MadAnalysis 5: Status & Plans

Jack Y. Araz
Stony Brook University

On behalf of Ma5 Team
RiF@CERN 2025

Jefferson Lab EXPLORING THE NATURE OF MATTER
U.S. DEPARTMENT OF ENERGY | Office of Science
MCnet
Stony Brook University

Search for new physics

Lagrangian

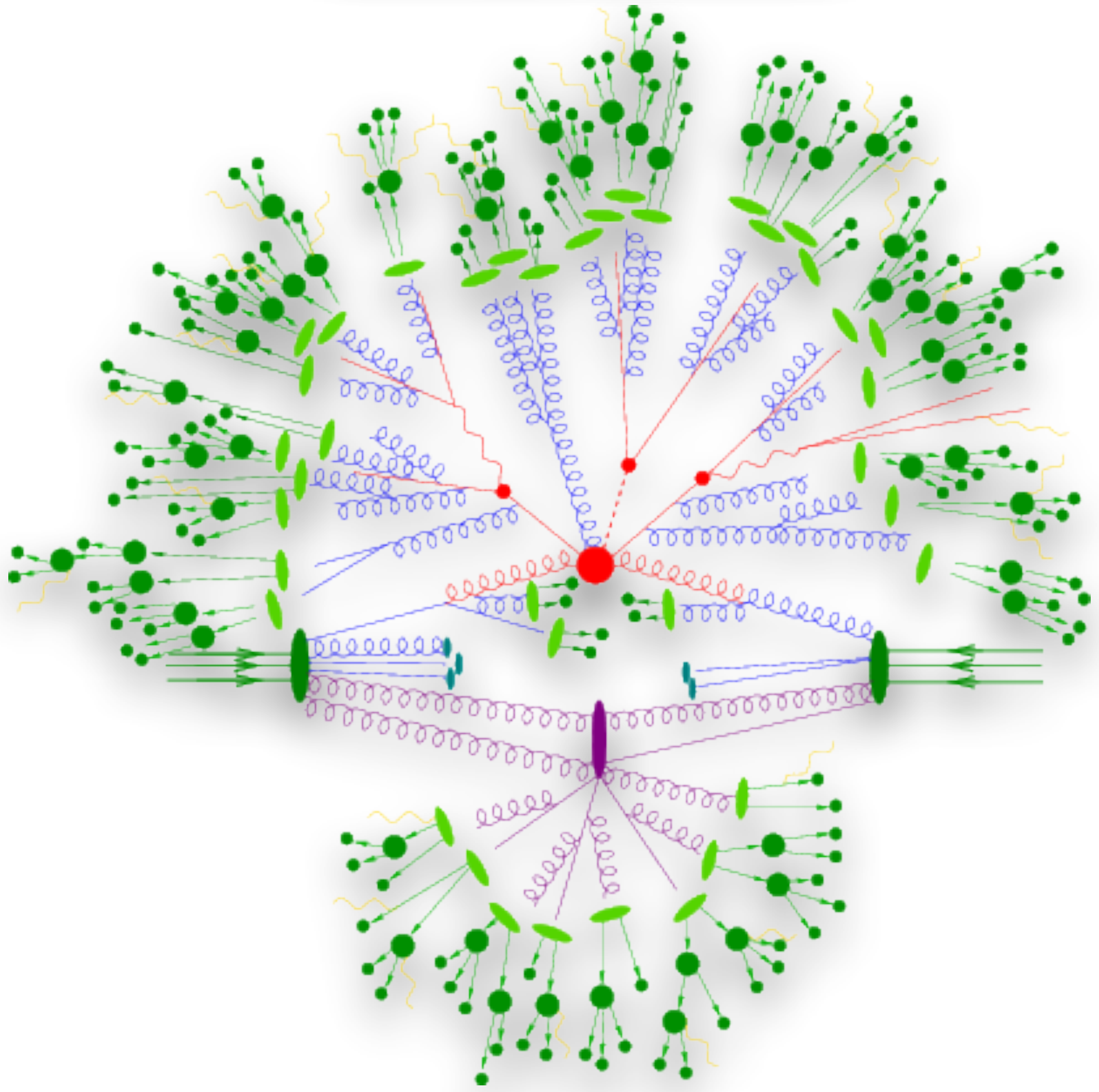
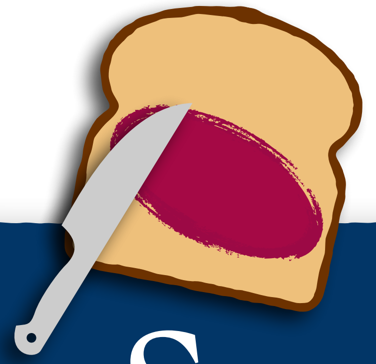
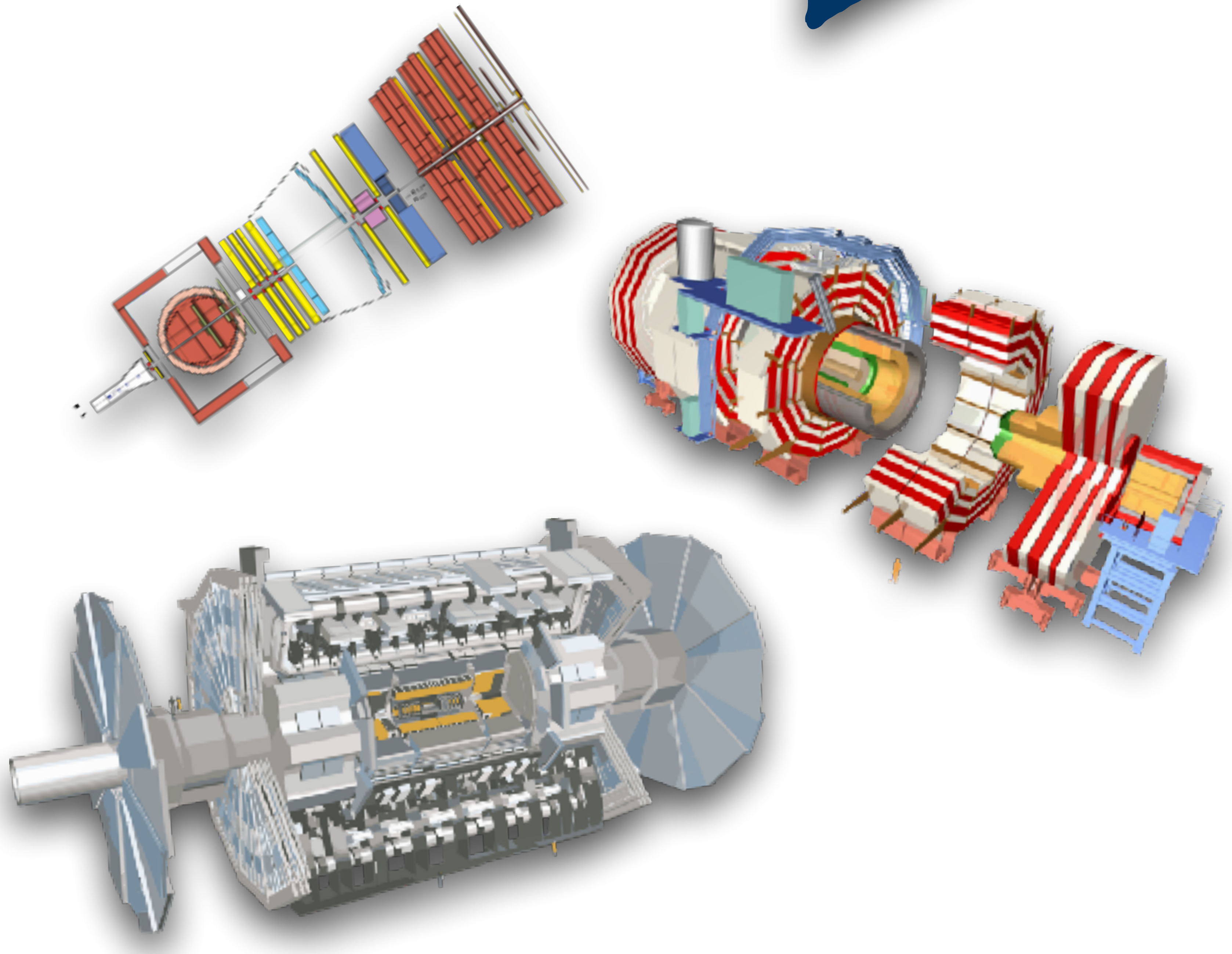
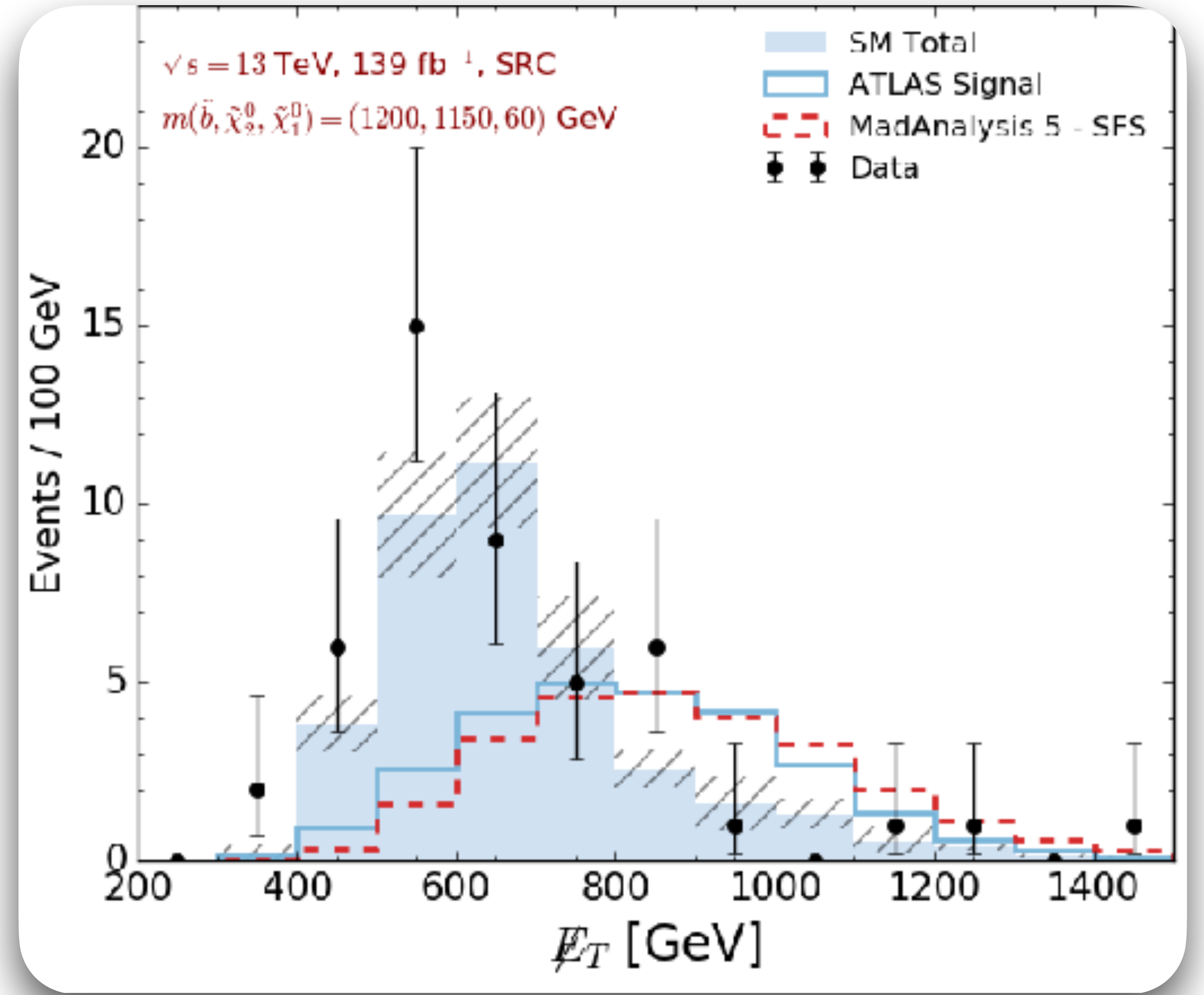


Image credit: Sherpa



Smearing





Smeared MC ⊕ observed data

MadAnalysis 5 in a nutshell

Conte, Fuks & Serret (CPC`13)

Conte & Fuks (IJMPA`19)

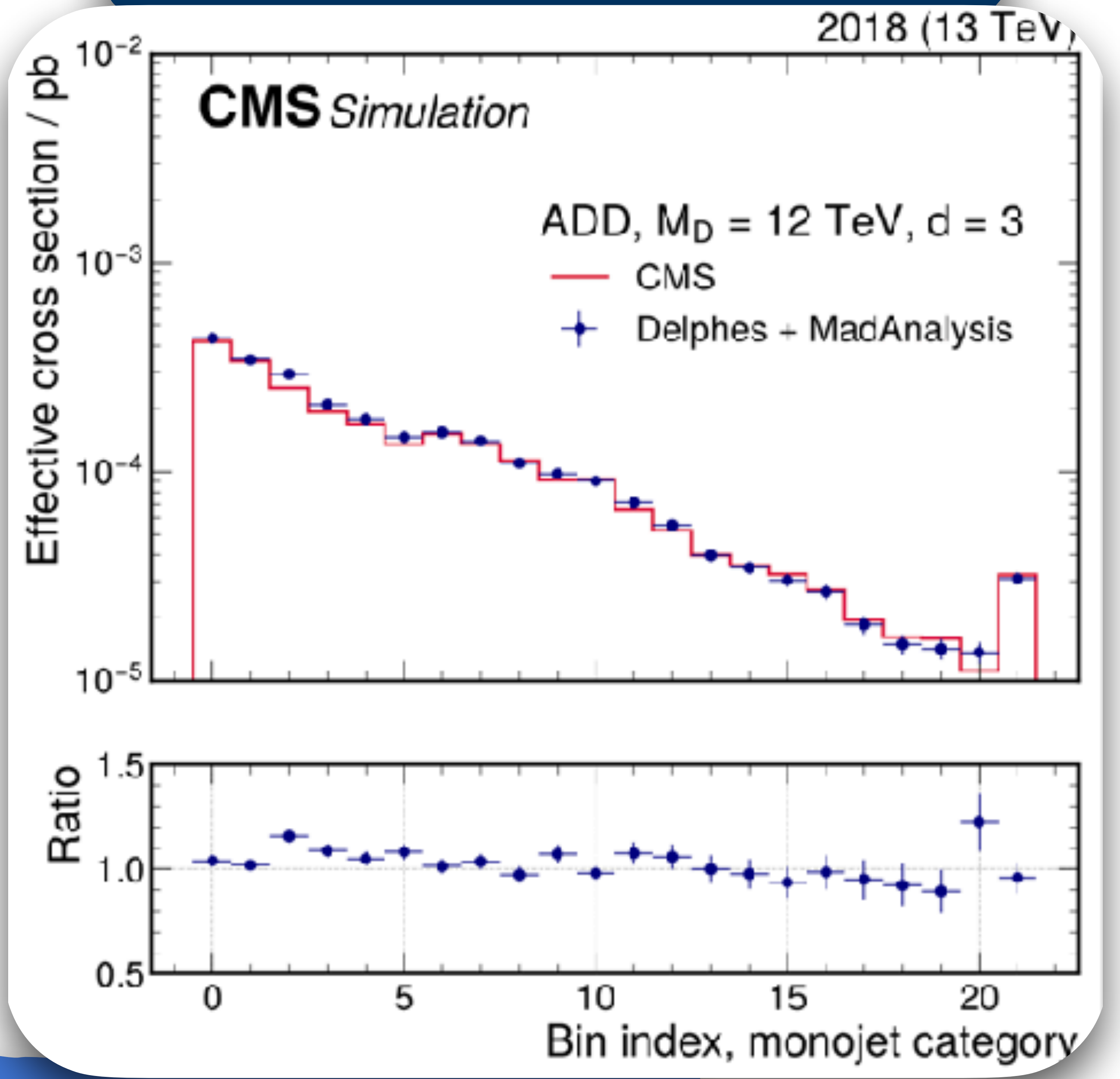
JYA & Fuks (in prep.)

Making the best of the LHC!

- Exploring the full potential of the LHC (for BSM)
  - Designing new analyses = probing new ideas
  - Recasting existing analyses = viability of models
- Data preservation beyond raw data
  - Analyses and their results = the LHC legacy

MadAnalysis 5 in a nutshell

Reinterpretation of CMS monojet search results with MA5



CMS (JHEP`21)

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!!! (Coming soon: even faster multi-core analysis!)
- Interfaces several HEP packages: MadGraph, FastJet, Delphes, pyhf, SFS

MadAnalysis 5 in a nutshell

Reinterpretation of CMS monojet search results with MA5

CMS (JHEP`21)

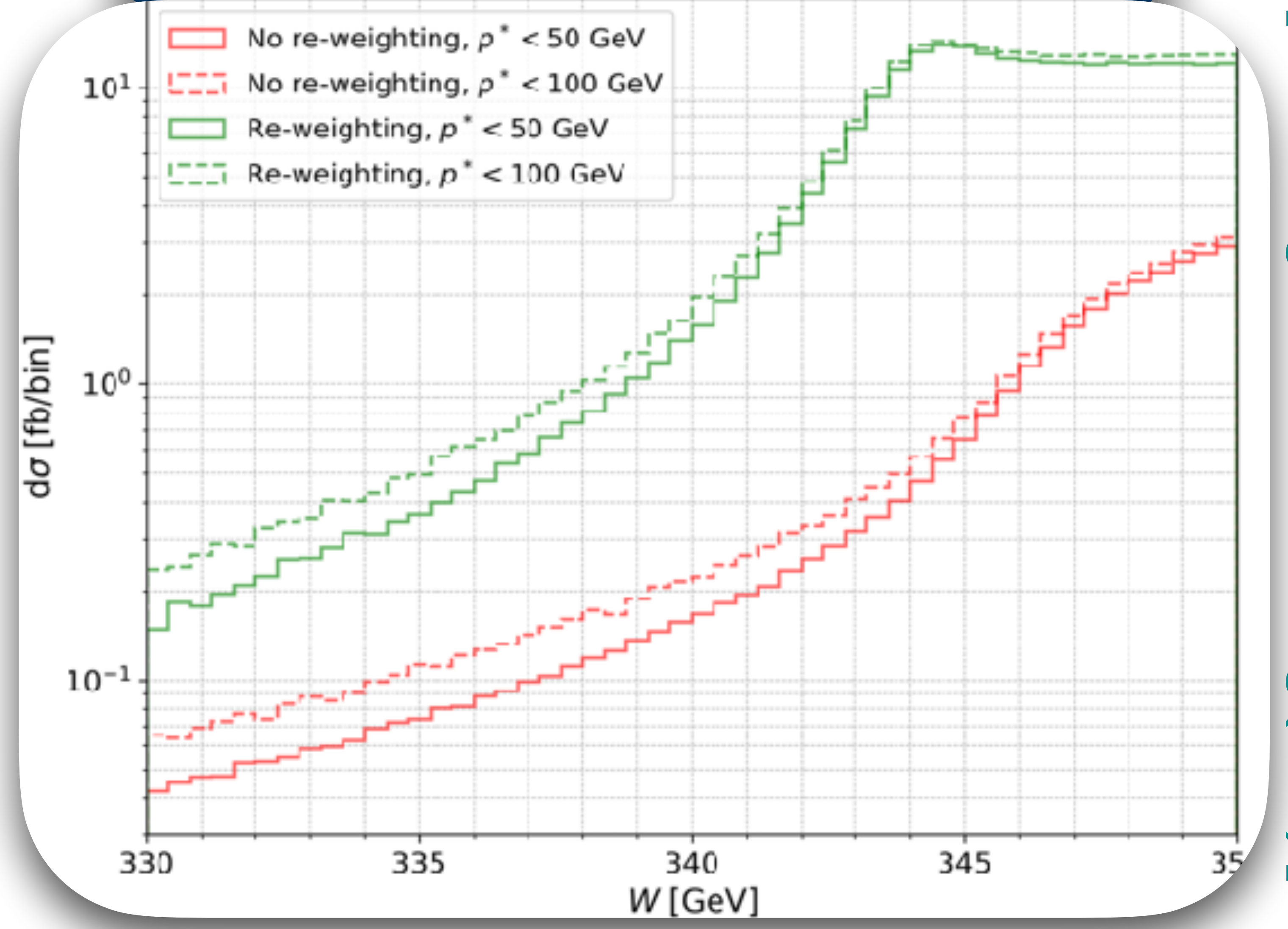
Normal Mode

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

Expert Mode

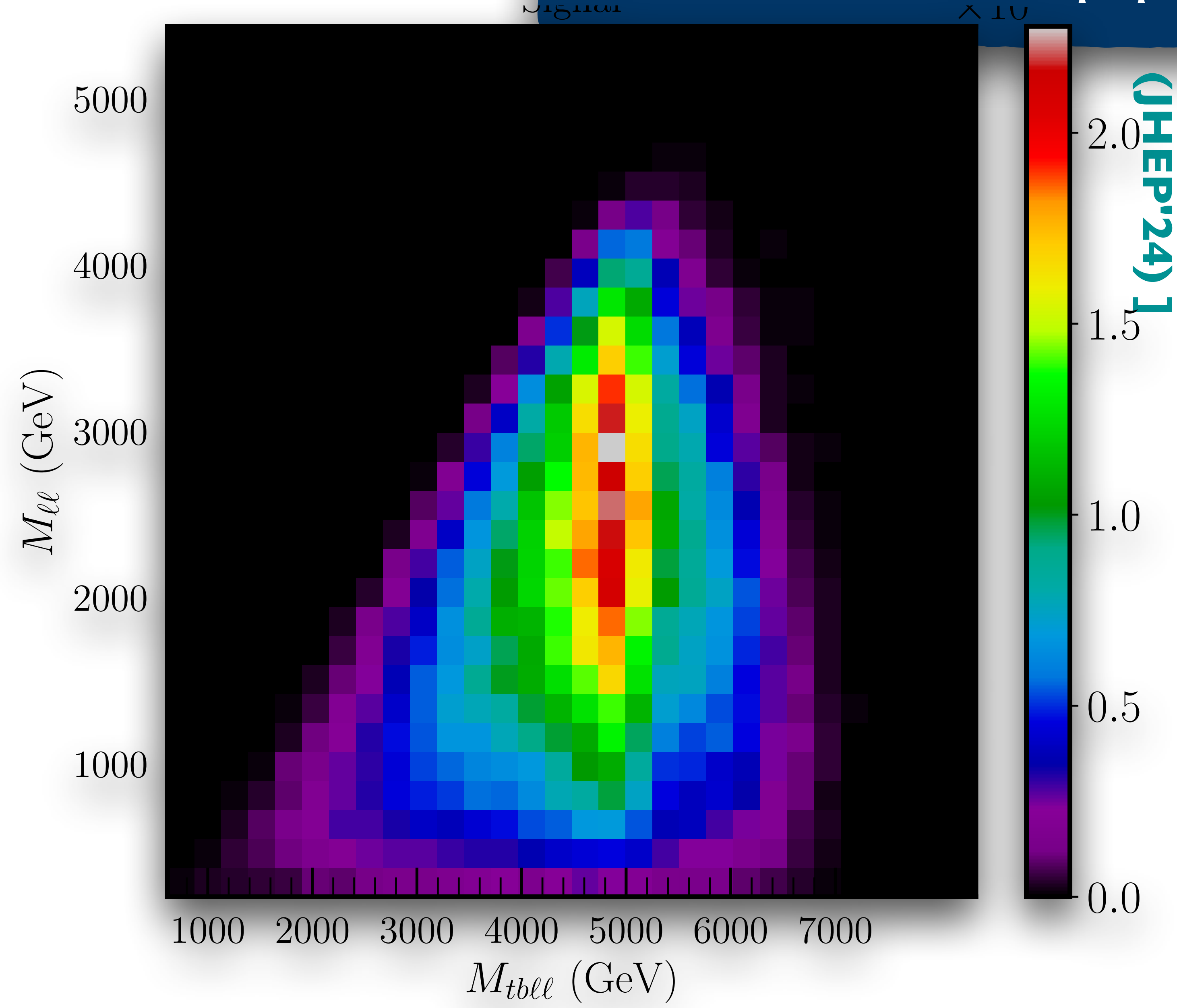
- C++ programming with the SampleAnalyzer framework
- Support for multiple sub-analyses, an efficient way for handling cuts and histograms, etc.

Top-antitop invariant mass distribution with toponium effects



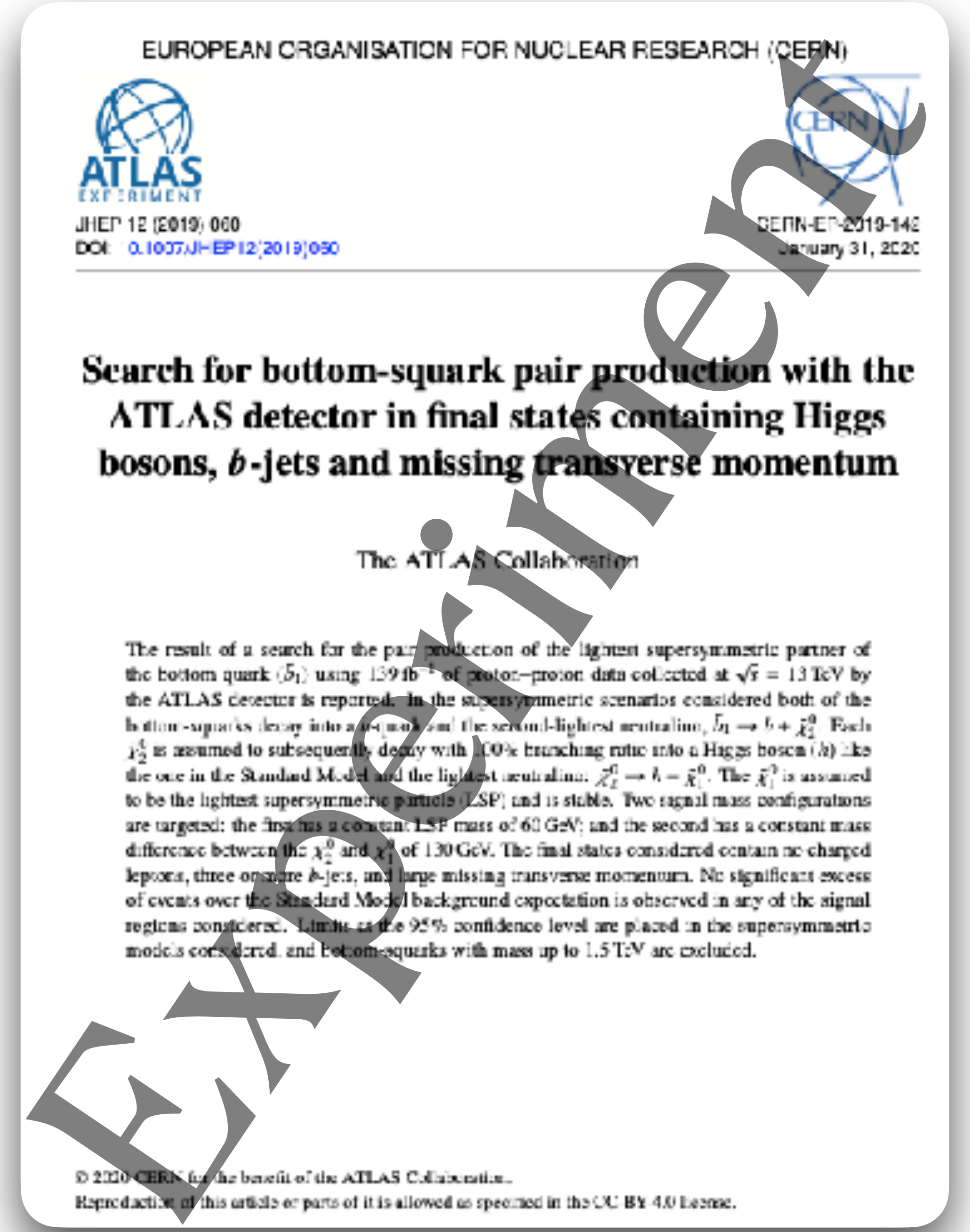
[Fuks, Hagiwara, Ma & Zheng (2024)]

KS mechanism with top quarks



[Frank, Fuks, Jueid, Moretti, Özdal (JHEP'24)]

(Re)interpretation of an analysis


EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

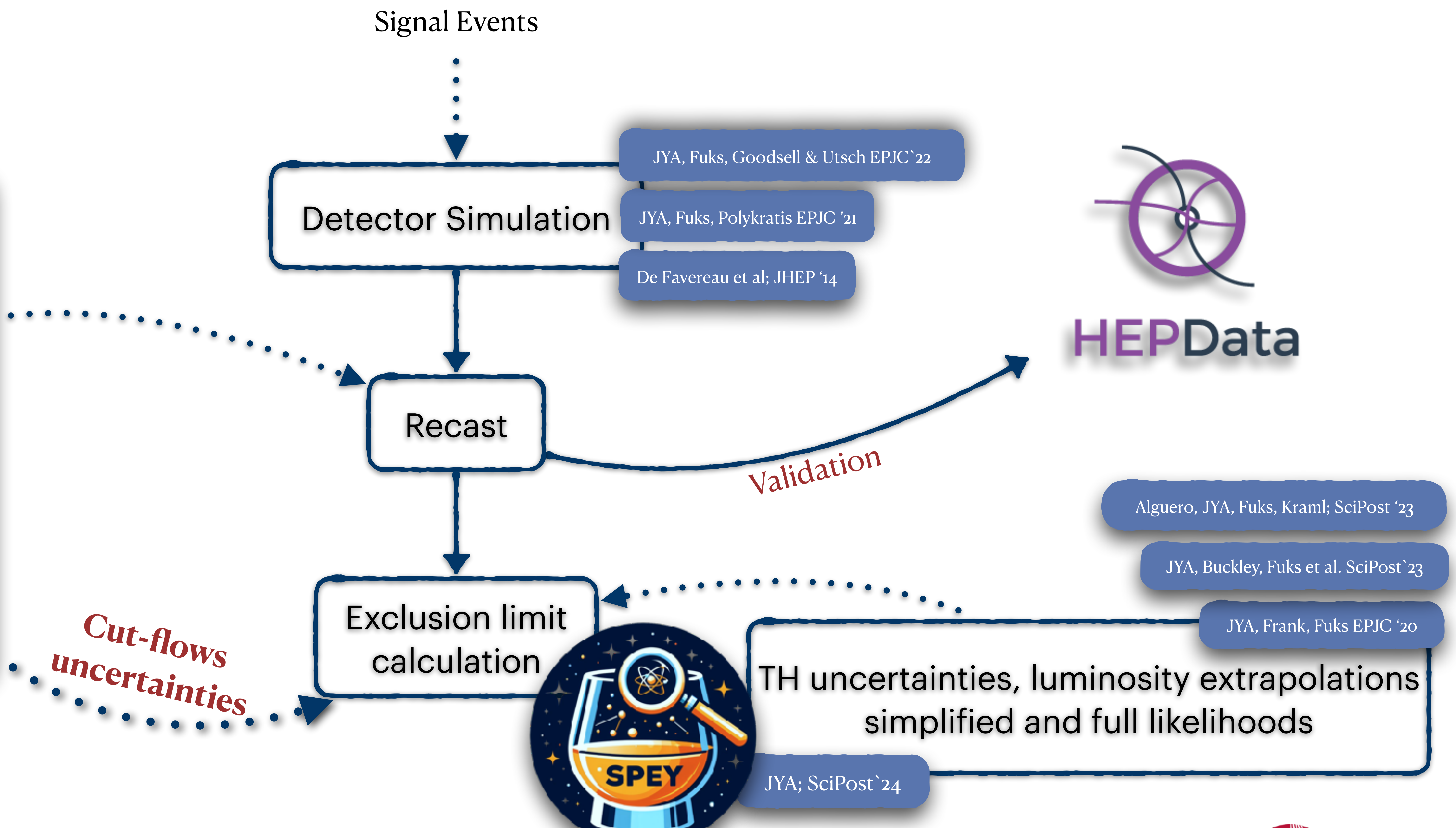
JHEP 12 (2019) 060
DOI: 10.1007/JHEP12(2019)060

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

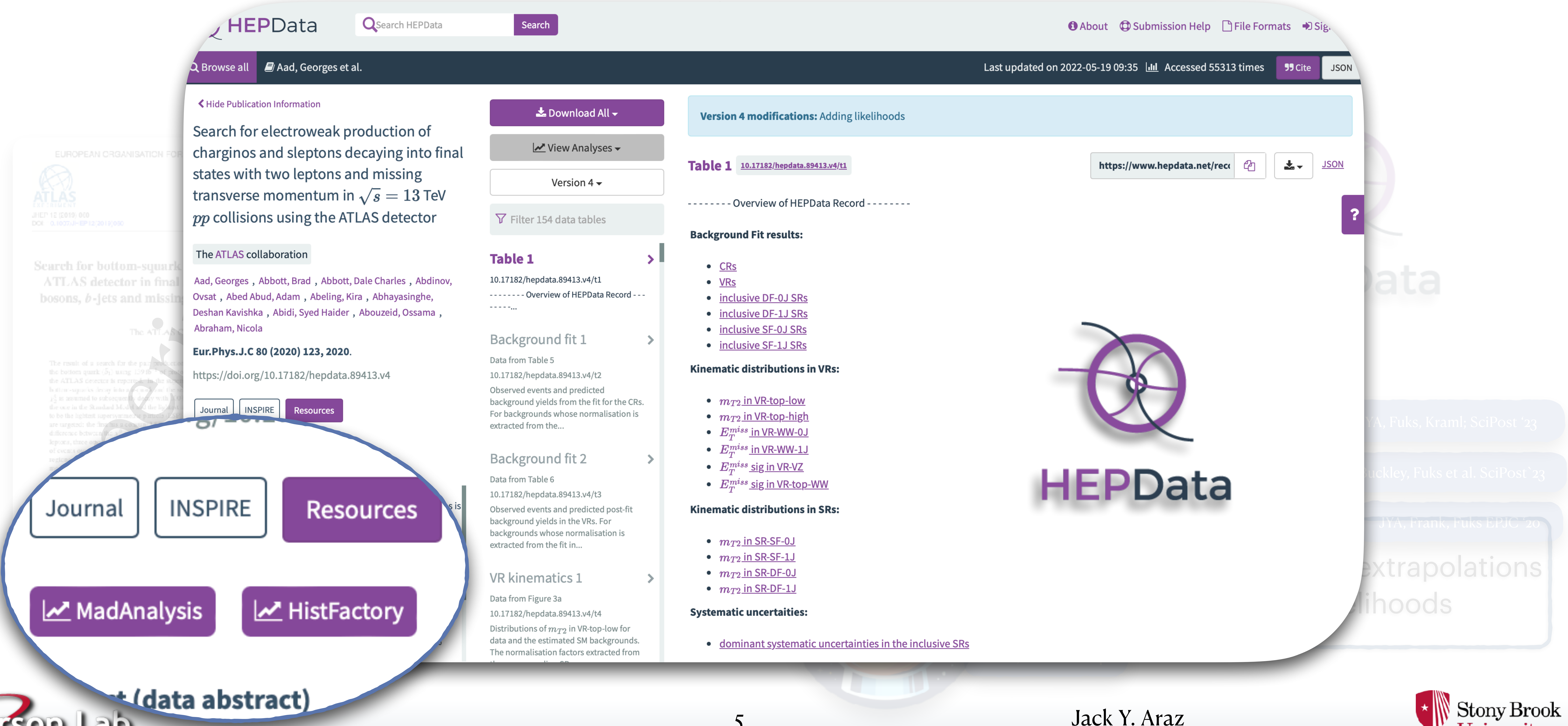
The ATLAS Collaboration

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.

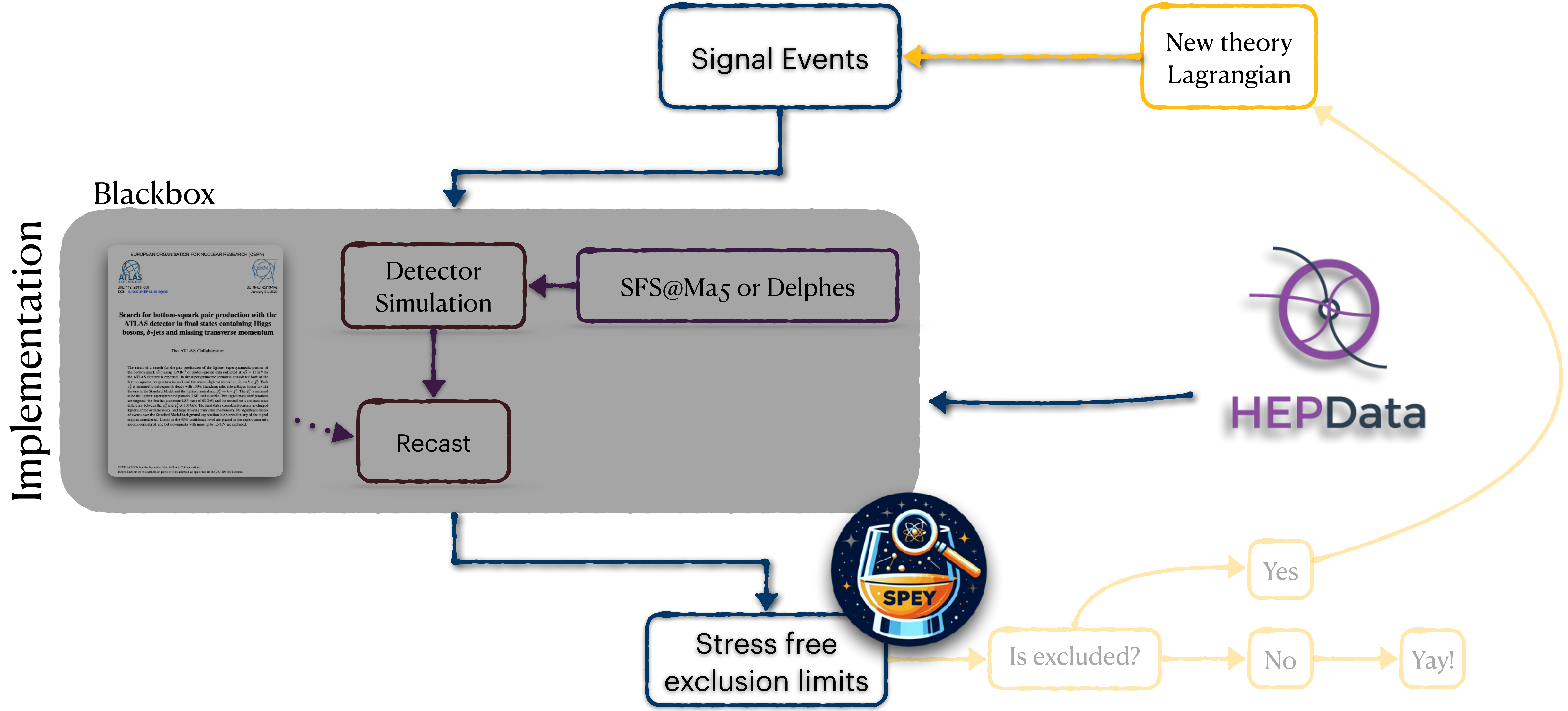


Jack Y. Araz

(Re)interpretation of an analysis



Reusable analysis framework



Reusable analysis framework

You DO NOT need to be a MadAnalysis 5 developer to contribute!

ADD YOUR RECAST TO THE PAD TODAY!

Blackbox

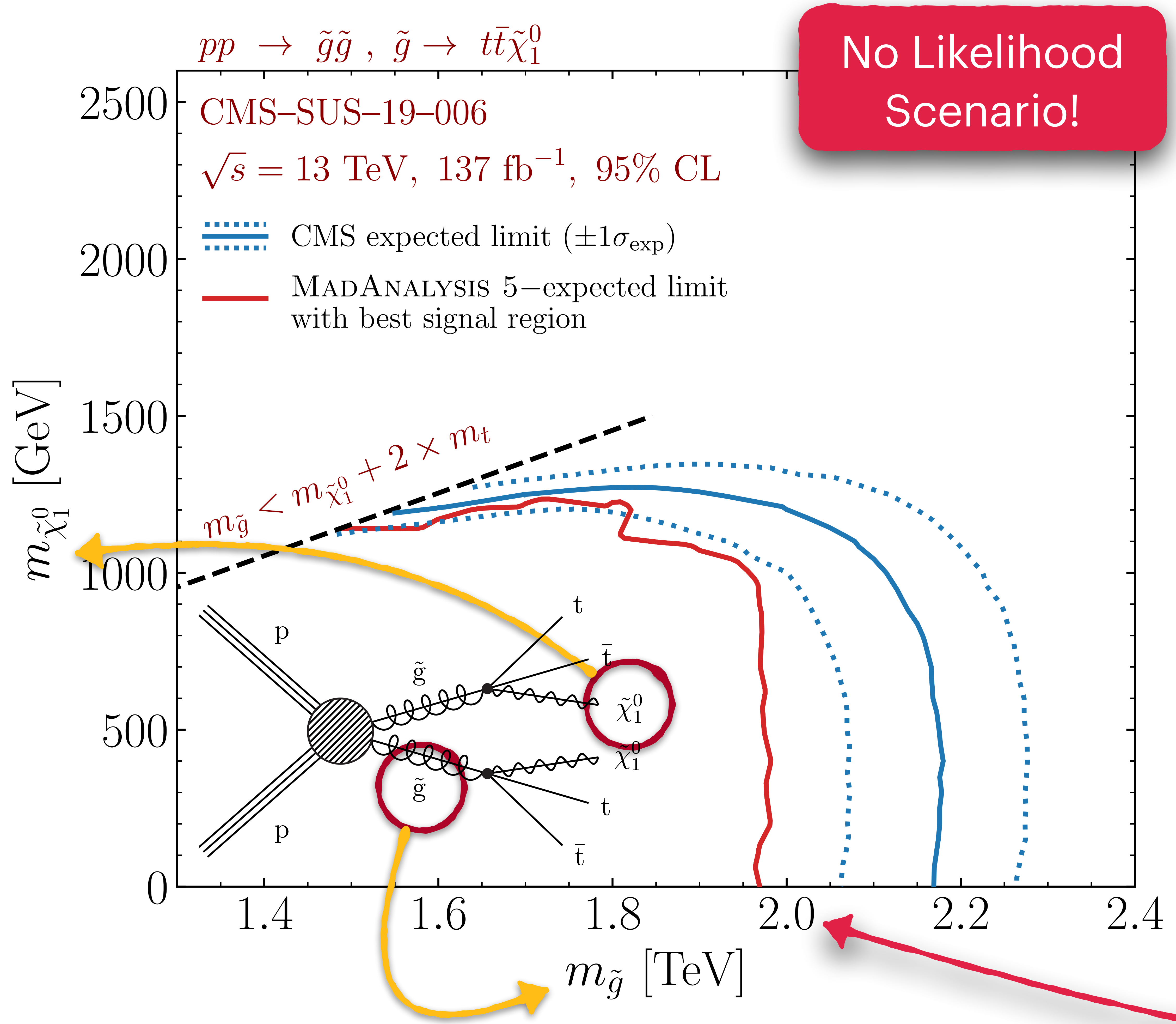
Stress free exclusion limits

Yay!

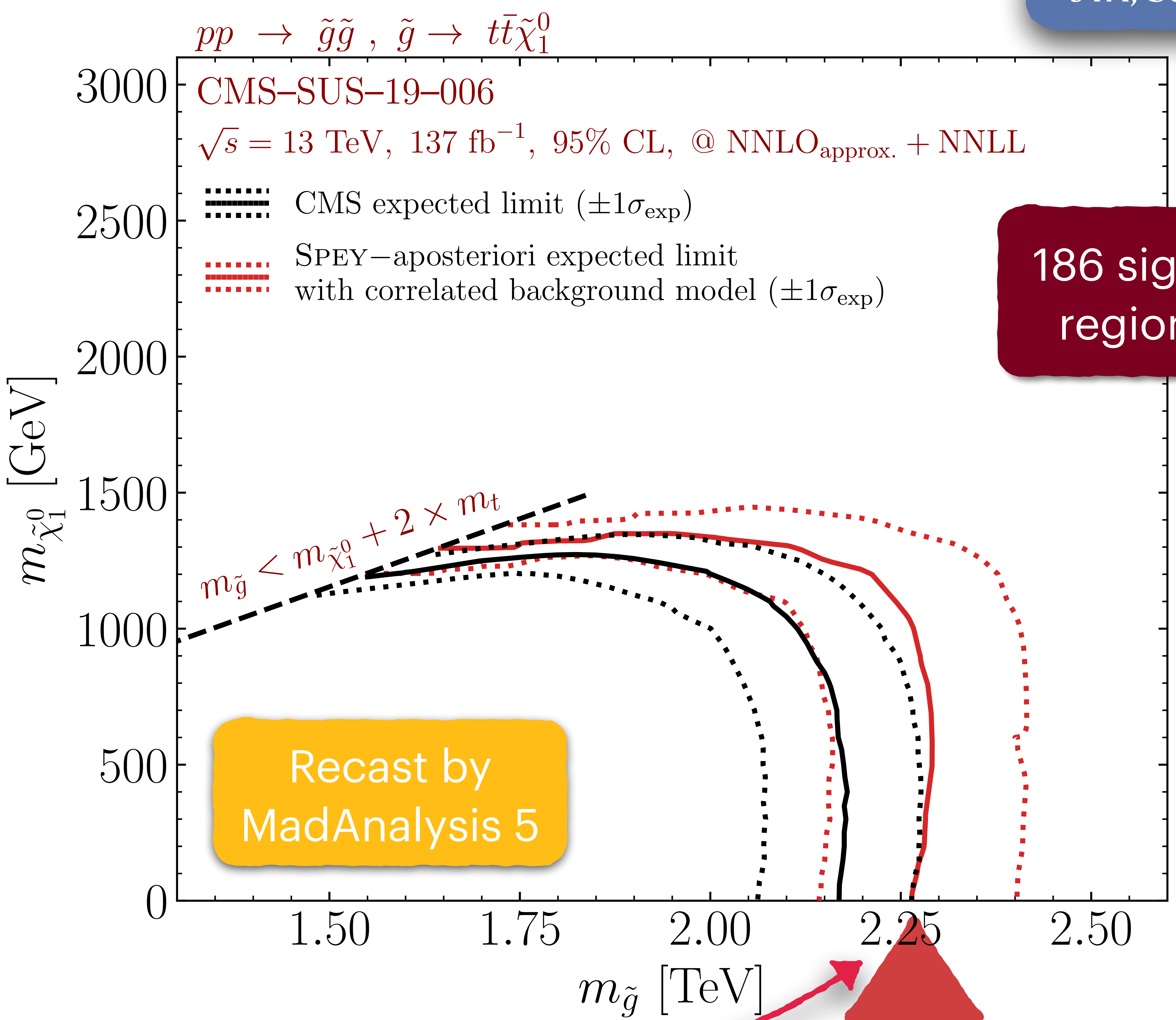
New Analyses & Techniques for PAD

Simplified likelihoods

JYA; SciPost`24



No Likelihood Scenario!



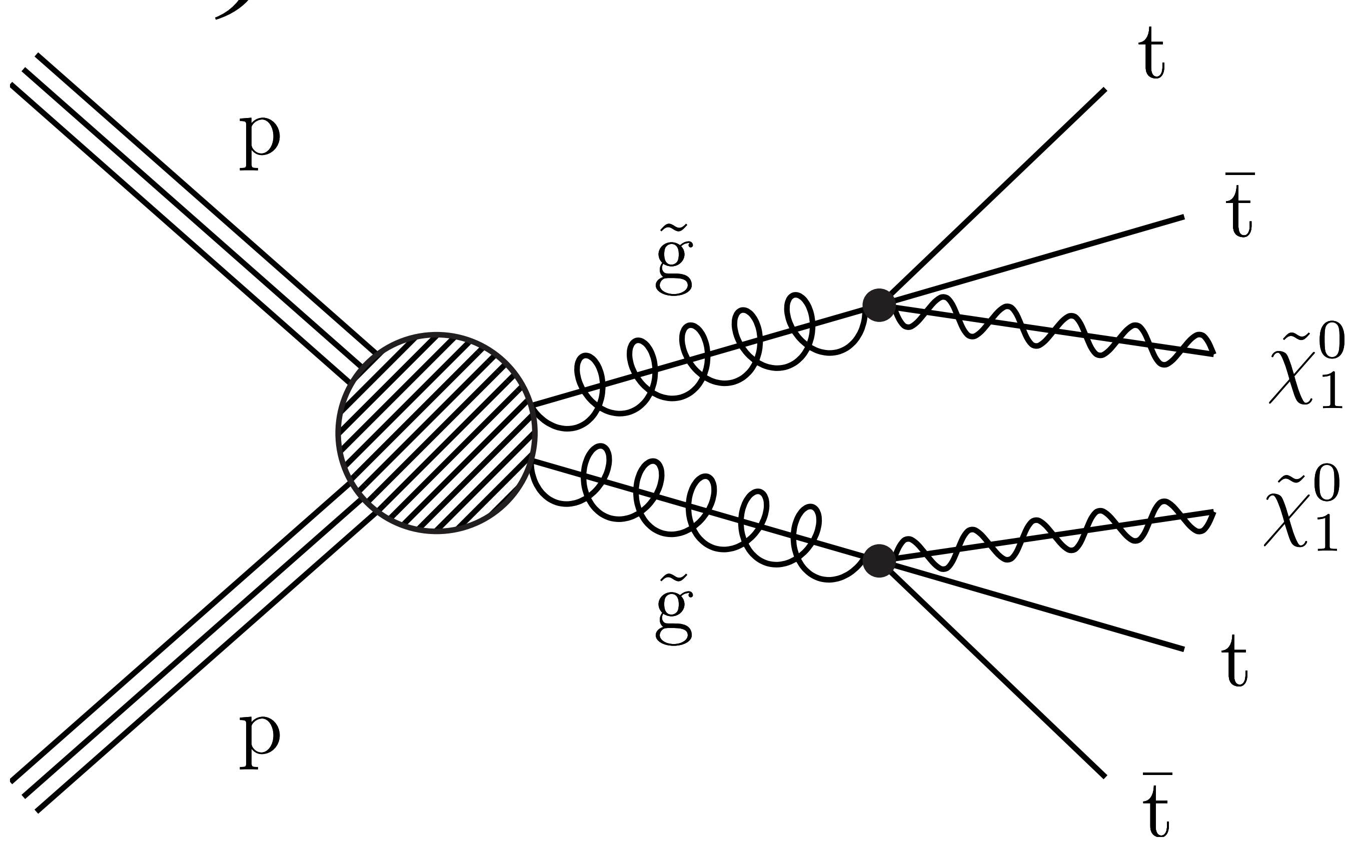
186 signal regions

Recast by MadAnalysis 5

Third moment expansion

JYA; SciPost`24

CMS-SUS-19-006

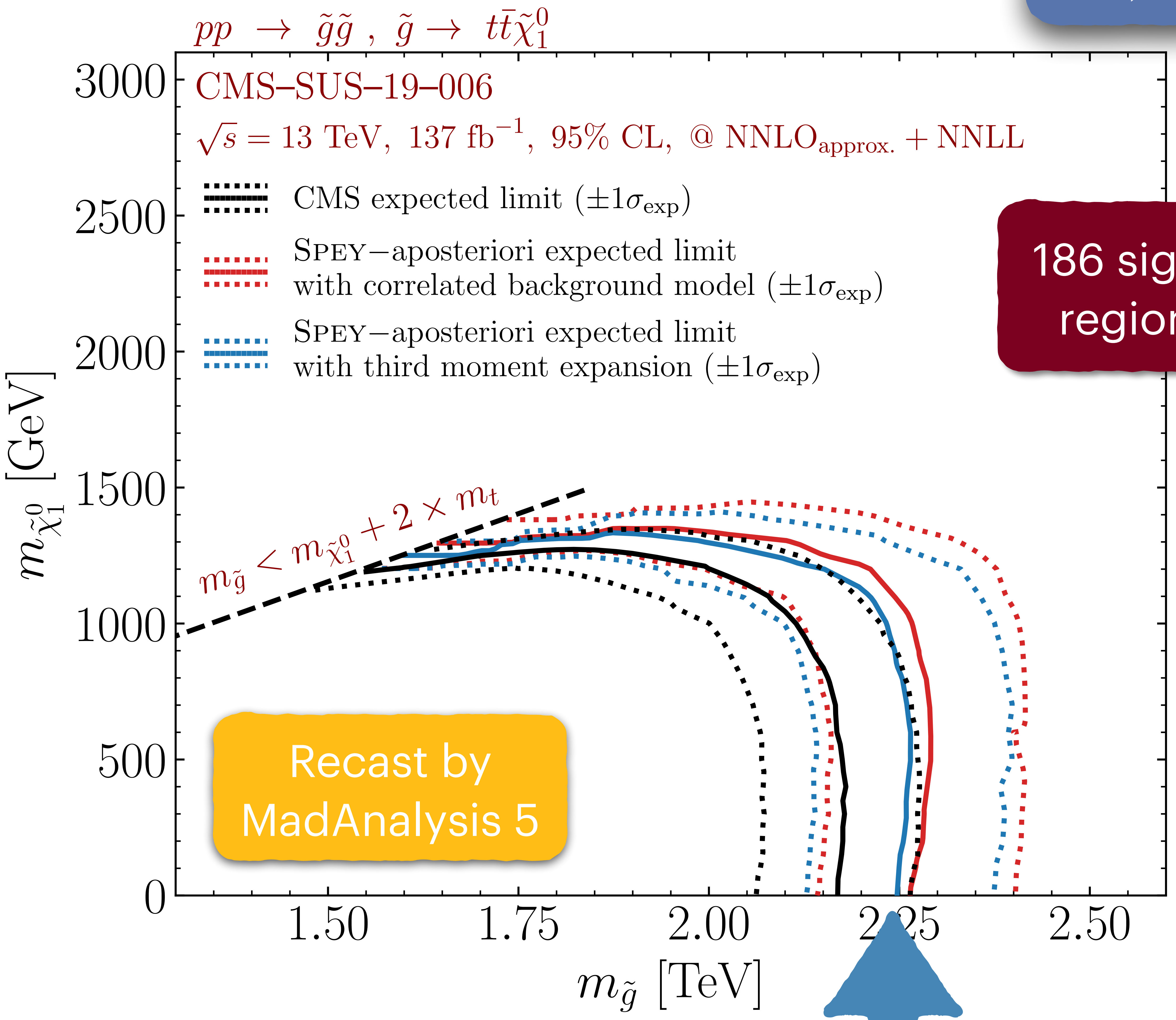


$$\mathscr{L}(\mu,\theta) = \left[\prod_{i\in\text{bins}} \text{Poiss}\left(n^i \,|\, \mu n_s^i + \bar{n}_b^i + A_i\theta_i + C_i\theta_i^2\right)\right] \cdot \mathscr{N}\left(\theta\,|\,0,\bar{\rho}\right)$$

$\bar{n}_b^i$:= the central value of the background

A_i := the effective sigma of the background uncertainty

C_i := asymmetry of the background uncertainty



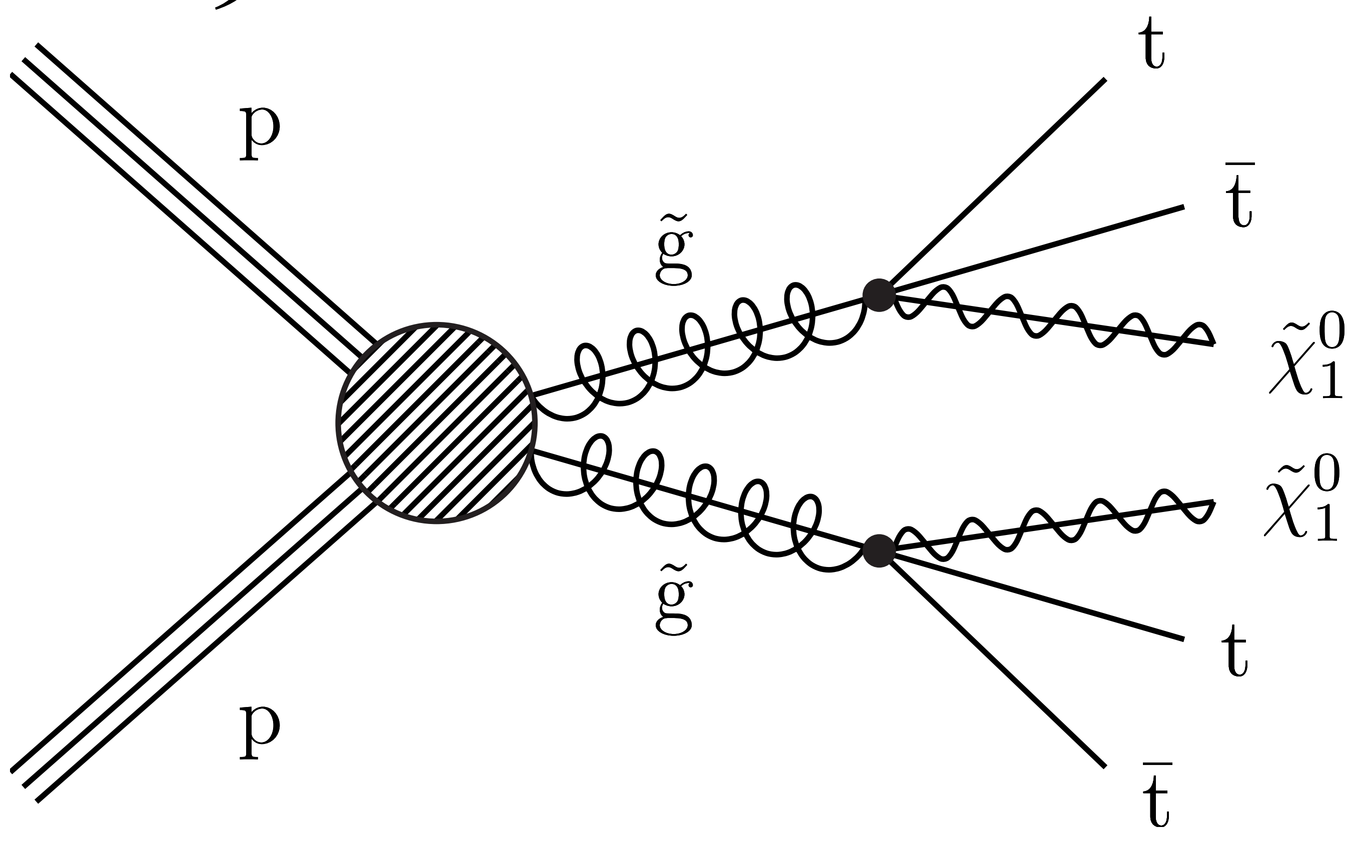
186 signal regions

Buckley, Citron, Fichet, Kraml, Waltenberger, Wardle; JHEP '18

Asymmetric Uncertainties

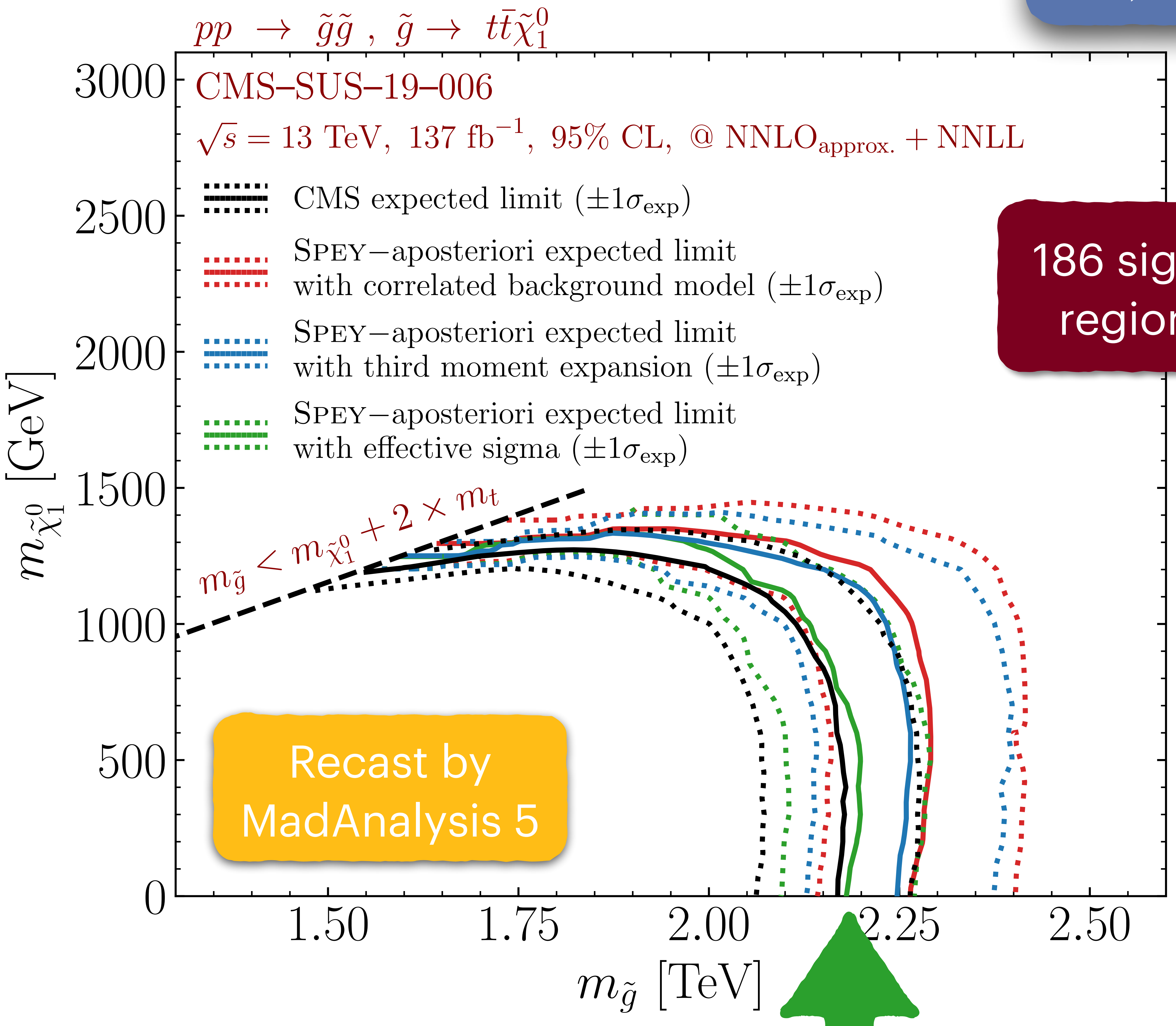
JYA; SciPost`24

CMS-SUS-19-006



$$\mathscr{L}(\mu,\theta) = \left[\prod_{i\in\text{bins}} \text{Poiss}\left(n^i \,|\, \mu n_s^i + n_b^i + \theta^i \sigma_{\text{eff}}^i(\theta^i)\right)\right] \cdot \mathscr{N}\left(\theta\,|\,0,\rho\right)$$

$$\sigma_{\text{eff}}^i(\theta^i) = \sqrt{\sigma_i^+\sigma_i^- + (\sigma_i^+ - \sigma_i^-)(\theta^i - n_b^i)}$$



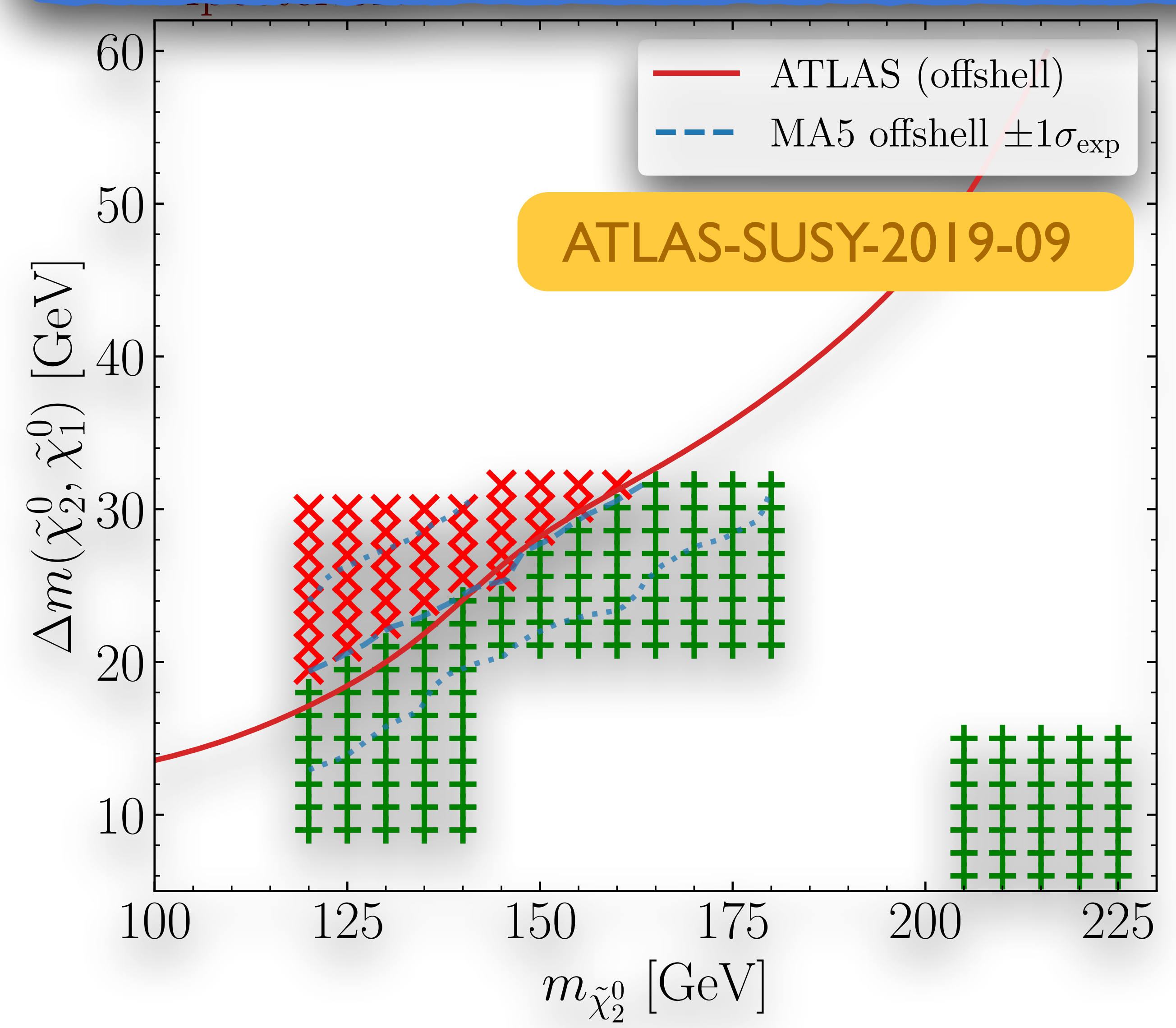
186 signal regions

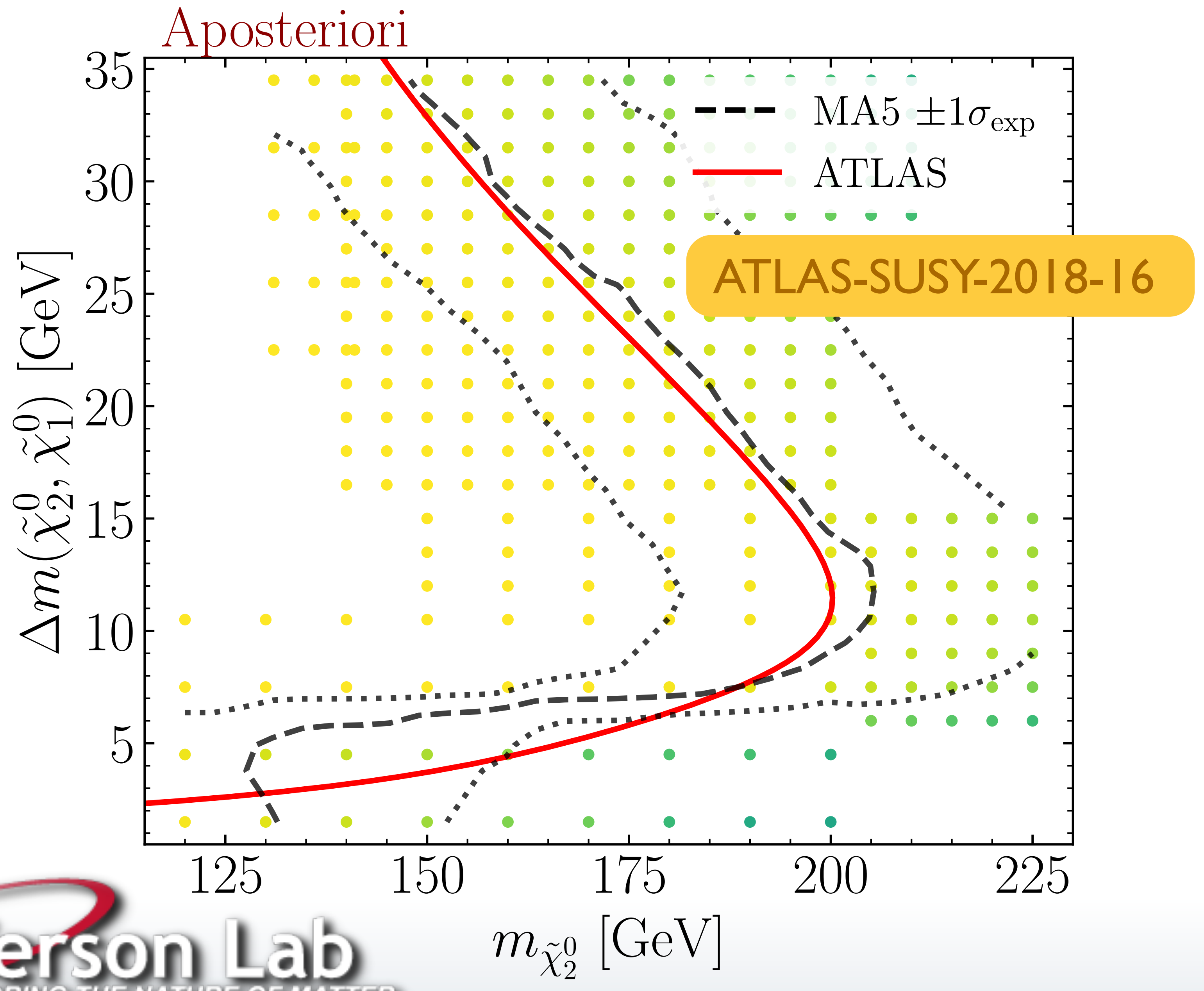
New analyses & techniques for PAD

JYA, Fuks, Goodsell & Murphy (in prep.)

ATLAS-SUSY-2018-16 (soft di-lepton + $\not\!\!E_T$) & ATLAS-SUSY-2019-09 (3 leptons + $\not\!\!E_T$)

- Standard selection
  - ➙ 1OSSF lepton pair, with/without jets
  - ➙ Object isolation, lepton properties, jigsaw, on-shell Z.; bins in $m_{\ell\ell}$
  - ➙ PYHF model file (full likelihoods for limit settings ≡ signal region combination) [interfaced in MA5 for more than three years]
- Recasting challenges
  - NEW
  - ➙ Using RESTFRAMES (*cf.* jigsaw variables) ≡ strong impact on exclusions
  - ➙ Event generation details important for compressed spectra





Many SRs; sub-percent/sub-permile efficiencies (per bin)

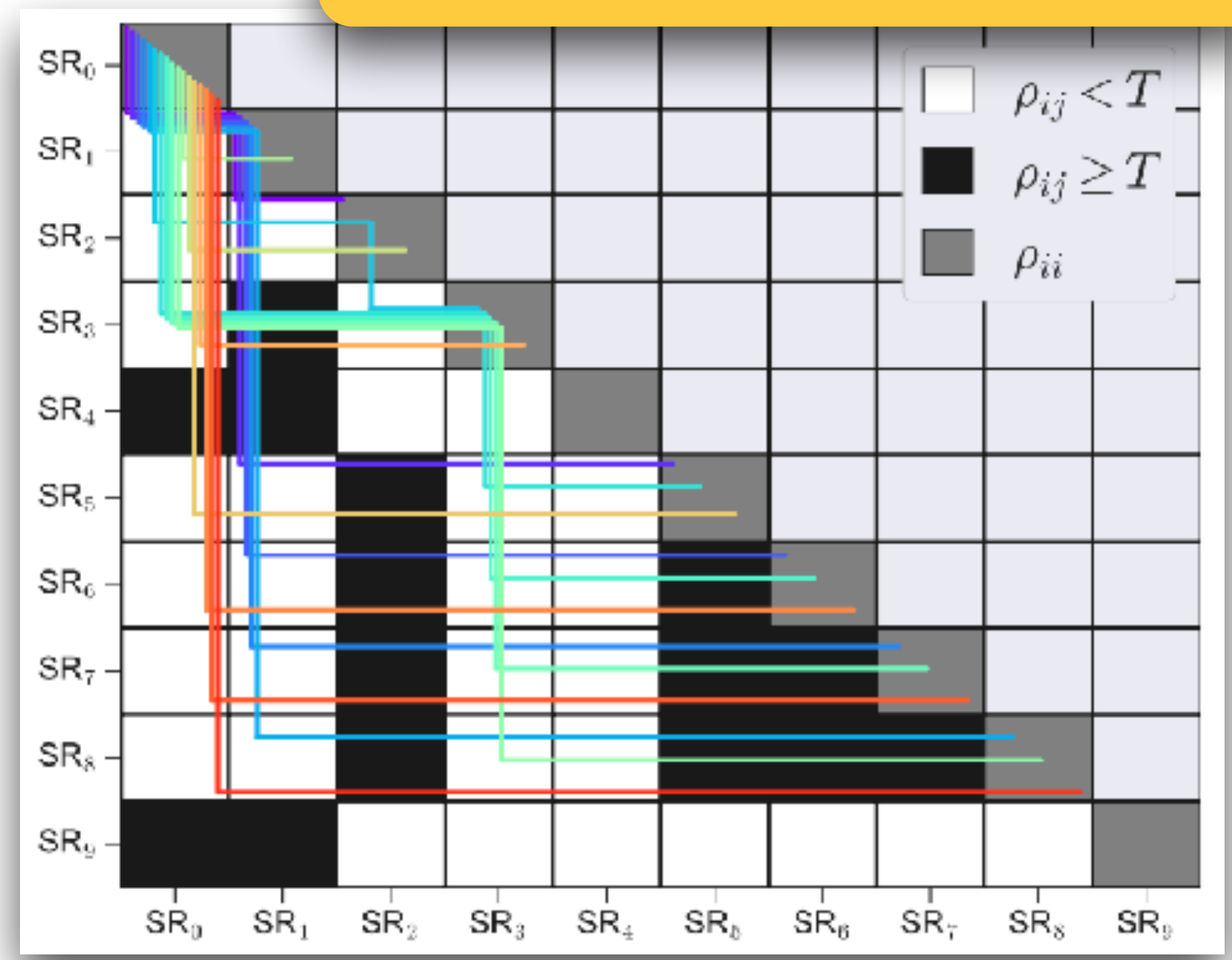
- 7.5 Mevents generated
  - ➙ numerical accuracy
- NEW
- Use of HEPDATA tabulated efficiencies
- Signal region combination crucial (dozens of SRs with low efficiencies)
- Signal rate and shape important (NLO+NLL; matrix-element merging)
- Outstanding agreement both for exp. and obs. results

Analysis Combination

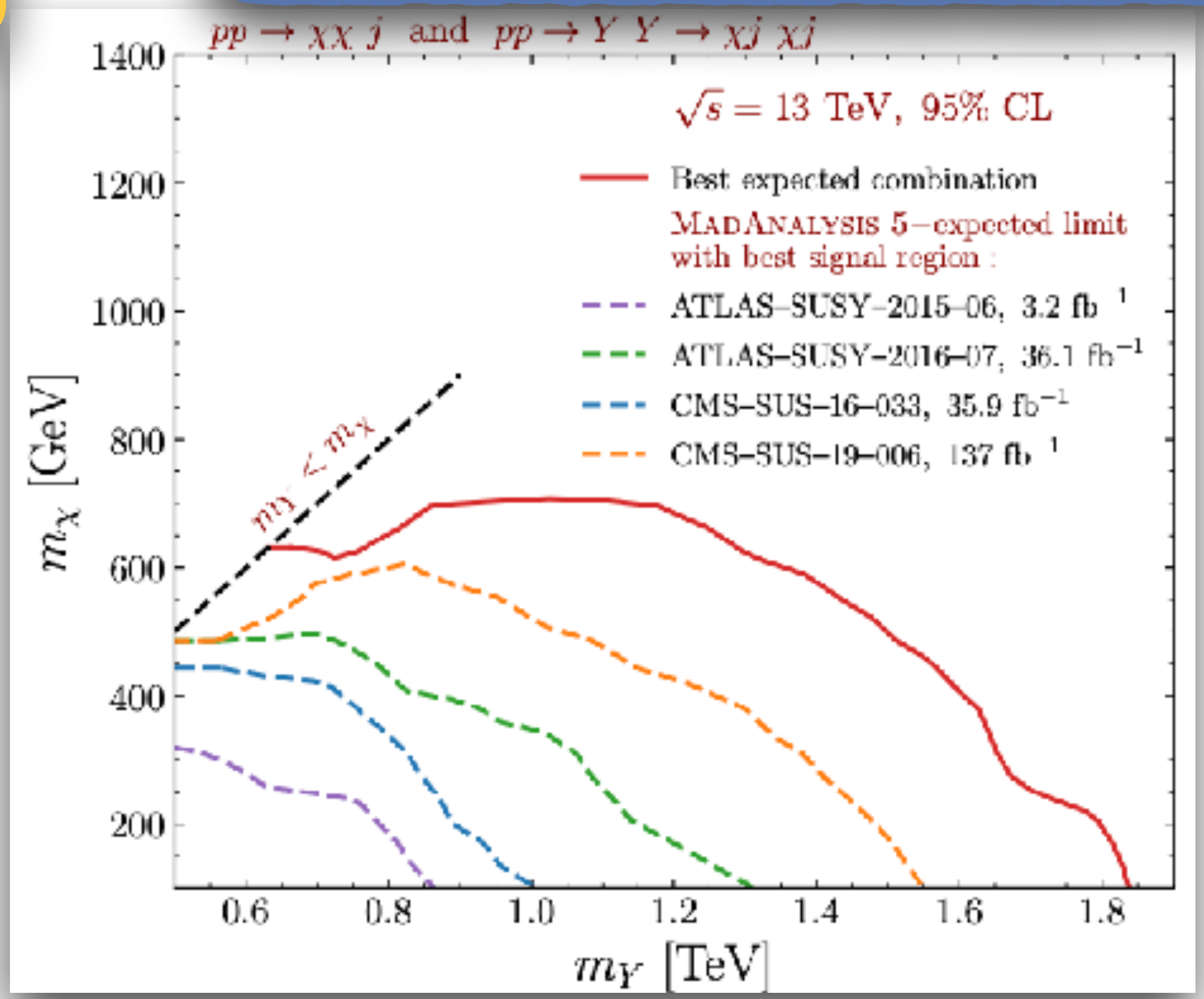
The TACO approach - testing analysis correlations

- Combination of analyses
  - ➙ Overlap matrix ≡ approximate correlation matrix
  - ➙ Path finding (set of non-overlapping regions) [weighted hereditary depth-first search algorithm]

Majorana DM+$\tilde{q}$ vs. jets+$\not{E}_T$



JYA, Buckley, Fuks et al. SciPost`23



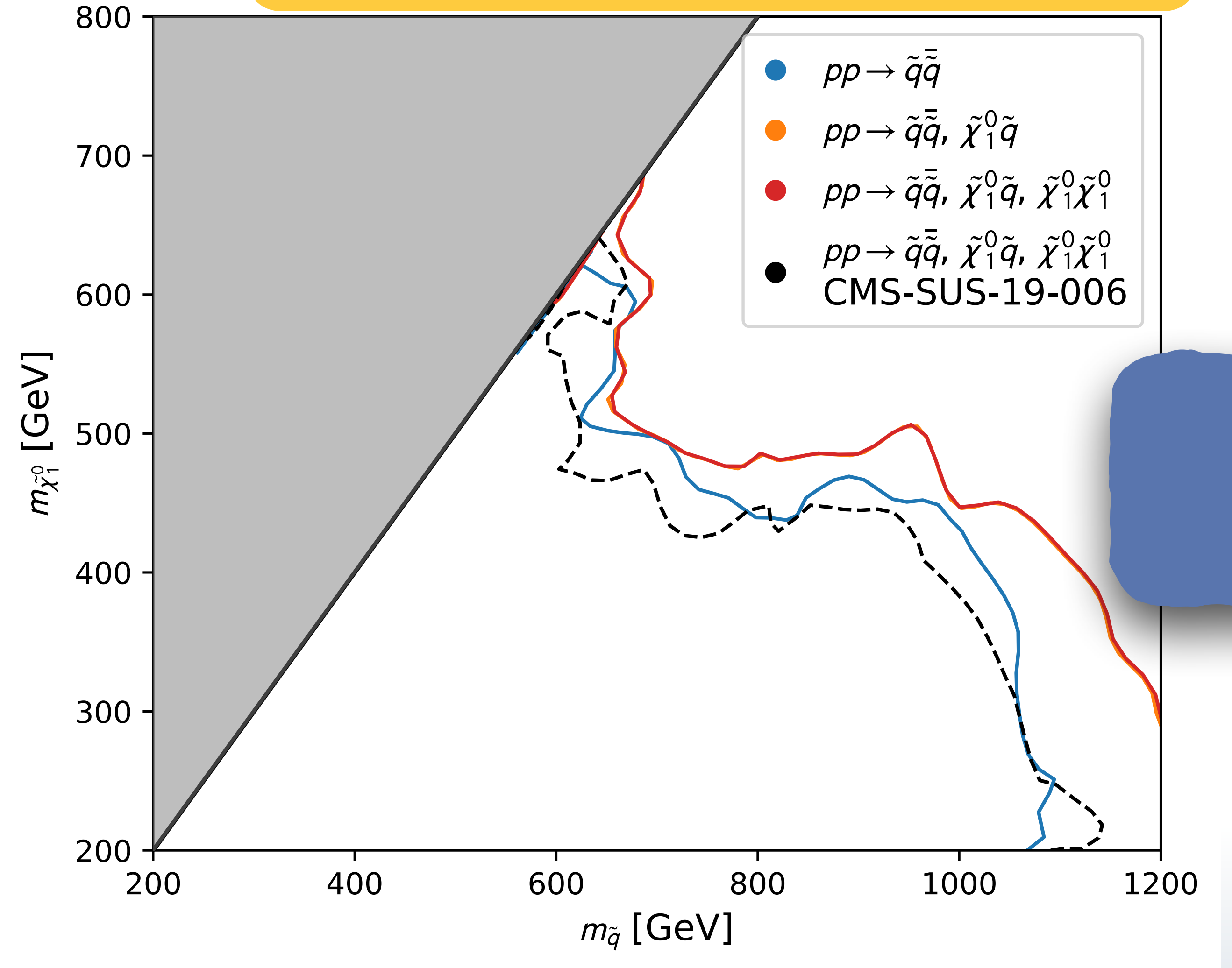
Application 1: the ATLAS 2015 pMSSM-19 scan

- 100s SRs: a few matter
- Going beyond ATLAS/CMS combinations
- Always a specific driving SR
  - ➙ Not powerful enough alone
  - ➙ *O(10)* regions combined

Application 2: MSSM

- Considered analyses: jets +$\not{E}_T$
  - ➙ Combination possible (non-overlapping SRs exist)
- Advantage of quantified measure of overlap
  - ➙ bounds more stringent by 100s GeV

Ewkino/squark vs. jets+$\not{E}_T$



Feike, Fiaschi, Fuks, Klasen & Neuwirth (JHEP`24)

Towards global sensitivity

JYA, SciPost '24

$\mathscr{L}' = \mathscr{L}_{\mathrm{ATLAS}} \oplus \mathscr{L}_{\mathrm{CMS}}$

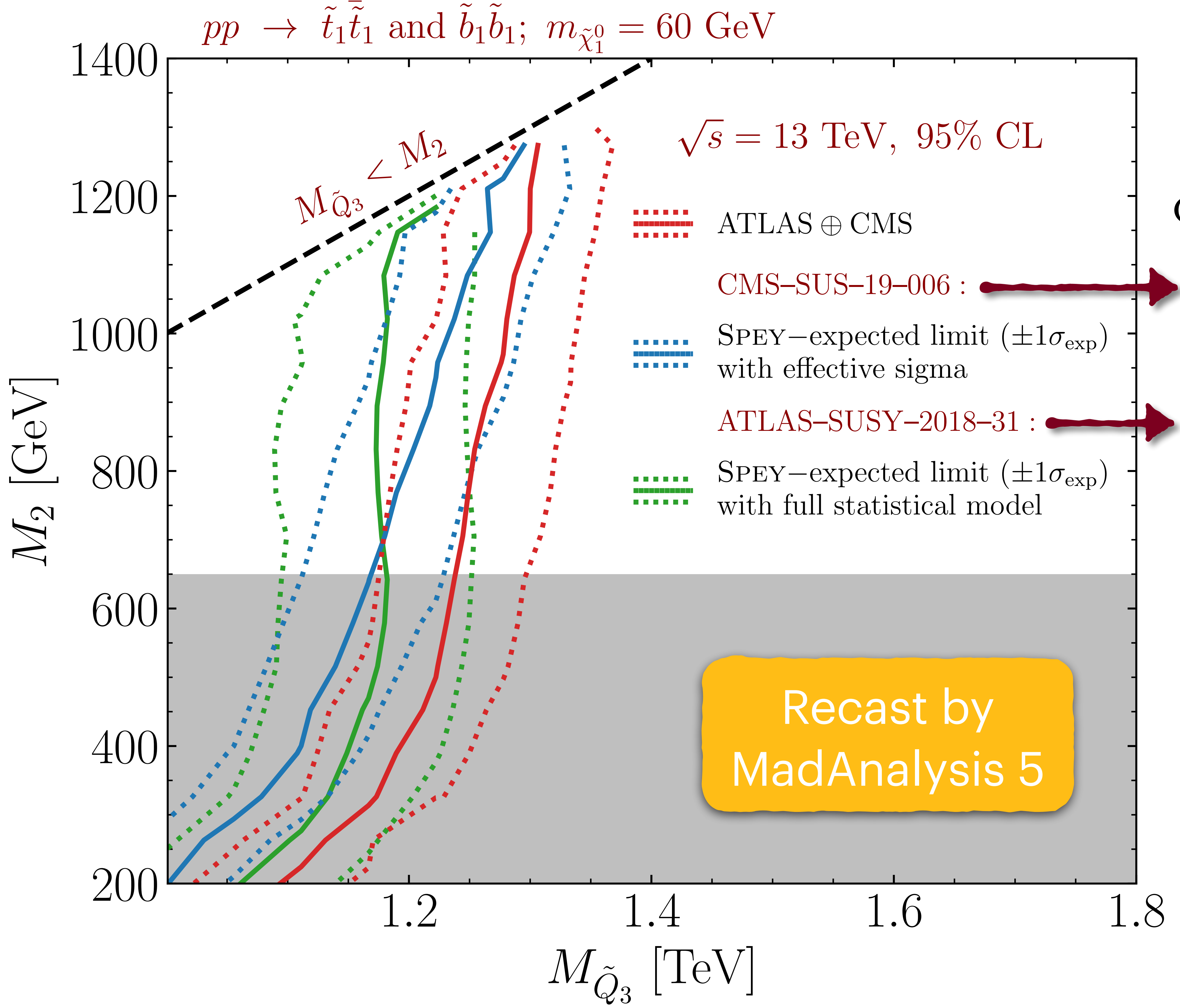
Full likelihood

Simplified likelihood with effective sigma model

A combination of analyses, rather than regions, contains much more information!

MSSM: $M_1 = M_2 = M_3 = M_{\tilde{Q}}$ at GUT scale

JYA; SciPost`24



Original analysis:

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

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

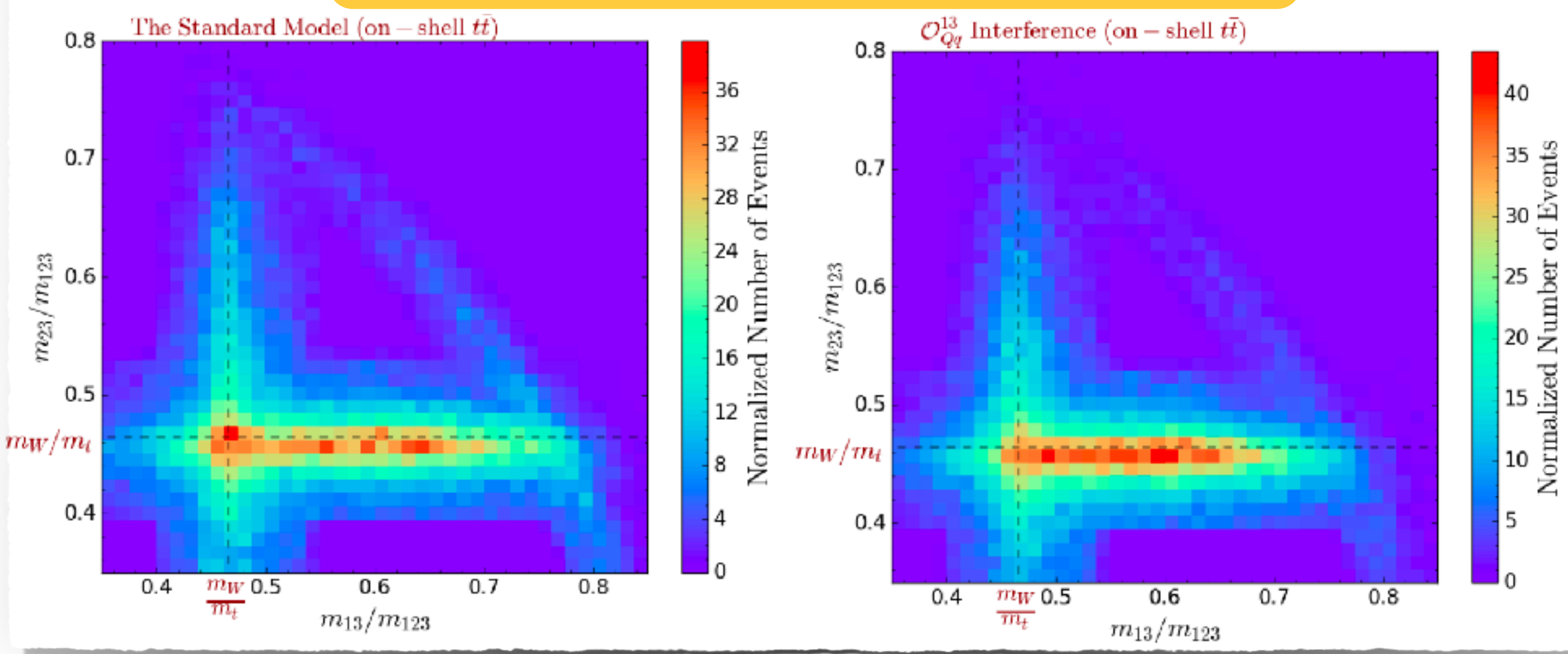
Upgrading SFS

JYA, Fuks (in progress)

Further SFS development: jet substructure tools

- Embedding of all jet substructure tools in the data format
- Interface with HEPTOPTAGGER
- Multiple definitions for a given type of objects
  → tight, medium, loose leptons, *etc.*
- Technical details to be released
- Follow us on GitHub (the *substructure* branch)

HEPTOPTAGGER with four-fermion operators

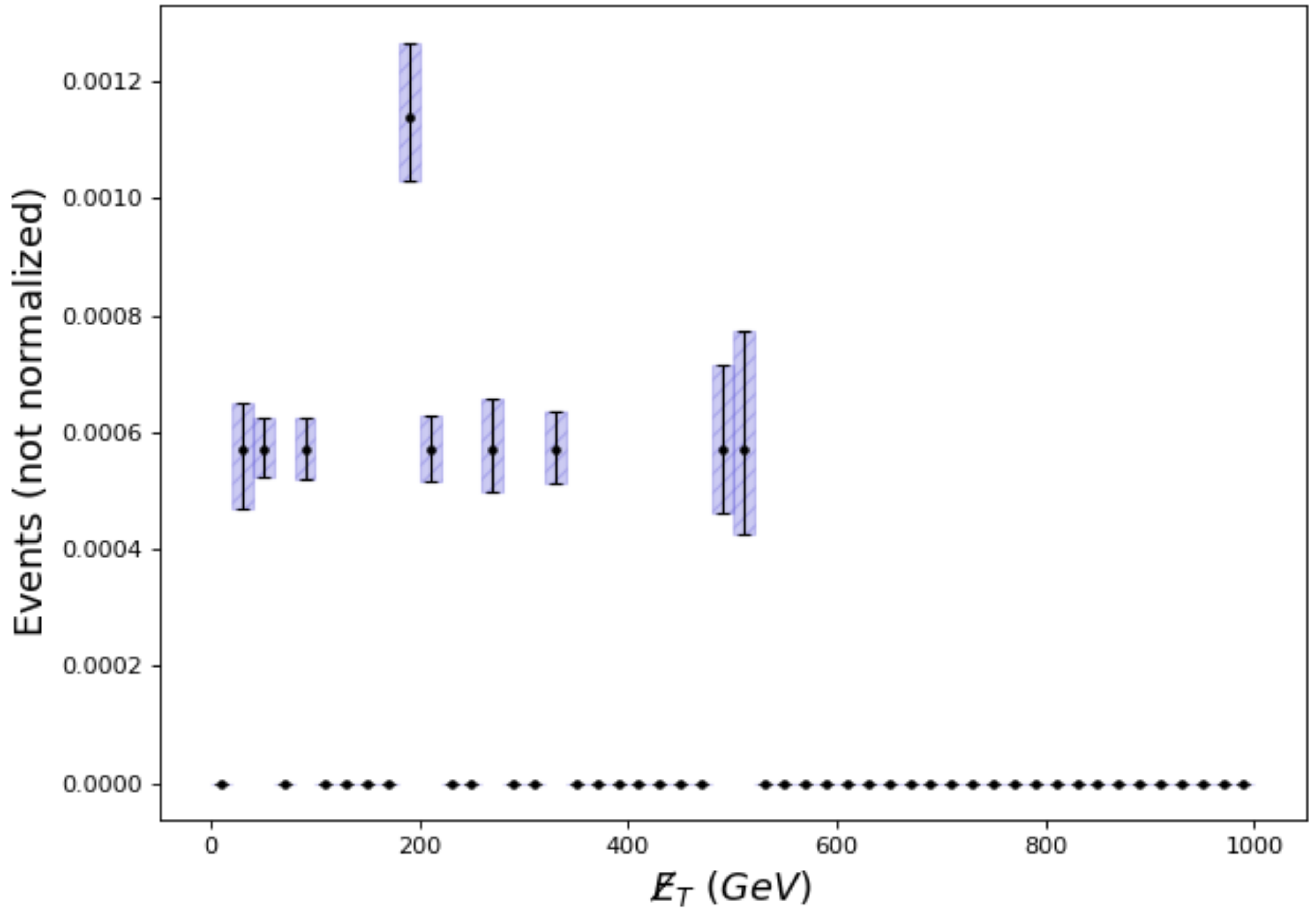


[JYA, Buckley & Fuks (EPJC'23)]

Uncertainties & Multi-weights

The long awaited upgrade: uncertainties

- Full treatment of multi-weighted events
- Weight propagation (cut-flows, histograms, *etc.*)
- Status for histograms
  - ➞ C++ data format updated (rewritten from scratch)
  - ➞ CLI rewritten
- Cut-flows: to be done
- Validation in progress
  - ➞ Scale uncertainties correctly handled in histograms
- Follow us on GitHub (the *multiweight* branch)



JYA, Arina, Fuks, Munoz, Panizzi & Tentori (in progress: 2025)

JYA, Fuks (in progress)

Machine Learning @Ma6

Machine learning in MADANALYSIS

- ONNX interface
  - ➙ NN export/import
  - ➙ ML-based ATLAS search recasting

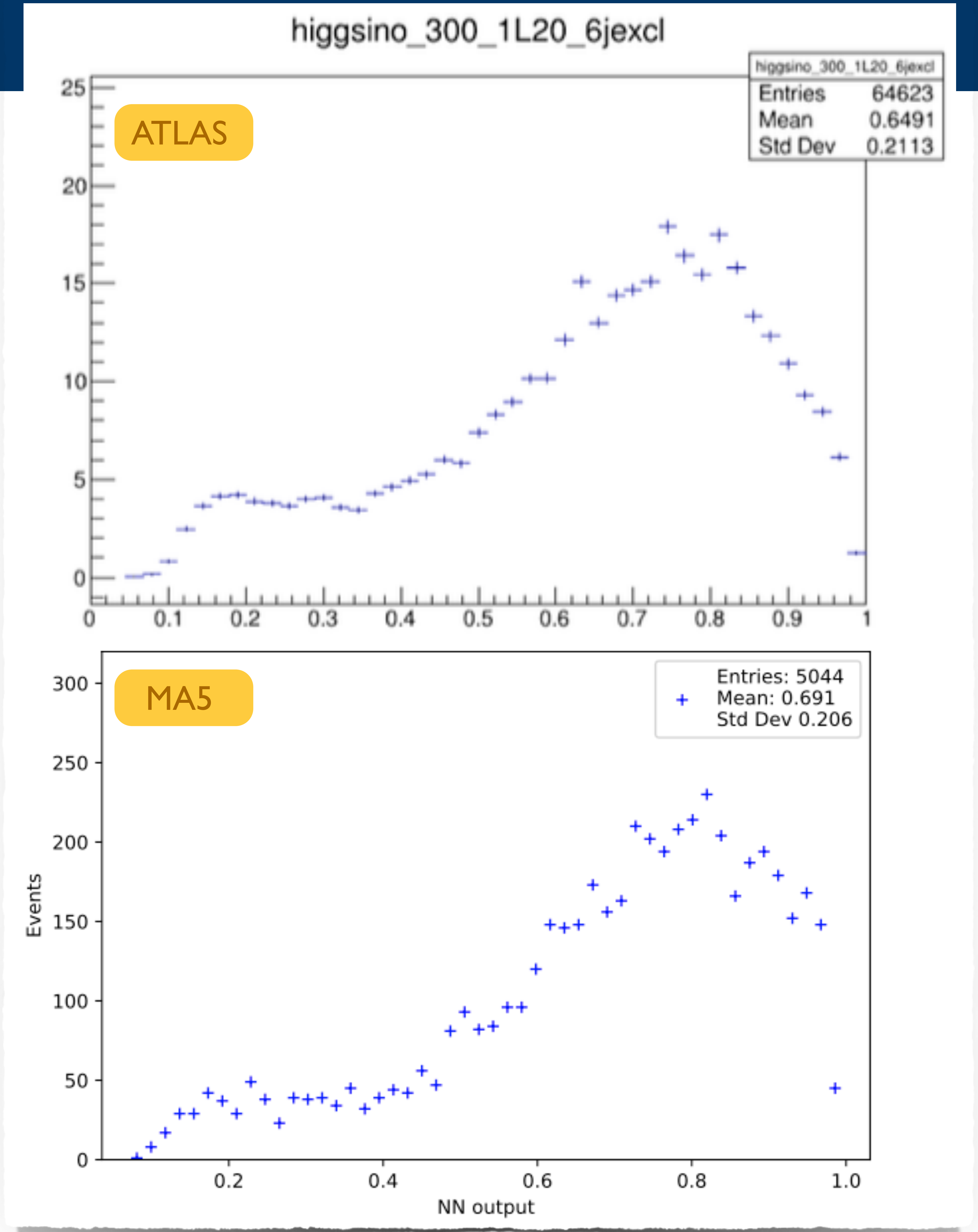
[Fuks, Goodsell, Laforge & Oudot (in progress)]

- Automatic S/B classifiers (BDTs, NNs)

[Cornell, Fuks, Goodsell & Ncube (in progress)]

- Improving the SFS with Lorenzetti Showers
  - ➙ NN-based efficiencies / tagging performance
  - ➙ Impact on LHC recasting

[Fuks, Goodsell & LPNHE group (in progress)]



Longer term future

Recasting developments

- Embedding SM measurements for LHC recasting (*à la* CONTUR, with different statistical treatment)
  - Full framework to constrain SMEFT
- Third MADANALYSIS 5 workshop in 2026 (CERN? Korea? China?)
  → Crucial for extending the database of recast analyses

Ma5 Workshop @ Korea 2020
(just before COVID)

An idea?
Feel free to reach out!

Jefferson Lab
EXPLORING THE NATURE OF MATTER

Stony Brook University

Thanks to the RiF community!

Andy Buckley

Tomasz Procter

Martin Habedank

Andre Lessa

HackAnalysis 2

Mark Goodsell

Joe Egan

Krzystof Rolbiecki

Matthew Feickert