the supersymmetric scenarios considered both of the bottom-squarks decay into a b-quark and the second-lightest neutralino, $\tilde{b}_1 \to b + \tilde{\chi}^0_2$. Each $\tilde{\chi}^0_2$ is assumed to subsequently decay with 100% branching ratio into a Higgs boson (h) like the one in the Standard Model and the lightest neutralino: $\tilde{\chi}^0_2 \to h + \tilde{\chi}^0_1$. The $\tilde{\chi}^0_1$ is assumed to be the lightest supersymmetric particle (LSP) and is stable. Two signal mass configurations are targeted: the first has a constant LSP mass of 60 GeV; and the second has a constant mass difference between the $\tilde{\chi}^0_2$ and $\tilde{\chi}^0_1$ of 130 GeV. The final states considered contain no charged leptons, three or more b-jets, and large missing transverse momentum. No significant excess of events over the Standard Model background expectation is observed in any of the signal regions considered. Limits at the 95% confidence level are placed in the supersymmetric models considered, and bottom-squarks with mass up to 1.5 TeV are excluded.

C++

Recast

Validation

Public Analysis Database++;

JYA, Frank, Fuks EPJC '20

Alguero, JYA, Fuks, Kraml; in progress

Exclusion limit calculation

Cut-flows
uncertainties

TH uncertainties, HL extrapolations
simplified and full likelihoods

Dumont, Fuks, Kraml, et. al. EPJC '15

Reimplementing an analysis in MadAnalysis 5

Recasting toolbox

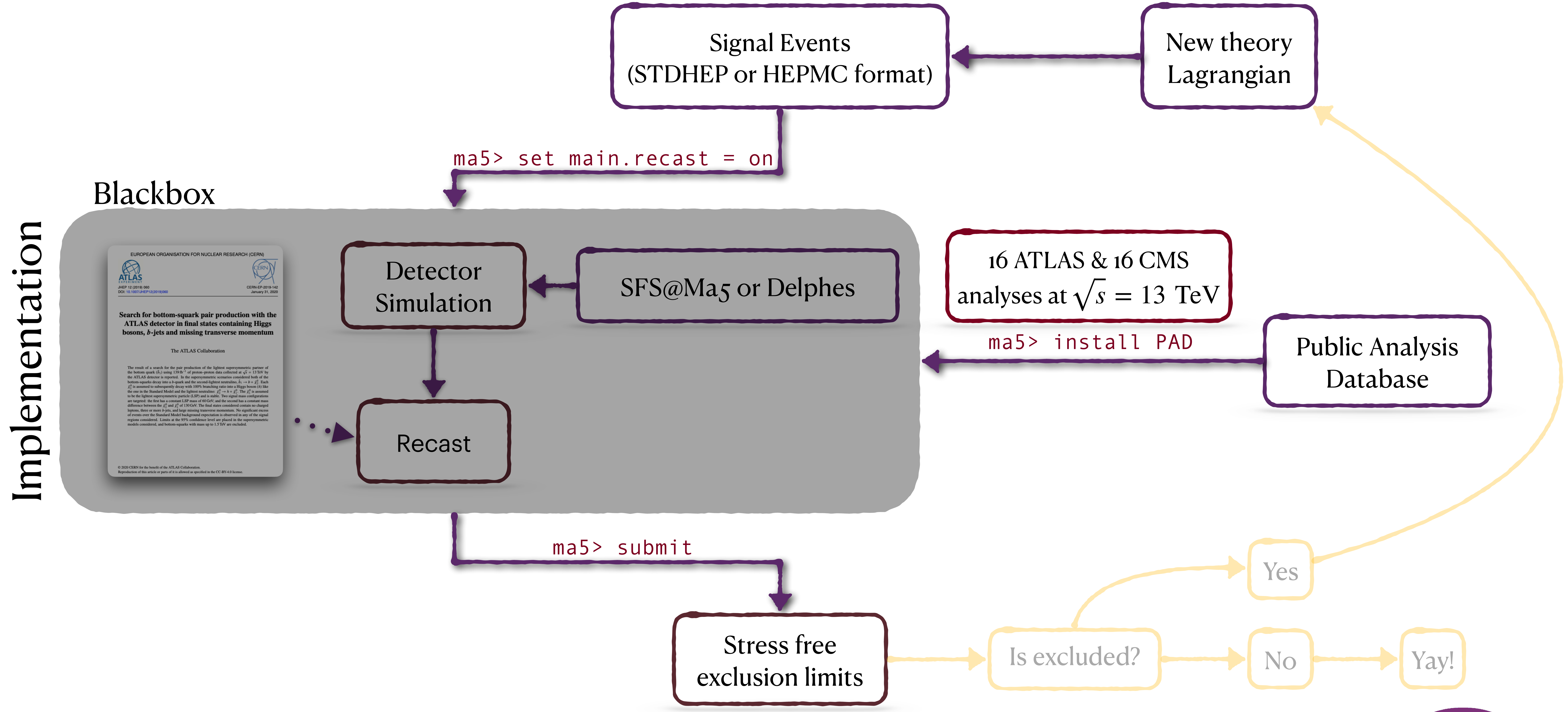
- Calculating exclusion limits, expected and observed excluded cross sections via uncorrelated signal regions.
- NEW Improved limits via full statistical models constructed from HistFactory-like likelihoods (ATLAS)
- NEW Improved limits via simplified likelihoods with CMS' correlation matrices

Alguero, JYA, Fuks, Kraml; 2206.14870

- Exclusion limits with theoretical uncertainties & higher luminosity extrapolations

JYA, Frank, Fuks EPJC '20

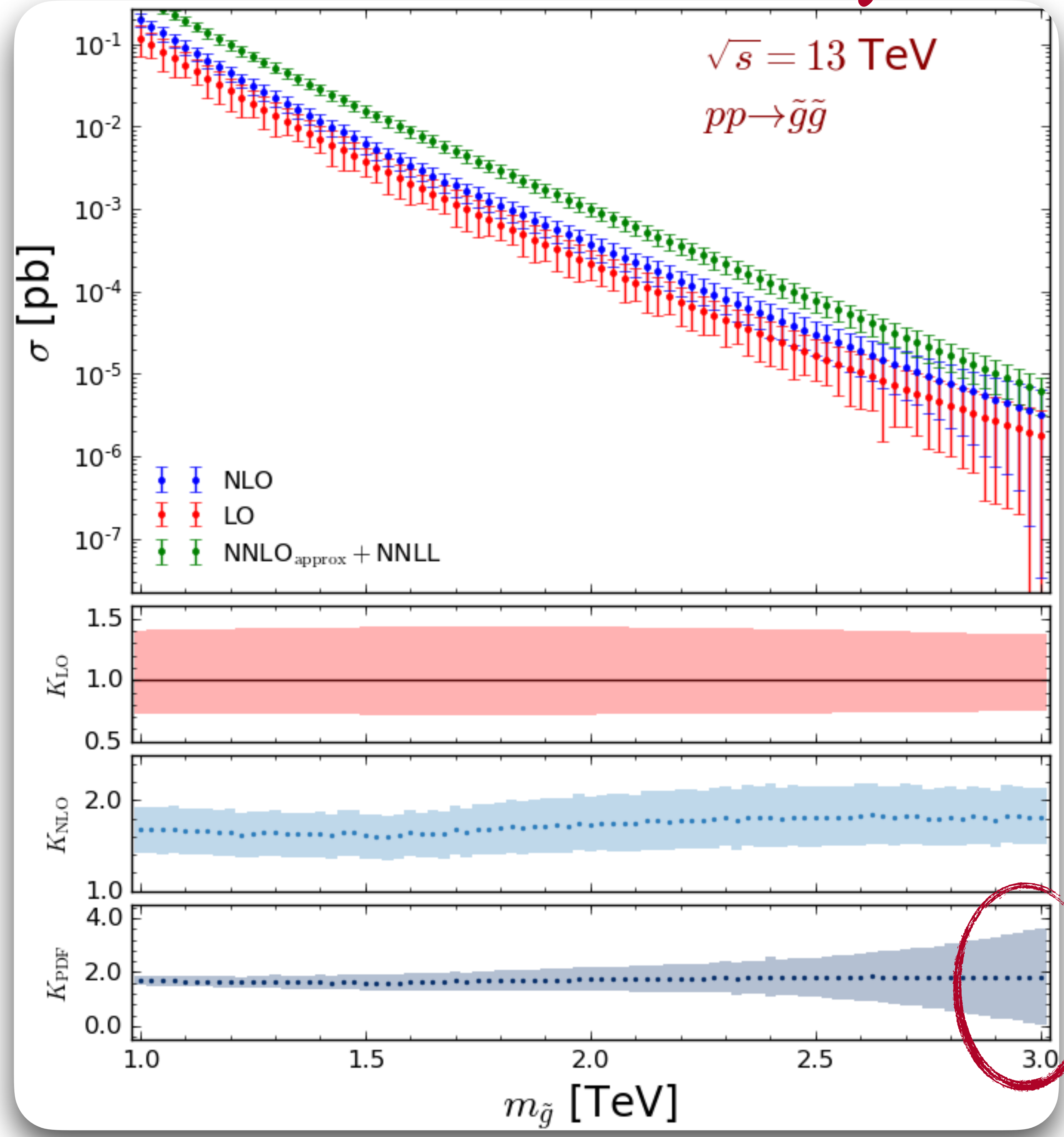
Reimplementing an analysis in MadAnalysis 5

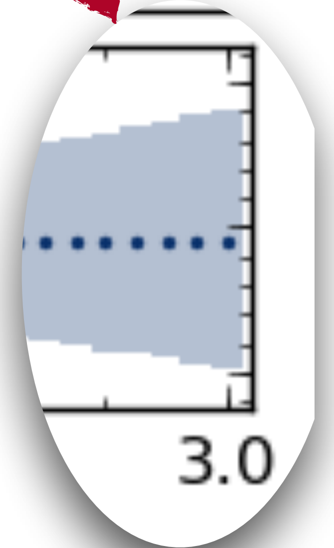


Dumont, Fuks, Kraml, et. al. EPJC '15

More on
Theoretical uncertainties & HL extrapolations

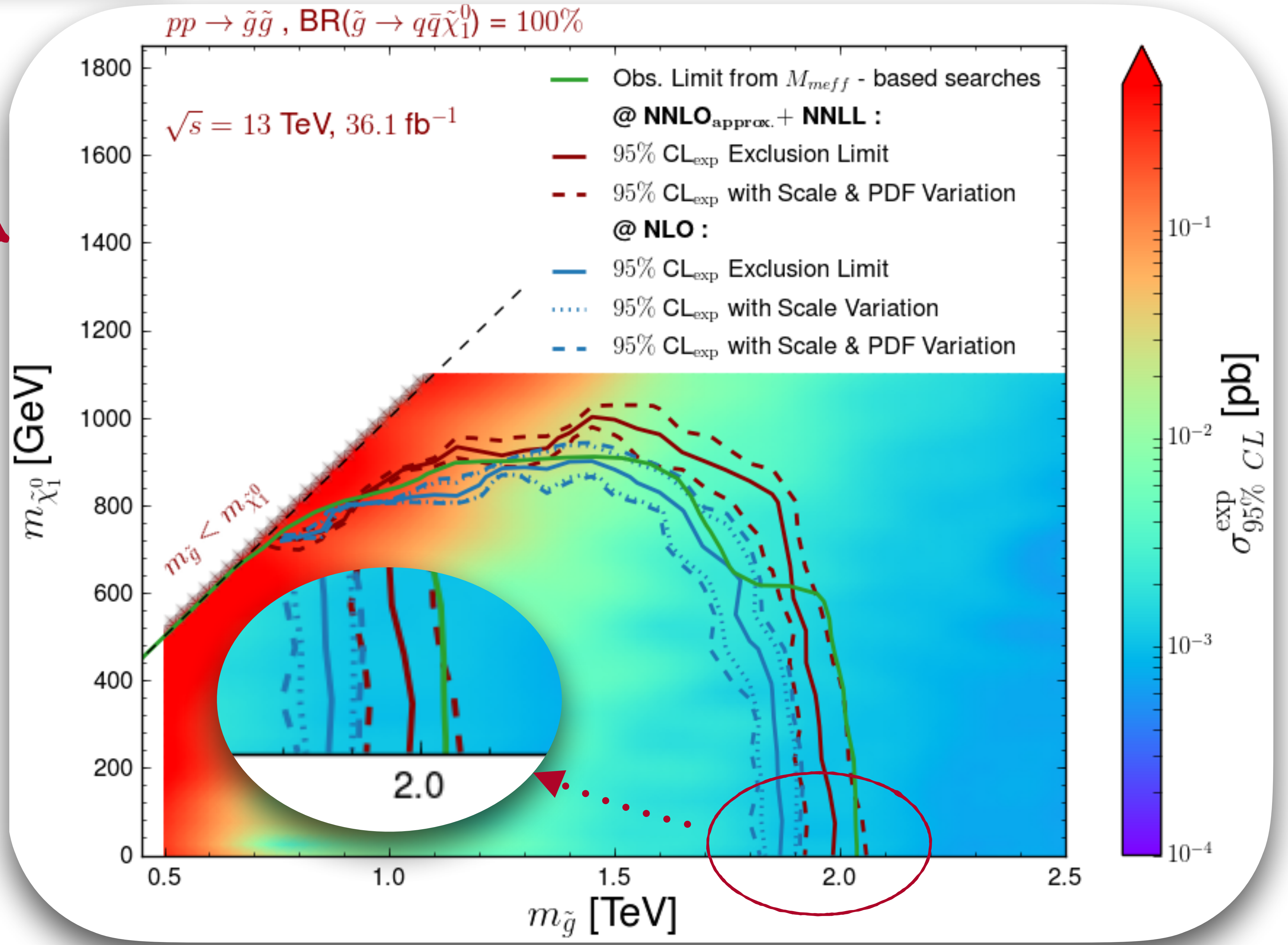
JYA, Frank, Fuks EPJC '20






```
ma5>set defaultset.xsection = 1.8524788535e-06
ma5>set defaultset.scale_up_variation = 0.388
ma5>set defaultset.scale_down_variation = 0.262
ma5>set defaultset.pdf_up_variation = 0.131
ma5>set defaultset.pdf_down_variation = 0.131
```

ATLAS-SUSY-2016-07

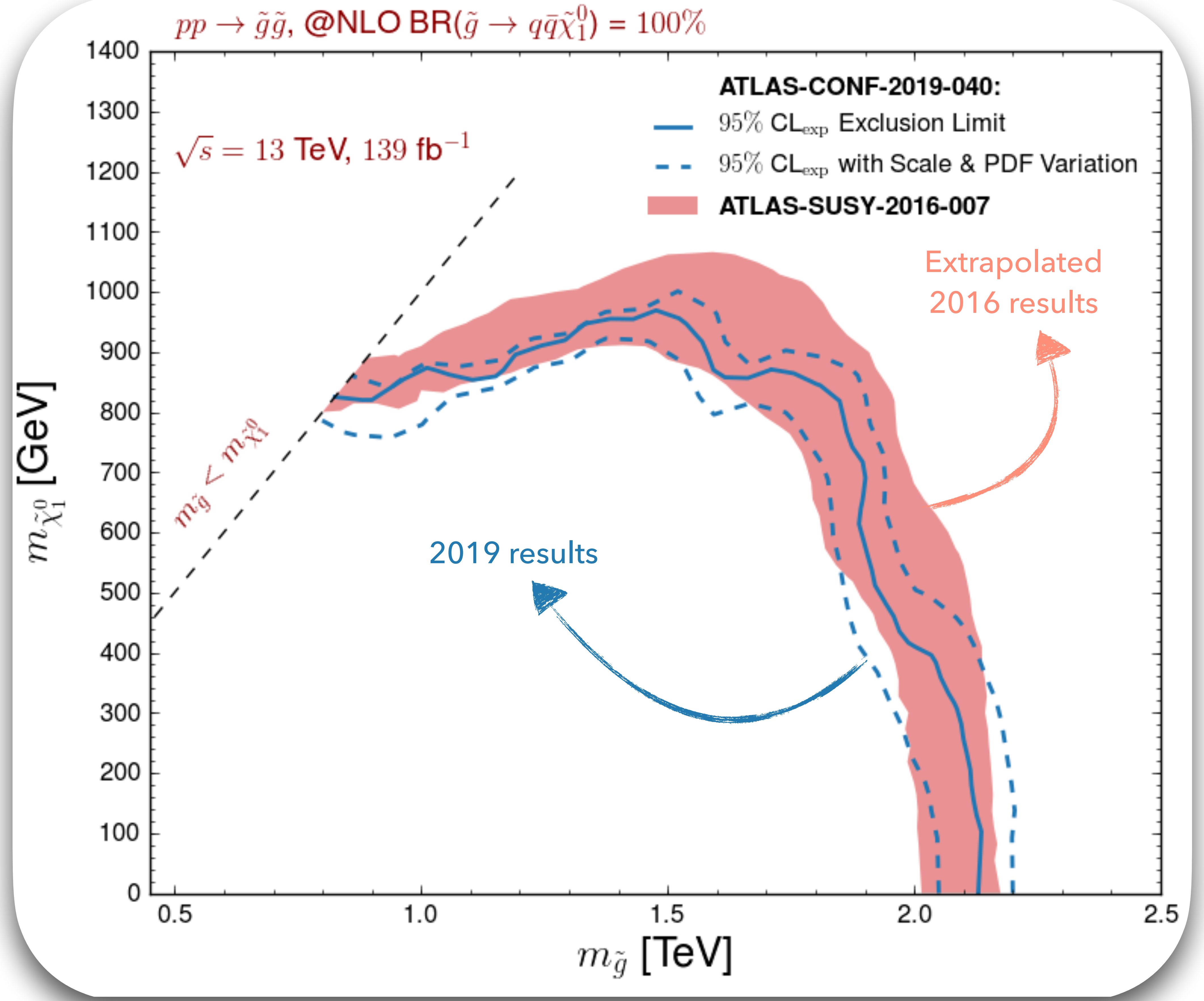


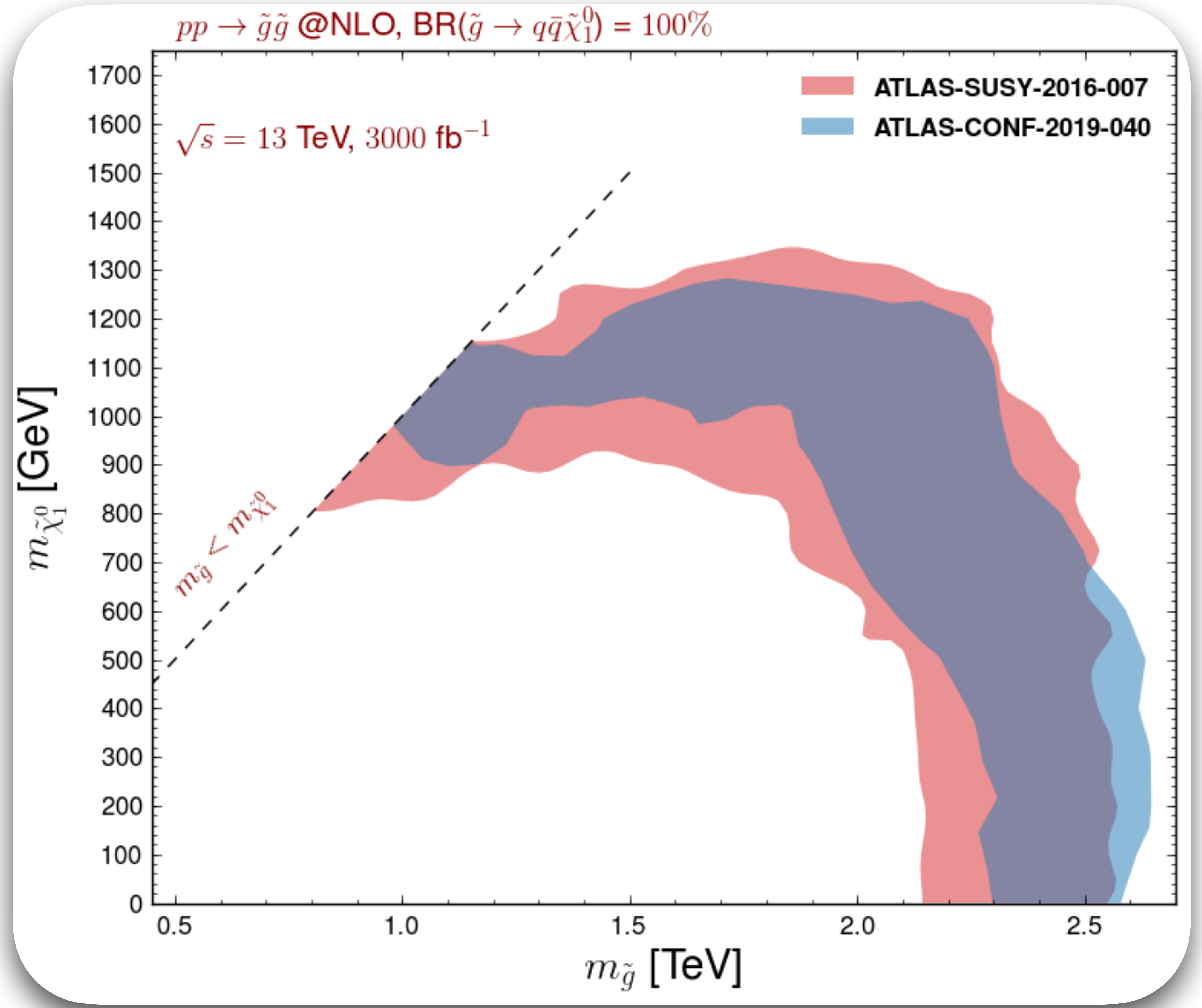
More on

Theoretical uncertainties & HL extrapolations

JYA, Frank, Fuks EPJC '20

```
ma5>set main.recast.add.extrapolated_luminosity = 300 3000
```





More on

Simplified & full likelihoods

Why important?

ATLAS SUSY and Exotics workshop
S. Kraml '20

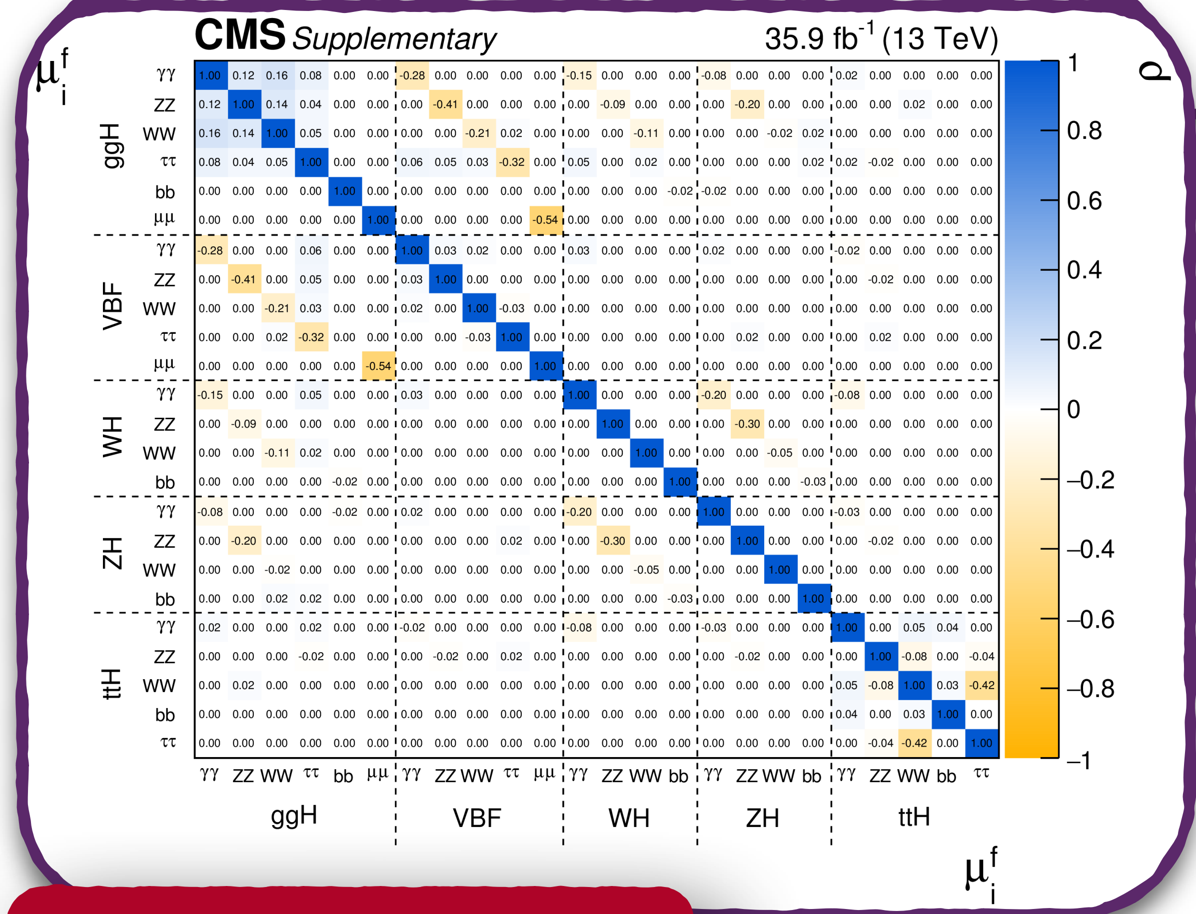
- The mathematical description of the analysis is provided within its statistical model.
- The likelihood profile enables the standard statistical approaches to extract information.
- *i.e.* how reasonably aligned the theoretical predictions with the experimental observations?

Les Houches Recommendations (EPJC '12)

SModelS: Alguero, Kraml, Waltenberg '20

Simplified likelihoods from CMS

CMS-NOTE-2017-001



For details see the "Publication of statistical models: hands on workshop"

Full likelihoods from ATLAS

pyhf
∂ifferentiable ℒikelihoods

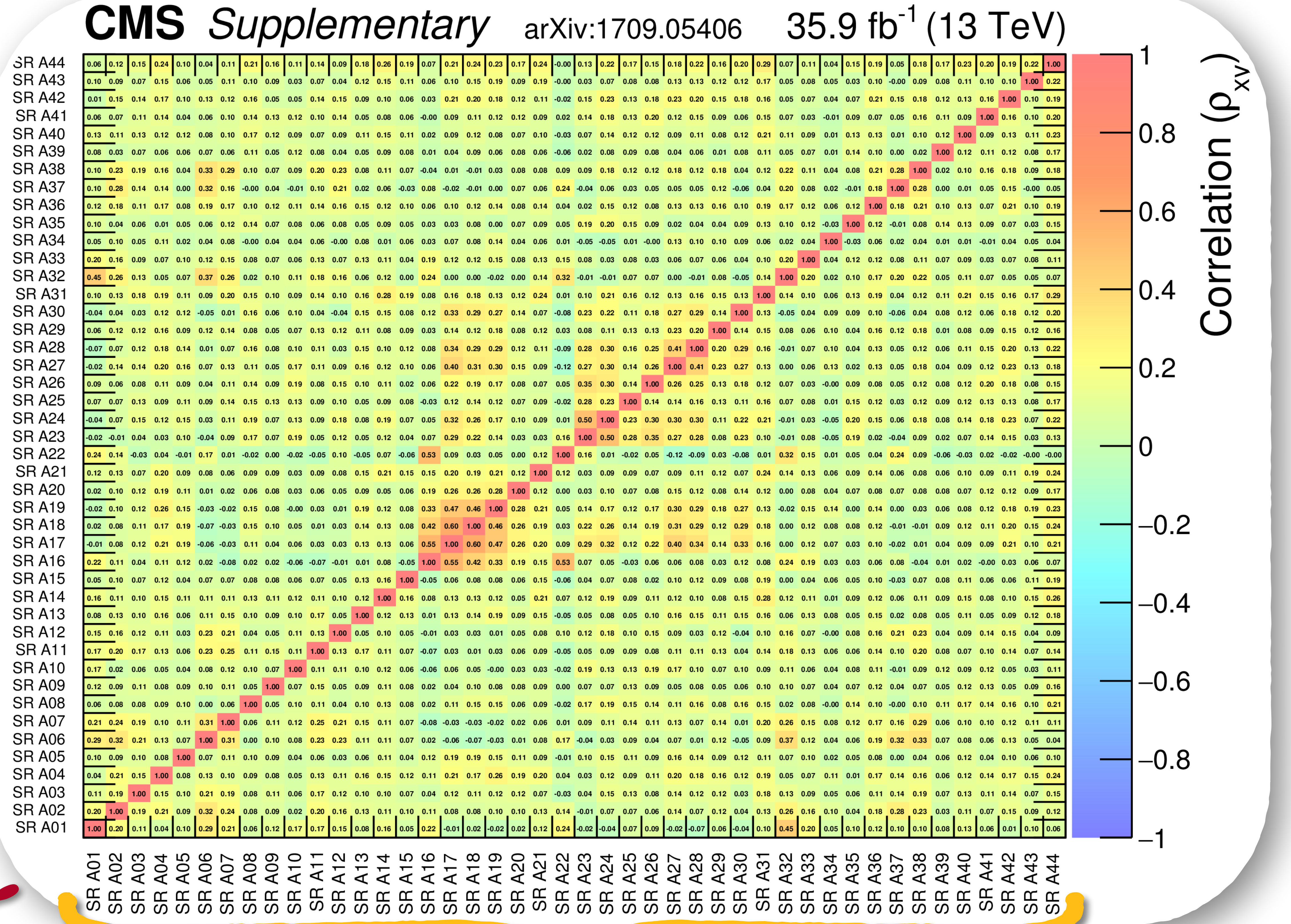
ATL-PHYS-PUB-2019-029

```
ma5> install pyhf
```

More on

Simplified likelihoods with CMS-SUS-16-039

- Correlation matrix allows us to form a covariance matrix which then used in a simplified likelihood under Gaussian approximation.



44 Signal Regions!

$$\mathscr{L}_S(\mu,\theta) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i + \theta_i)^{n_i}\, e^{-(\mu s_i + b_i + \theta_i)}}{n_i!} \exp\left(-\frac{1}{2}\theta^T \mathbf{V}^{-1}\theta\right)$$

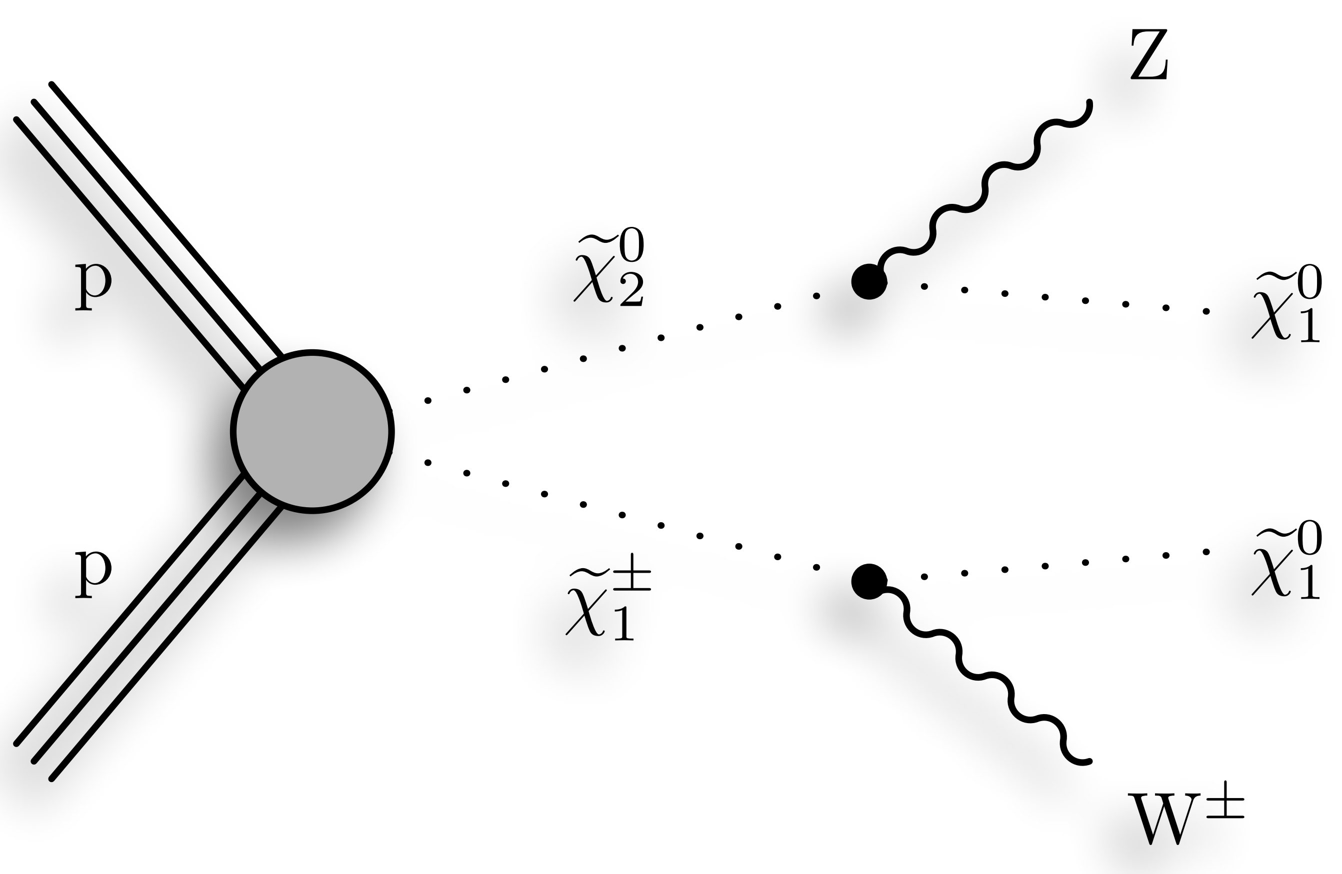
$$\mathbf{V} = \mathbb{E}[\theta_i \times \theta_j]$$

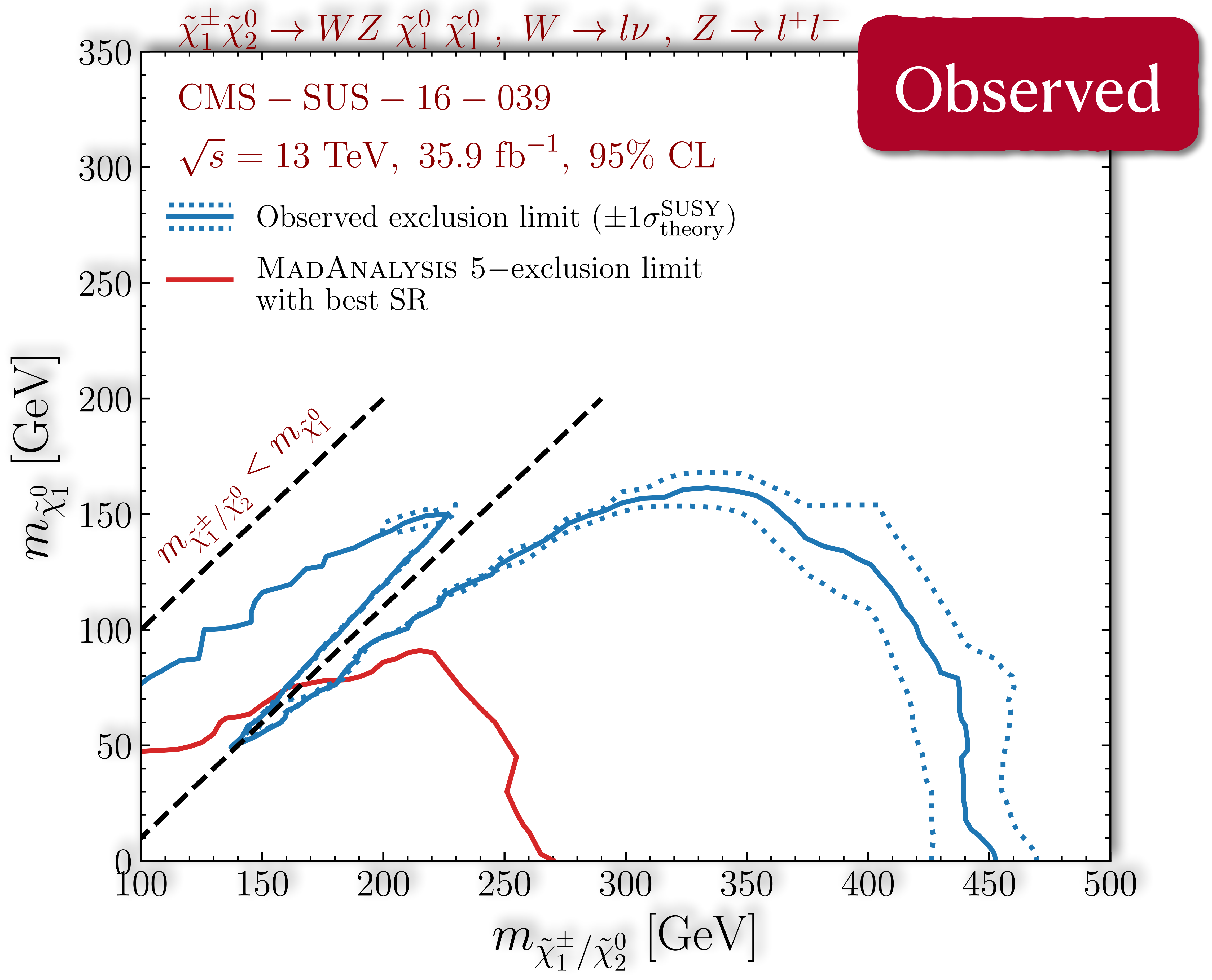
More on

Simplified likelihoods with CMS-SUS-16-039

Alguero, JYA, Fuks, Kraml; 2206.14870

Electroweakino searches





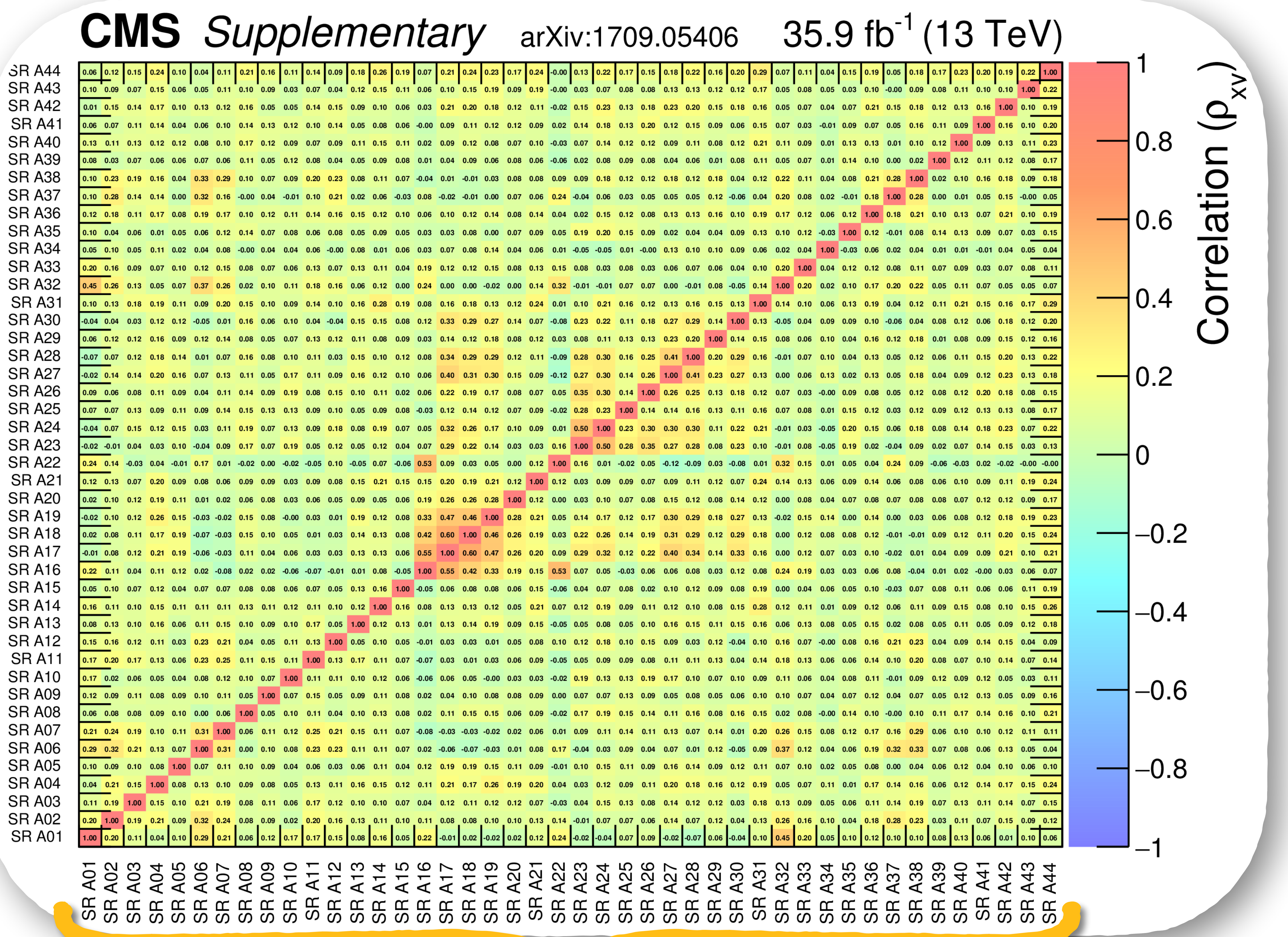
- Wino-like electroweakinos
- Choosing the most sensitive signal region with respect to expected exclusion cross section.

More on

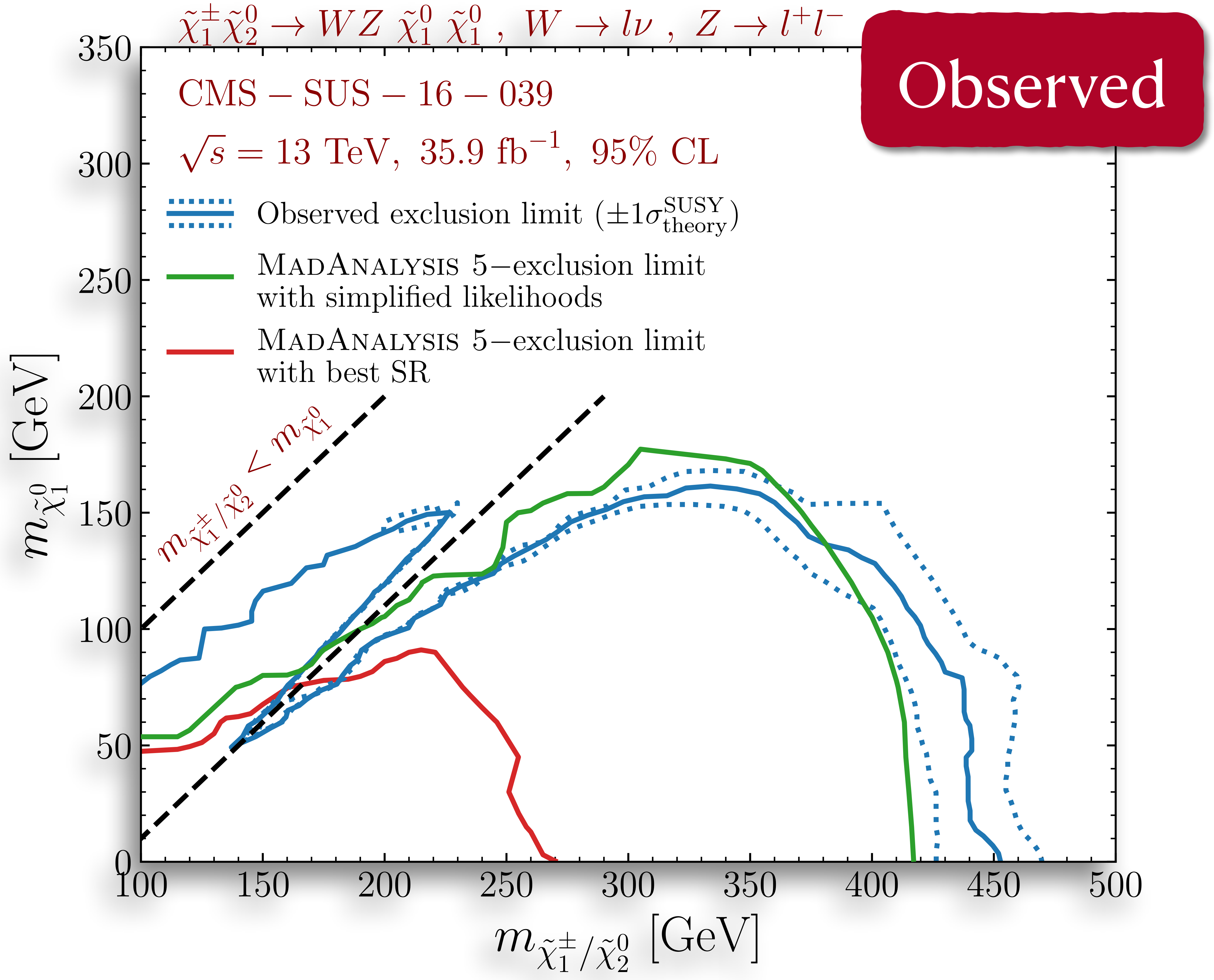
Simplified likelihoods with CMS-SUS-16-039

Alguero, JYA, Fuks, Kraml; 2206.14870

Electroweakino searches



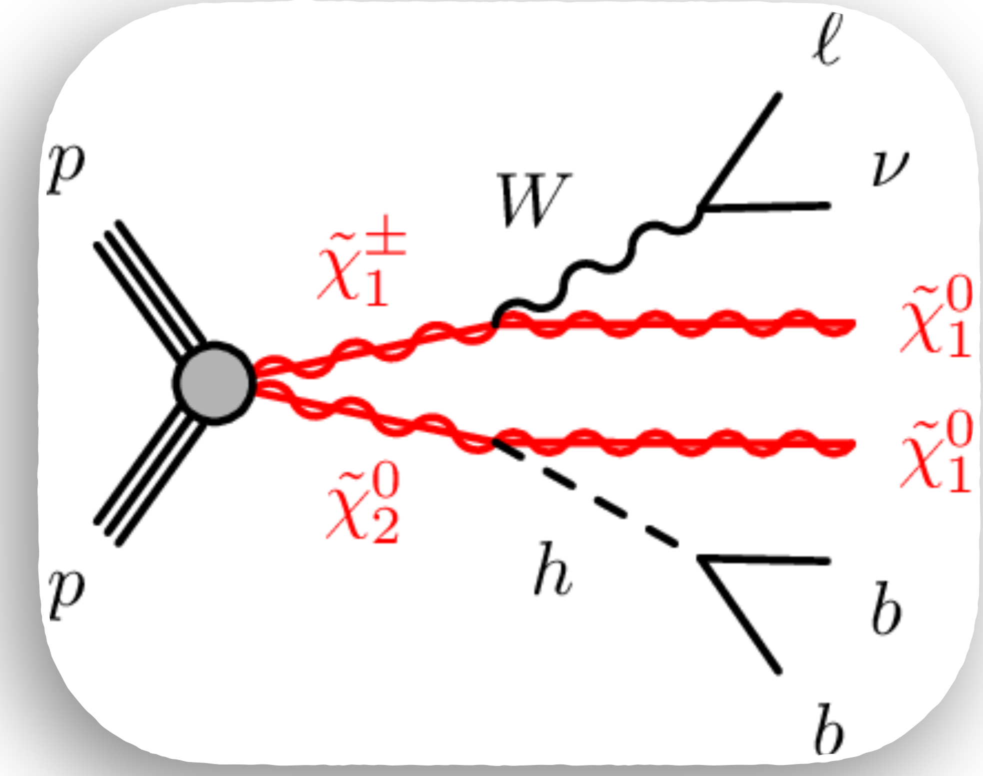
44 Signal Regions!



More on

Full likelihoods with ATLAS-SUSY-19-08

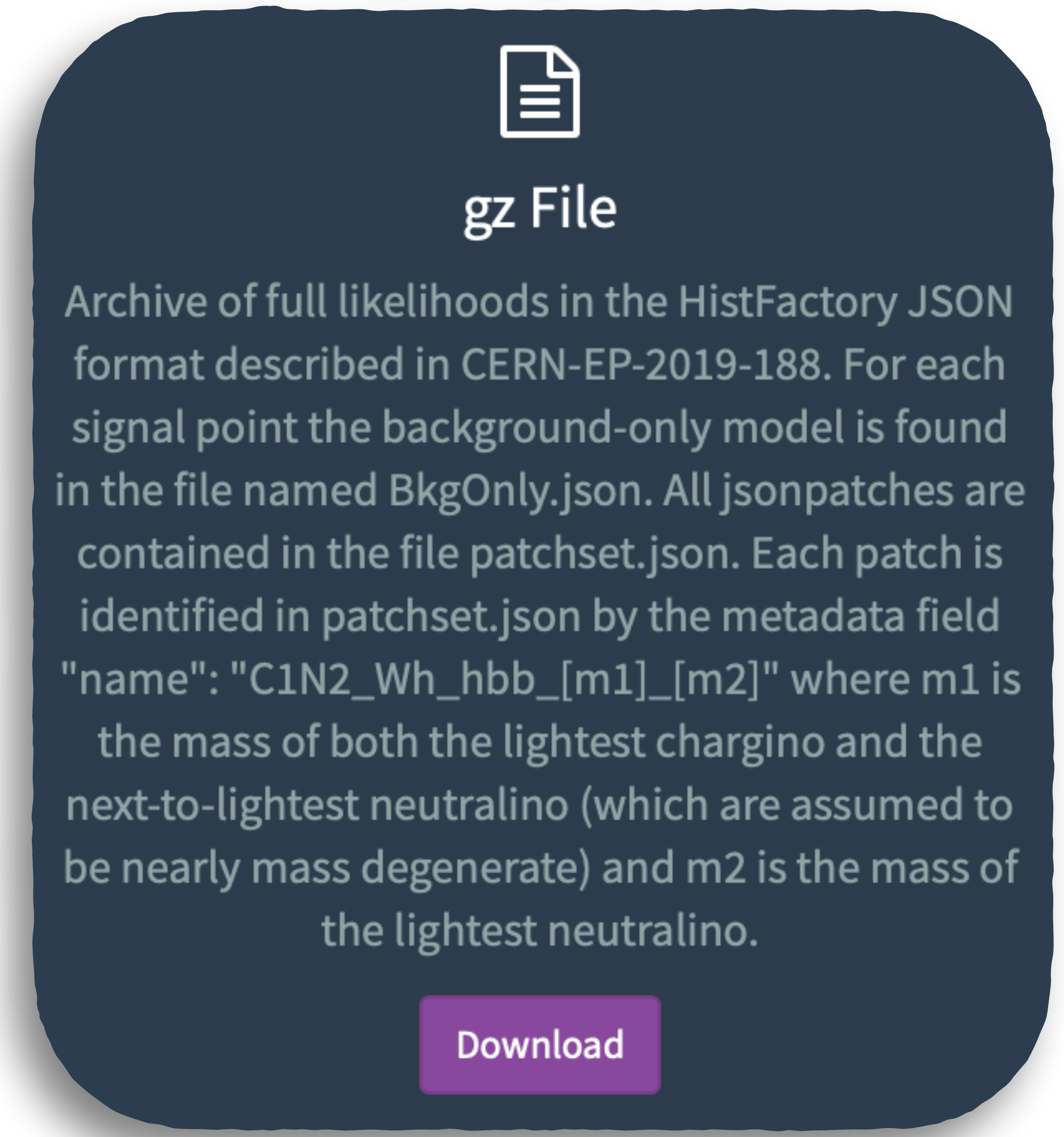
Alguero, JYA, Fuks, Kraml; 2206.14870



Electroweakino searches

- Wino-like electroweakinos
- ATLAS shares HistFactory like json files to form full profile likelihoods.
- Each file includes detailed information on backgrounds and corresponding nuisance parameters.
- Simplified likelihoods are achieved by compressing all the information into single nuisance parameter.

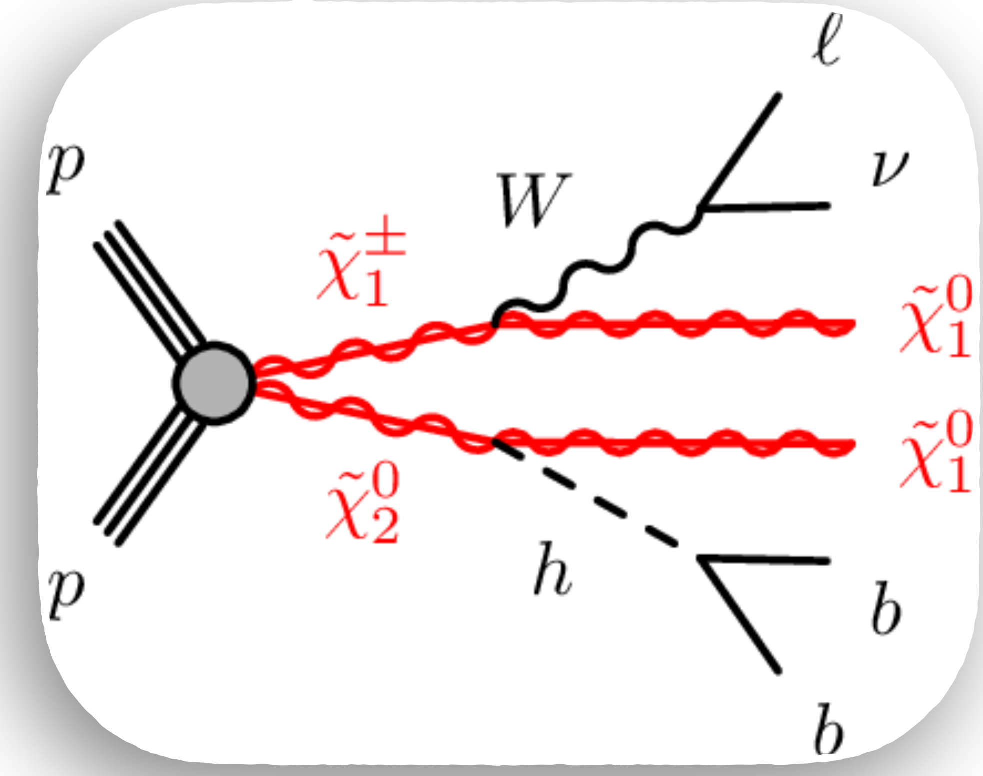
HEPData



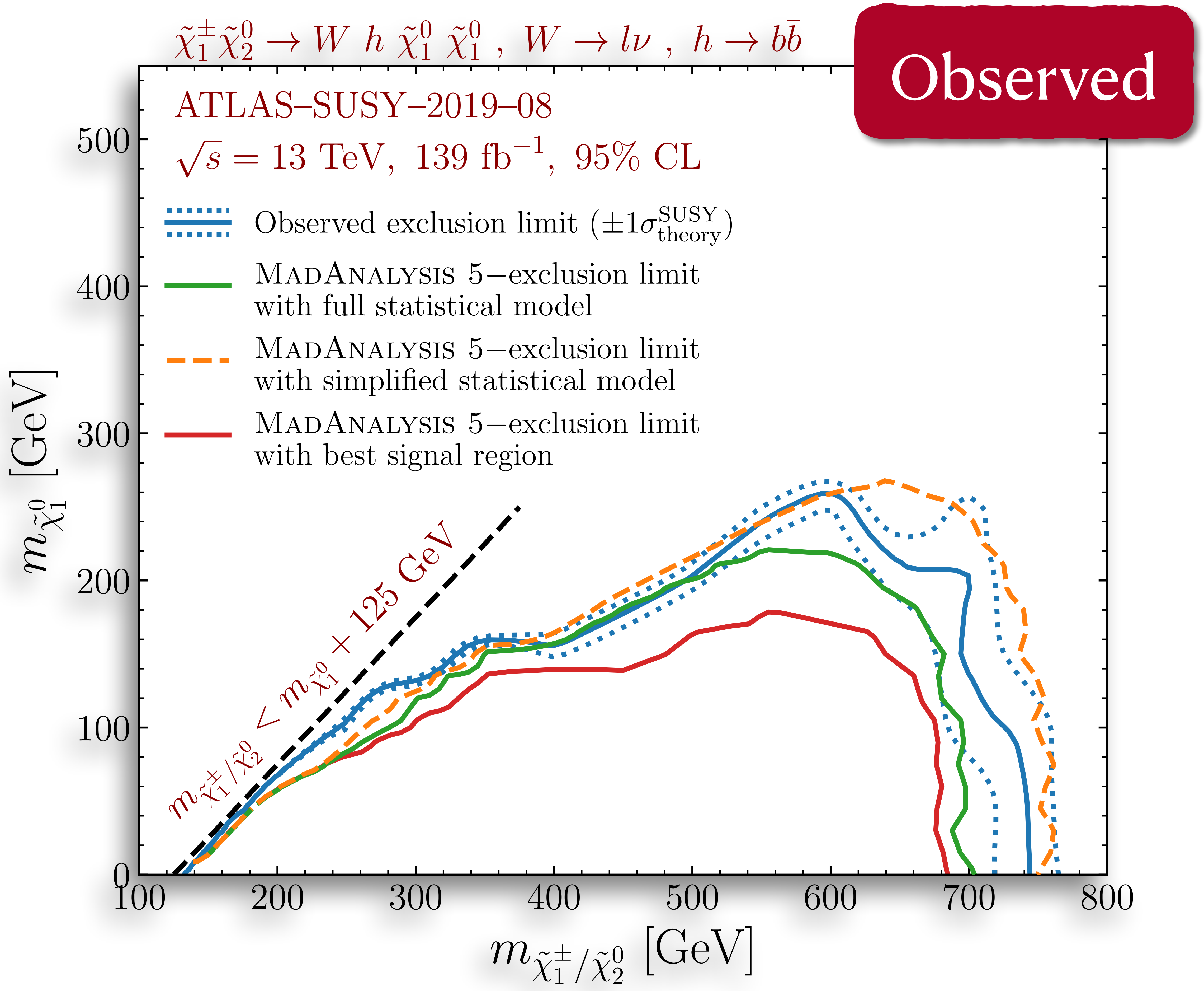
More on

Simplified & full likelihoods in action!

Alguero, JYA, Fuks, Kraml; 2206.14870



Electroweakino searches



- Wino-like electroweakinos
- ATLAS shares HistFactory like json files to form full profile likelihoods.
- Each file includes detailed information on backgrounds and corresponding nuisance parameters.
- Simplified likelihoods are achieved by compressing all the information into single nuisance parameter.

Simplify module has been used for orange curve: ATL-PHYS-PUB-2021-038

More on

Simplified & full likelihoods in action!

Alguero, JYA, Fuks, Kraml; 2206.14870

Simplified likelihoods and pyhf interface has been fully integrated in MadAnalysis 5 v1.10.4

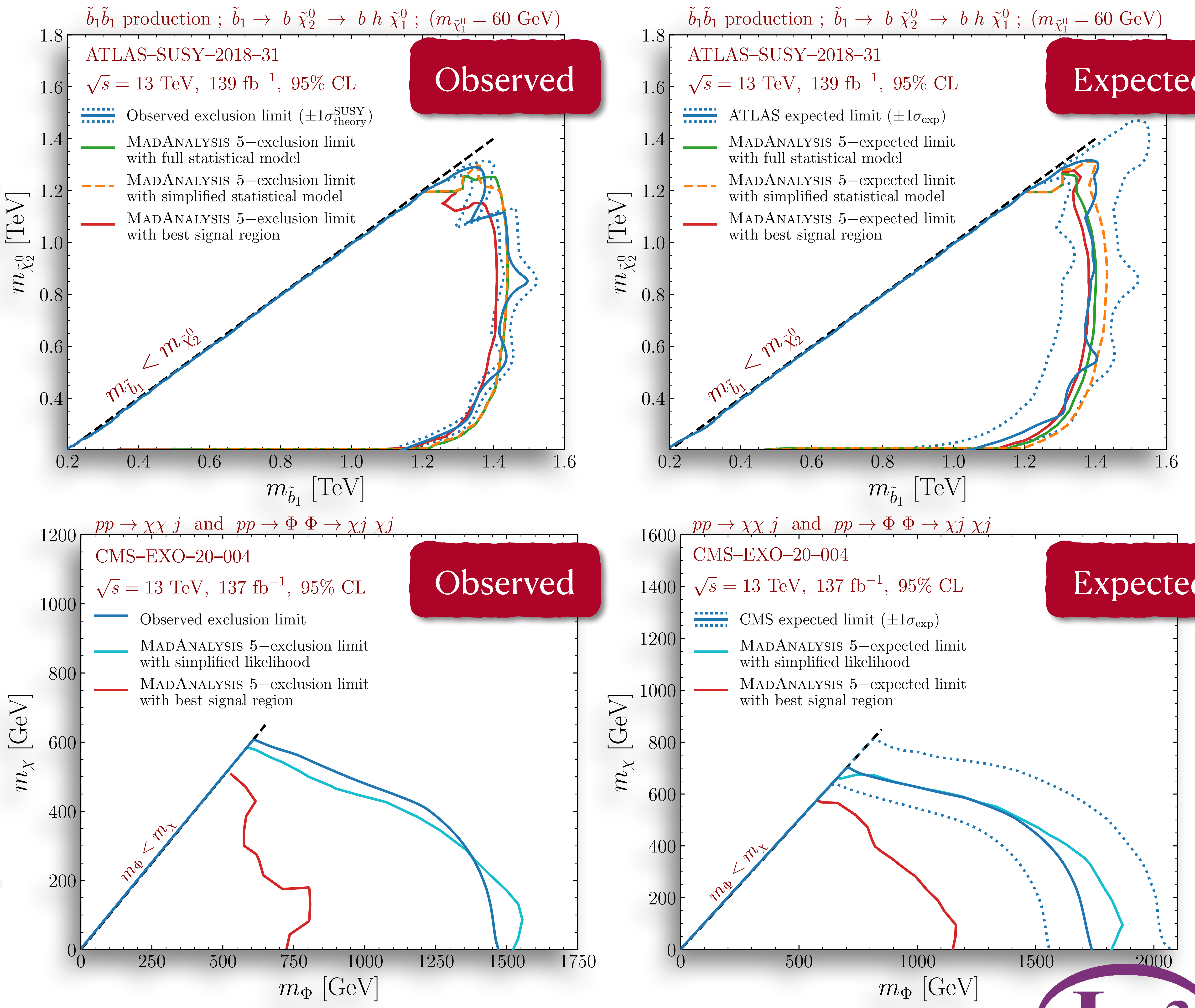
Analyses using full/simplified likelihoods

- ✅ ATLAS - SUSY - 2018 - 31
- ✅ ATLAS - SUSY - 2018 - 04
- ✅ ATLAS - SUSY - 2019 - 08
- ✅ ATLAS - SUSY - 2018 - 06
- ✏️ ATLAS - SUSY - 2018 - 14

- ✅ CMS - SUS - 2016 - 39
- ✅ CMS - SUS - 2017 - 01
- ✅ CMS - SUS - 2019 - 06
- ✅ CMS - EXO - 20 - 004

NEW

By Andreas Albert; see his talk at RAMP#3



Durham University

Well, I can get those curves from original papers!

Alguero, JYA, Fuks, Kraml; 2206.14870

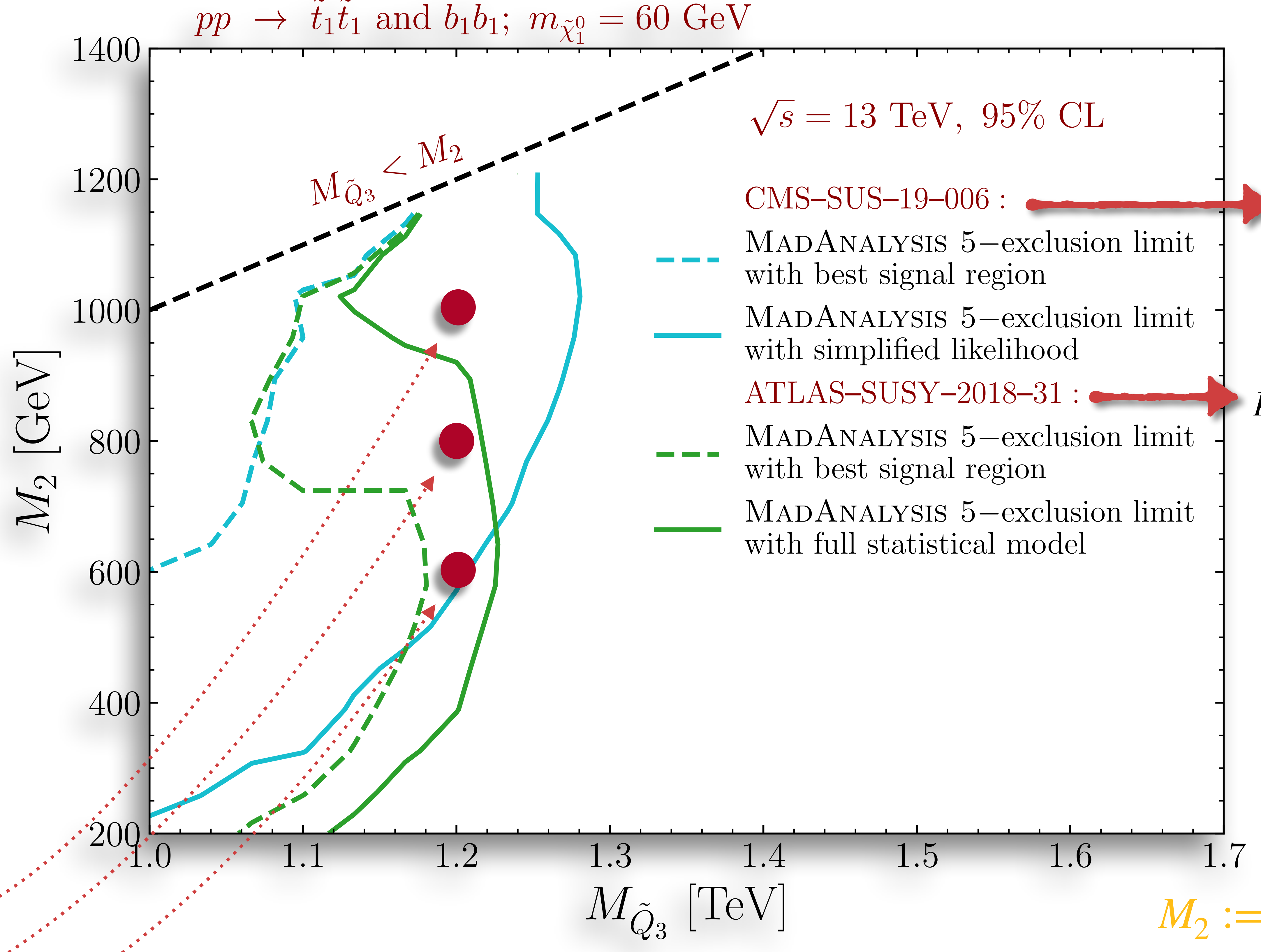
Note that experimental analyses are generally using simplified models.

- The goal of reinterpretation is not just reproducing the experimental results but testing the new theories we come up with in the shower!

$1 - \mathrm{CL}_s$ (obs)

analysis	method	$M_2 = 600$ GeV $\tilde{b}_1\tilde{b}_1^*$	$\tilde{t}_1\tilde{t}_1^*$	total	$M_2 = 800$ GeV $\tilde{b}_1\tilde{b}_1^*$	$\tilde{t}_1\tilde{t}_1^*$	total	$M_2 = 1$ TeV $\tilde{b}_1\tilde{b}_1^*$	$\tilde{t}_1\tilde{t}_1^*$	total
ATLAS	best-SR	0.71	0.66	0.94	0.70	0.59	0.91	0.29	0.21	0.57
	combined	0.83	0.80	**0.98**	0.84	0.74	**0.97**	0.80	0.56	**0.92**
CMS	best-SR	0.31	0.37	0.62	0.38	0.45	0.73	0.29	0.38	0.70
	combined	0.79	0.71	**0.96**	0.89	0.83	**0.99**	0.93	0.82	**0.99**

Realistic MSSM simulation with proper RGE evolution



Original analysis:

$pp \to \tilde{g}\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$

$pp \to \tilde{b}_1\bar{\tilde{b}}_1, \tilde{b}_1 \to bh\tilde{\chi}_2^0$

M_2 := Wino mass

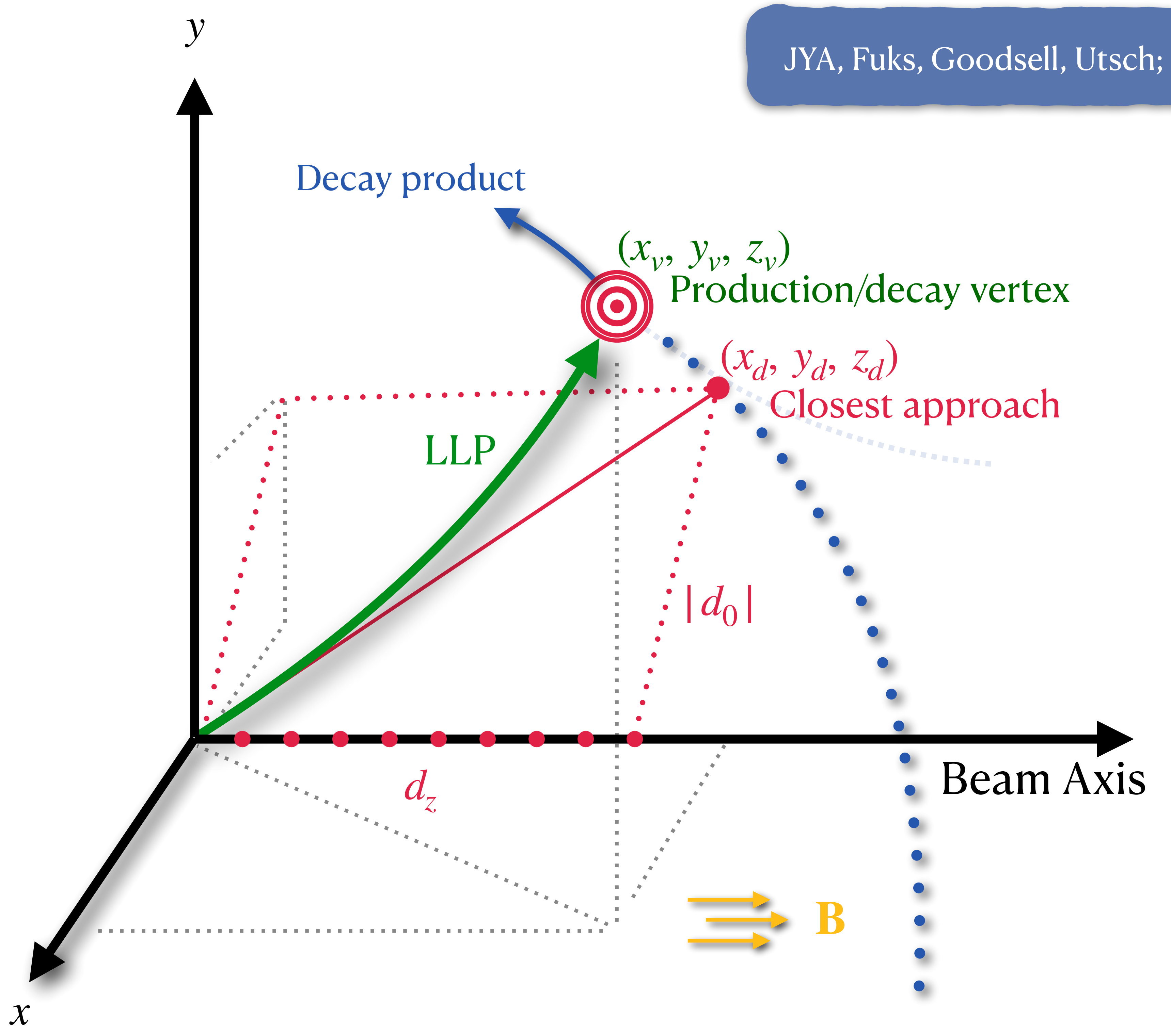
$M_{\tilde{Q}_3}$:= left squark soft-mass for the third generation

Particle propagation in SFS

JYA, Fuks, Polykratis EPJC '21

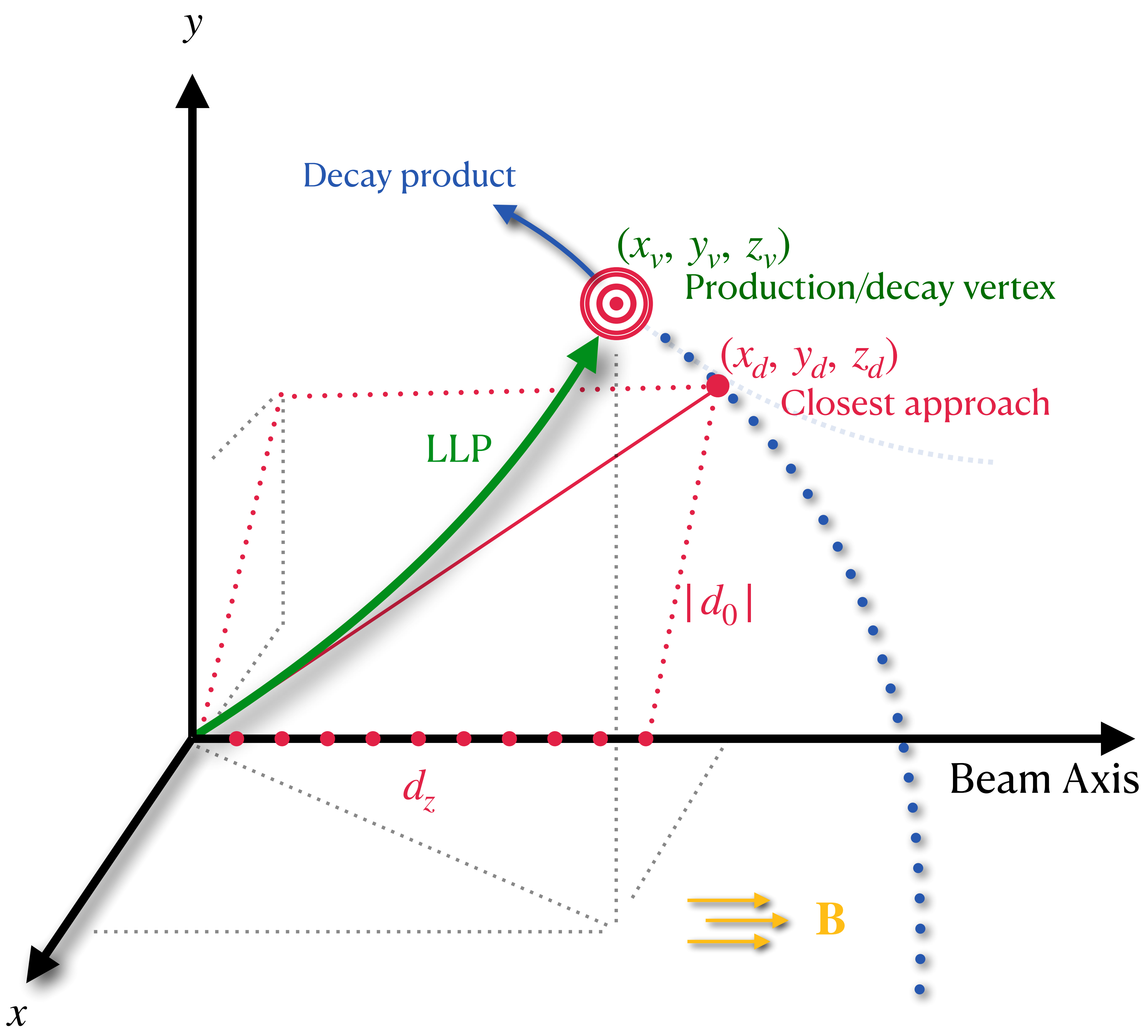
- SFS module allows for simple observable smearing based on transfer functions.
- Transverse impact parameter (d_0) and longitudinal impact parameter (d_z) can be calculated with straight trajectory assumption (default behaviour for other recasting softwares).
- Modification of particle trajectories under constant magnetic field can provide relevant effects for unusual particle signatures.

JYA, Fuks, Goodsell, Utsch; EPJC '22

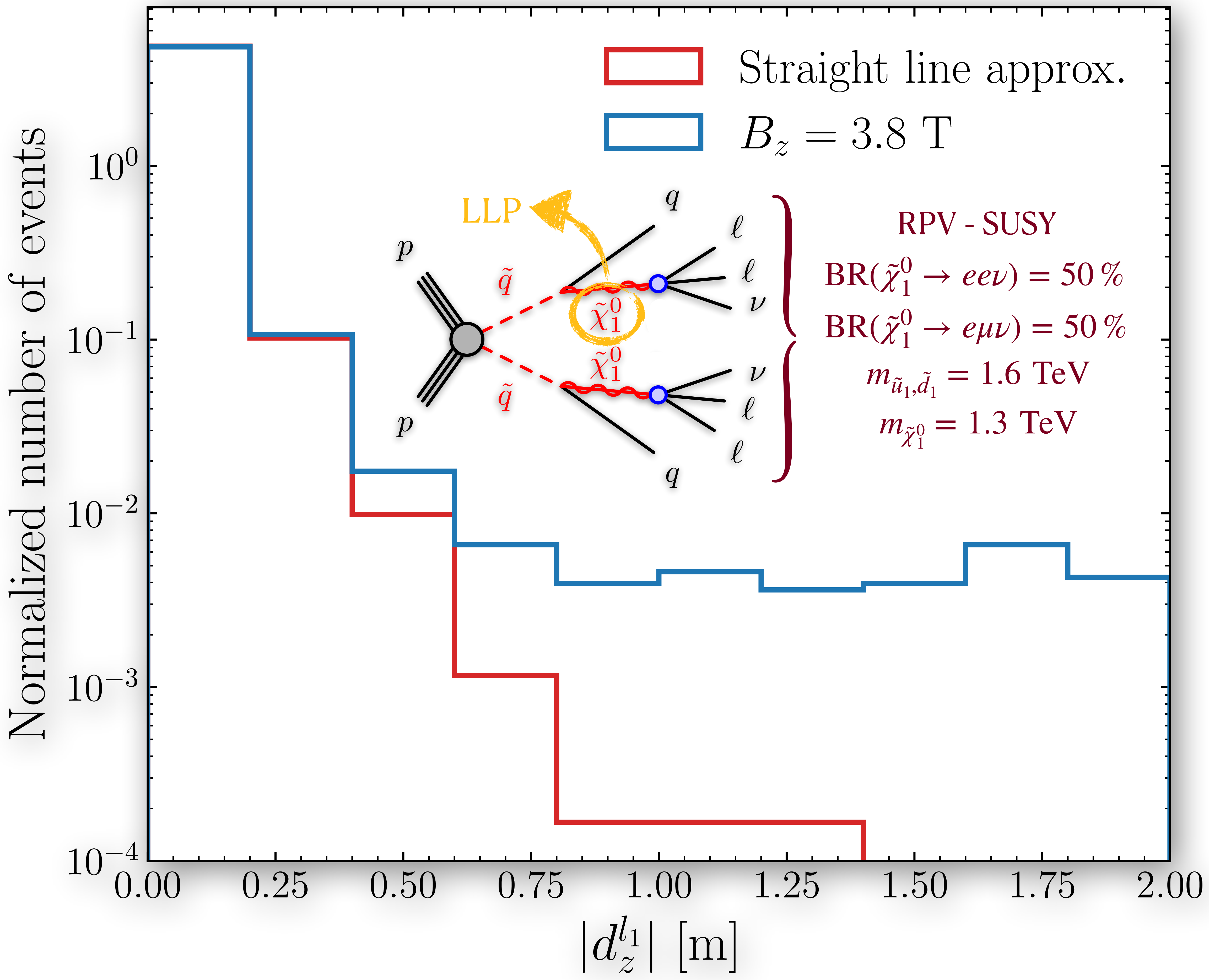


Particle propagation in SFS

JYA, Fuks, Goodsell, Utsch; EPJC '22

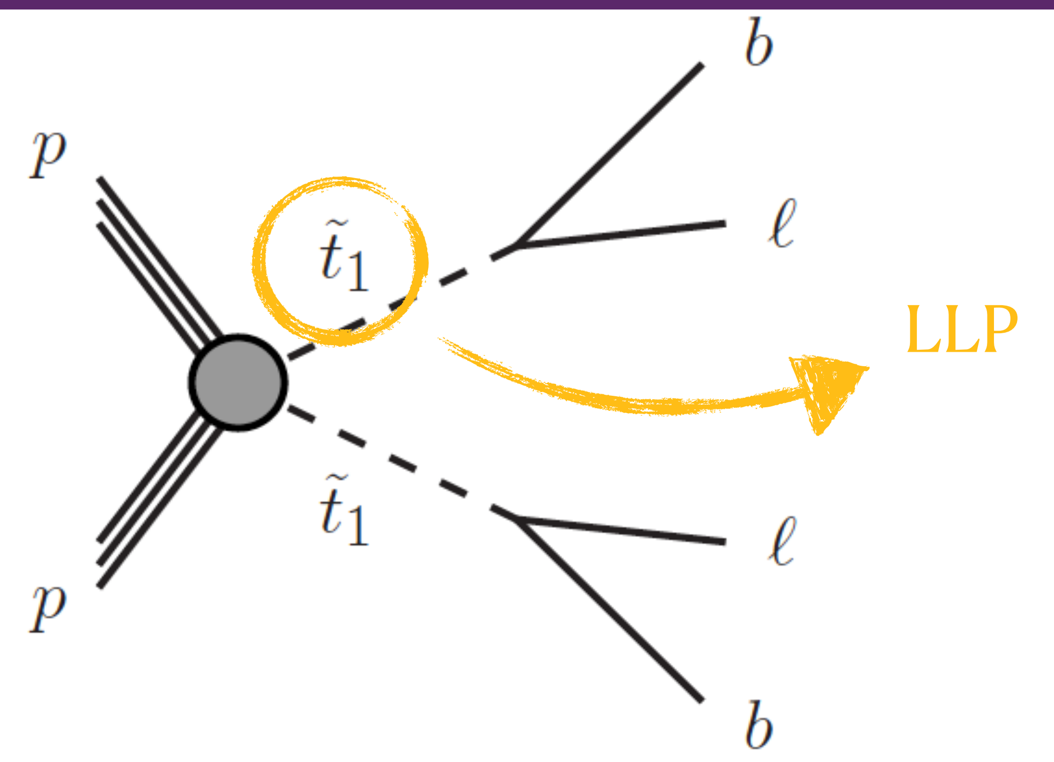



```
ma5> set main.fastsim.particle_propagator = on
ma5> set main.fastsim.magneticfield = 3.8
```

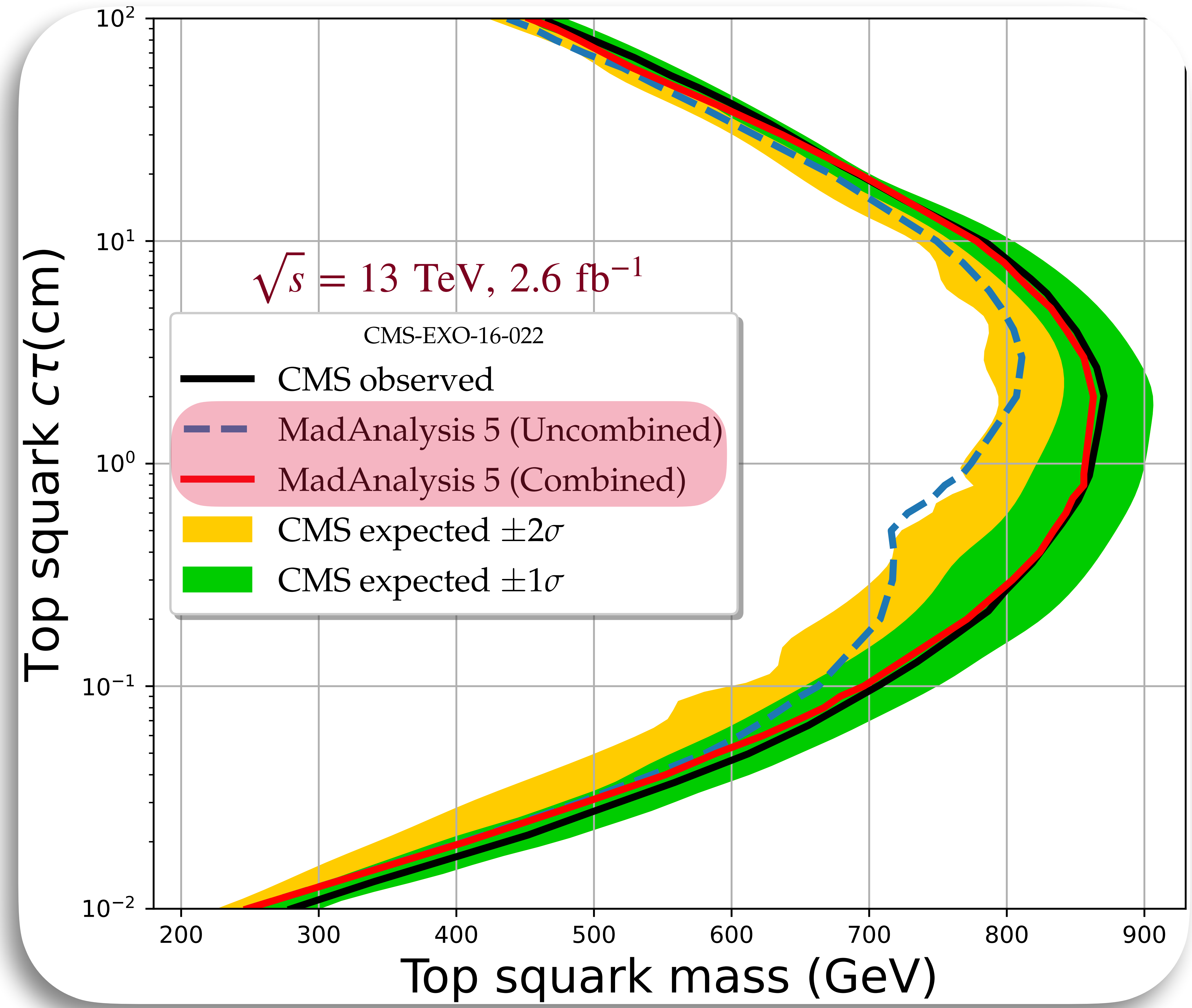


CMS-EXO-16-022: displaced leptons

Recasted by Manuel Utsch

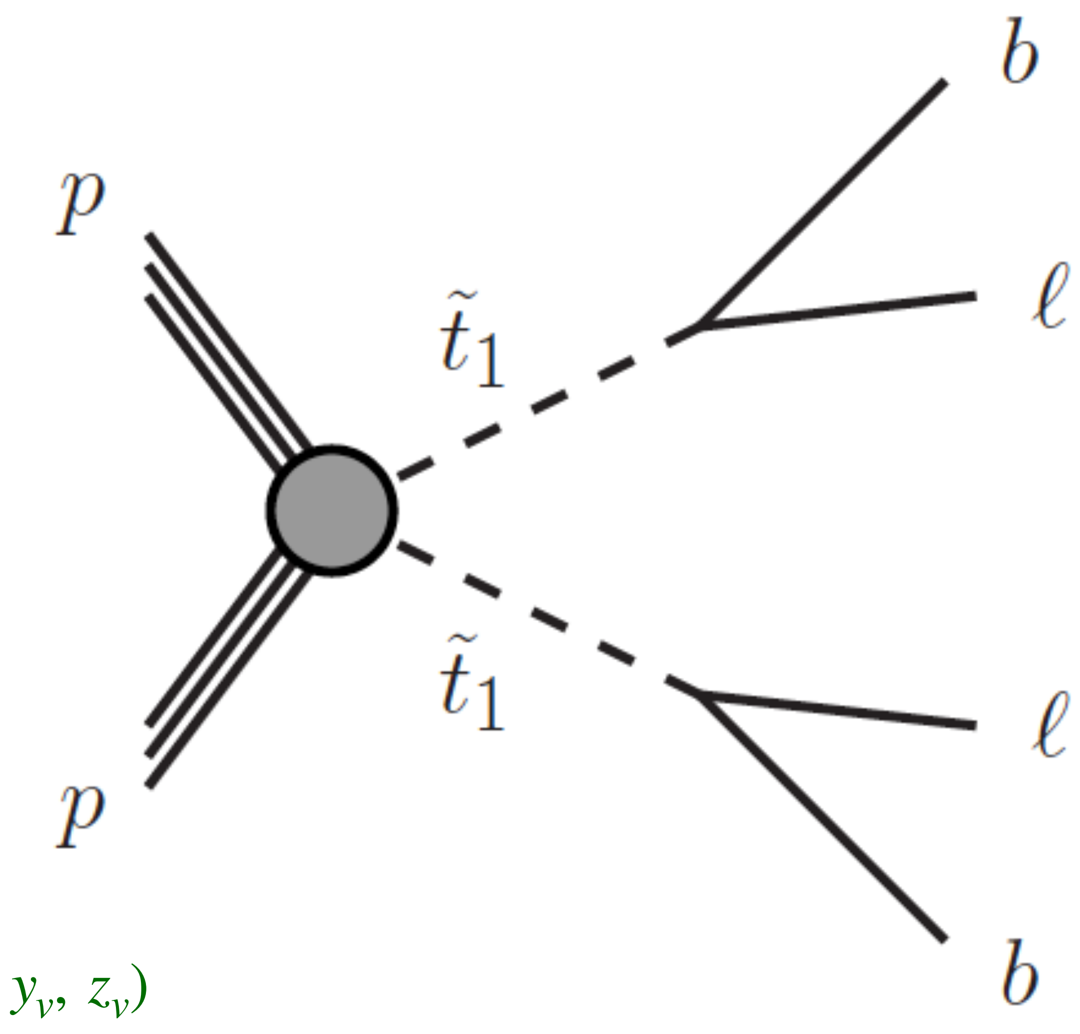


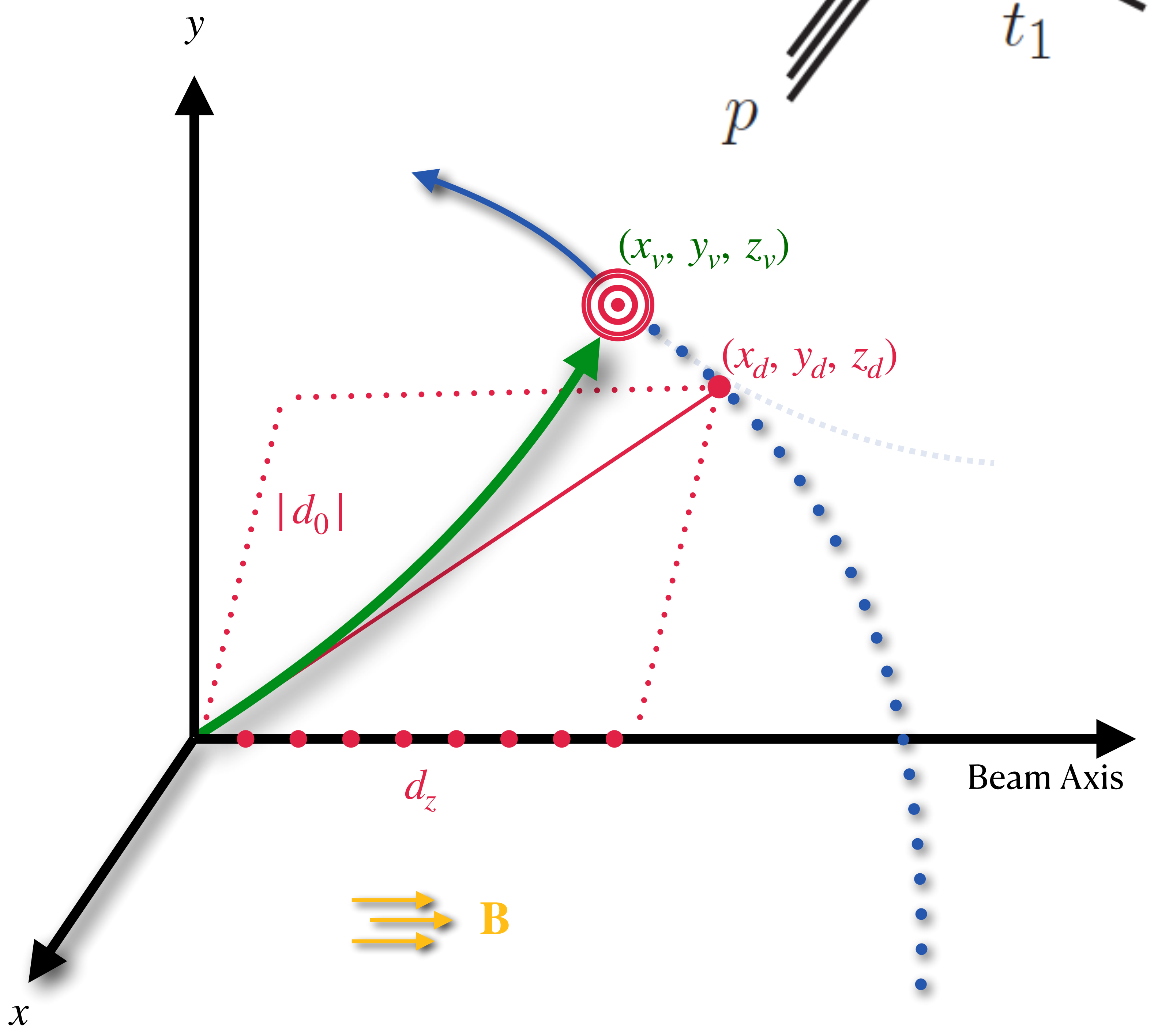
- Step I) Validation of the SFS module with particle propagator
- Existing recast from 2018 adapted to the SFS with particle propagation module.
- Improved track based isolation cones.
- Very scarce validation material.
- No available statistics!



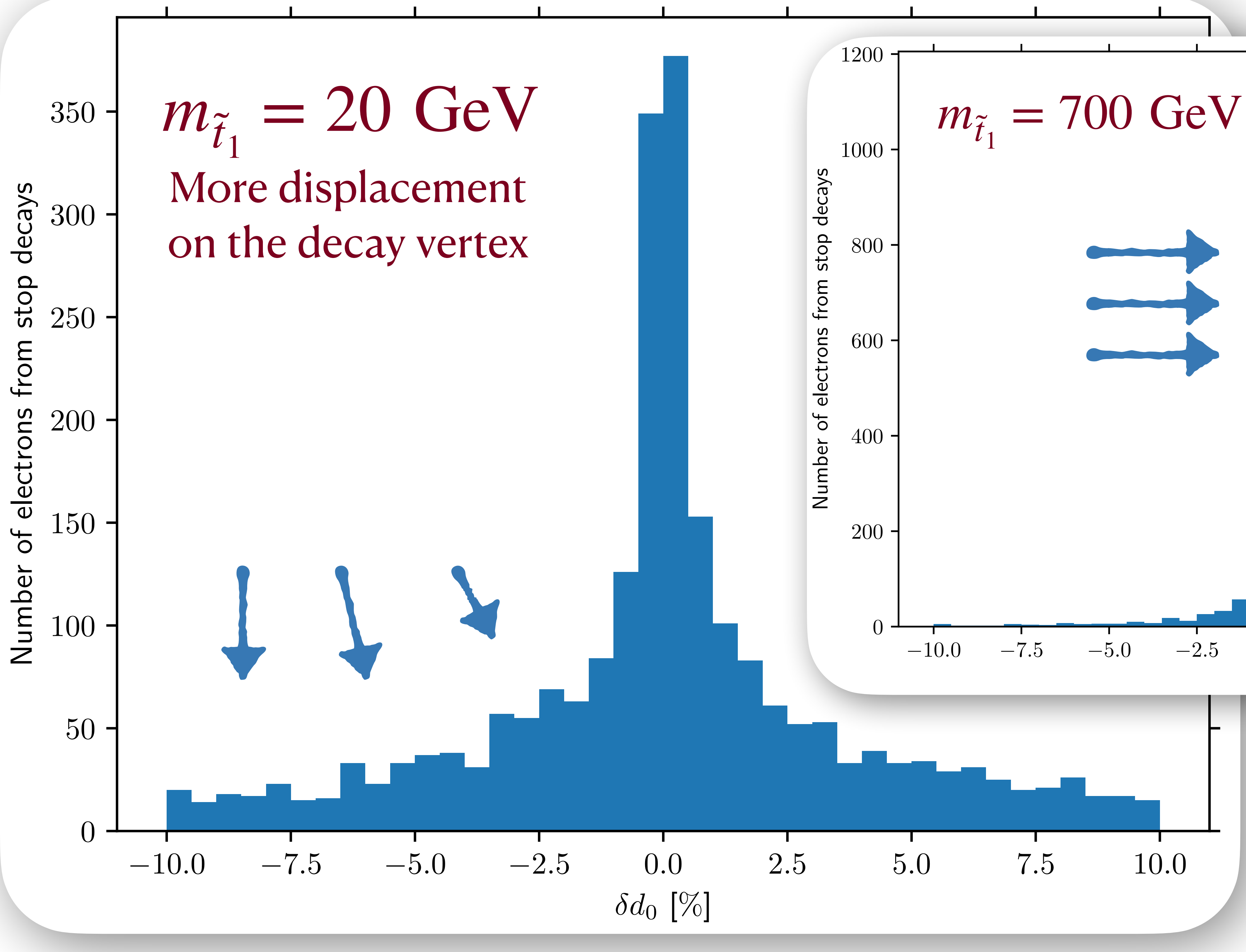
CMS-EXO-16-022: Impact of the particle propagator

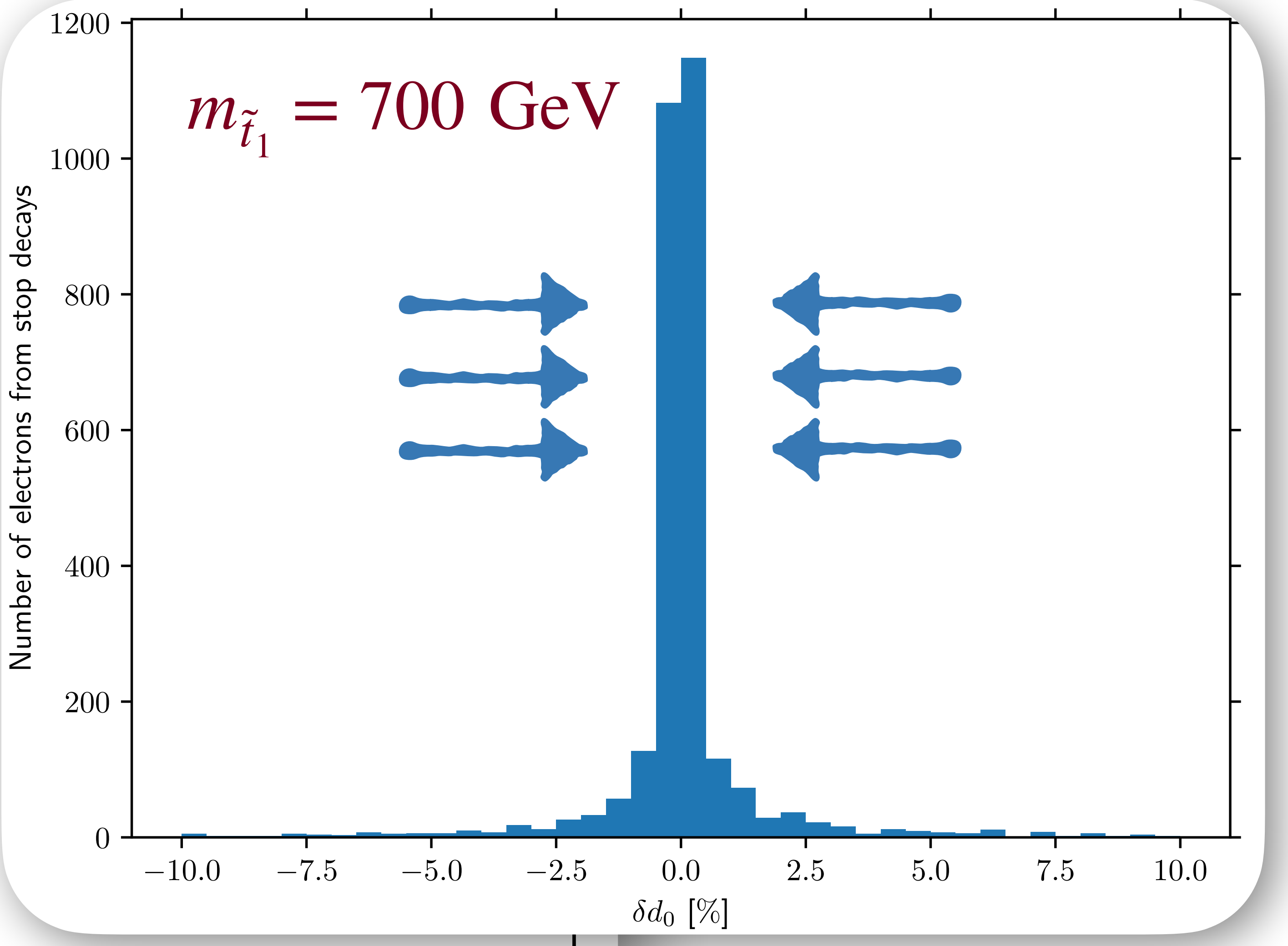
JYA, Fuks, Goodsell, Utsch; EPJC '22





$$\delta d_0 = \frac{d_0 - d_0^{\rm approx}}{d_0^{\rm approx}}$$



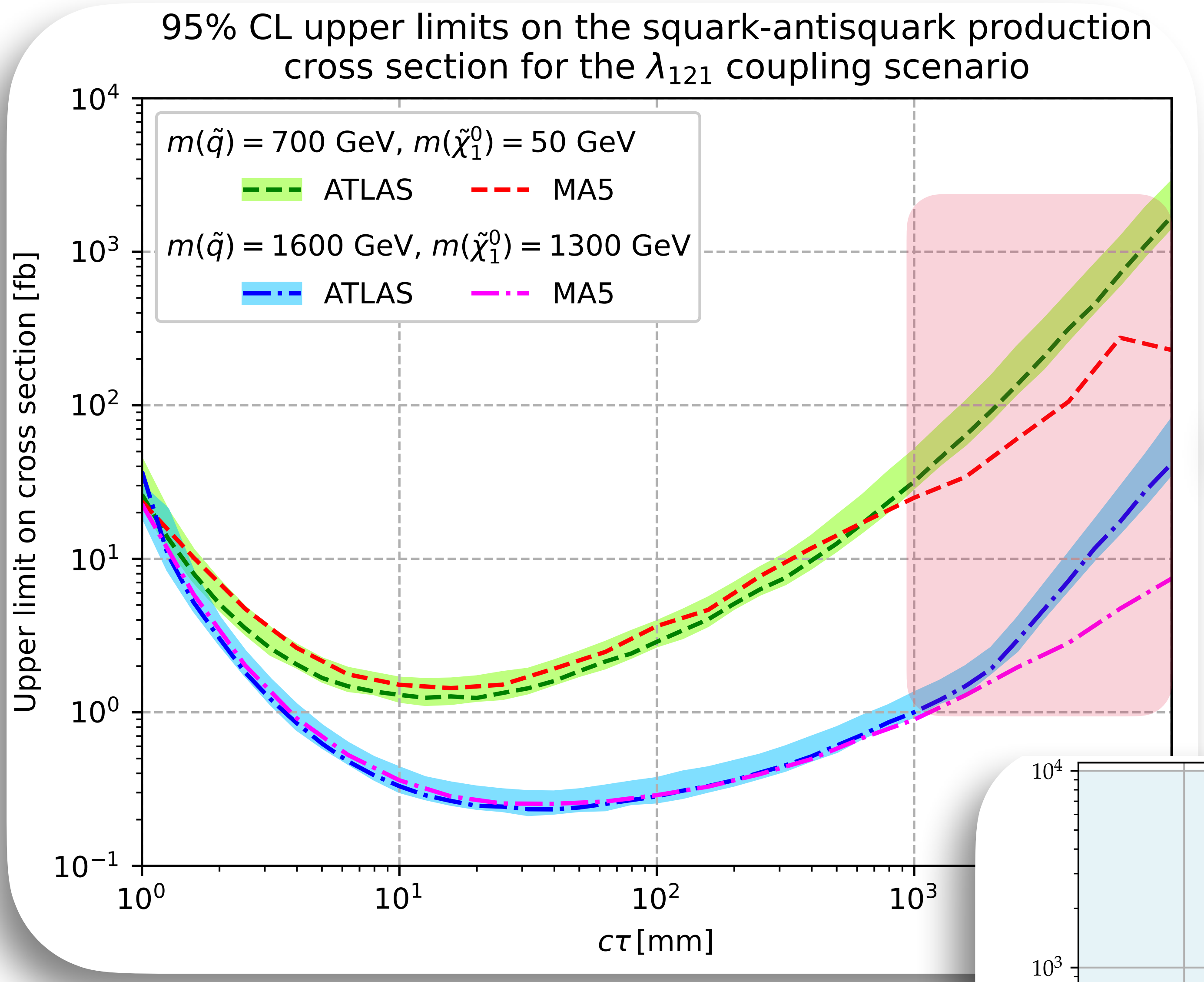


Simplified & full likelihoods in action!

JYA, Fuks, Goodsell, Utsch; EPJC '22

Particle propagator module and fresh LLP recasts are available in MadAnalysis 5 - v1.9_beta

ATLAS - SUSY - 2017 - 04



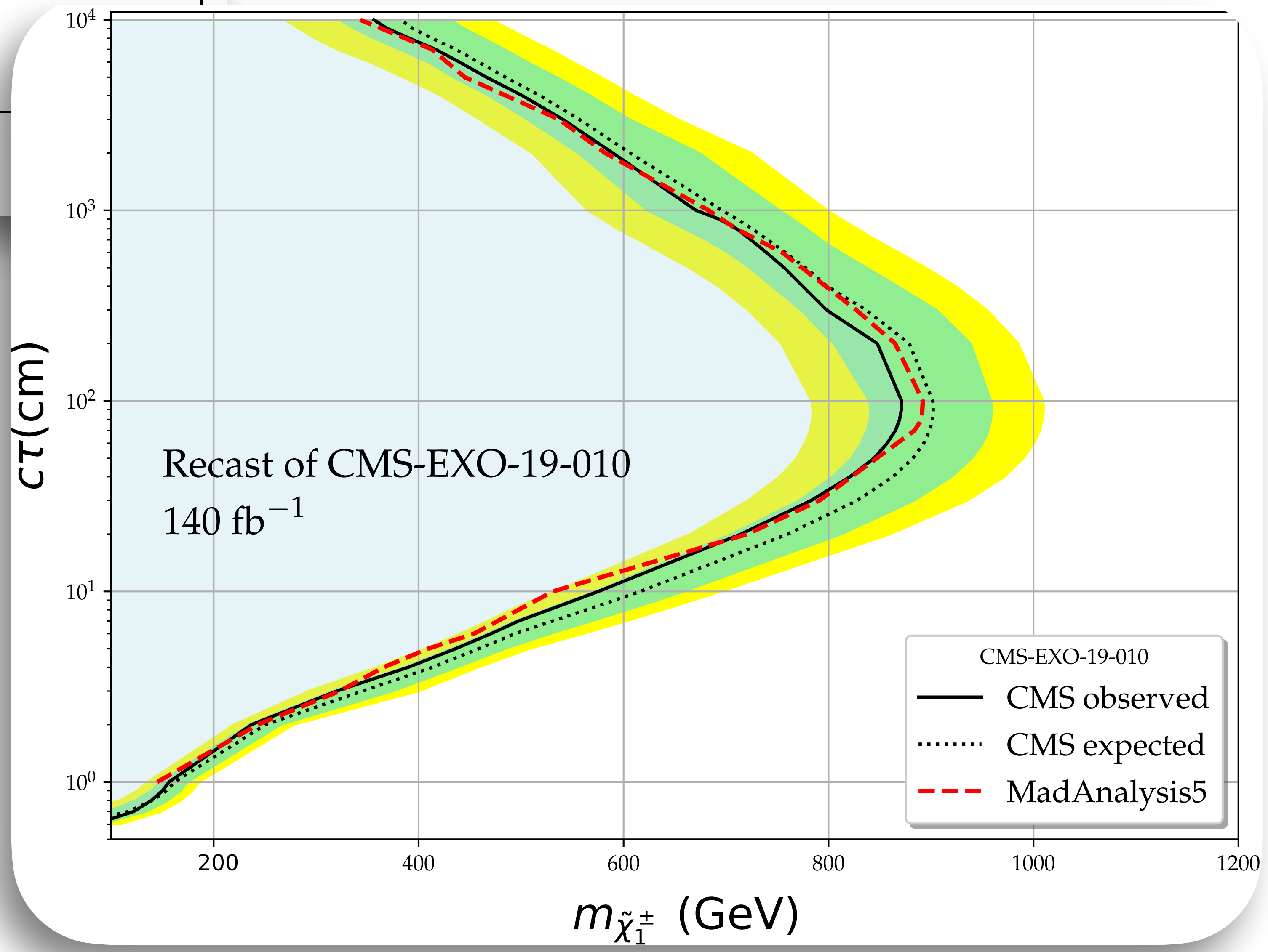
$c\tau > 1$ m

CMS - EXO - 19 - 010

CMS - EXO - 16 - 022

CMS - EXO - 18 - 003

For details see the MadAnalysis talk at "Tenth workshop of the LLP Community"



Conclusion

Conclusion

- Particle propagation can have a significant effect on analysis outcome, depending on the theory behind it.
- Uncorrelated signal regions do not represent the statistical model of the analysis well enough. Full or simplified statistical models are essential for better reinterpretation.
- NEW MadAnalysis 5 is fully capable of using correlation matrices and full likelihood profiles to improve exclusion limits.
- NEW The particle propagation module is available with MadAnalysis v1.9 alongside various LLP recasts.

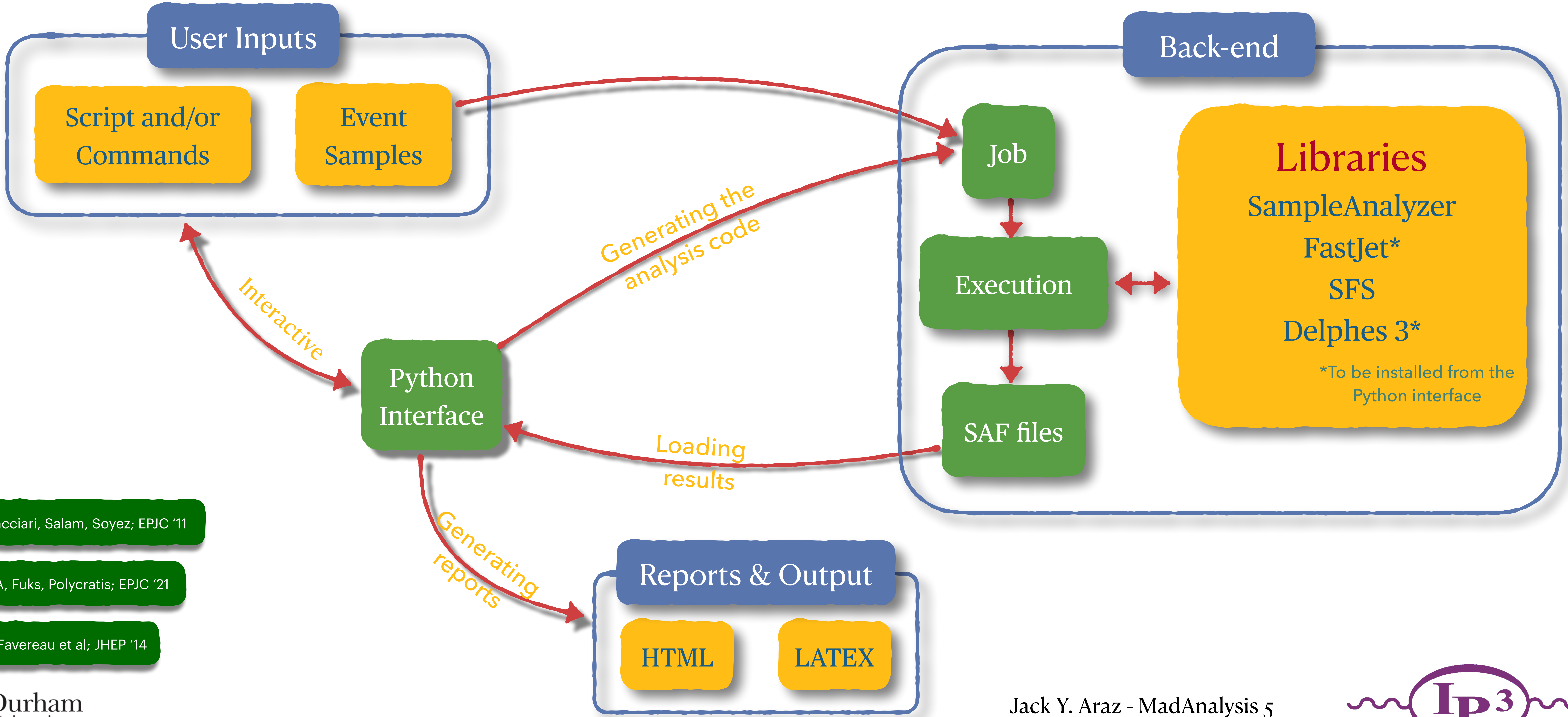
WANTED: Analysis codes

Scientific reproducibility and data preservation solely depend on preserving analysis logic in a reinterpretable form. You can contribute to the HEP community by sharing the LHC recast you have implemented in the MadAnalysis 5 framework through Public Analysis Database! Please send us your analysis code, detector card, info file and validation note to be included in PAD for public use.

More information and examples can be found in the proceedings of the second MadAnalysis 5 Workshop on LHC recasting in Korea. Analysis codes have been published, documented and got a DOI so that they can now be cited.

Hands on!

MadAnalysis 5: Normal Mode

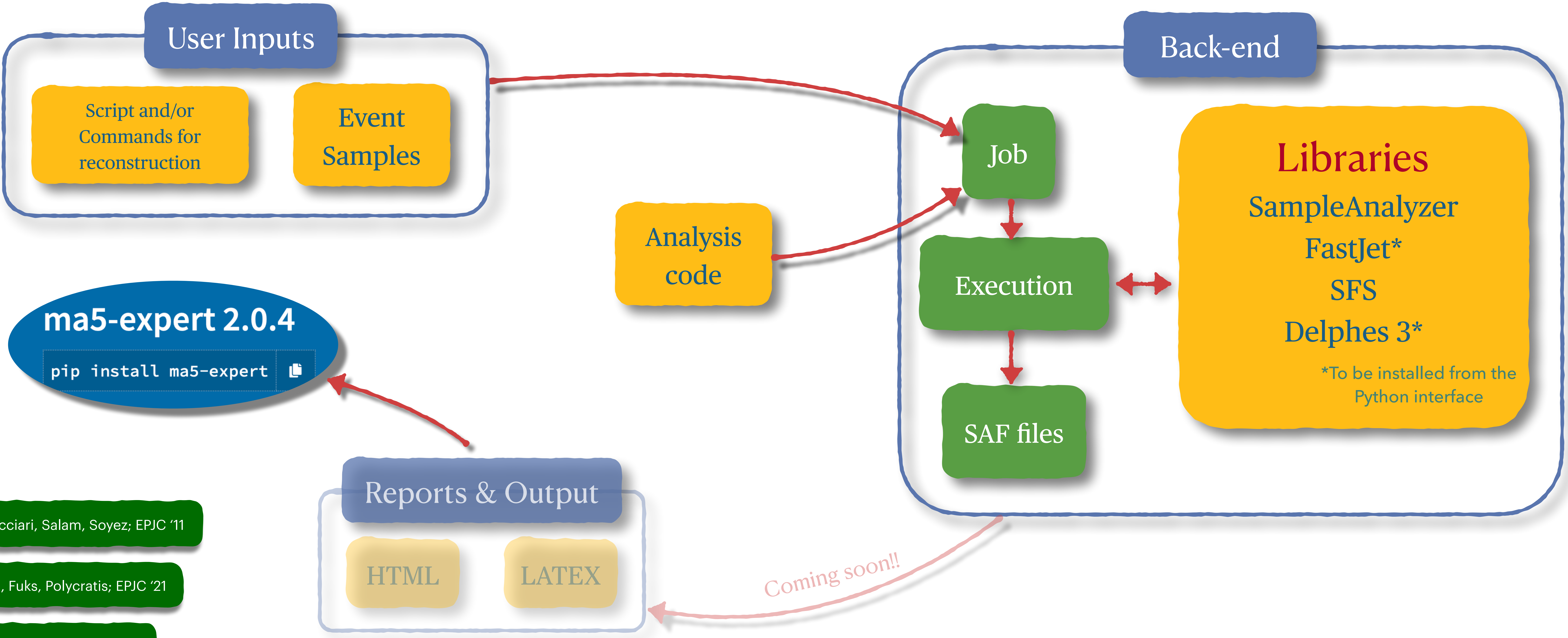


Cacciari, Salam, Soyez; EPJC '11

JYA, Fuks, Polycratis; EPJC '21

De Favereau et al; JHEP '14

MadAnalysis 5: Expert Mode

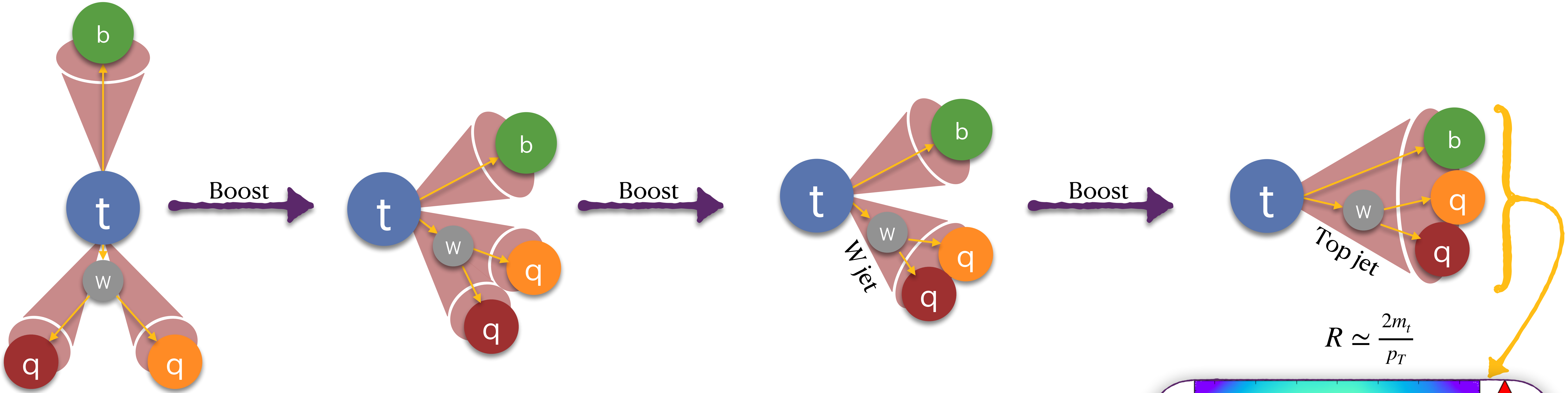


Cacciari, Salam, Soyez; EPJC '11

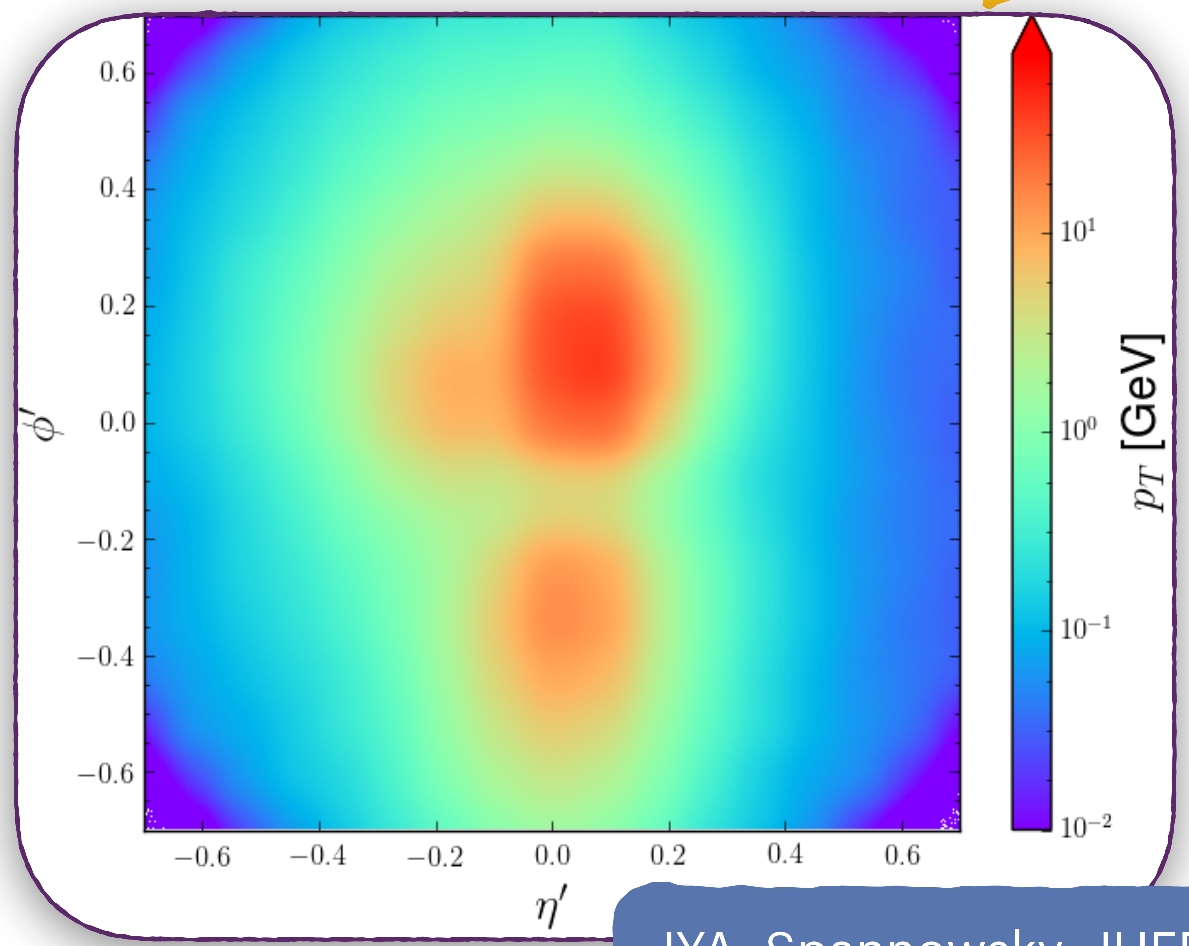
JYA, Fuks, Polycratis; EPJC '21

De Favereau et al; JHEP '14

Boosted top tagging



$$R \simeq \frac{2m_t}{p_T}$$

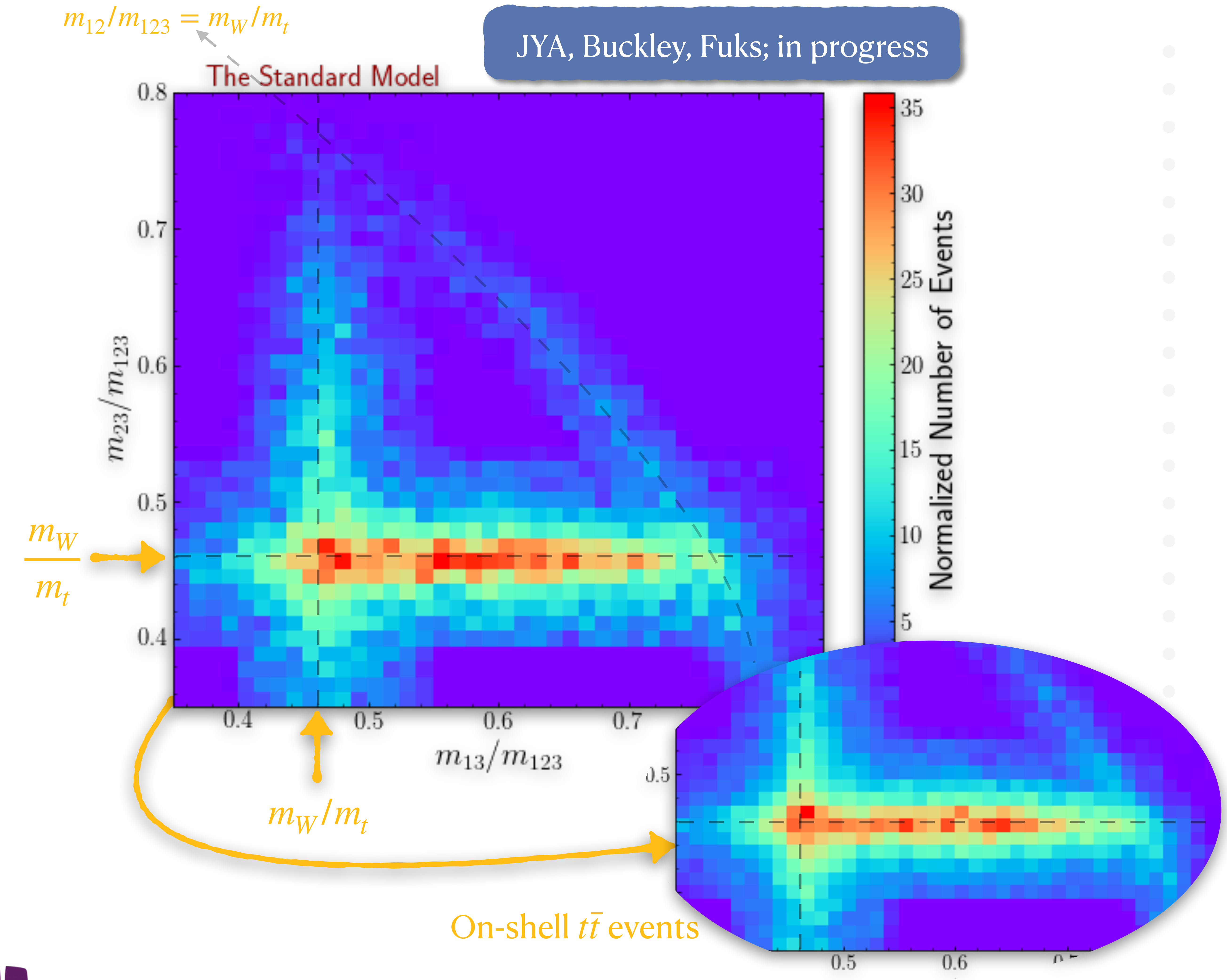


JYA, Spannowsky; JHEP '21

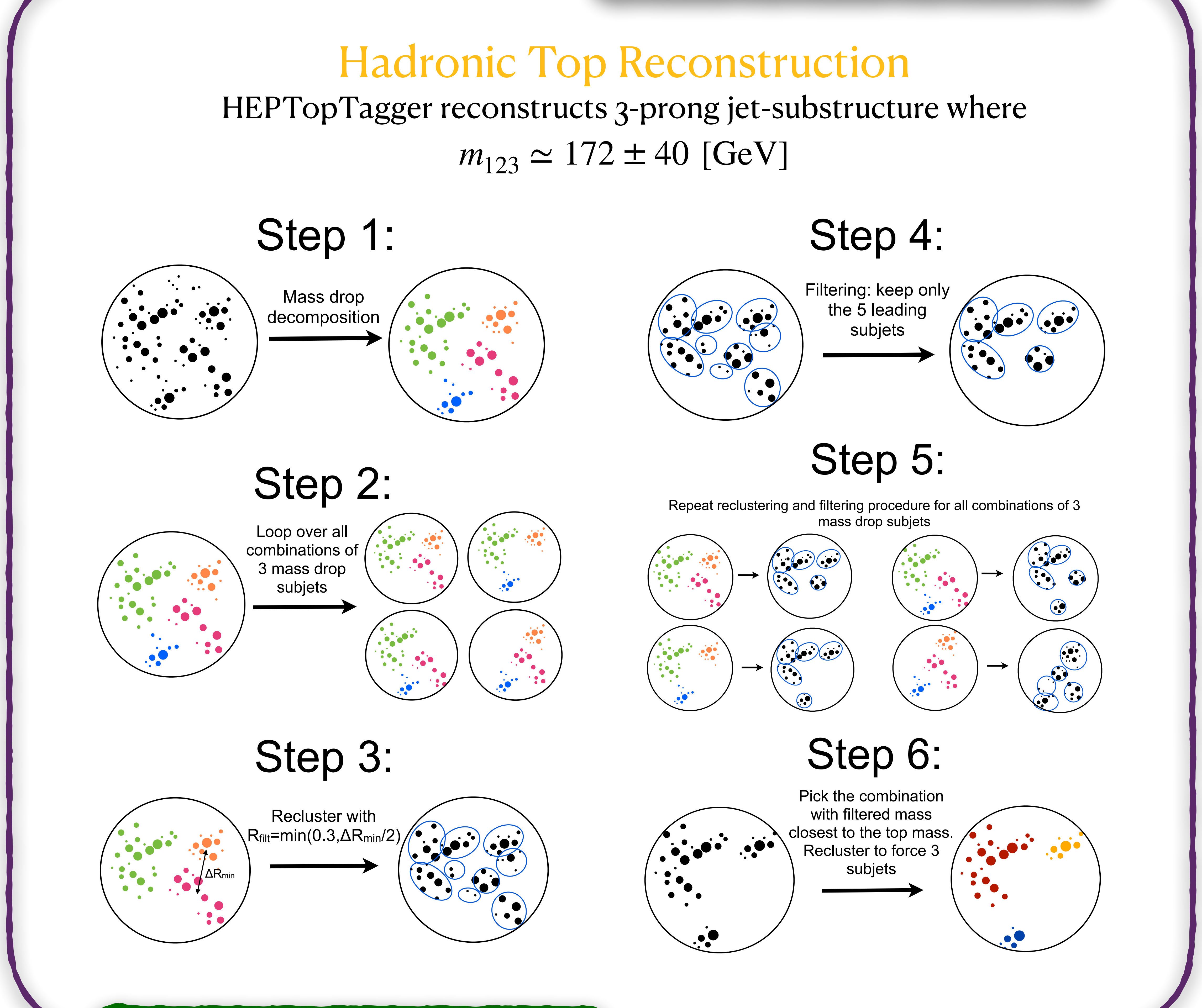
- ❖ With the increased boost factor, jets (top decay products) are getting more collimated.
- ❖ Hadronic top tagging tools: Mass grooming and filtering, Pruning, Trimming, Soft Drop Tagger, Mass Drop Tagger, HEPTopTagger, Machine Learning

HEPTopTagger on $t\bar{t}$ events

JYA, Buckley, Fuks; in progress



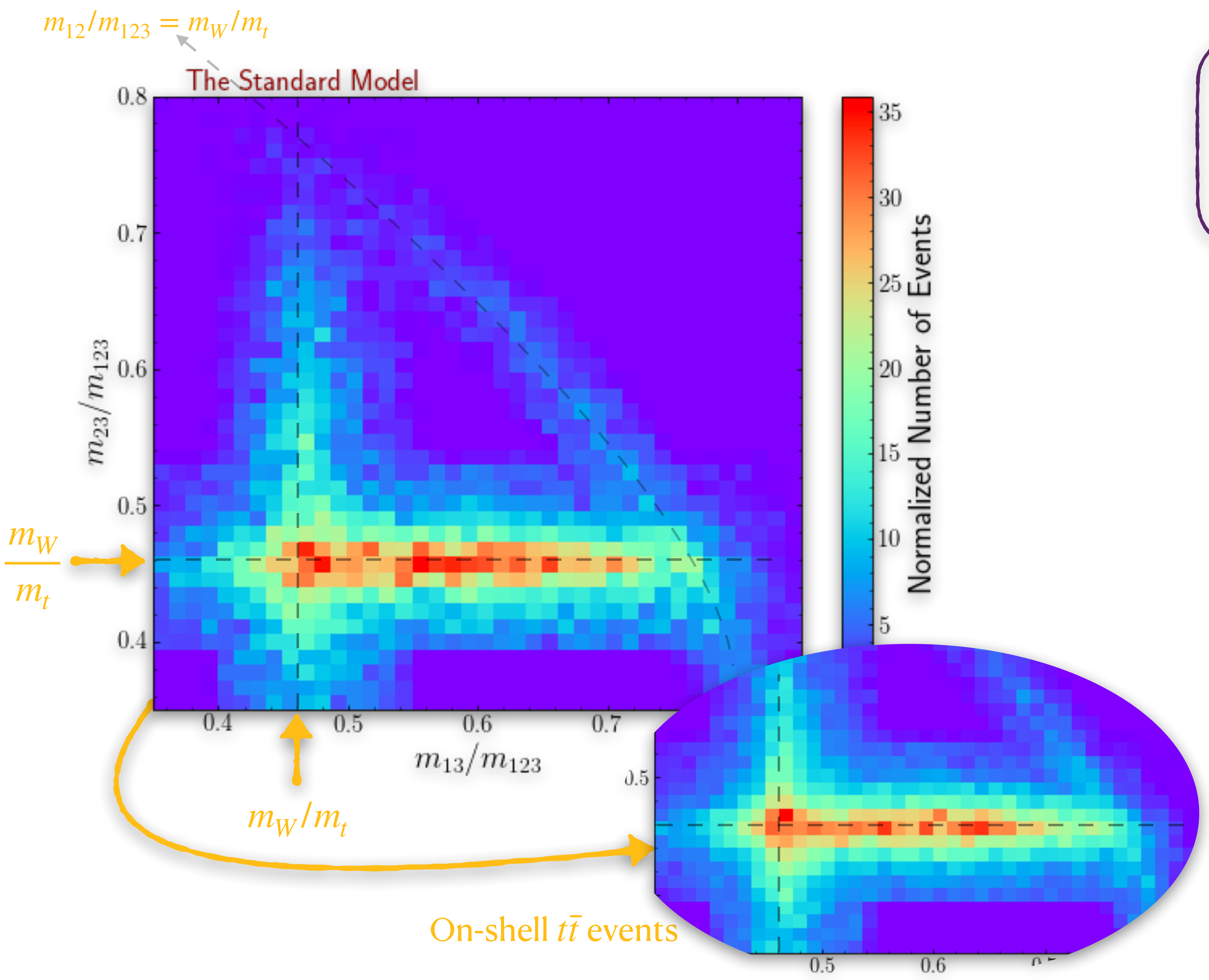
Plehn, Spannowsky, Takeuchi; PRD '11



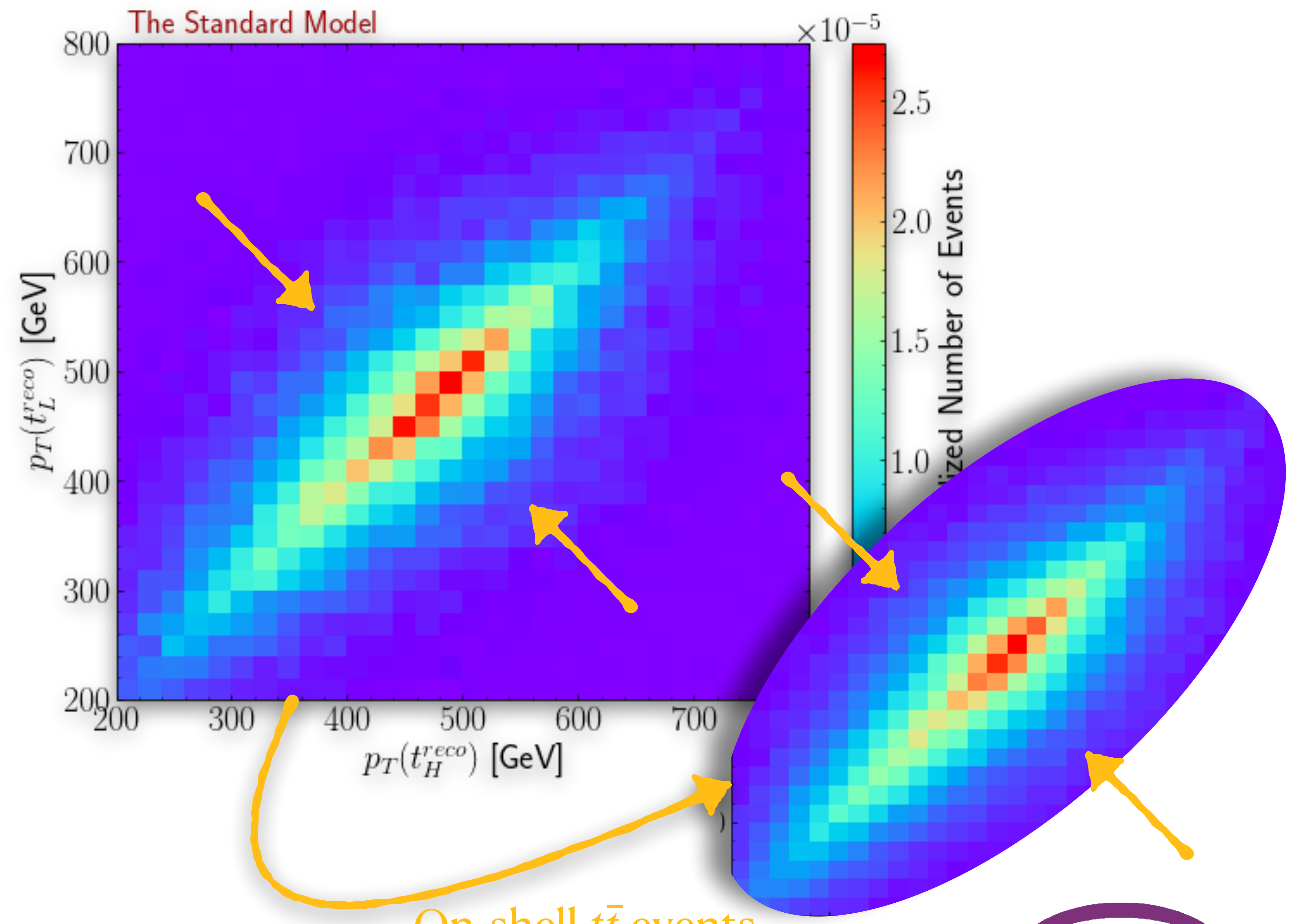
Marzani, Gregory, Spannowsky '19

HEPTopTagger on $t\bar{t}$ events

JYA, Buckley, Fuks; in progress



HEPTopTagger is looking for the resonance signature of W and top, which can be seen from the Dalitz plot.



Lets get started