Spey: A wee technical update

Jack Y. Araz

Monte-Carlo support tools
March 18, 2026

OpenMapp Collaboration

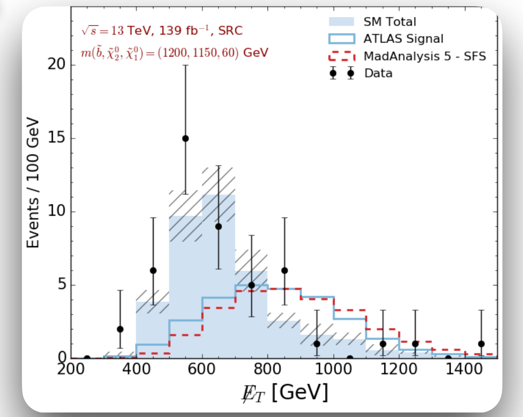
Outline

- What is Spey?
- What are we up to?
  - Signal uncertainties.
  - Spey-HS3 plug-in
  - Multi-POI Profiling and Confidence Intervals
- Future plans

What is Spey?

Backend agnostic inference

F A I R
FINDABLE ACCESSIBLE INTEROPERABLE REUSABLE



HS^3

Simplified likelihoods

pyhf
differentiable likelihoods

Machine Learned Models

CMS Combine

Backend agnostic inference

F A I R
FINDABLE ACCESSIBLE INTEROPERABLE REUSABLE

Backend agnostic inference

F A I R

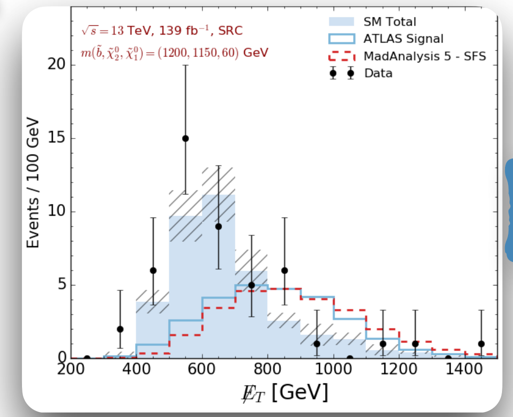
FINDABLE ACCESSIBLE INTEROPERABLE REUSABLE

rivet

MA5

SModelS

CHECKMATE

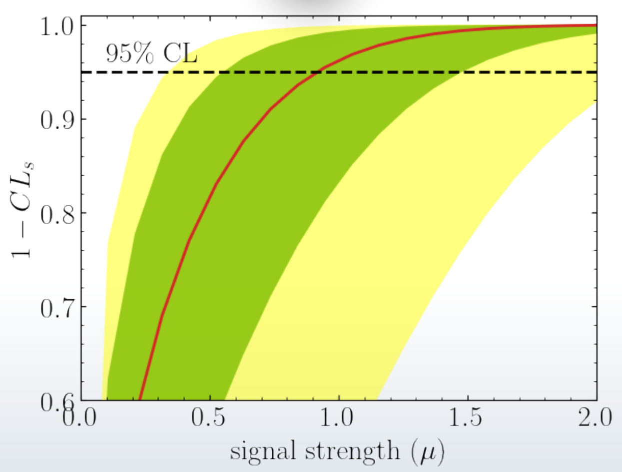


SPEY

Significance

Combine likelihoods

Multi-POI profiling



Exclusion limit

HS^3

Simplified likelihoods

pyhf

differentiable likelihoods

Machine Learned Models

CMS Combine

UCL

CITY ST GEORGE'S UNIVERSITY OF LONDON

Adoption by the community

arXiv > hep-ph > arXiv:2507.08565

High Energy Physics – Phenomenology

[Submitted on 11 Jul 2025 (v1), last revised 5 Aug 2025 (this version, v2)]

Implementation of full and simplified likelihoods in CheckMATE

Iñaki Lara, Krzysztof Rolbiecki

We present the implementation of simplified and full likelihood models for multibin signal regions in CheckMATE. A total of 13 searches are included from ATLAS and CMS, and several methods are presented for the implementation and evaluation of likelihood functions. Statistical combinations increase the sensitivity of searches and open up the possibility of combining orthogonal search channels in the CheckMATE framework.

arXiv > hep-ph > arXiv:2505.09272

High Energy Physics – Phenomenology

[Submitted on 14 May 2025]

Constraints On New Theories Using Rivet : CONTUR version 3 release note

Andy Buckley, Jon Butterworth, Joseph Egan, Christian Gutschow, Sihyun Jeon, Martin Habedank, Tomasz Procter, Peng Wang, Yoran Yeh, Luzhan Yue

The CONTUR toolkit exploits RIVET and its library of more than a thousand energy-frontier differential cross-section measurements from the Large Hadron Collider to allow rapid limit-setting and consistency checks for new physics models. In this note we summarise the main changes in the new CONTUR 3 major release series. These include additional statistical treatments, efficiency improvements, new plotting utilities and many new measurements and Standard Model predictions.

arXiv > hep-ph > arXiv:2507.08927

High Energy Physics – Phenomenology

[Submitted on 11 Jul 2025]

Deciphering compressed electroweakino excesses with MadAnalysis 5

Jack Y. Araz, Benjamin Fuks, Mark D. Goodsell, Taylor Murphy

We present version 1.11 of MadAnalysis 5, which extends the software package in several major ways to improve the handling of efficiency tables, the computation of observables in different reference frames and the calculation of statistical limits and/or significance. We detail how these improvements, whose development was motivated by the desire to implement two Run 2 LHC analyses targeting signatures with soft leptons and missing energy and exhibiting mild excesses (ATLAS-SUSY-2018-16 and ATLAS-SUSY-2019-09), have been implemented by both direct extensions of the code and integrations with third-party software. We then document the implementation and validation of these analyses, demonstrating their utility along with the improved statistics capabilities of MadAnalysis 5 through an investigation of the Next-to-Minimal Supersymmetric Standard Model in the context of a larger set of overlapping excesses in channels with soft leptons/jets and missing transverse energy.

Adoption by the community


arXiv > hep-ex > arXiv:2503.19836

High Energy Physics – Experiment

[Submitted on 25 Mar 2025 (v1), last revised 20 Oct 2025 (this version, v2)]

A measurement of the high-mass $\tau\bar{\tau}$ production cross-section at $\sqrt{s}$ = 13 TeV with the ATLAS detector and constraints on new particles and couplings

The ATLAS Collaboration

The production cross-section of high-mass τ-lepton pairs is measured as a function of the dilepton visible invariant mass, using 140 fb^{-1} of $\sqrt{s}$ = 13 TeV proton–proton collision data recorded with the ATLAS detector at the Large Hadron Collider. The measurement agrees with the predictions of the Stand... invariant mass distribution is performed as a function of b-jet multiplicity, to constrain the non-resonant production of new particles described by an effective field theory or in models containing leptoquarks or Z'... third-generation fermions. The constraints on new particles improve on previous results, and the constraints on effective operators include those affecting the anomalous magnetic moment of the τ...

arXiv > hep-ex > arXiv:2602.18611

High Energy Physics – Experiment

[Submitted on 20 Feb 2026]

Combined measurements and interpretations of Higgs boson production and decay in proton–proton collisions at $\sqrt{s}$ = 13 TeV

CMS Collaboration

Combined measurements of Higgs boson production and decay rates are reported, representing the most comprehensive study performed by the CMS Collaboration to date. The included analyses use proton–proton collision data recorded by the CMS experiment at $\sqrt{s}$ = 13 TeV from 2016 to 2018, corresponding to an integrated luminosity of 138 fb^{-1}. The statistical combination is based on analyses that measure the following decay channels: H → γγ, H → ZZ, H → WW, H → ττ, H → bb, H → μμ, and H → Zγ → ℓℓγ (ℓ = e,μ). Information in the events from each decay channel is used to target multiple Higgs boson production processes. Searches for invisible Higgs boson decays are also considered, as well as an analysis that measures off-shell Higgs boson production in the H → ZZ → 4ℓ decay channel. The best fit inclusive signal yield is measured to be $1.014^{+0.055}_{-0.053}$ times the standard model expectation, for a Higgs boson mass of 125.38 GeV. Measurements in kinematic regions defined by the simplified template cross section framework are also provided, as well as interpretations in the coupling modifier and standard model effective field theory frameworks. The coupling modifier interpretation is further used to place constraints on various two-Higgs-doublet models. The results show good compatibility with the standard model predictions for the majority of the measured parameters.

arXiv > hep-ph > arXiv:2507.08565

High Energy Physics – Phenomenology

[Submitted on 11 Jul 2025 (v1), last revised 5 Aug 2025 (this version, v2)]

Implementation of full and simplified likelihoods in CheckMA...

Iñaki Lara, Krzysztof Rolbiecki

We present the implementation of simplified and full likelihood models for ... evaluation of likelihood functions. Statistical combinations increase ...

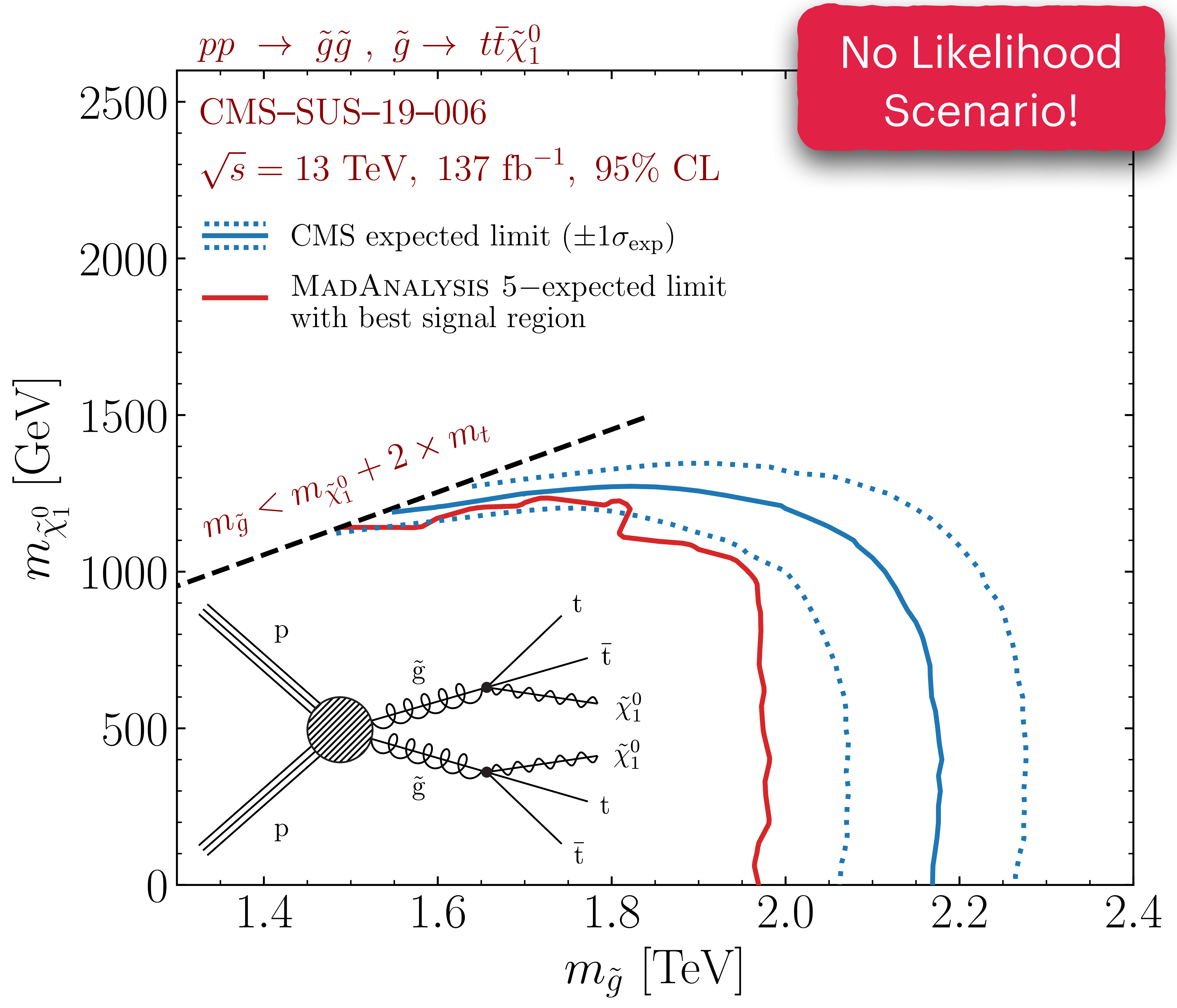
Deciph... electroweakino excesses with MadAna...

Jack Y. Araz, ... Mark D. Goodsell, Taylor Murphy

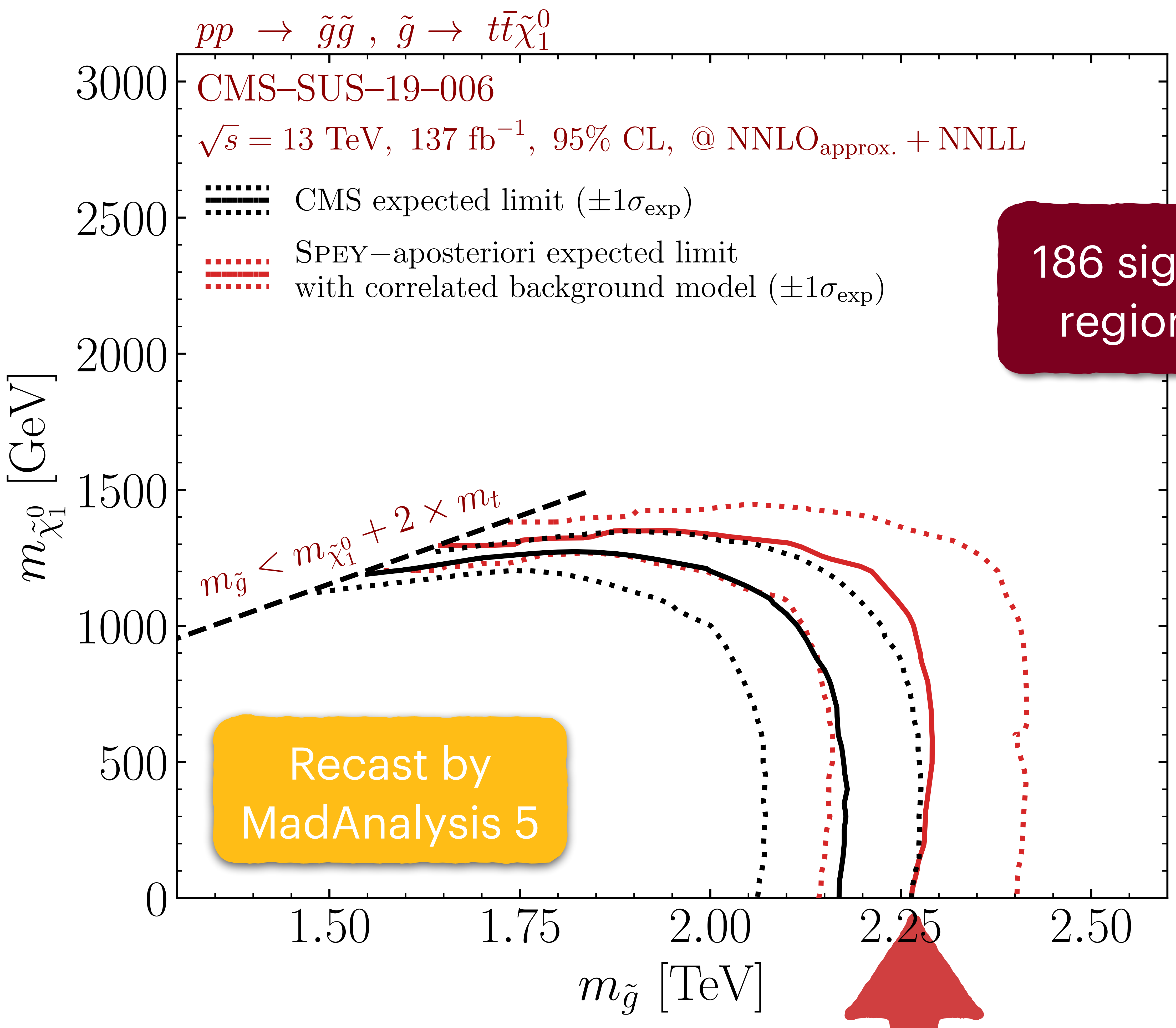
We present version 1.11 of MadAnalysis 5, which extends the software package in several major ways to improve ... statistical limits and/or significance. We detail how these improvements, whose development was motivated by the ... excesses (ATLAS-SUSY-2018-16 and ATLAS-SUSY-2019-09), have been implemented by both direct extensions of ... analyses, demonstrating their utility along with the improved statistics capabilities of MadAnalysis 5 through an inv... channels with soft leptons/jets and missing transverse energy.

What is the impact?

[JYA; SciPost '24]



No Likelihood Scenario!

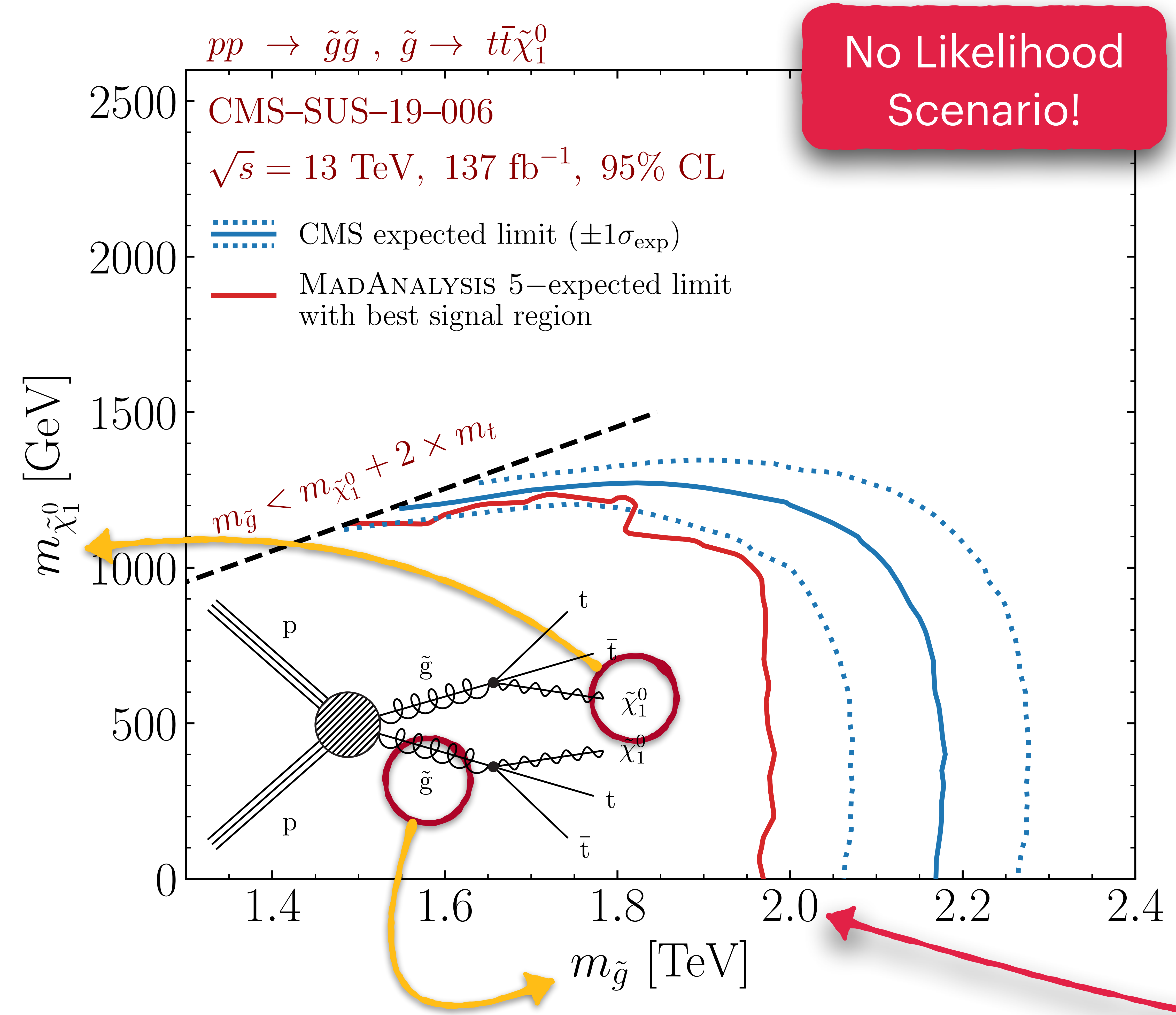


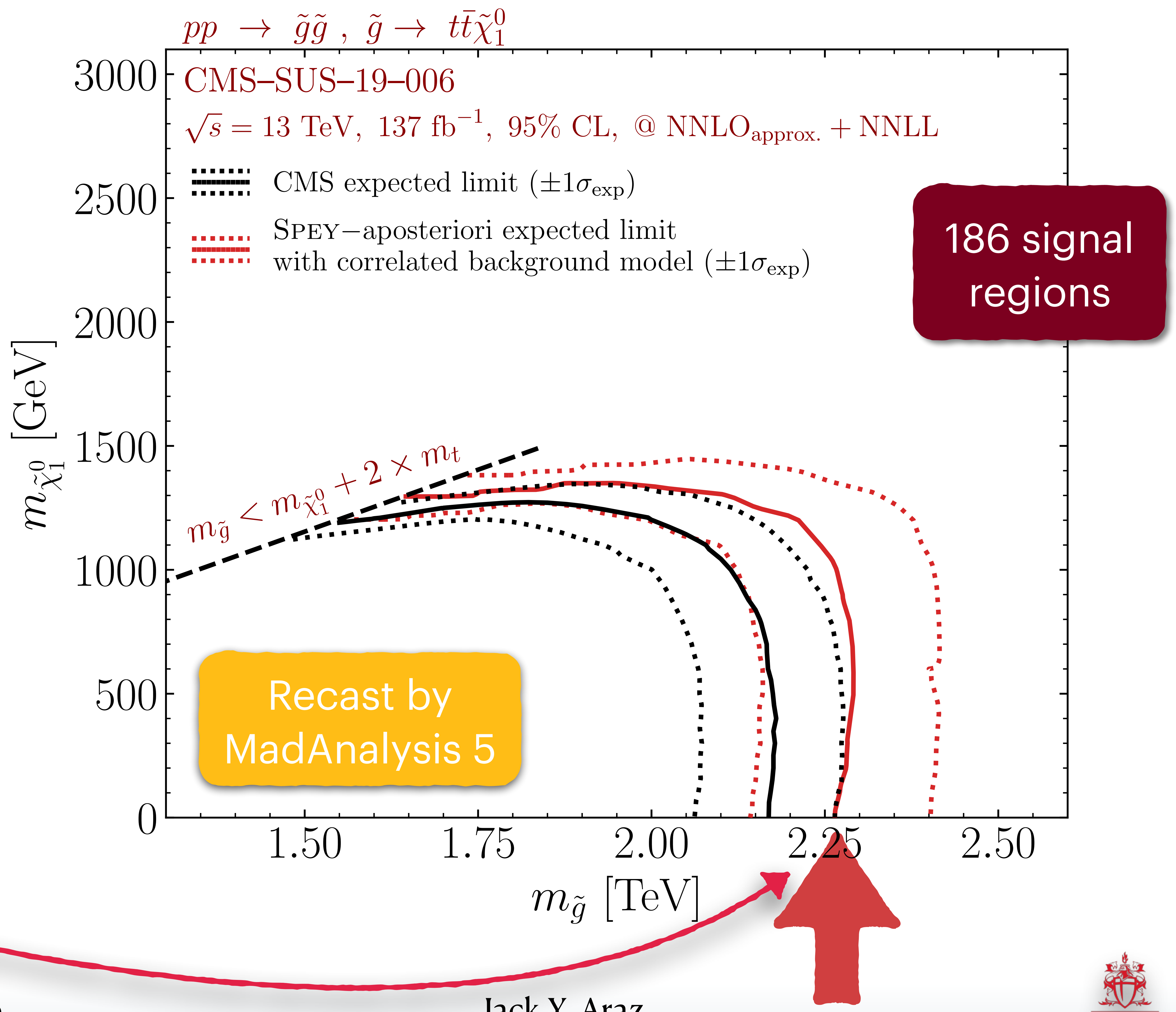
186 signal regions

Recast by MadAnalysis 5

What is the impact?

[JYA; SciPost '24]

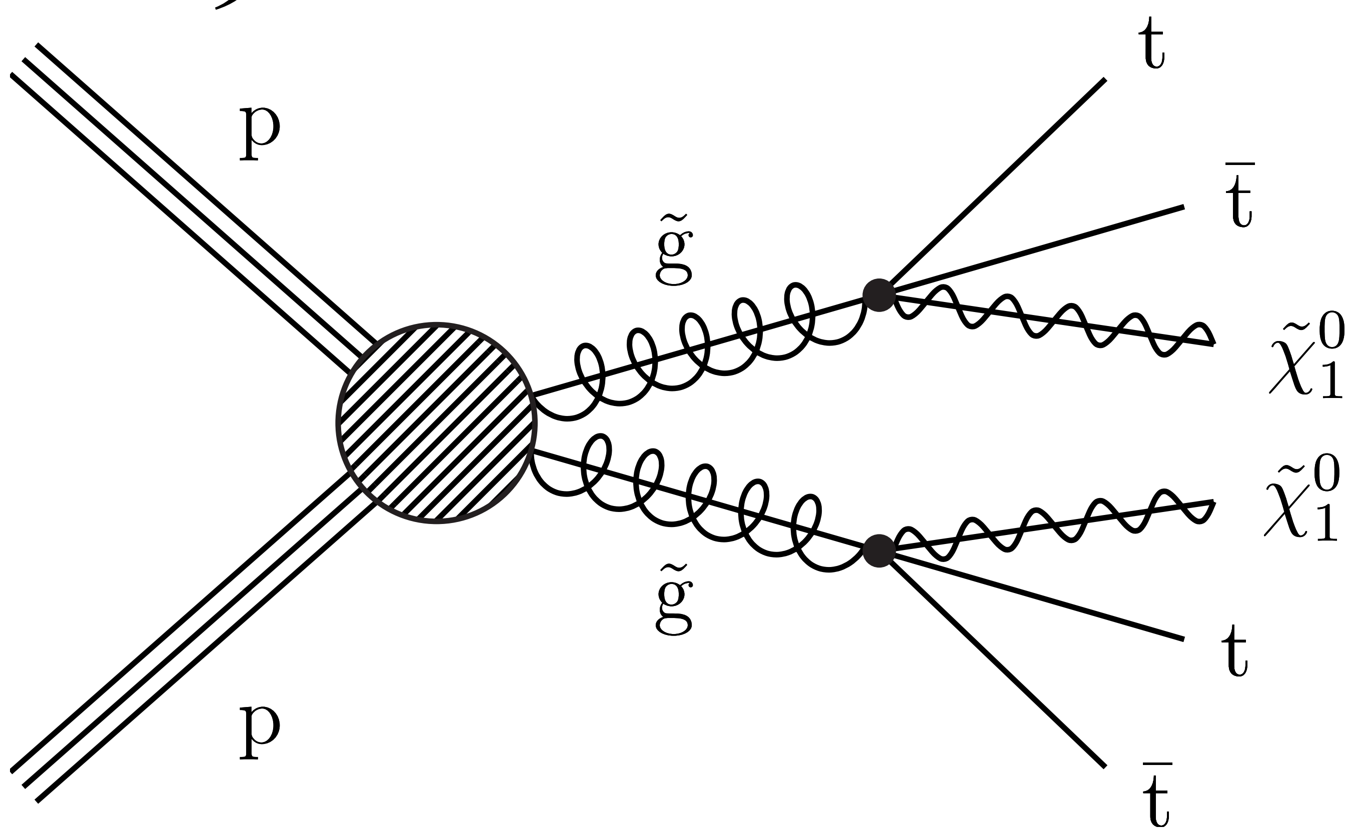




Third moment expansion

[JYA; SciPost '24]

CMS-SUS-19-006

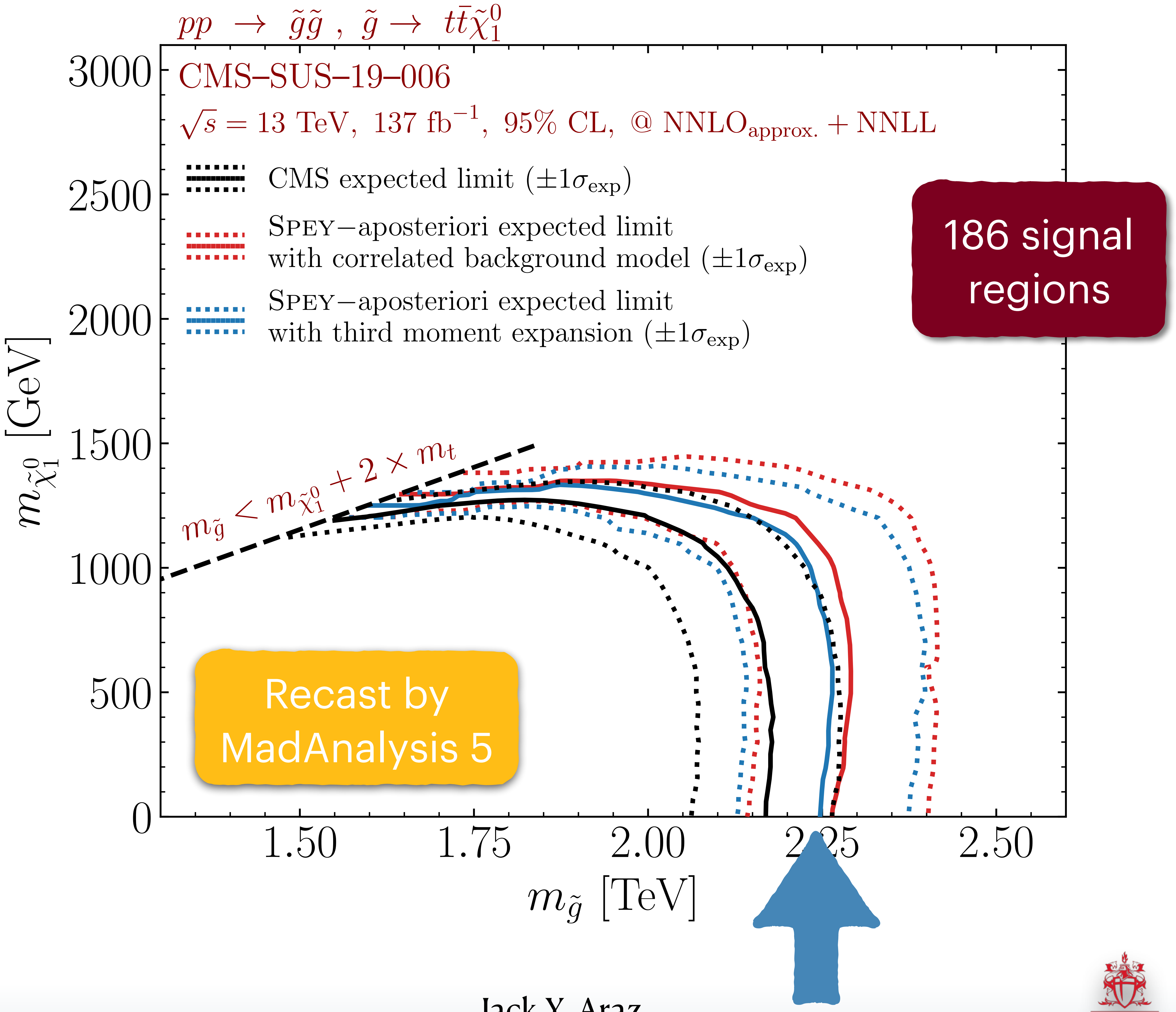


$$\mathscr{L}(\mu,\theta) = \left[\prod_{i\in\text{bins}} \text{Poiss}\left(n^i \,|\, \mu n_s^i + \bar{n}_b^i + A_i\theta_i + C_i\theta_i^2\right)\right] \cdot \mathcal{N}\left(\theta \,|\, 0, \bar{\rho}\right)$$

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

A_i := the effective sigma of the background uncertainty

C_i := asymmetry of the background uncertainty

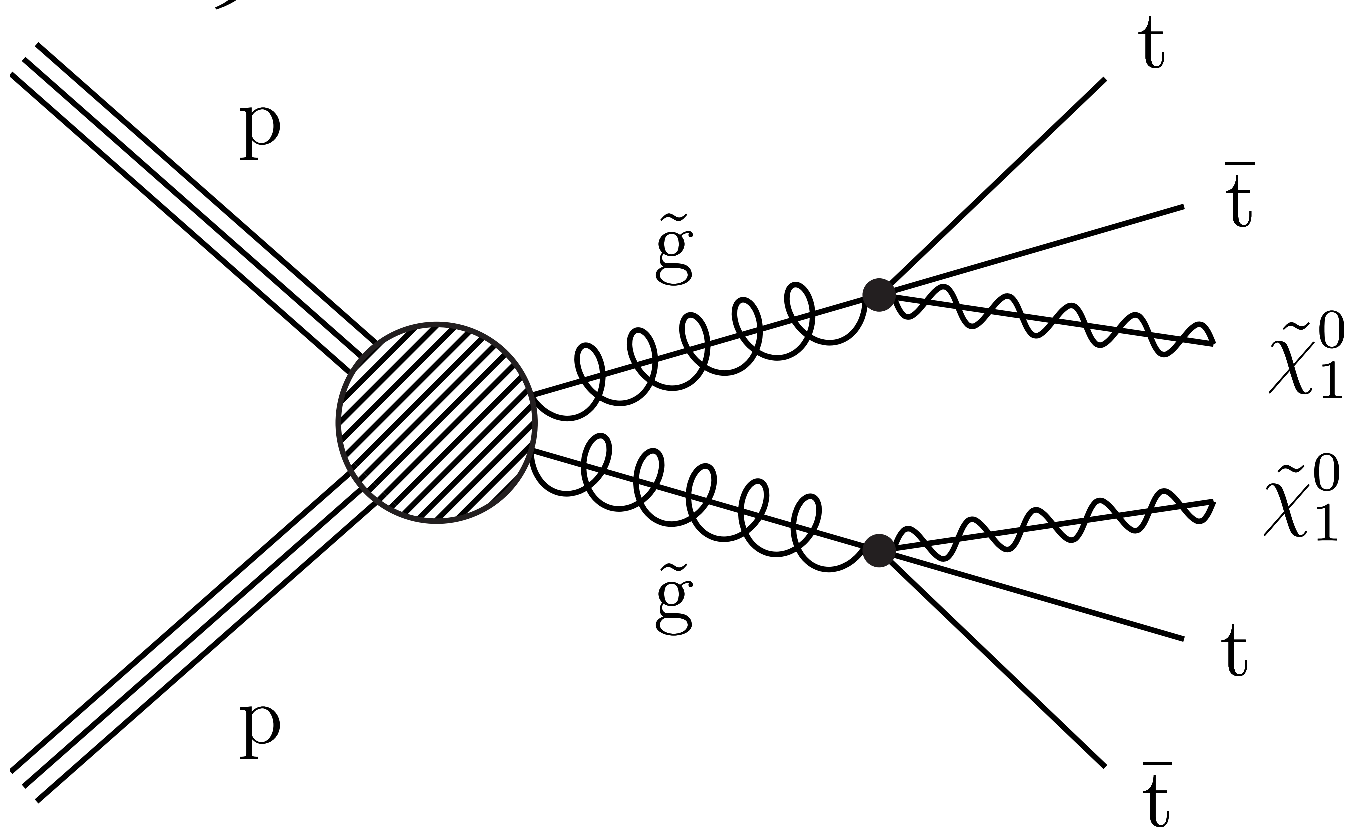


Buckley, Citron, Fichet, Kraml, Waltenberger, Wardle; JHEP '18

Asymmetric Uncertainties

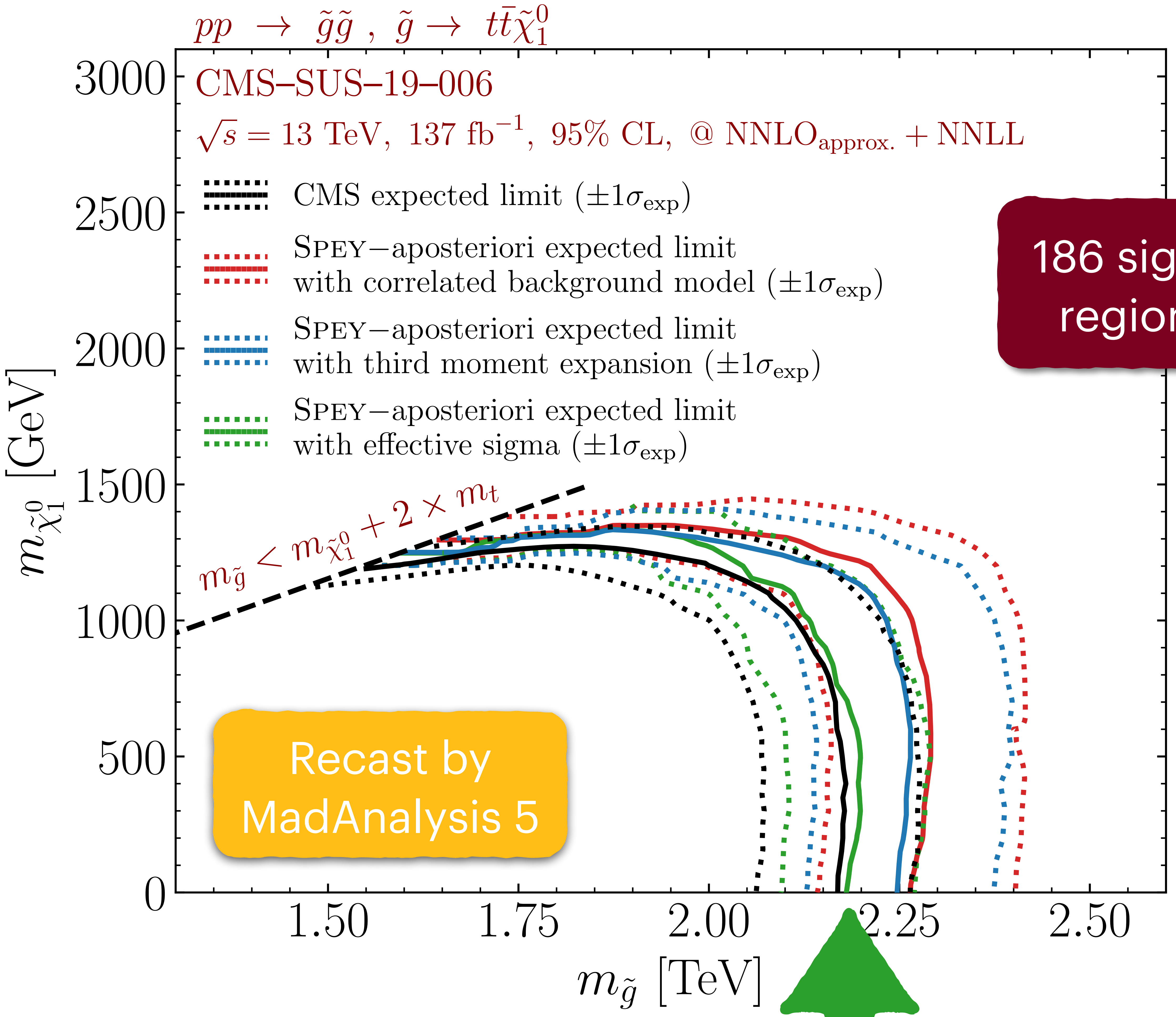
[JYA; SciPost '24]

CMS-SUS-19-006



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

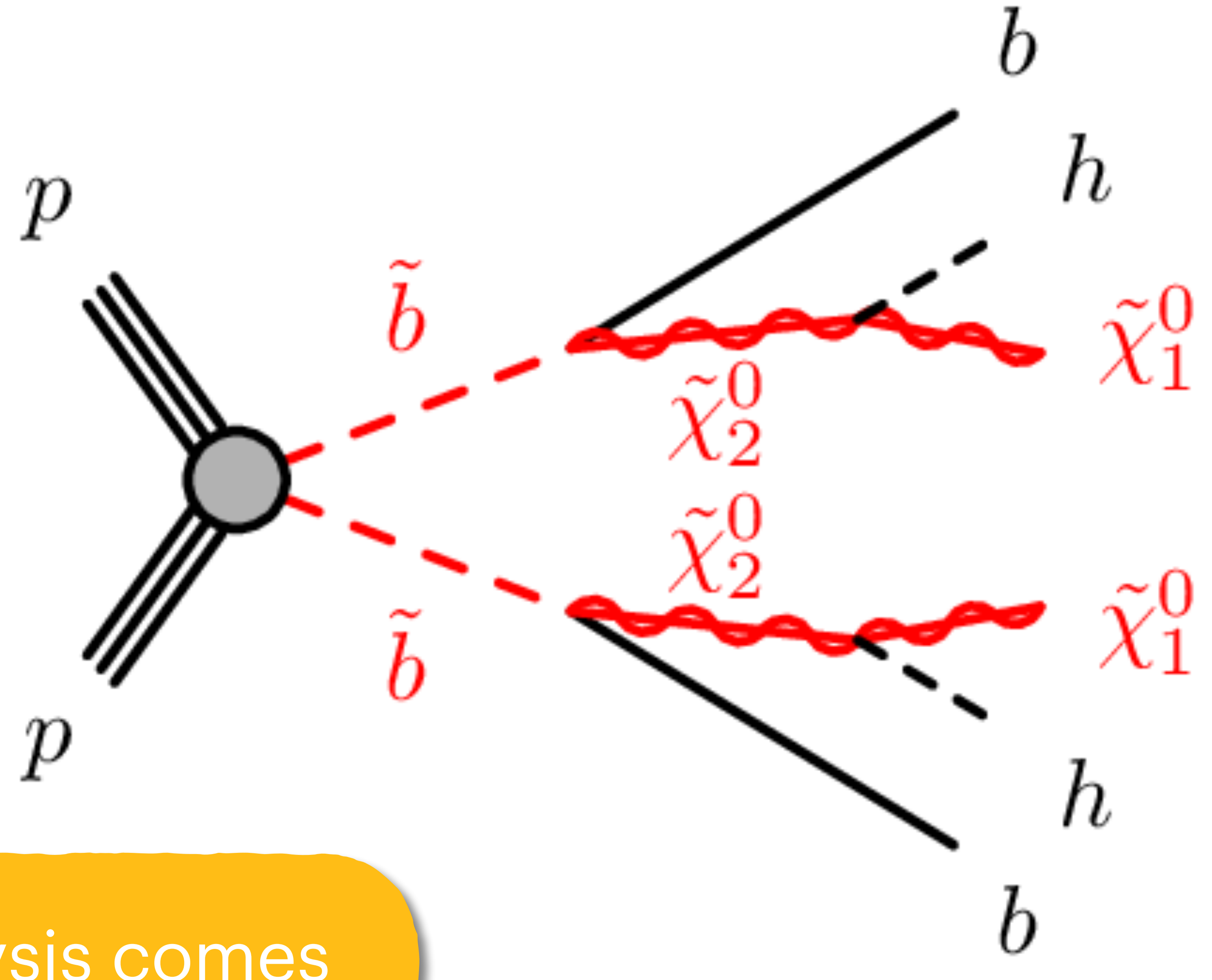
$$\sigma_{\text{eff}}^i(\theta^i) = \sqrt{\sigma_i^+\sigma_i^- + (\sigma_i^+ - \sigma_i^-)(\theta^i - n_b^i)}$$



186 signal regions

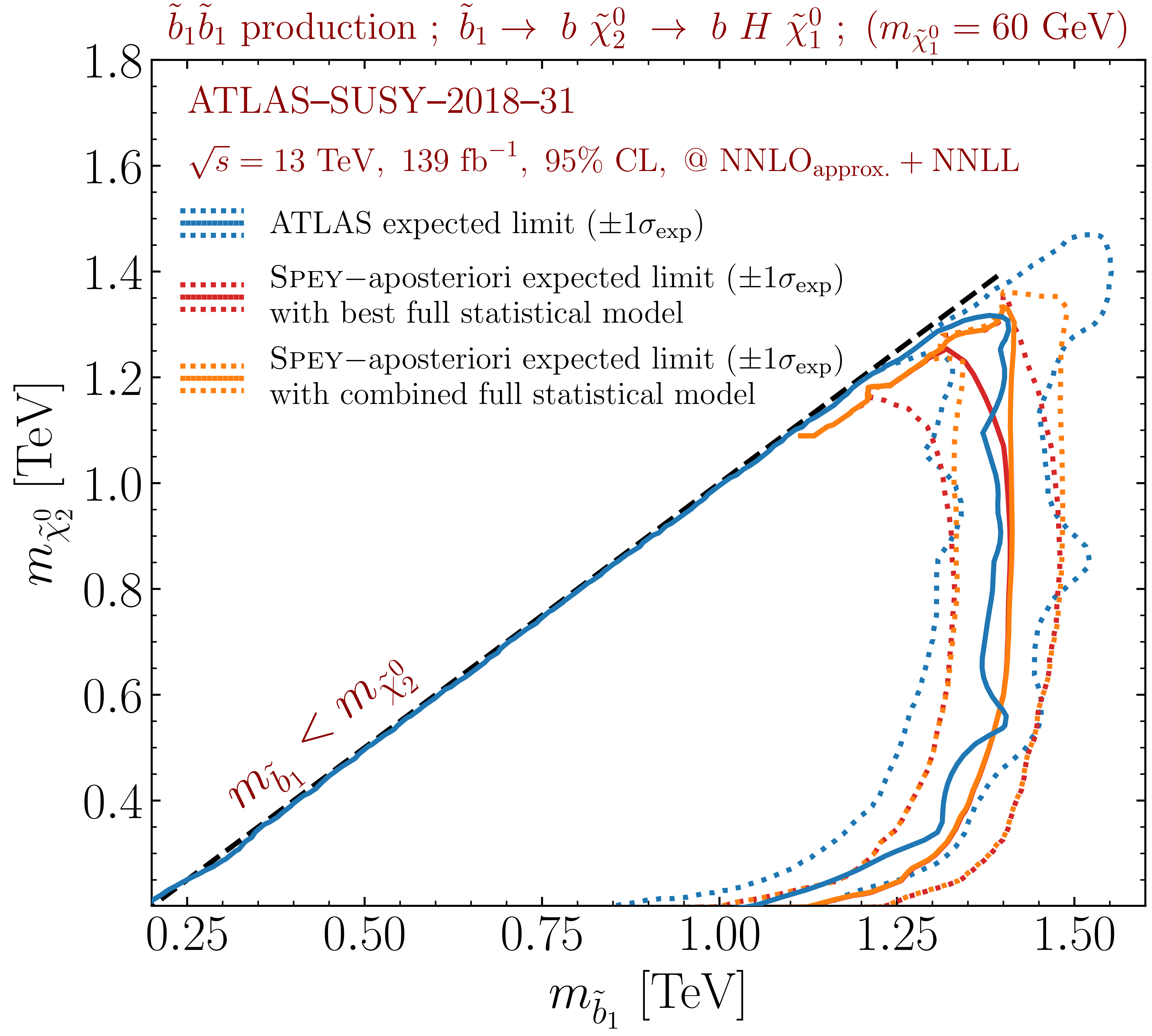
Full likelihoods

[JYA; SciPost '24]



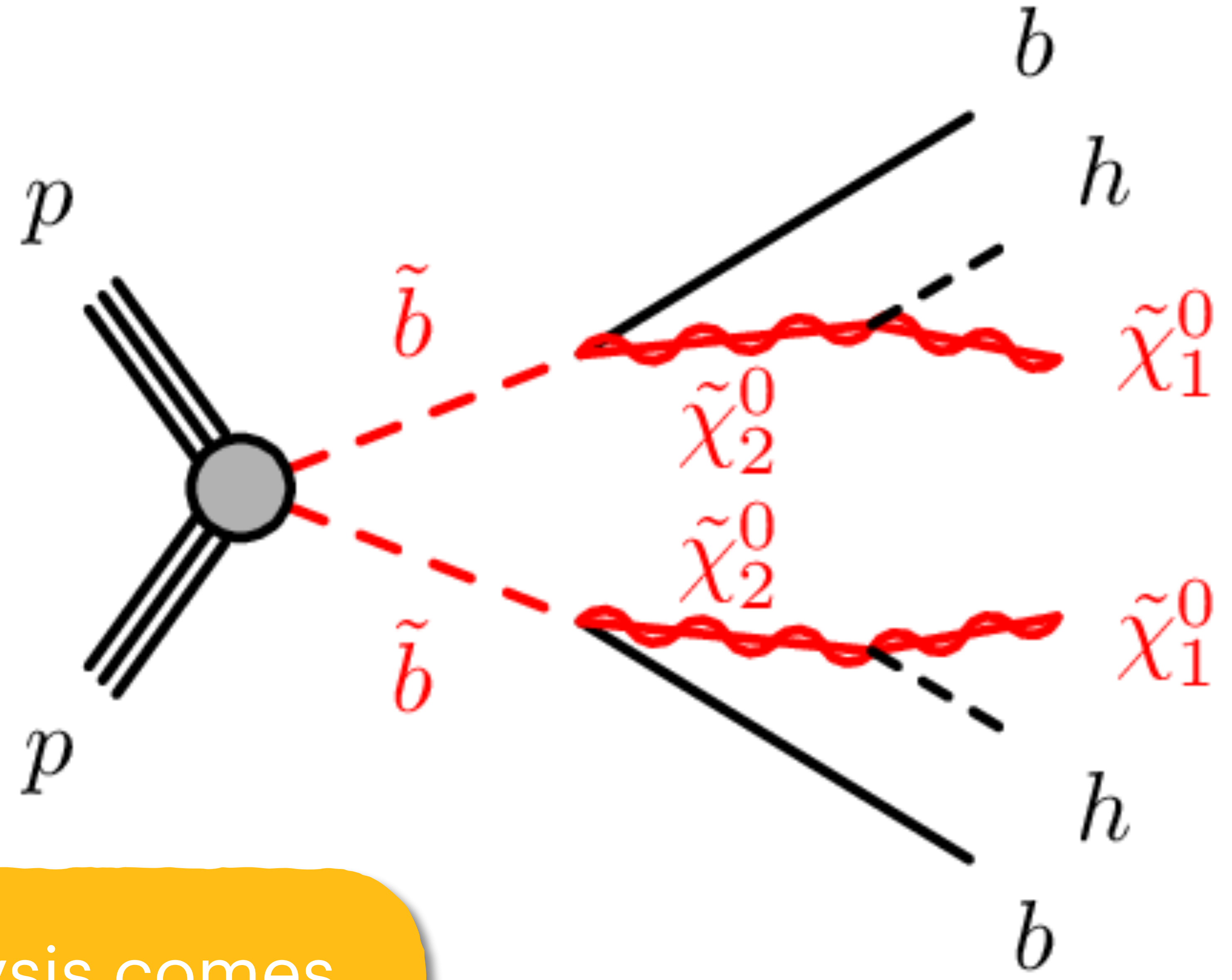
This analysis comes with three different super regions!

Full likelihoods include all the necessary information to mix and match nuisance parameters to combine them!



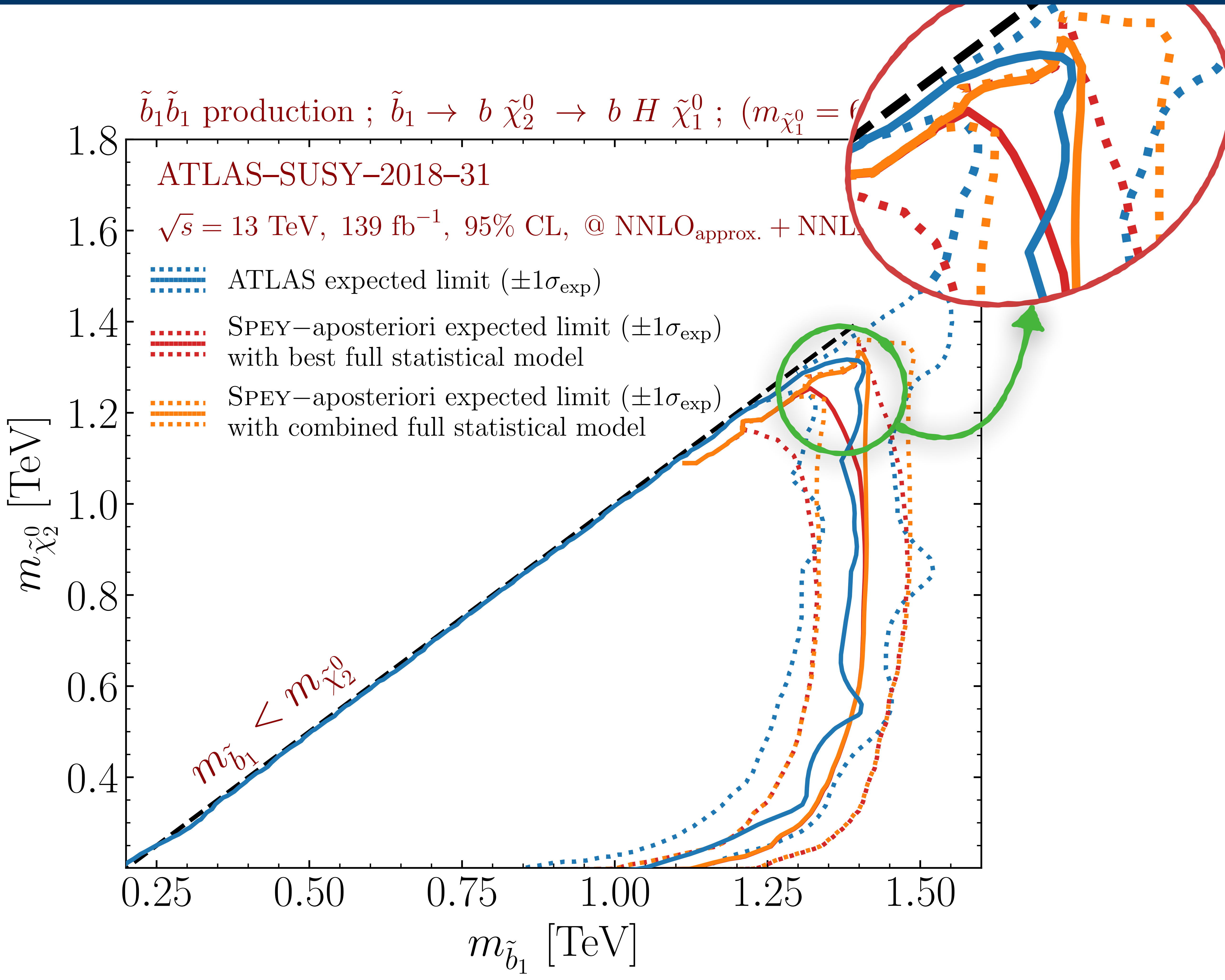
Full likelihoods

[JYA; SciPost '24]



This analysis comes with three different super regions!

Full likelihoods include all the necessary information to mix and match nuisance parameters to combine them!



Towards global sensitivity

[JYA; SciPost '24]

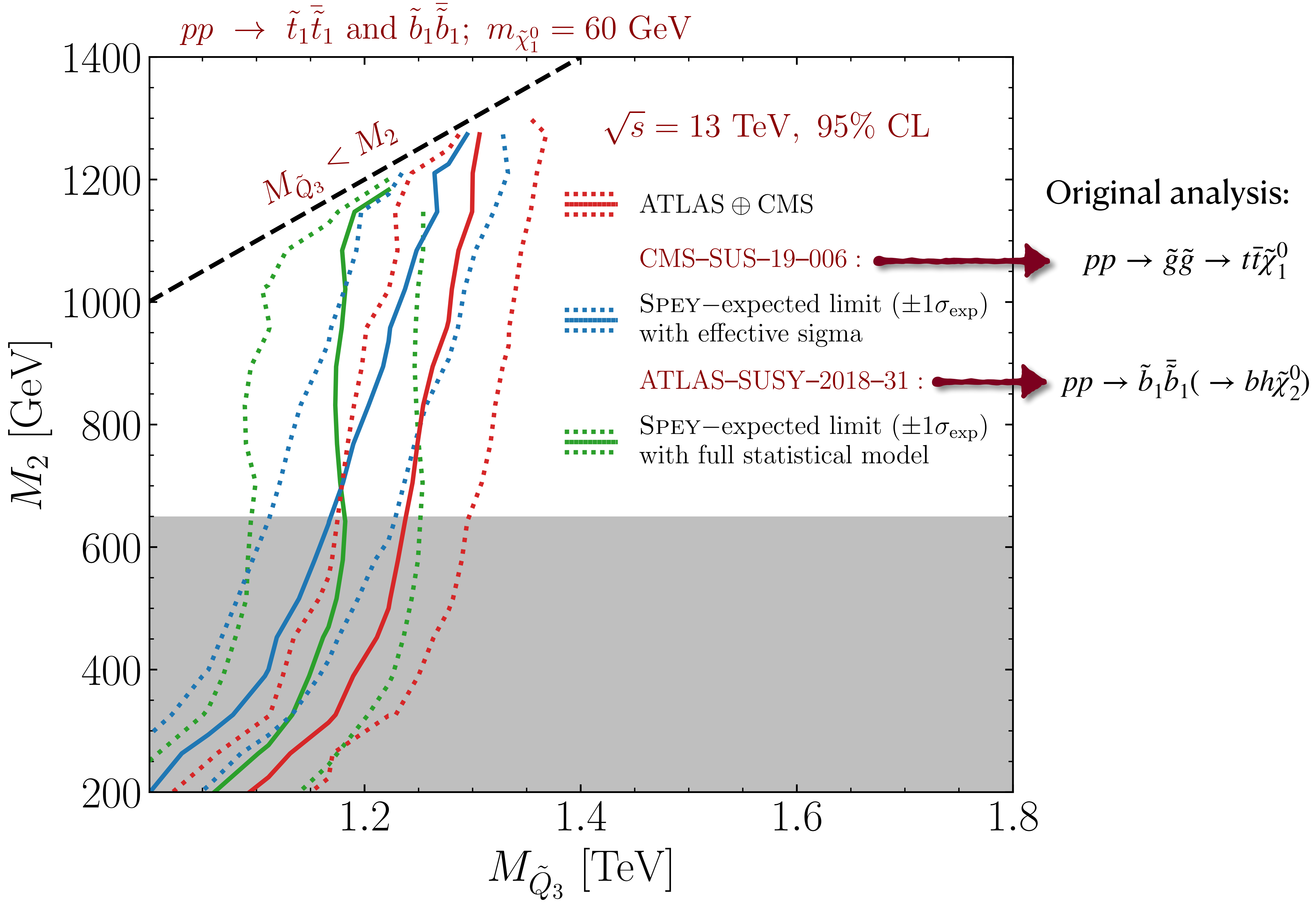
$$\mathscr{L}' = \mathscr{L}_{\text{ATLAS}} \oplus \mathscr{L}_{\text{CMS}}$$

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



What are we up to?

Theoretical Uncertainties

Current implementation

$$\mathcal{L}(\mu,\theta) = \prod_{i\in\text{bins}} \text{Poiss}(n_i \,|\, \mu n_i^{(s)} + \theta_i^{(s)}\sigma_i^{(s)} + n_i^{(b)} + \theta_i^{(b)}\sigma_i^{(b)}) \cdot \prod_{j\in\text{nui}} \mathcal{N}(\theta_j^{(b)} \,|\, 0,\rho) \cdot \prod_{j\in\text{nui}} \mathcal{N}(\theta_j^{(s)} \,|\, 0,1)$$

➡Does not work for theoretical uncertainties, leads to under exclusion

Theoretical Uncertainties

Current implementation

$$\mathscr{L}(\mu,\theta) = \prod_{i\in\text{bins}} \text{Poiss}(n_i \,|\, \mu n_i^{(s)} + \theta_i^{(s)}\sigma_i^{(s)} + n_i^{(b)} + \theta_i^{(b)}\sigma_i^{(b)}) \cdot \prod_{j\in\text{nui}} \mathcal{N}(\theta_j^{(b)}\,|\,0,\rho) \cdot \prod_{j\in\text{nui}} \mathcal{N}(\theta_j^{(s)}\,|\,0,1)$$

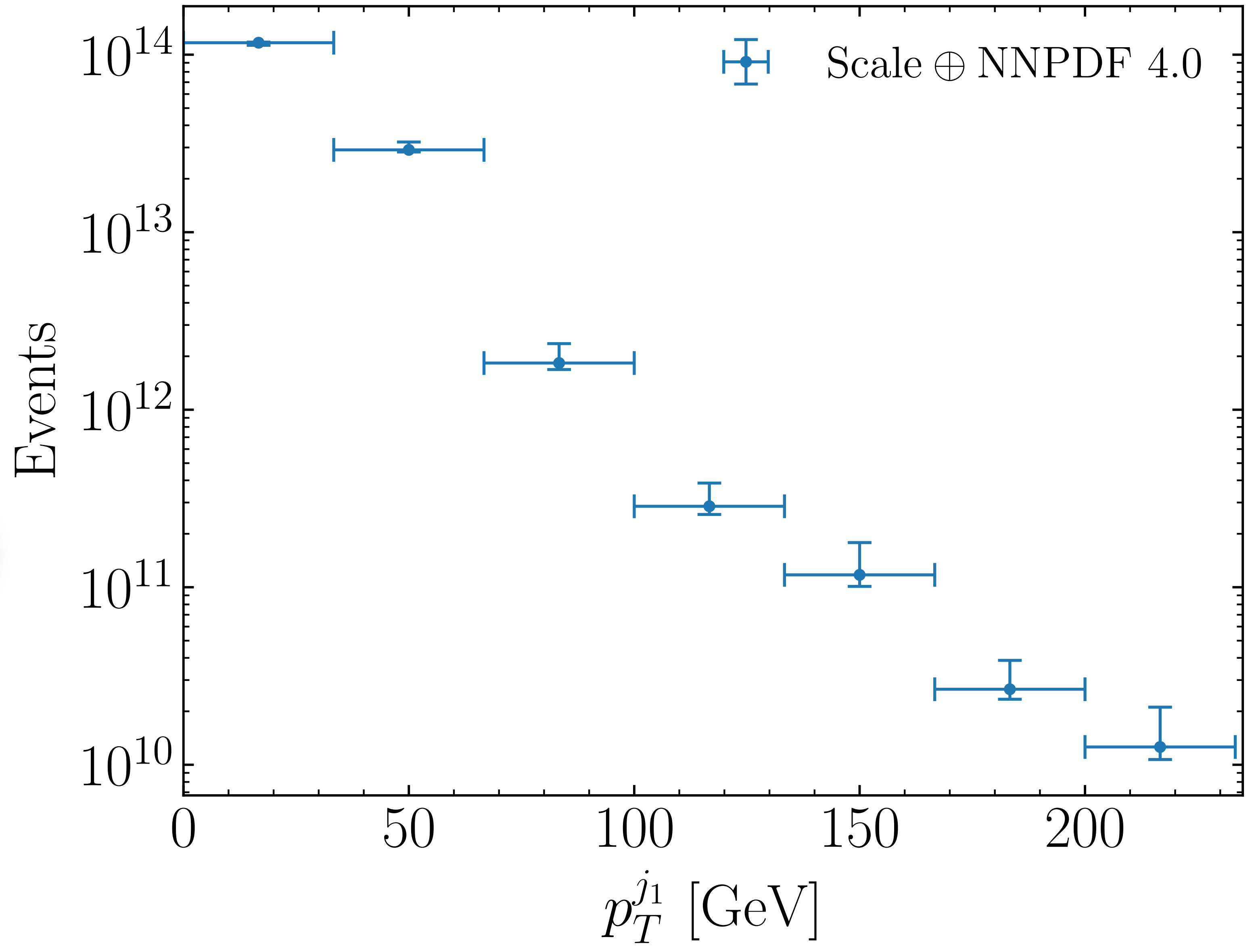
➡Does not work for theoretical uncertainties, leads to under exclusion

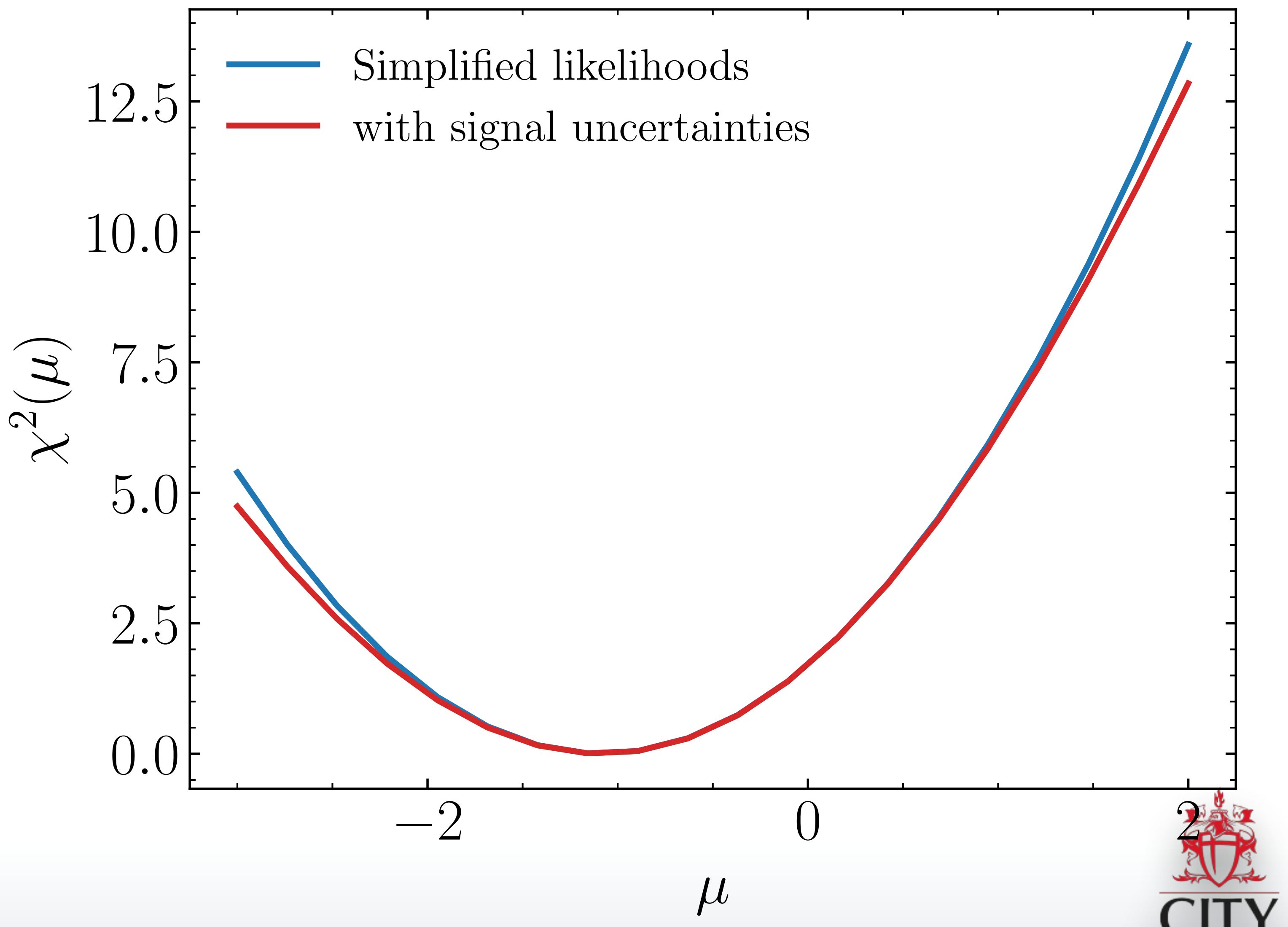
$$\mathscr{L}(\mu,\theta) = \prod_{i\in\text{bins}} \text{Poiss}\left(n \,\middle|\, \mu n_s^{(i)} \prod_j f_s^j(\theta_s^{j,i}, \sigma_s^{(i)}) + n_b^i + \Theta\sigma_b\right) \cdot \mathcal{N}(\Theta\,|\,0,\rho) \cdot \prod_j \mathcal{N}(\theta_s^j\,|\,0,1)$$

$$f_{i,k}(\theta_{i,k}) = \begin{cases} \theta_{i,k} \log\left(1+\Delta_{i,k}^{+}\right), & \text{if } \theta_k \geq 0 \\ \theta_{i,k} \log\left(1+\Delta_{i,k}^{-}\right), & \text{otherwise} \end{cases}$$

✦ One nuisance parameter per source of uncertainty e.g. PDF, scale etc.

Coming with v0.2.7, available in spey/uncertainties branch





HEP Statistics Serialisation Standard (HS³)

With Carsten Bugard, Simon Cello & Giordon Holtsberg

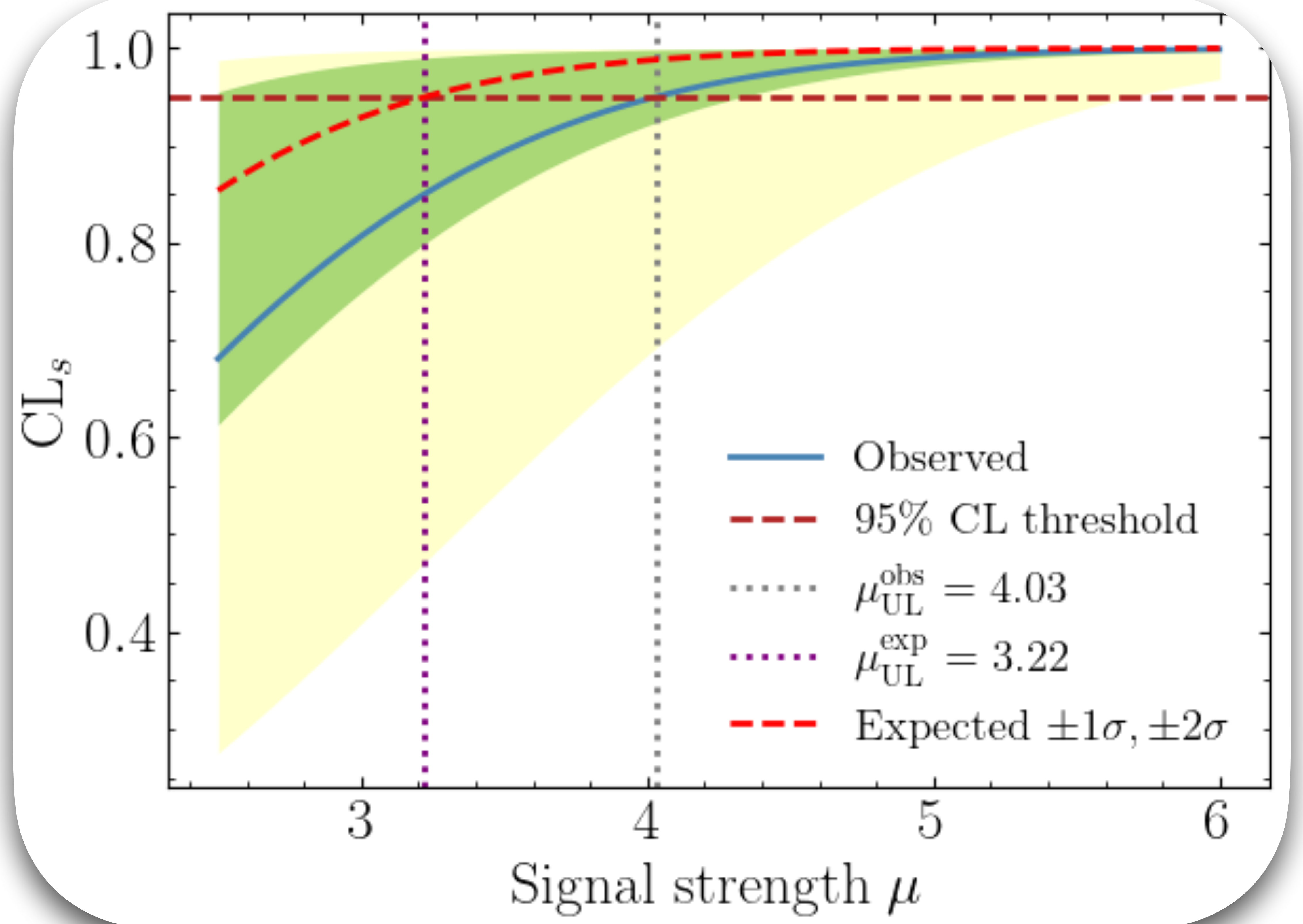
Example: Distributions

```
"distributions":[
    {
        "name":"gauss1",
        "type":"gaussian_dist",
        "mean":1.0,
        "sigma":"param_sigma",
        "x":"param_x"
    },
    {
        "name":"exp1",
        "type":"exponential_dist",
        "c":-2,
        "x":"data_x"
    },
    ...
]
```

pip install spey-hs3

```
HS3 = spey.get_backend("hs3")

simple_model = HS3(
    hs3_dict=simple_hs3,
    signal_yields=simple_signal,
    mode="FAST_COMPILE",
)
```



HEP Statistics Serialisation Standard (HS³)

With Carsten Bugard, Simon Cello & Giordon Holtsberg

```python
from spey_hs3 import WorkspaceInterpreter

interp = WorkspaceInterpreter(hs3_dict)
interp.summary()
```

```
=============================================
HS3 Workspace Summary
  hs3_version : 0.2
  analyses           : 1
  likelihoods        : 1
  histfactory dists  : 1
  data entries       : 1

Analysis : demo
  POIs       : ['mu']
  Likelihood: L
  Distributions (1):
     1.      channel  (2 bins)  obs=2
=============================================
```

Property	Description
`interp.distributions`	Names of all `histfactory_dist` distributions
`interp.analyses`	Names of all analyses
`interp.poi_names`	POIs per analysis: `{analysis: [poi, ...]}`
`interp.samples`	Sample names per distribution
`interp.modifier_types`	Modifier types per sample per distribution
`interp.bin_map`	Number of bins per distribution
`interp.expected_background_yields`	Total background per bin per distribution
`interp.observed_data`	Observed data per distribution
`interp.parameters`	Parameter metadata from the nominal parameter point

HS3 Translator

```python
# Single distribution
interp.inject_signal("channel", "signal", [5.0, 5.0])

# Multiple distributions at once
interp.inject_signals({
    "model_SR_0j": {"Signal": [3.0, 5.0, 2.0]},
    "model_SR_1j": {"Signal": {"contents": [1.0, 2.0], "errors": [0.1, 0.2]}},
})

# Retrieve patched workspace and pass to HS3Interface
stat_model = spey.get_backend("hs3")(hs3_dict=interp.patch)
```

Multi-POI profiling & confidence intervals

Functional Signal

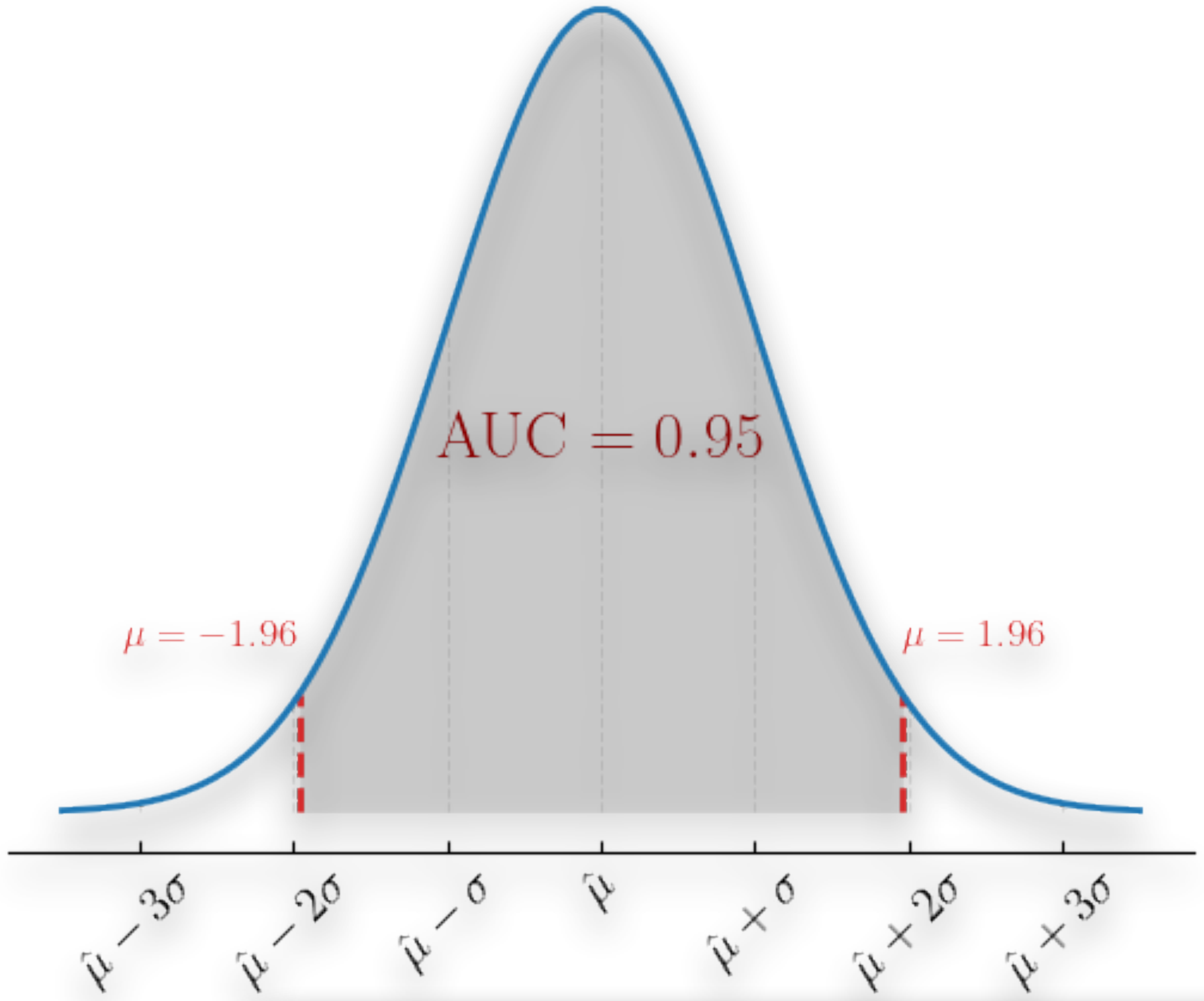
```python
def signal_yields_fn(pars: np.ndarray) -> np.ndarray:
    """EFT signal in all 7 M(WZ) bins as a function of (c_W, c_WWW)."""
    c_W_, c_WWW_ = pars[0], pars[1]
    return A_W * c_W_ + B_W * c_W_**2 + A_WWW * c_WWW_ + B_WWW * c_WWW_**2


# —— Construct the model ——————————————————————————————
pdf_wrapper = spey.get_backend("default.normal")

stat_model = pdf_wrapper(
    signal_yields=signal_yields_fn,  # callable (c_W, c_WWW) → shape(7,)
    background_yields=sm_yields,
    data=observed,
    absolute_uncertainties=sigma_eff,
    n_signal_parameters=2,  # declares c_W and c_WWW
    # signal_parameter_bounds=[(-5.0, 3.0), (-3.0, 3.0)],
    analysis="CMS-SMP-20-014-WZ-EFT",
)
```

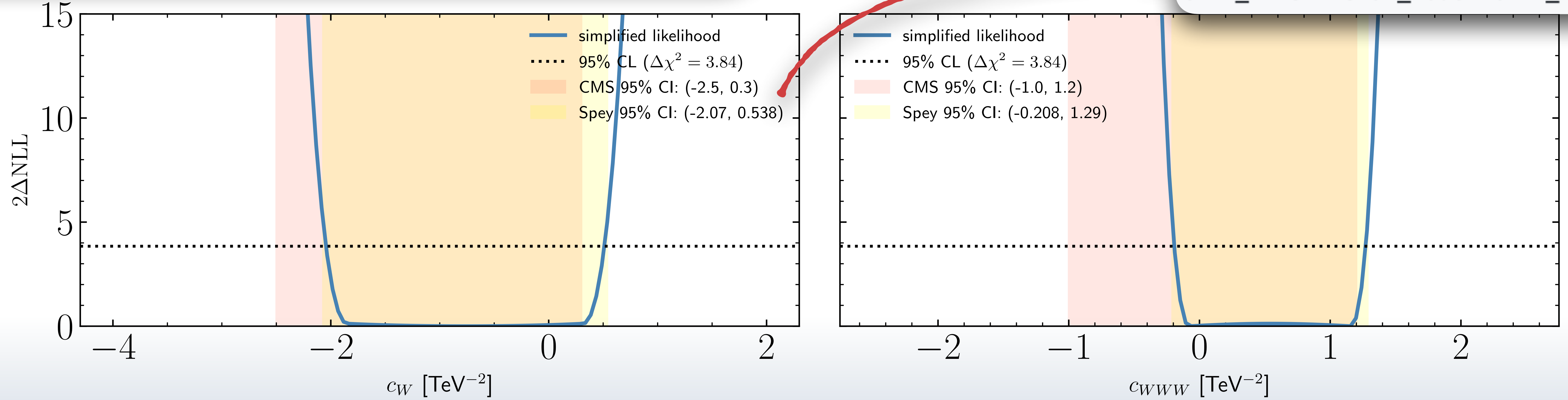
How many parameters does your signal definition require?

Bounds on parameters




```python
cw_lims = stat_model.chi2_test(parameter="c_w")
cwww_lims = stat_model.chi2_test(parameter="c_www")
```

Coming with v0.2.7 or v0.2.8



UCL

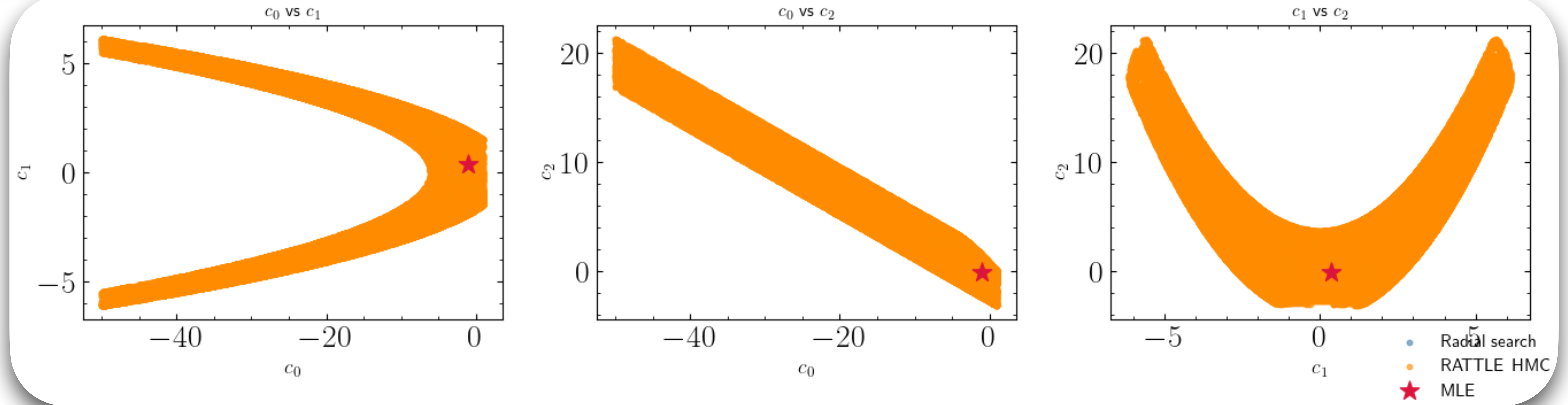
CITY ST GEORGE'S UNIVERSITY OF LONDON

Multi-POI profiling & confidence intervals

```python
result = find_contour(
    stat_model,
    confidence_level=0.95,
    n_radial=2000,  # Stage 2: random radial rays
    n_hmc_chains=100,  # Stage 4: RATTLE chains (gap seeds)
    n_hmc_steps=4000,  # Stage 4: leapfrog steps per chain
    hmc_step_size=0.05,  # Stage 4: leapfrog step size ε
    n_gap_candidates=3000,  # Stage 3: candidate directions for gap detection
    random_seed=42,
    bounds=[(-50, 1), (None, None), (None, None)],
    n_jobs=8,
)
```

```python
def signal_yields(param: np.ndarray) -> np.ndarray:
    """Combined signal as a function of shape parameters (c0, c1, c2).

    c1 enters quadratically: reflects a squared-coupling dependence
    typical in EFT or BSM amplitude-squared contributions.
    """
    return (
        param[0] * signal_yields1
        + param[1] ** 2 * signal_yields2
        + param[2] * signal_yields3
    )
```



Coming with v0.2.7 or v0.2.8

In the pipeline

With Gianfranco De Castro

With Carsten Bugard, Simon Cello & Giordon Holtsberg

- Scalable SMEFT fits with measurements [with Jon Butterworth & Joe Egan]
- Usage of Machine Learned likelihoods [with Sabine Kraml et al.]
- Including theoretical uncertainties [with Luca Pannizi & Benjamin Fuks]