UC SANTA CRUZ

CERN

State-of-the-art Statistics

pyhf

Dr. Giordon Stark
Belle-II Physics Week
October 16th, 2024
indico://e/12273
giordonstark.com

ATLAS EXPERIMENT

Run: 300800

if you can read this, you're too close

Event: 2418777995

2016-06-04 03:47:03

About Me

- B.S Caltech 2012 (LIGO)
  - Brownian Thermal Noise
- Working on ATLAS since 2014
- PhD UChicago 2018 (ATLAS)
  - Search for new (hadronic) physics and instrumentation upgrades (hardware filtering)
- Currently project scientist at SCIPP, UC Santa Cruz since Aug. 2024
  - Search for new (electroweak) physics, large-scale physics analysis combinations, Standard Model measurements, software development, and instrumentation upgrades
- Lots of outreach/teaching/DEI experience (bootcamps, workshops, committees)

The motivation

- How do we make sure that our analysis results are still **interpretable** with new phenomenology today?
- How do we **combine** different analysis results to constrain the allowed (new physics?) models?

not a real experimentalist

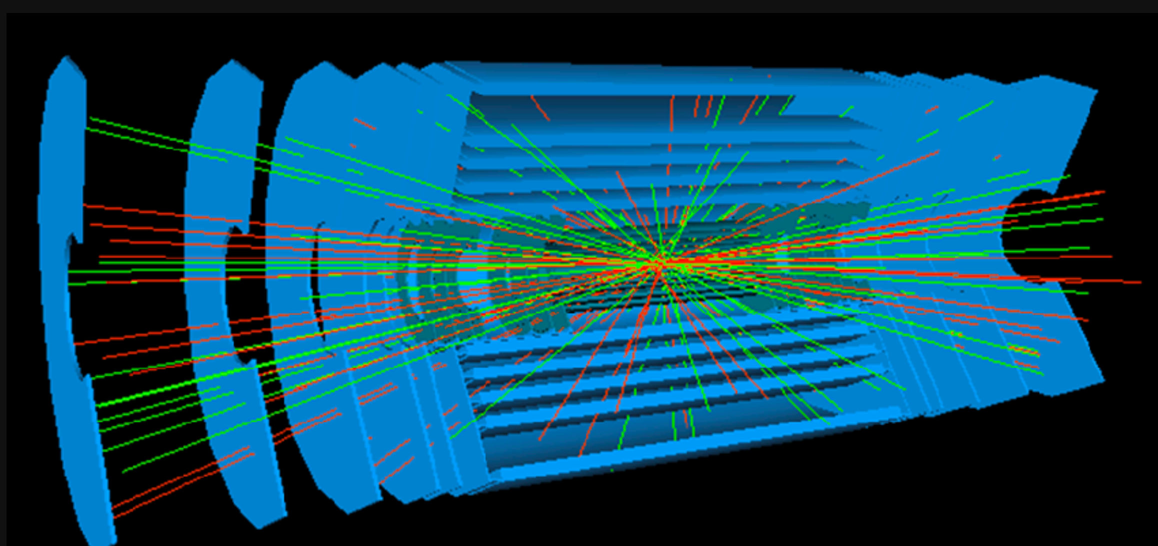
Statistical Techniques

How do experimentalists count?

"Do you know why they call me the Count? Because I love to count! Ah-hah-hah!" — The Count (of Sesame Street)

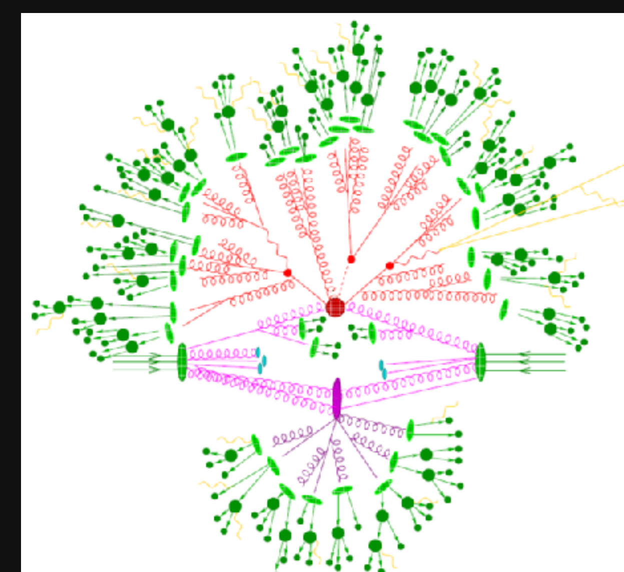
The Big Picture

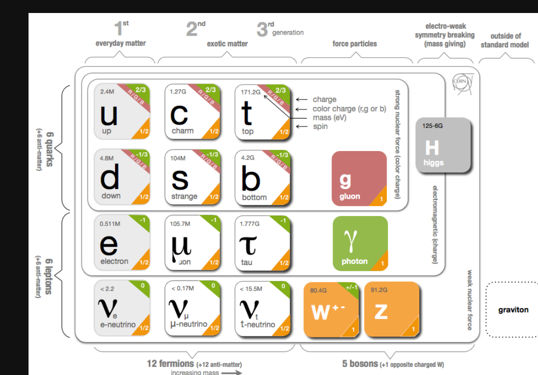
observations

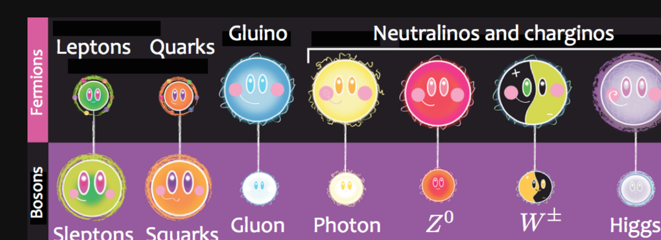


$p(\text{data} = \text{observed})$

model
(SM + SUSY)







$p(\text{theory})$

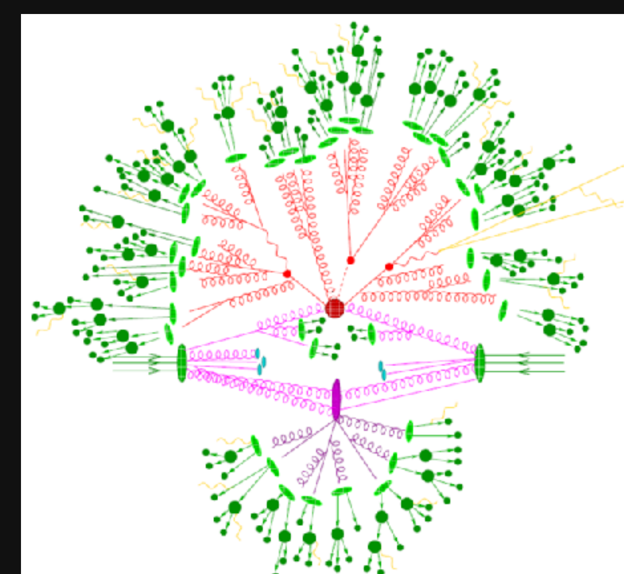
❗ **A likelihood function** encodes everything we know about the **detector**, the **theory**, and the **data**

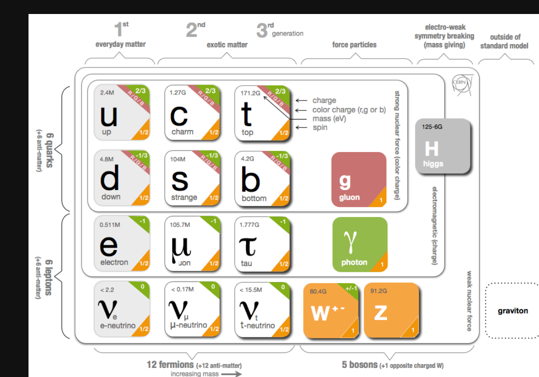
How good is our theory?

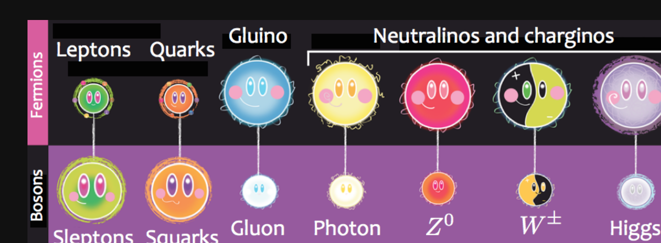
The Big Picture (I)

- Rely on our best **understanding of the theory**
  - Standard Model / Beyond the Standard Model (e.g. QED, QCD, MSSM)
  - Matrix Elements
  - Parton Distribution Functions
  - Finite Order in Perturbative Calculations (e.g. NLO, NNLO, etc…)
  - Parton Showering and Hadronization
- … and best **simulation of our detector**
  - Material Interactions
  - In-time and out-of-time pile-up
  - Calorimeter efficiency
  - Tracking efficiency
  - Magnetic field mapping
  - Beamspot origin
  - etc…

model
(SM + SUSY)





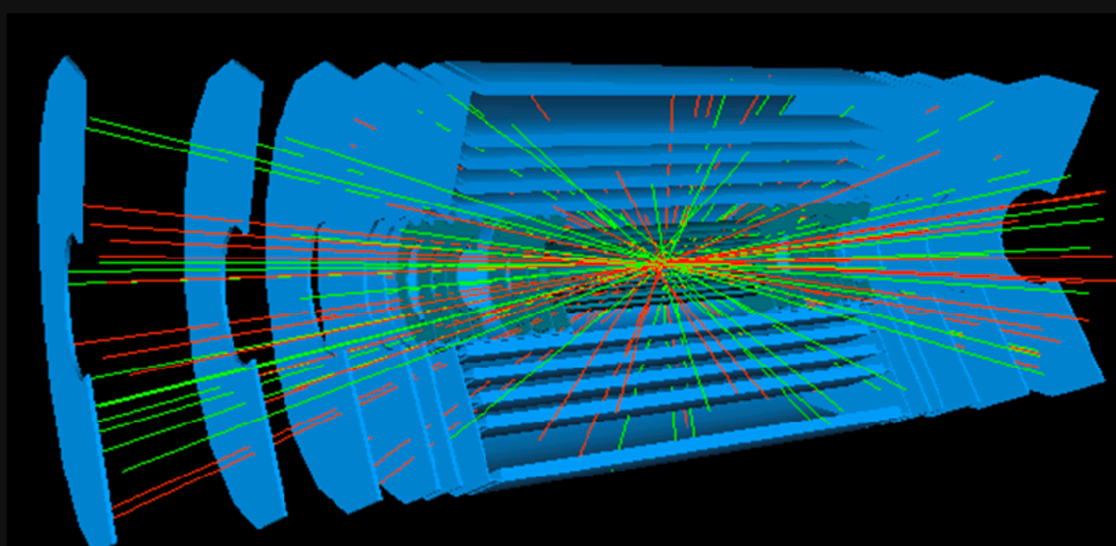


$p(\text{theory})$

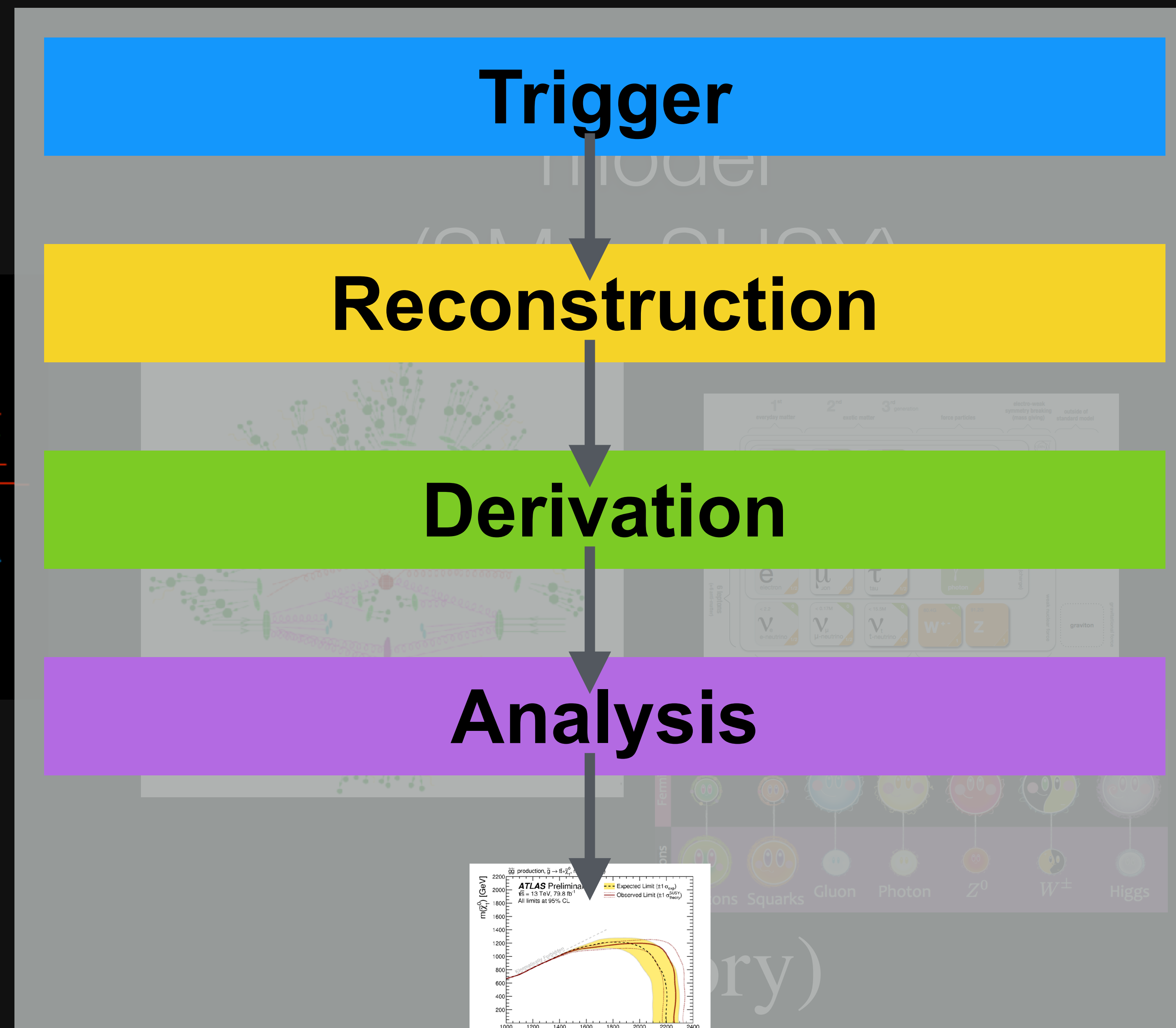
How good is our modeling?

The Big Picture (II)

observations



$$p(\text{data} = \text{observed})$$

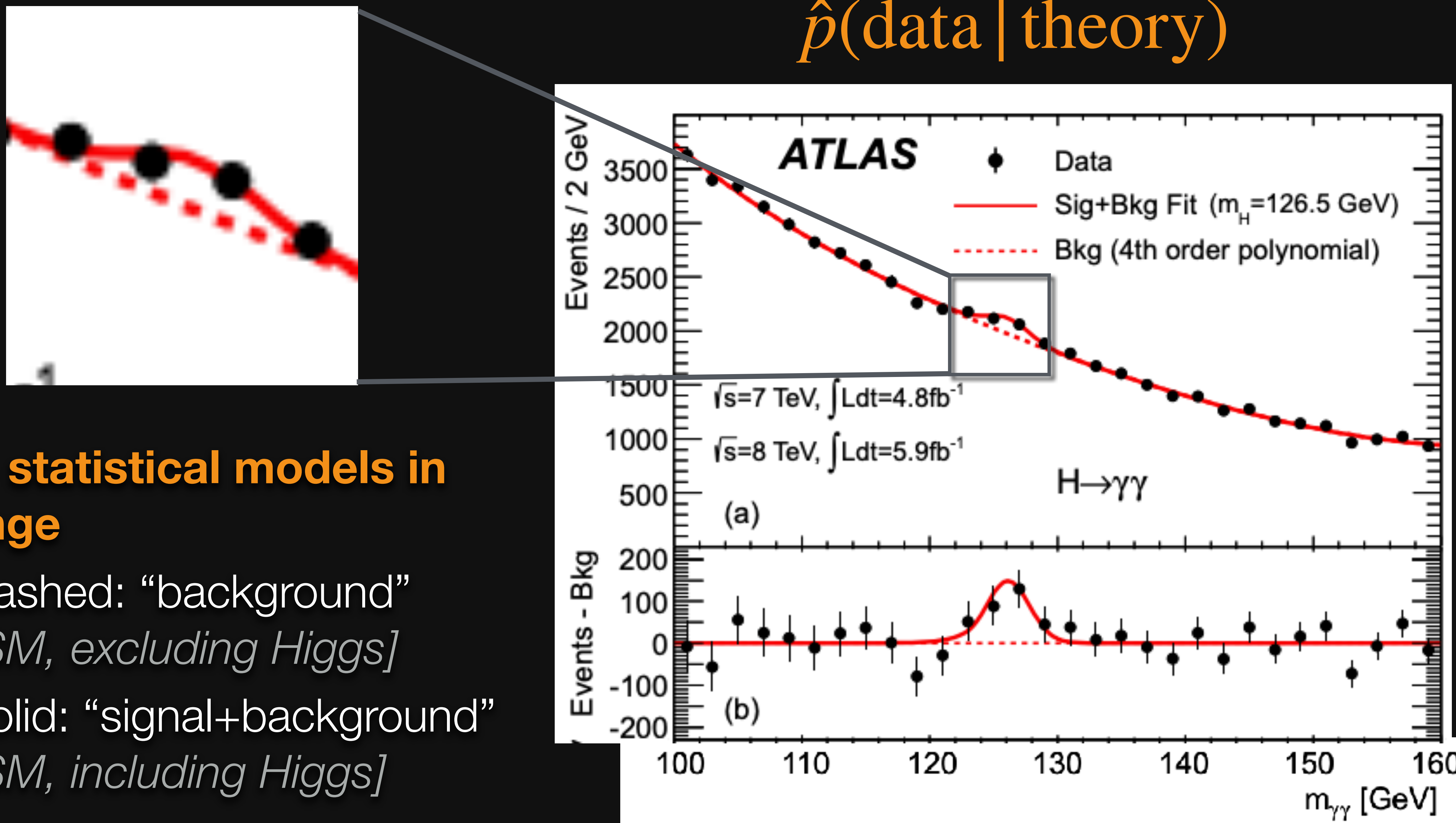


arXiv:1207.7214
[Cowan (Statistical Data Analysis)]()
[James (Statistical Methods in Experimental Physics)]()

*Statistical model **inferences** can provide: p-values, confidence intervals, limits, expected yields, data/MC comparisons, etc…*

What is a statistical model?

$$\hat{p}(\text{data}\,|\,\text{theory})$$



- **Two statistical models in orange**
  - dashed: "background" *[SM, excluding Higgs]*
  - solid: "signal+background" *[SM, including Higgs]*

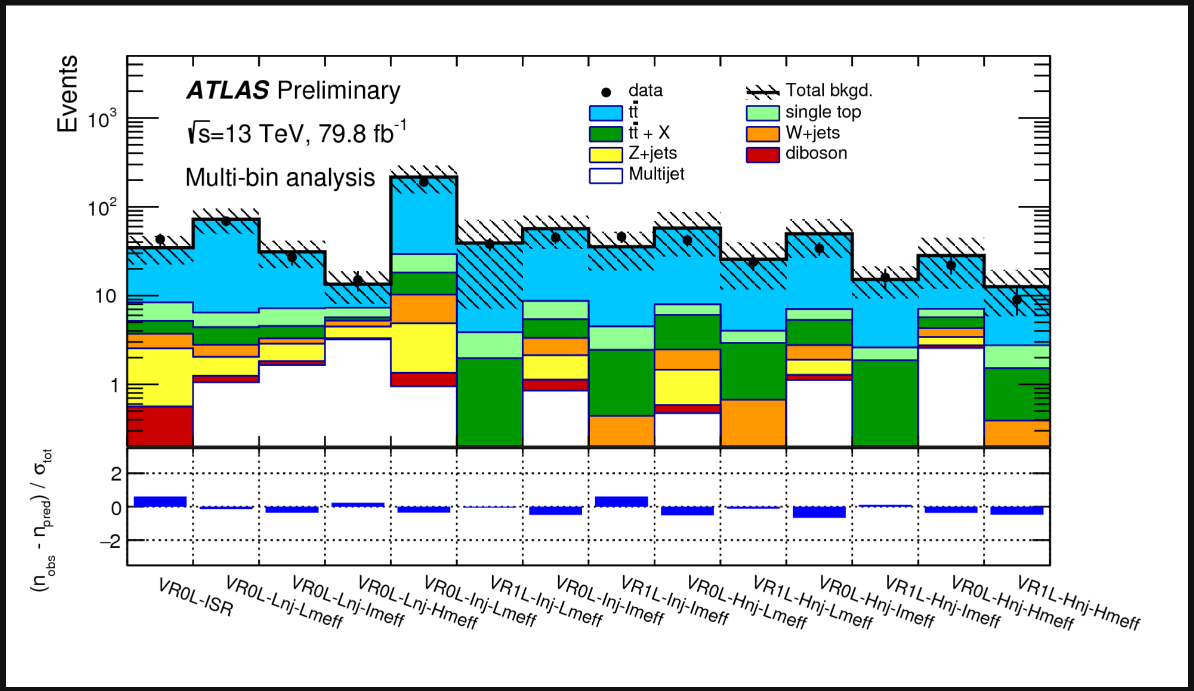
Hypothesis: is the Higgs boson part of the Standard Model?

[CERN-OPEN-2012-016]

HistFactory

- A flexible **p.d.f template specification** for the building of statistical models from binned distributions and data
- Developed by Cranmer, Lewis, Moneta, Shibata, and Verkerke
- Widely used by the HEP community for standard model measurements and BSM searches

Calculated using HistFactory →



K. Cranmer

HistFactory is partially independent of its implementation in ROOT

HistFactory? It's just math!

$$f(\boldsymbol{n}, \boldsymbol{a} \mid \boldsymbol{\eta}, \boldsymbol{\chi}) = \underbrace{\prod_{c \in \text{channels}} \prod_{b \in \text{bins}_c} \text{Pois}\left(n_{cb} \mid \nu_{cb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right)\right)}_{\text{Simultaneous measurement of multiple channels}} \underbrace{\prod_{\chi \in \boldsymbol{\chi}} c_{\chi}(a_{\chi} \mid \chi)}_{\text{constraint terms for "auxiliary measurements"}},$$

Multiple, disjoint channels of binned distributions with multiple samples contributing to each with additional (shared[?]) systematics between sample estimates

- An XML specification with data stored in ROOT files — it's been the ***only implementation*** of this calculation
  - **Poisson p.d.f.** for bins observed in all channels
  - **Constraint p.d.f.** (and data) for auxiliary measurements (systematics: normalization, shape, etc)
  - ⚠ Tied to ROOT ecosystem
  - ⚠ How do we scale? (No multi-threading for larger workspaces e.g. combinations)
  - ⚠ How do we preserve?
  - ⚠ What if there's a bug in ROOT's HistFactory implementation? No cross-check!

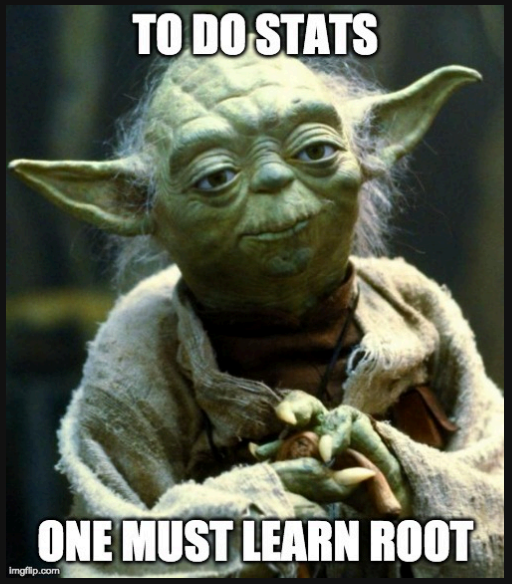
$$\nu_{cb}\left(\boldsymbol{\phi}\right) = \sum_{s \in \text{samples}} \nu_{scb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right) = \sum_{s \in \text{samples}} \underbrace{\left(\prod_{\kappa \in \boldsymbol{\kappa}} \kappa_{scb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right)\right)}_{\text{multiplicative modifiers}} \underbrace{\left(\nu_{scb}^{0}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right) + \sum_{\Delta \in \boldsymbol{\Delta}} \Delta_{scb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right)\right)}_{\text{additive modifiers}}.$$

HistFactory? It's just math!

$$f(\boldsymbol{n}, \boldsymbol{a} \,|\, \boldsymbol{\eta}, \boldsymbol{\chi}) = \underbrace{\prod_{c \in \text{channels}} \prod_{b \in \text{bins}_c} \text{Pois}\left(n_{cb} \,|\, \nu_{cb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right)\right)}_{\text{Simultaneous measurement of multiple channels}} \underbrace{\prod_{\chi \in \boldsymbol{\chi}} c_{\chi}(a_{\chi} \,|\, \chi)}_{\text{constraint terms for "auxiliary measurements"}},$$

Multiple, disjoint channels of binned distributions with multiple samples contributing to each with additional (shared[?]) systematics between sample estimates

- An XML specification with data stored in ROOT files — it's been the ***only implementation*** of this calculation
  - **Poisson p.d.f.** for bins observed in all channels
  - **Constraint p.d.f.** (and data) for auxiliary measurements (systematics: normalization, shape, etc)
  - ⚠ Tied to ROOT ecosystem
  - ⚠ How do we scale? (No multi-threading for larger workspaces e.g. combinations)
  - ⚠ How do we preserve?
  - ⚠ What if there's a bug in ROOT's HistFactory implementation? No cross-check!



$$\nu_{cb}(\boldsymbol{\phi}) = \sum_{s \in \text{samples}} \nu_{scb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right) = \sum_{s \in \text{samples}} \underbrace{\left(\prod_{\kappa \in \boldsymbol{\kappa}} \kappa_{scb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right)\right)}_{\text{multiplicative modifiers}} \underbrace{\left(\nu_{scb}^{0}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right) + \sum_{\Delta \in \boldsymbol{\Delta}} \Delta_{scb}\left(\boldsymbol{\eta}, \boldsymbol{\chi}\right)\right)}_{\text{additive modifiers}}.$$

https://indico.cern.ch/event/100458/

Back in 2000…

Origins I: The First "Statistics in HEP" conference

WORKSHOP ON CONFIDENCE LIMITS

CERN, Geneva, Switzerland
17–18 January 2000

CERN 2000–005

Massimo Corradi

Does everybody agree on this statement, to publish likelihoods?

Louis Lyons

Any disagreement? Carried unanimously. That's actually quite an achievement for this Workshop.

…[Fred James wants to be able to calculate coverage, Don Groom wants to able to calculate goodness of fit]…

Cousins

I thought the point of unanimity was that publishing the likelihood function was a *necessary* condition, not a sufficient condition.

But a practical problem remained: *How* to communicate multi-D likelihood?

ATLAS agreed to publish likelihoods!

In 2019, we did it:

THE MAIN DEVELOPERS

G. Stark

M. Feickert

L. Heinrich

pyhf ∂ifferentiable ℒikelihoods

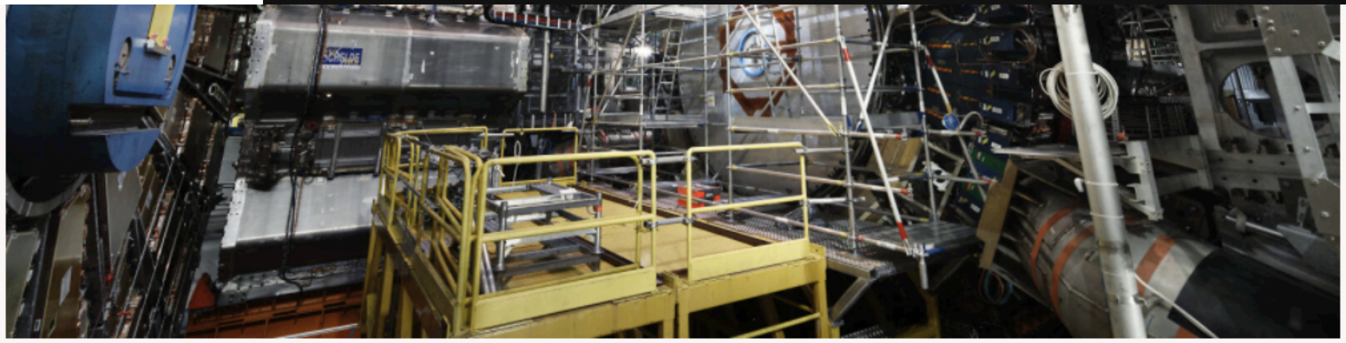
New open release streamlines interactions with theoretical physicists

The ATLAS Collaboration has released the first open likelihoods from an LHC experiment.

12th December 2019 | By Katarina Anthony

Explore ATLAS open likelihoods on the HEPData platform. (Original image: Ahmet Anil Sen/Behance)

https://atlas.cern/updates/news/new-open-likelihoods



Courtesy of CERN

ATLAS releases 'full orchestra' of analysis instruments

01/14/21 | By Stephanie Melchor

The ATLAS collaboration has begun to publish likelihood functions, information that will allow researchers to better understand and use their experiment's data in future analyses.

Meyrin, Switzerland, sits serenely near the Swiss-French border, surrounded by green fields and the beautiful Rhône river. But a hundred

https://www.symmetrymagazine.org/article/atlas-releases-full-orchestra-of-analysis-instruments

Easy to install

it would be useful to **run statistical analysis outside of ROOT**, RooFit, RooStats framework

```
pip install pyhf
```

A **python-only** (scipy, numpy) implementation of the HistFactory model + profile likelihood hypothesis tests

For free: a single plain-text file (JSON) specifies the entire workspace

https://scikit-hep.org/pyhf/

Easy to use ("hello world")

```
>>> import pyhf
>>> pdf = pyhf.simplemodels.uncorrelated_background(signal=[12.,11.],
... bkg=[50.,52.], bkg_uncertainty=[3.,7.])

>>> results = pyhf.utils.runOnePoint(1.0, [51, 48] + pdf.config.auxdata, pdf,
... return_tail_probs=True, return_expected=True, return_expected_set=True)

>>> print('Observed: {} Expected: {}'.format(results[-2], results[-1][2]))
Observed: [0.05290116] Expected: [0.06445521]
```

CLs

- Want to use…
  - tensorflow? `pip install pyhf[tensorflow]`
  - pytorch? `pip install pyhf[pytorch]`
  - jax? `pip install pyhf[jax]`
- If the JSON workspace is online, can pipe and calculate CLs instantly

```
$ curl http://url-to-json/workspace.json | pyhf cls
```

```
$ curl pdf.json | pyhf cls
```

Demo (I) — Simple CLs


```json
{
    "channels": [{
        "name": "singlechannel",
        "samples": [{
                "name": "sig",
                "data": [12.0, 11.0],
                "modifiers": [{ "name": "mu", "data": null, "type": "normfactor" }]
            },
            {
                "name": "bkg",
                "data": [50.0, 52.0],
                "modifiers": [{ "name": "uncorr_bkguncrt", "data": [3.0, 7.0], "type": "shapesys" }]
            }
        ]
    }],
    "observations": [
        {"name": "singlechannel", "data": [51.0, 48.0]}
    ],
    "measurements": [{
            "config": { "poi": "mu", "parameters": [] },
            "name": "singlechannel"
    }],
    "version": "1.0.0"
}
```


JSON defining a single channel, two bin counting experiment with systematics

```
$ curl -sL https://git.io/fj1yb | pyhf cls | jq .CLs_obs
0.052515541856109835
```

```
$ curl pdf.json | pyhf cls --patch patch.json
```

Demo (II) — Simple Re-use

```
{
    "channels": [{
        "name": "singlechannel",
        "samples": [{
                "name": "sig",
                "data": [12.0, 11.0],
                "modifiers": [{ "name": "mu", "data": null, "type": "normfactor" }]
            },
            {
                "name": "bkg",
                "data": [50.0, 52.0],
                "modifiers": [{ "name": "uncorr_bkguncrt", "data": [3.0, 7.0], "type": "shapesys" }]
            }
        ]
    }],
    "observations": [
        {"name": "singlechannel", "data": [51.0, 48.0]}
    ],
    "measurements": [{
            "config": { "poi": "mu", "parameters": [] },
            "name": "singlechannel"
    }],
    "version": "1.0.0"
}
```

- Let's patch the pyhf JSON spec provided with a different signal and recalculate!

```
# new_signal.json
[{
    "op": "replace",
    "path": "/channels/0/samples/0/data",
    "value": [5.0, 6.0]
}]
```

```
$ curl pdf.json | pyhf cls --patch patch.json
```

Demo (II) — Simple Re-use

```
$ curl -sL https://git.io/fj1yb | pyhf cls | jq .CLs_obs
0.052515541856109835

# reinterpretation time
$ curl -sL https://git.io/fj1yb | pyhf cls --patch <(curl -sL https://git.io/fj1yN)
| jq .CLs_obs
0.3365054427336307б
```

Patch with JSONPatch (http://jsonpatch.com/)

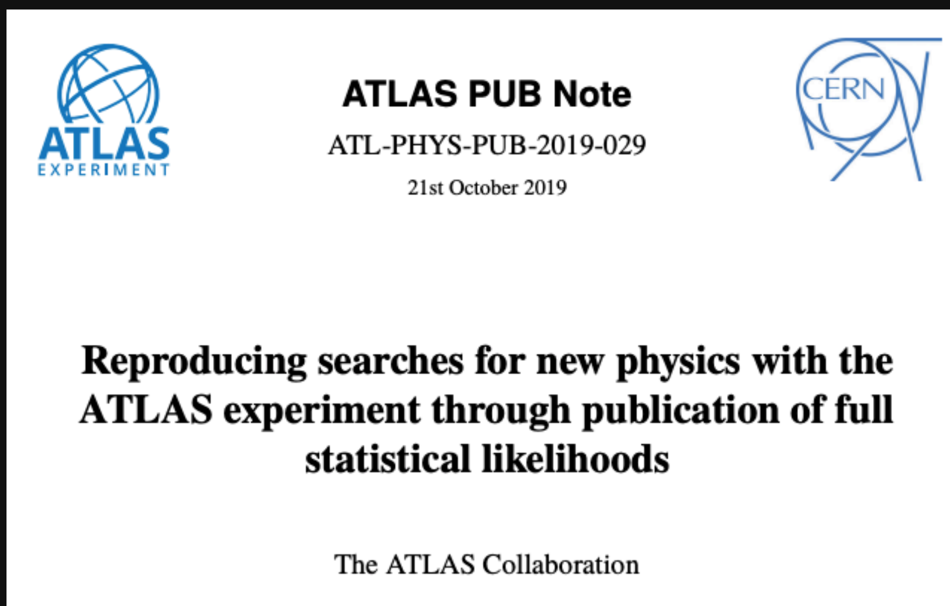
- Let's patch the pyhf JSON spec provided with a different signal and recalculate!

```
# new_signal.json
[{
    "op": "replace",
    "path": "/channels/0/samples/0/data",
    "value": [5.0, 6.0]
}]
```

[ATLAS-CONF-2019-011](#) [JHEP12(2019)060](#) [ATL-PHYS-PUB-2019-029](#)
https://www.hepdata.net/record/ins1748602

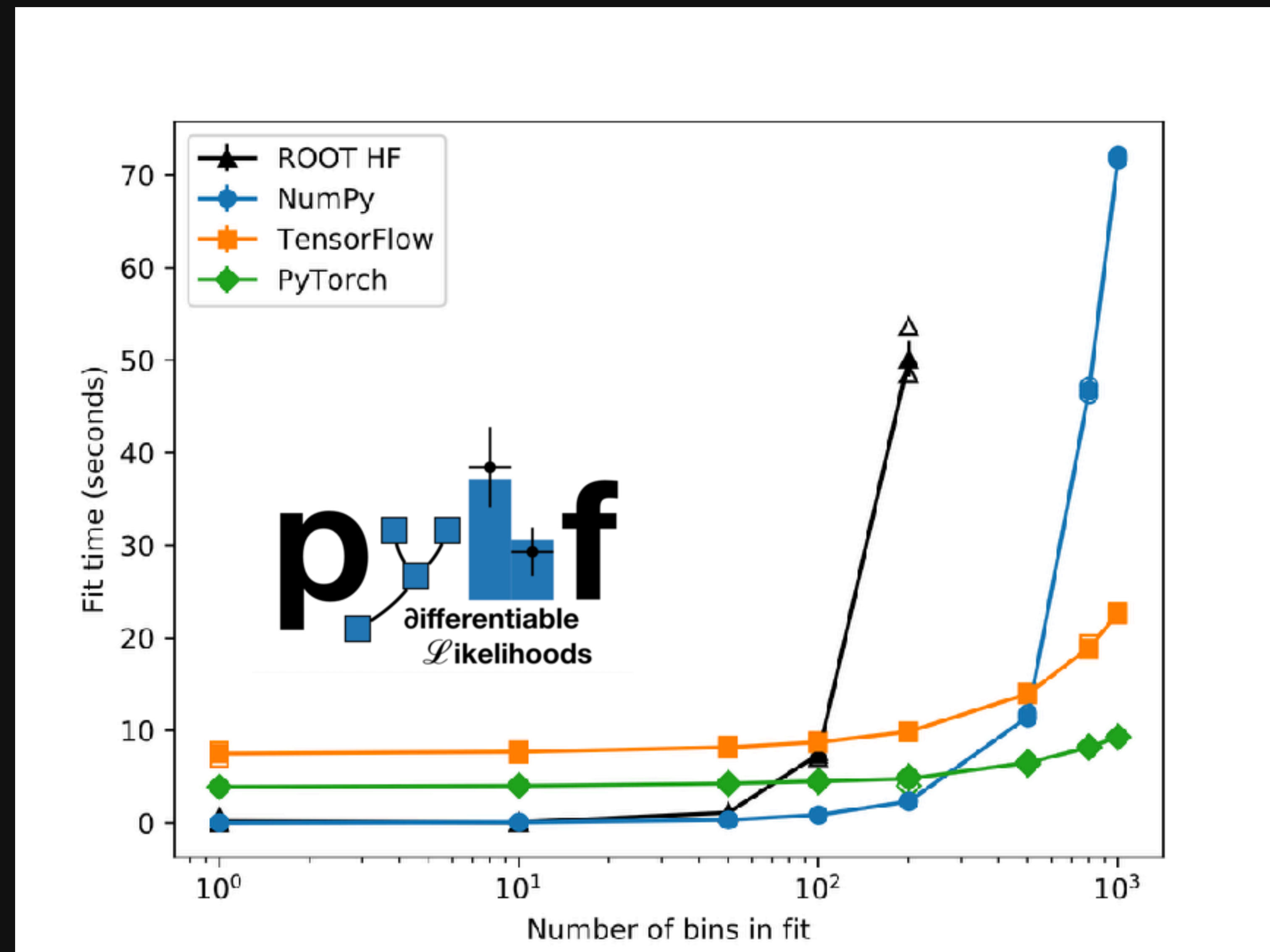
Computationally efficient!

launch binder


ATLAS PUB Note
ATL-PHYS-PUB-2019-029
21st October 2019

Reproducing searches for new physics with the ATLAS experiment through publication of full statistical likelihoods

The ATLAS Collaboration

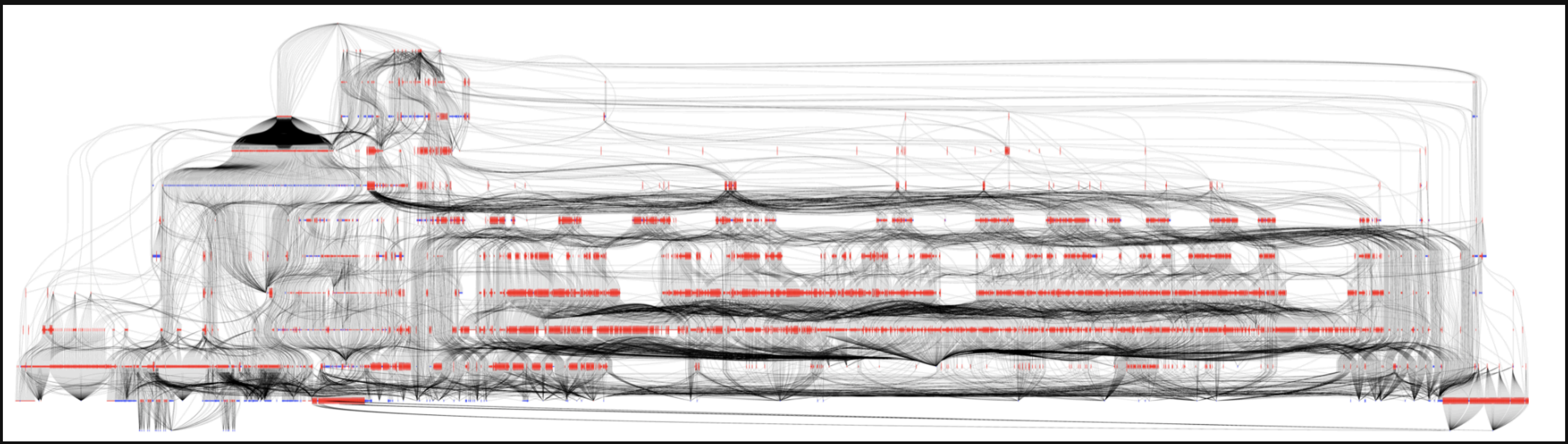


PyTorch TensorFlow NumPy

“python -m pip install pyhf”

orders-of-magnitude faster inference

Why the speed up? (I)



somewhat complex with lots of structure

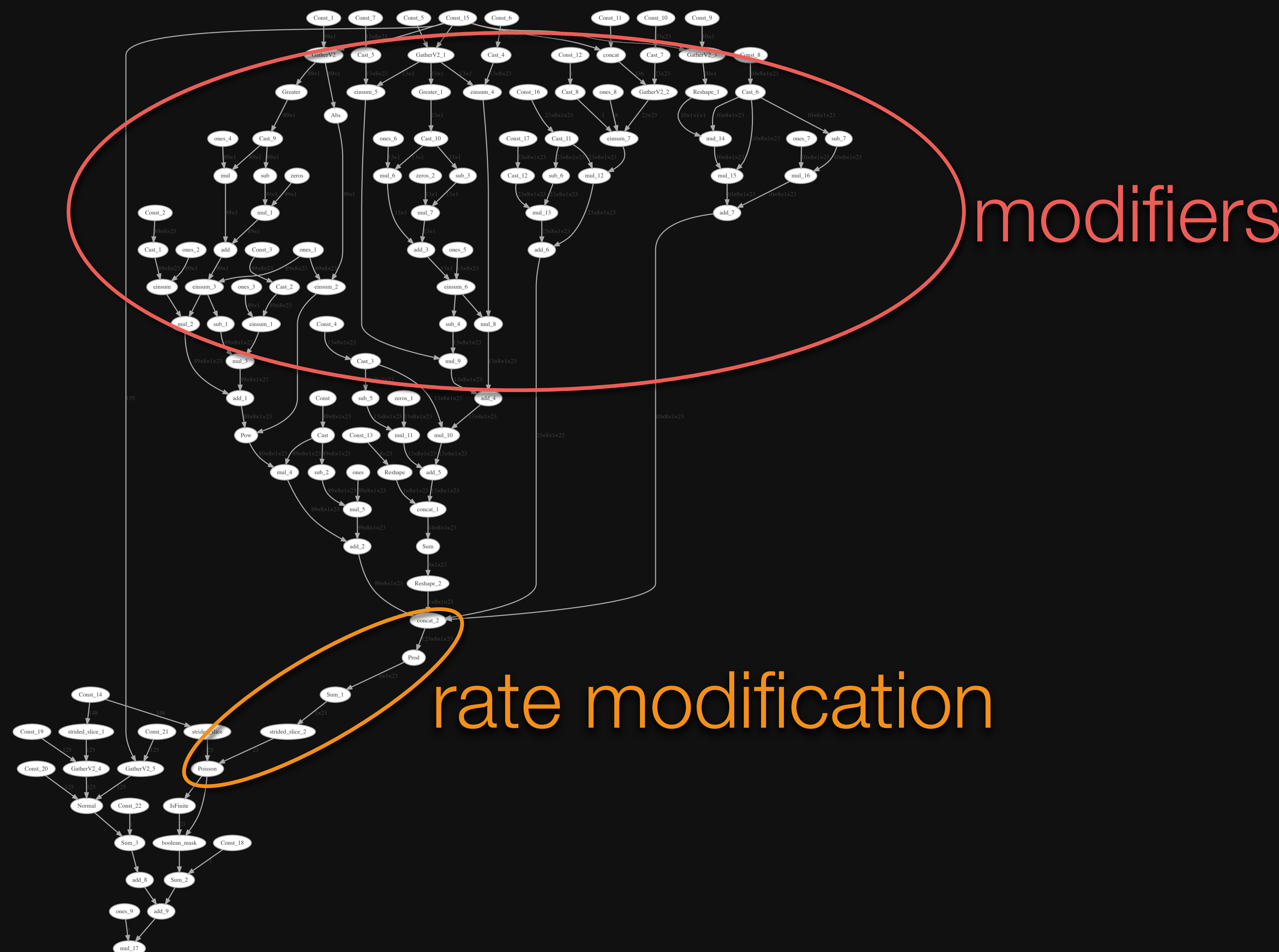
ROOT-HistFactory computational tree

tensor dimensionality (rather than computational graph complexity) correlates with statistical model complexity

Why the speed up? (II)

$$\frac{\partial L}{\partial u}, \frac{\partial L}{\partial \theta_i}$$

simpler, tensor-based, with automatic differentiation

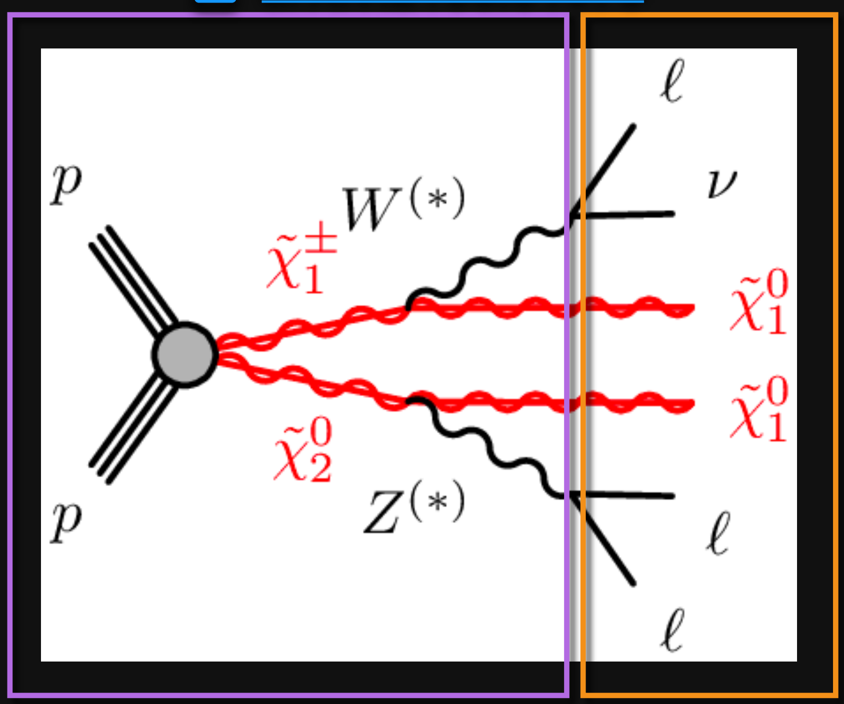


python-HistFactory computational tree

E.G.: Stat. Combination (I)

- Multiple analyses with different **signatures** can still target the same **model**

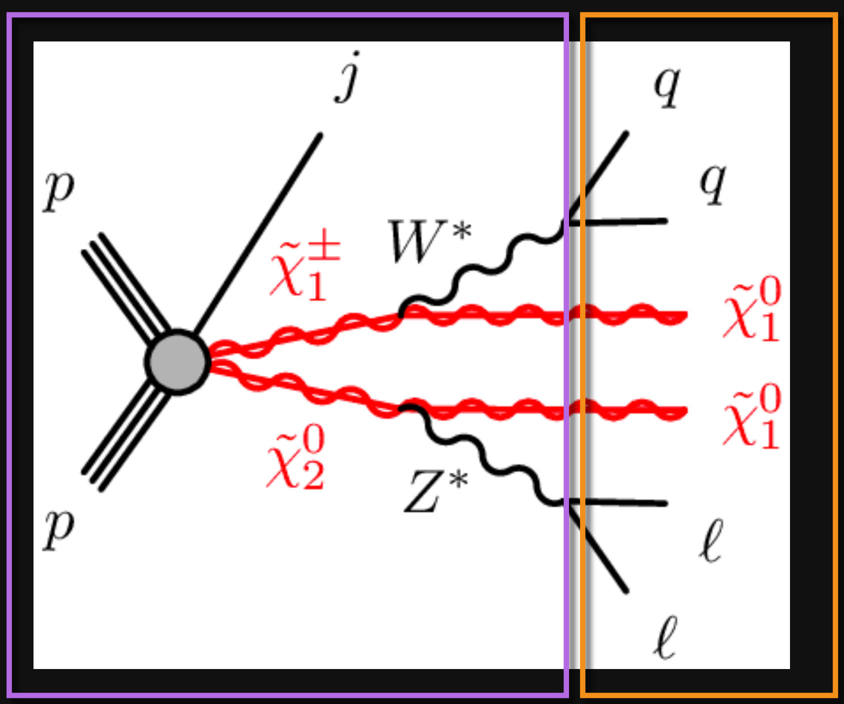
ATLAS-CONF-2020-015
arXiv:2106.01676



signature: $3\ell + 0j + E_T^{miss}$

"3-lepton"

arXiv:1911.12606



signature: $2\ell + 3j + E_T^{miss}$

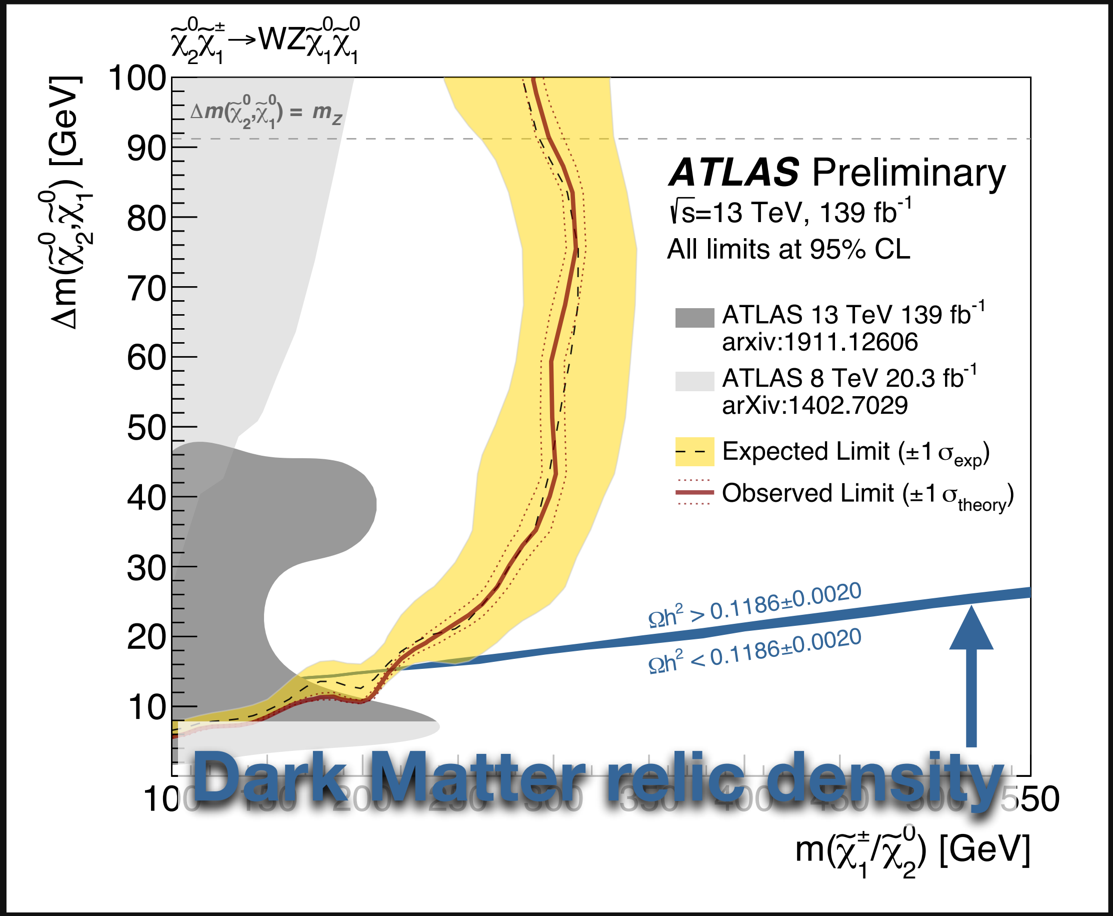
"soft 2-lepton"

arXiv:1911.12606 ATLAS-CONF-2020-015 arXiv:2106.01676 ATL-PHYS-PUB-2021-019
atlas.cern/updates/briefing/new-higgsino-limits

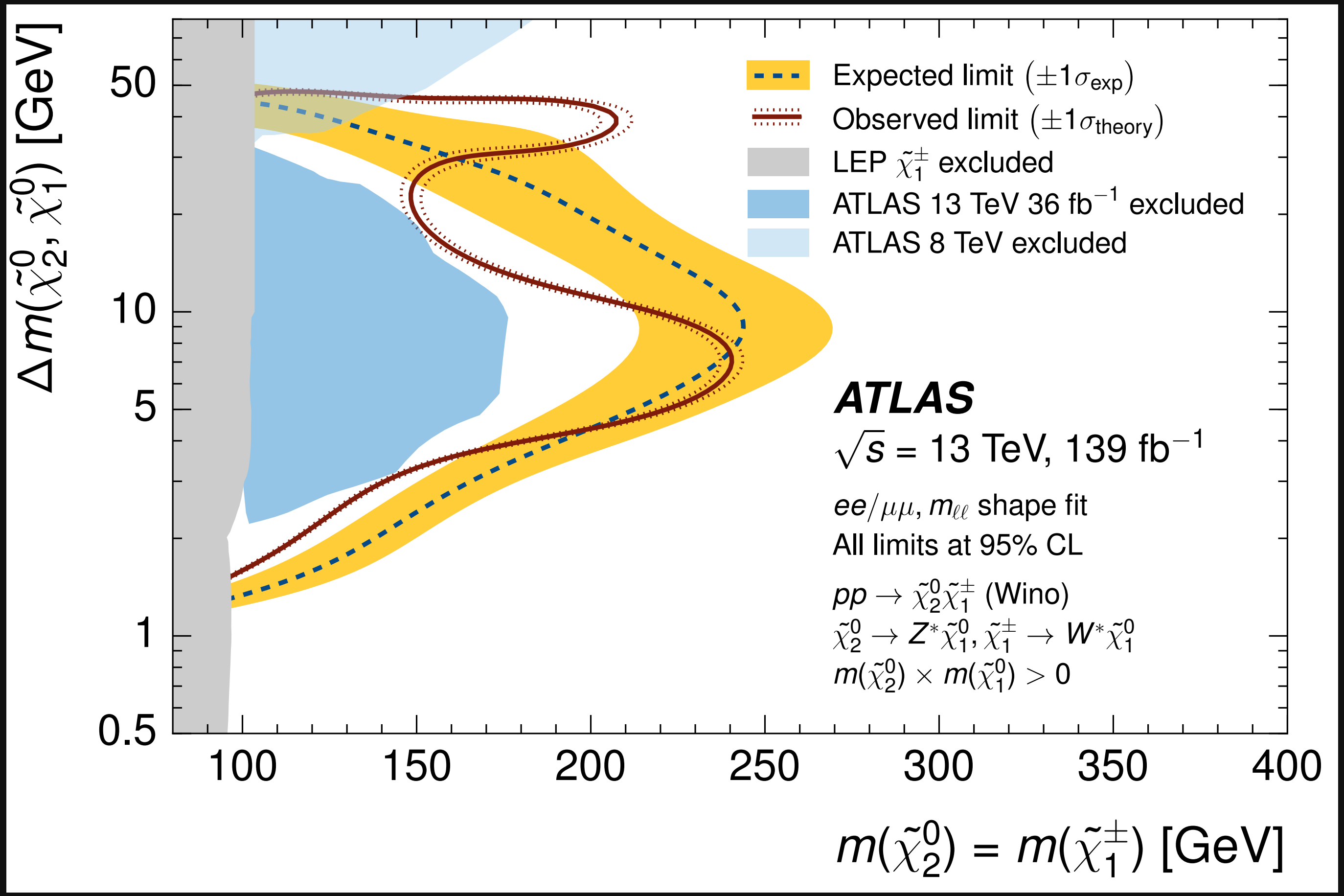
E.G.: Stat. Combination (II)

- Goal: **combine multiple searches** to paint a tapestry of our sensitivity to a set of simplified models (**electroweak-production**) which decay to on-shell/off-shell Standard Model bosons (W/Z)

three lepton



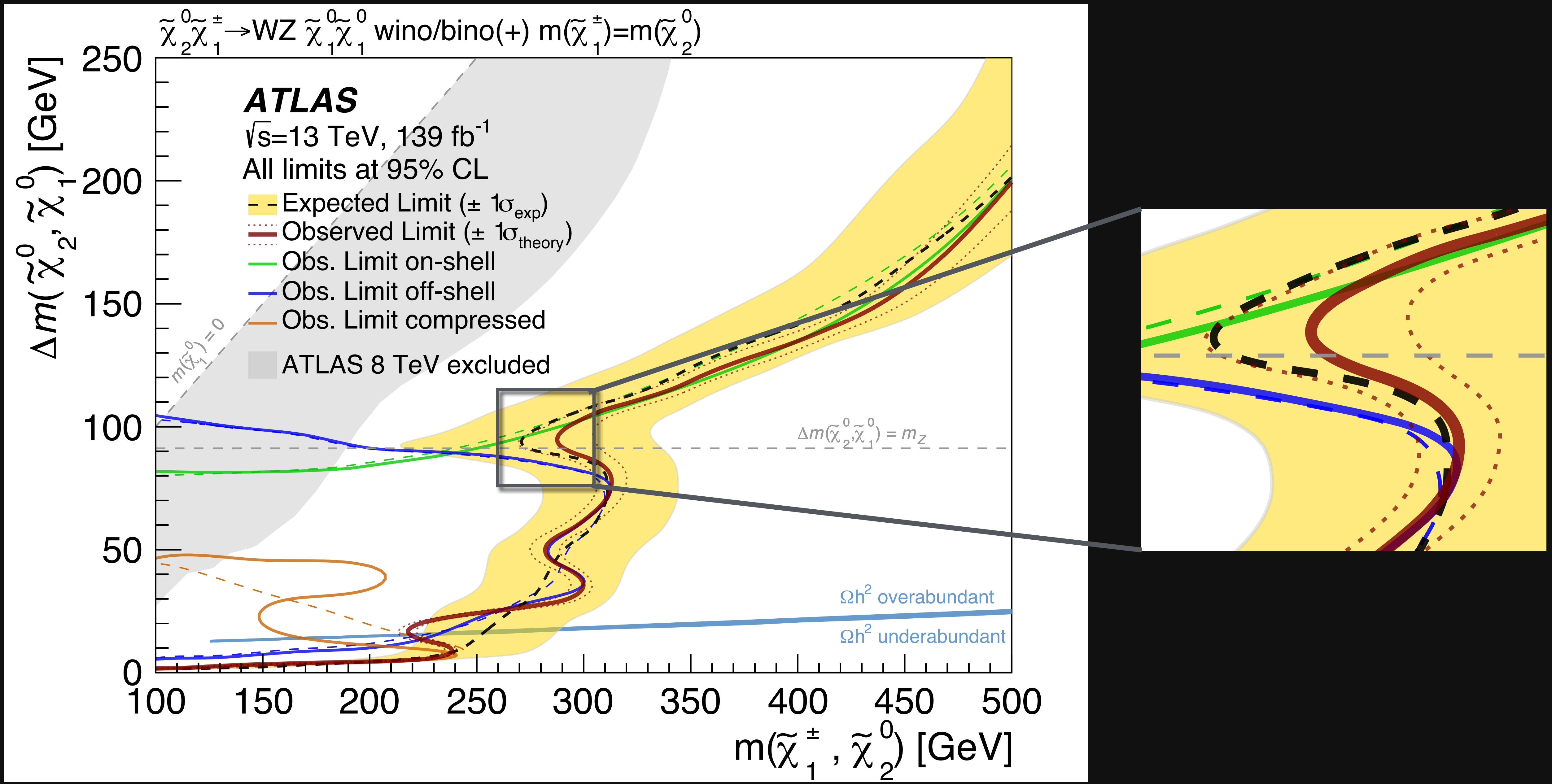
soft two lepton



$m(\tilde{\chi}^\pm_1,\tilde{\chi}^0_2)$ $m(\tilde{\chi}^\pm_1,\tilde{\chi}^0_2)$

arXiv:1911.12606 ATLAS-CONF-2020-015 arXiv:2106.01676 ATL-PHYS-PUB-2021-019
atlas.cern/updates/briefing/new-higgsino-limits

E.G.: Stat. Combination (III)



! Not possible without my work externally on pyhf and reproducibility, but also driving it internally in the ATLAS collaboration

Phys. Rev. Lett. 133 (2024) 031802

Full Run-2 Paper

- **Statistical combination** of existing (public) ATLAS searches for electroweak SUSY $(\tilde{\chi}^{\pm}, \tilde{\chi}^{0}, \tilde{G})$ production in **four decay channels**

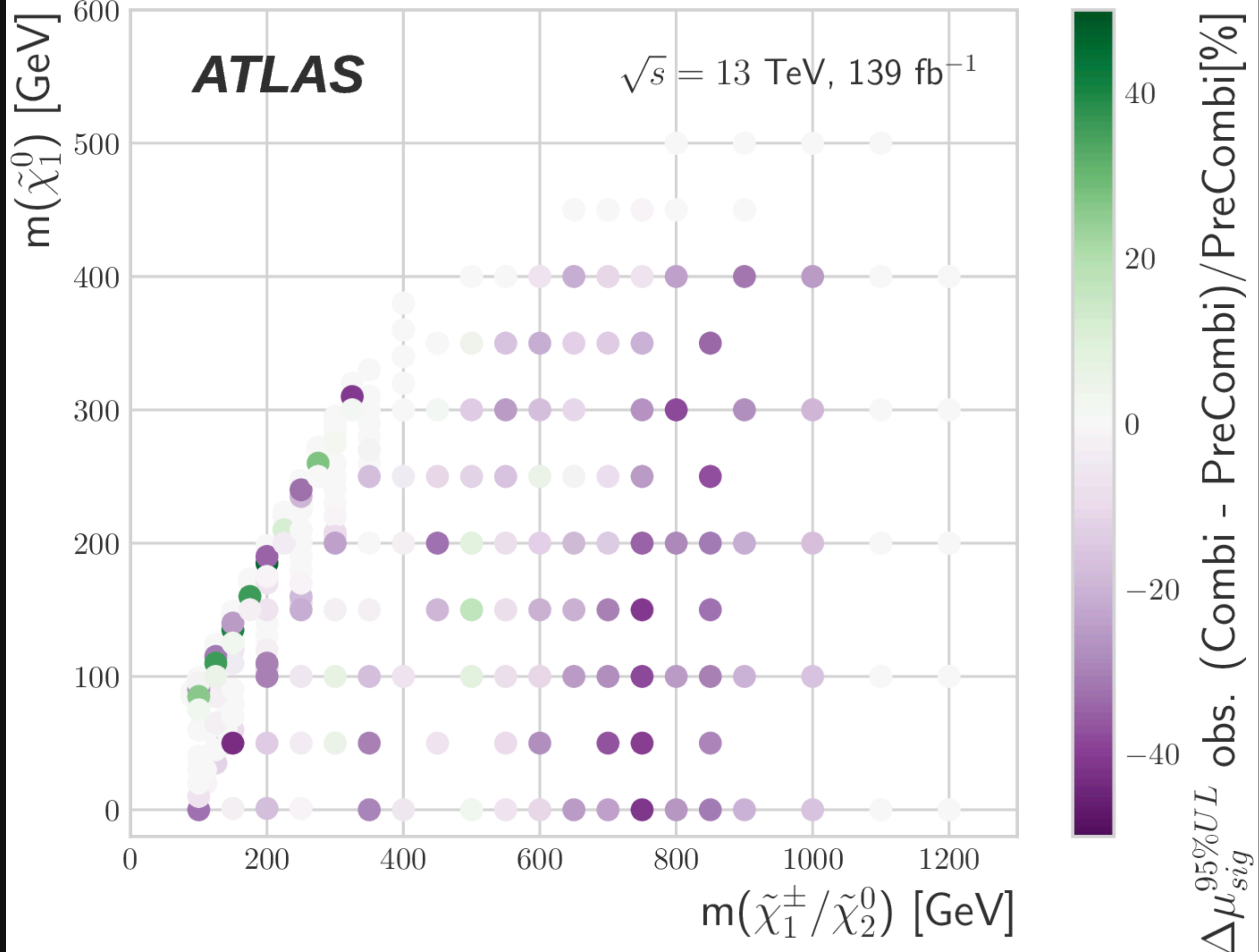
Production mode	Wino $\tilde{\chi}_1^{+}\tilde{\chi}_1^{+}$	Wino $\tilde{\chi}_1^{\pm}\tilde{\chi}_2^{0}$	Wino $\tilde{\chi}_1^{\pm}\tilde{\chi}_2^{0}$	Higgsino GGM $\tilde{\chi}_1^{+}\tilde{\chi}_1^{+}, \tilde{\chi}_1^{\pm}\tilde{\chi}_{1,2}^{0}, \tilde{\chi}_1^{0}\tilde{\chi}_2^{0}$
Decay mode	$\tilde{\chi}_1^{\mp} \to W\tilde{\chi}_1^{0}$	$\tilde{\chi}_1^{\pm} \to W\tilde{\chi}_1^{0}$, $\tilde{\chi}_2^{0} \to Z\tilde{\chi}_1^{0}$	$\tilde{\chi}_1^{\pm} \to W\tilde{\chi}_1^{0}$, $\tilde{\chi}_2^{0} \to h\tilde{\chi}_1^{0}$	$\tilde{\chi}_1^{0} \to Z/h\tilde{G}$
Searches				
All Hadronic [19]	✓	✓	✓	✓
1L [**SUSY-2019-19**]	✓	✓		
1Lbb [20]			✓	
2L Compressed [21]		✓		
2L0J $\Delta m > m(W)$ [22]	✓			
2L0J $\Delta m \sim m(W)$ [23]	✓			
2L2J [24]		✓		✓
2tau [25]			✓	
3L [26]		✓	✓	
SS/3L [27]		✓	✓	
4L [28]				✓
Multi-b [**SUSY-2020-16**]				✓

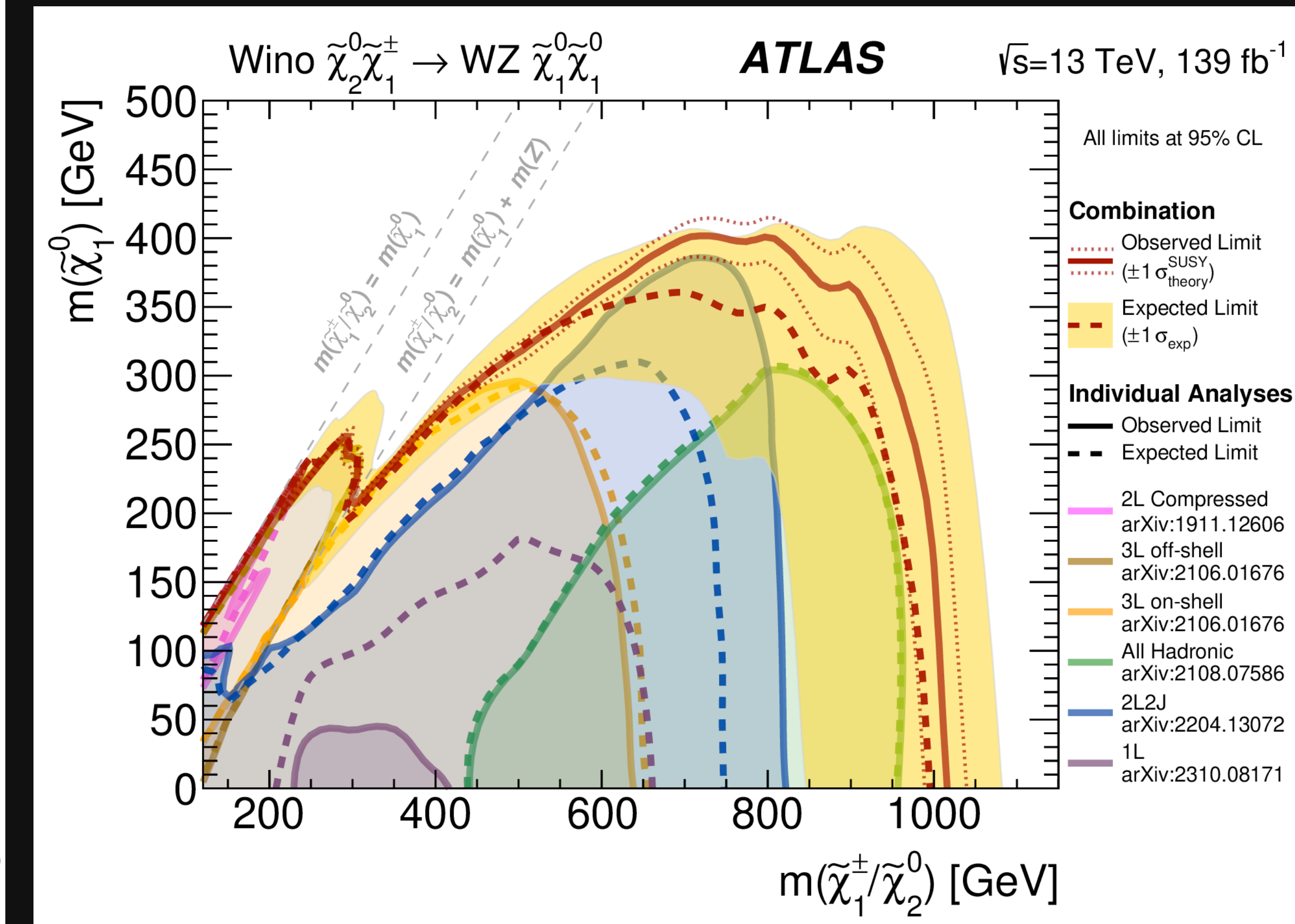
WW **WZ** **Wh** **GGM**

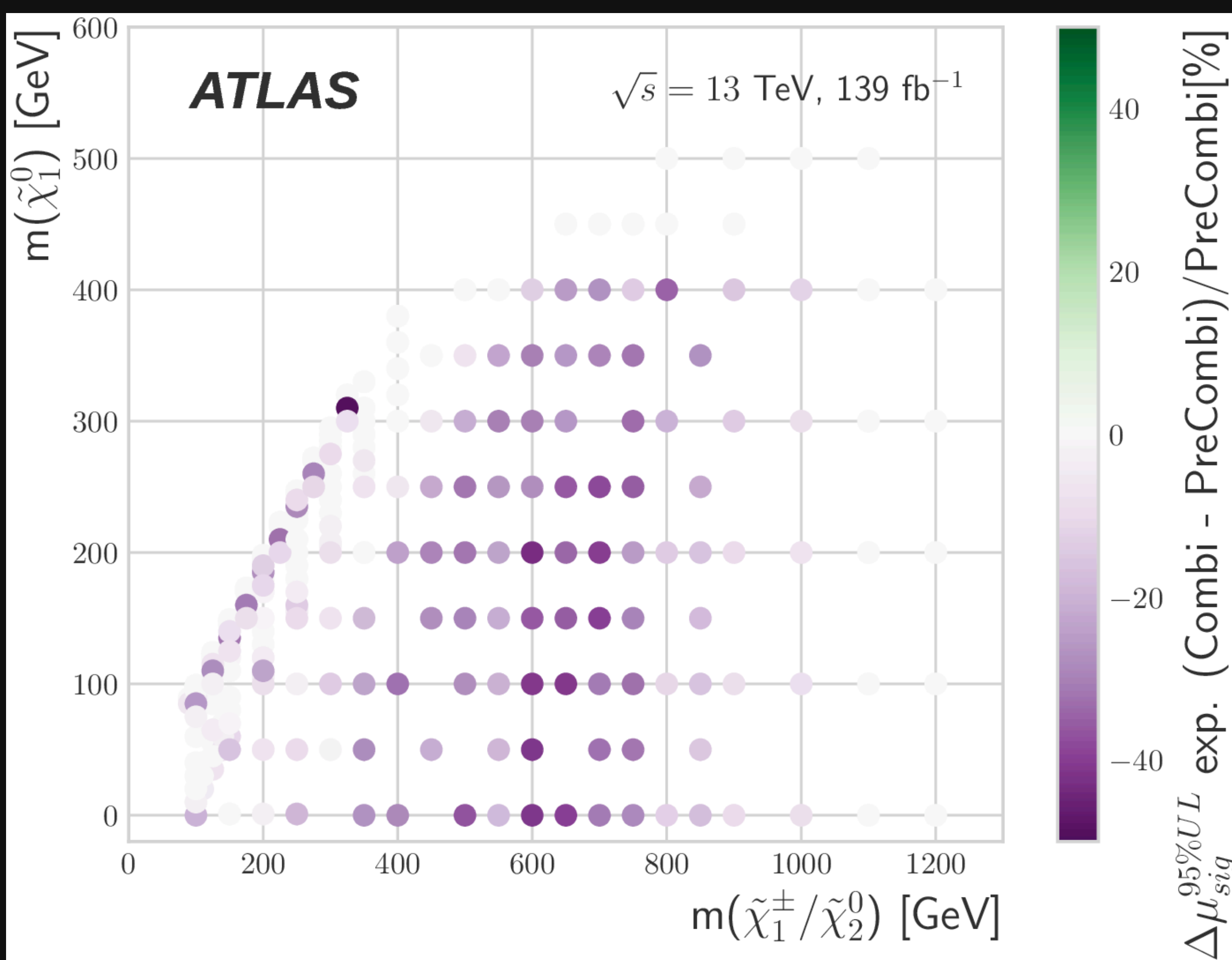
Long game — a culmination of five years of work!

- Required many pieces in place:
  - harmonized object definitions
  - serialized likelihoods
  - fully-preserved analysis workflows
  - unblinded (and public) analyses
  - grid extensions (non-stat.)
  - understanding statistical overlaps
  - (re)defining uncertainty correlations

★ **Improve sensitivity to EWK SUSY beyond current limits**







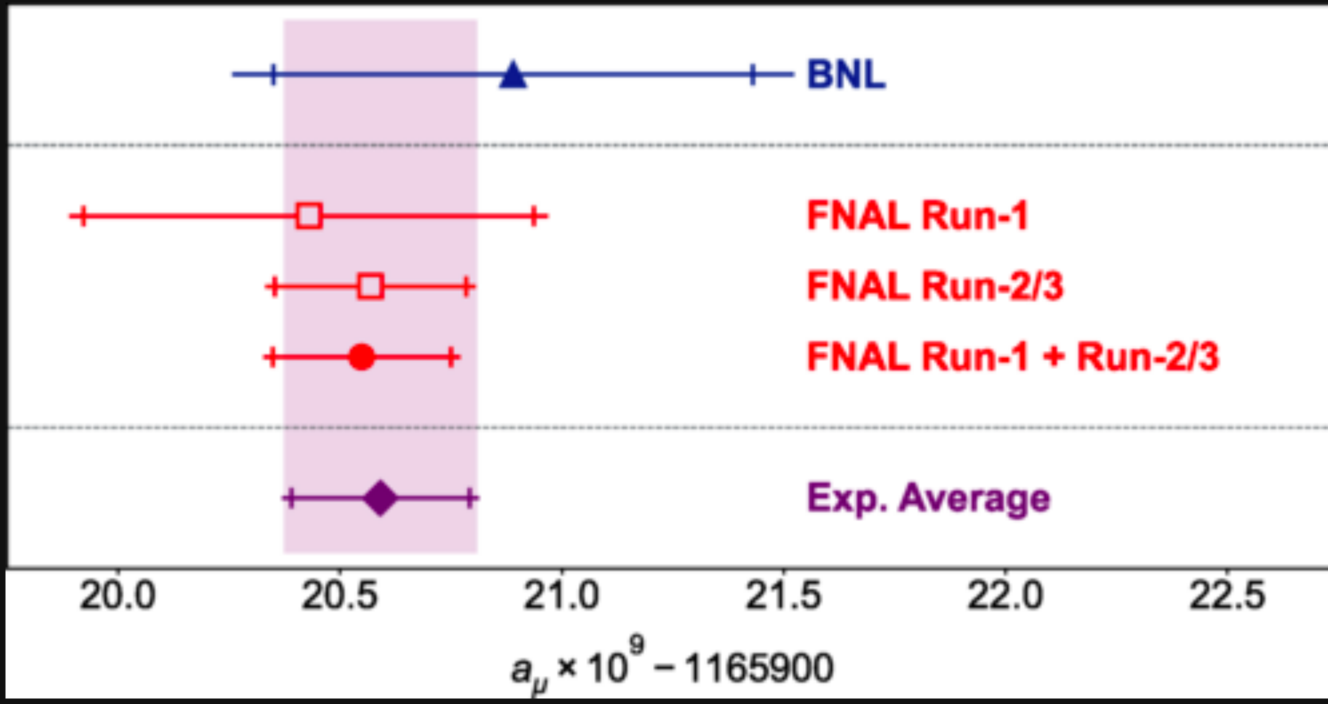
Not just about sensitivity!

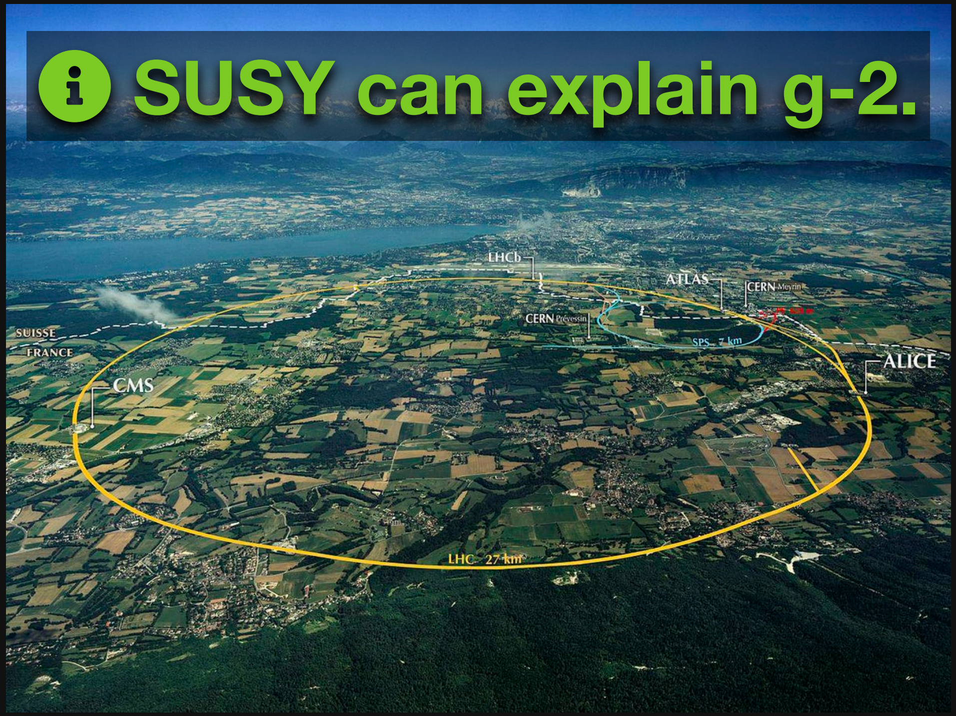
- Theorists don't understand experimental correlations across analyses
  - provide a combined probability model that they can use instead to represent the combined contour of this model (signature-independent!)
- Depth of exclusion allows us to make significantly stronger statements about where new physics shouldn't not exist

ATL-PHYS-PUB-2021-019 arXiv:1403.5294 arXiv:1908.08215 arXiv:1911.12606
10.1103/PhysRevLett.126.141801

Example: global fit (g-2) (I)

combine!



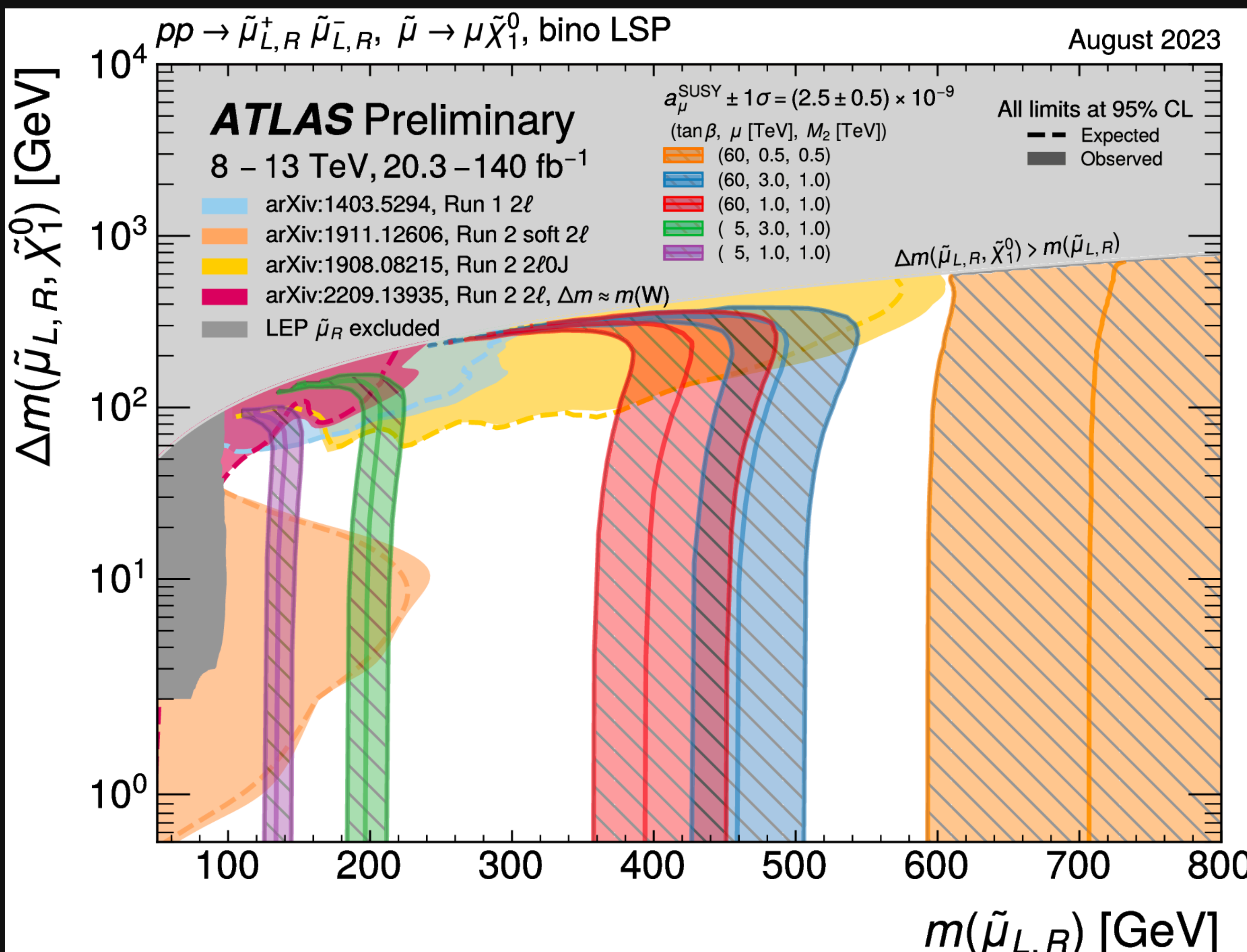


August 10th, 2023

Can we use experimental measurements of g-2 to inform our ATLAS physics program?

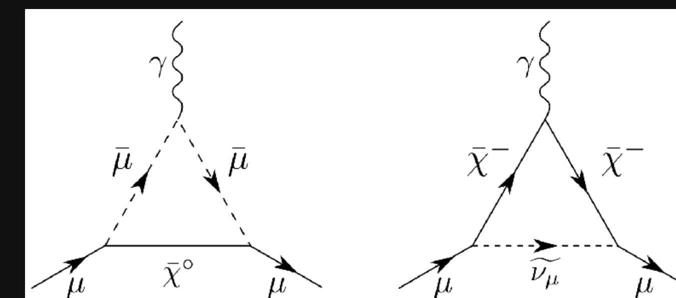
ATL-PHYS-PUB-2023-025 arXiv:1403.5294 arXiv:1908.08215 arXiv:1911.12606
10.1103/PhysRevLett.126.141801

Example: global fit (g-2) (II)



center line = combined g-2 measurement

hatching = $\pm 1\sigma$ uncertainty on experimental measurement



$\tilde{\mu}^\pm/\tilde{\chi}^0$ or $\tilde{\chi}^\pm/\tilde{\nu}_\mu$

- Exclusion limits in the $(\tilde{\mu}, \tilde{\chi}^0_1)$ mass plane
  - Overlay ATLAS analyses that are sensitive to the smuon model with muon/bino-like $\tilde{\chi}^0_1$ signature
- Hatched bands are compatible with the observed g-2 anomaly measured by Fermilab and BNL experiments to $\pm 1\sigma$, for a variety of pMSSM parameters

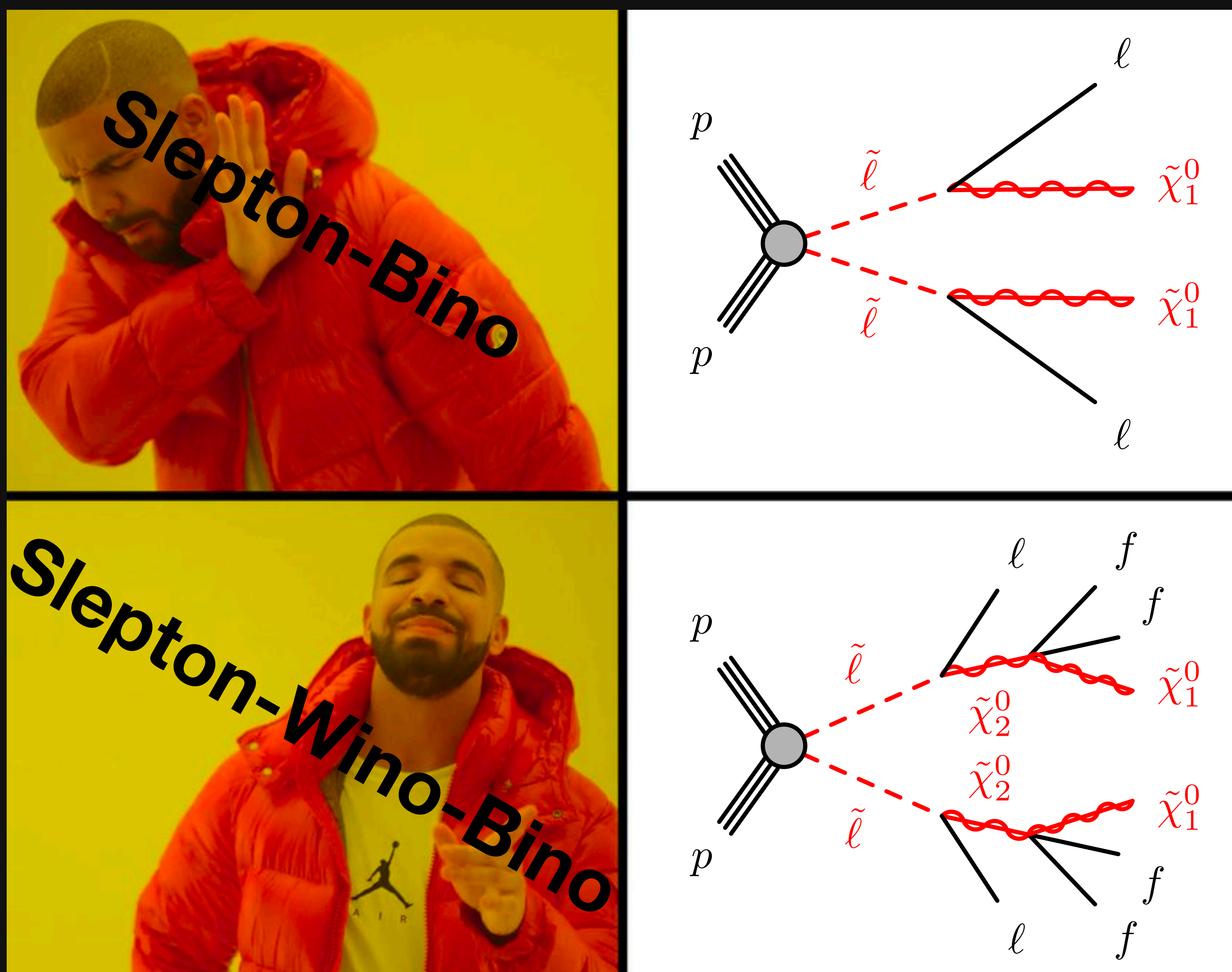
Light smuon/light neutralino needed!

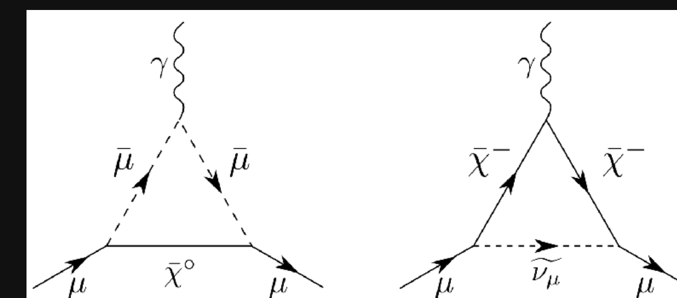
arXiv:2306.11055 arXiv:1911.12606

⚠ Ties in with g-2 on previous slide!

Example: reinterpretation

mapyde





$\tilde{\mu}^{\pm}/\tilde{\chi}^0$ or $\tilde{\chi}^{\pm}/\tilde{\nu}_\mu$

- ATLAS did a "simplified model" search — is it too simple?

ℹ See my tutorial at RiF2022

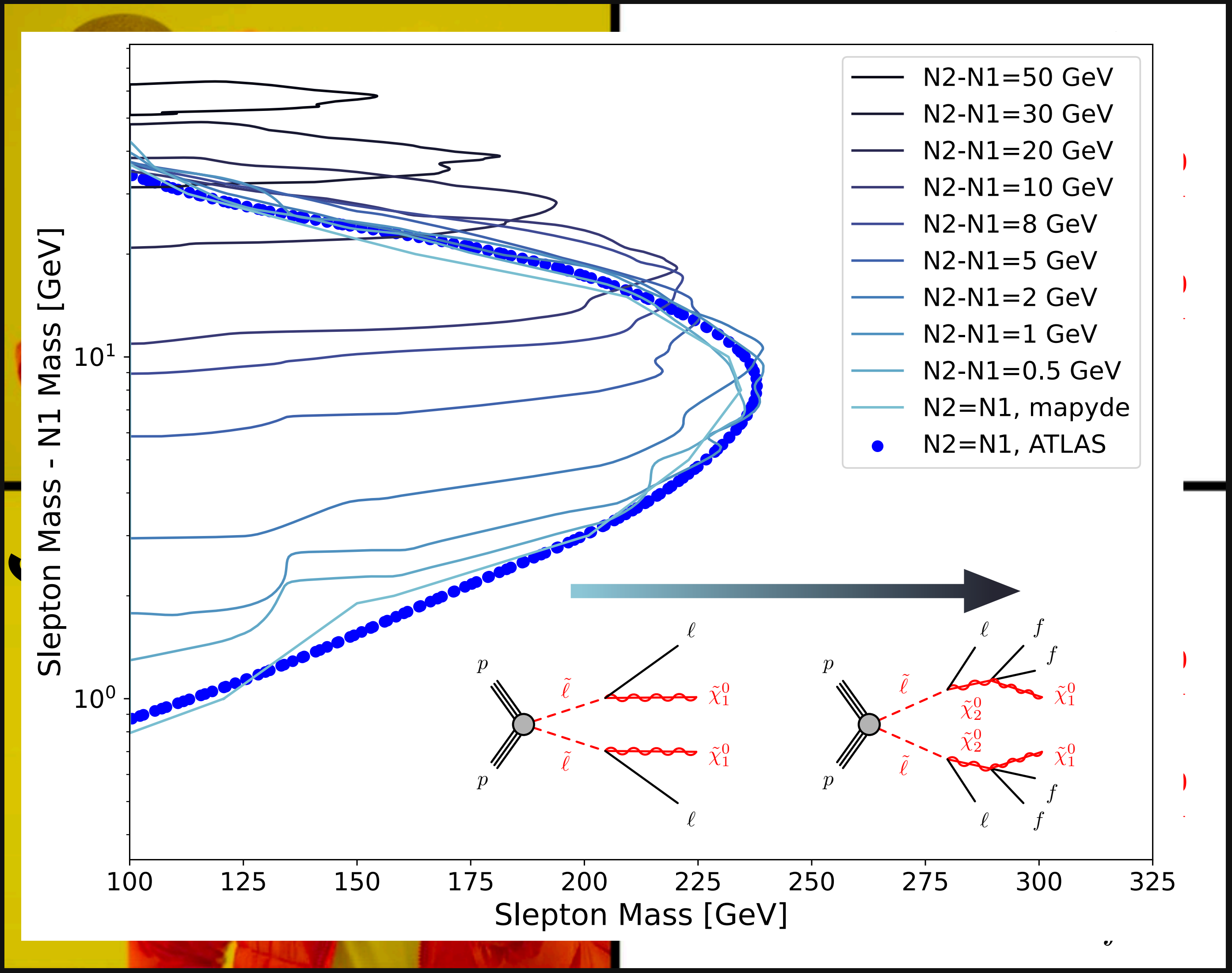
approachable pheno tooling for students(!)

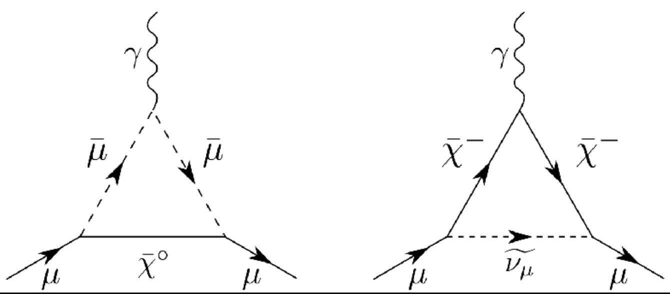
arXiv:2306.11055 arXiv:1911.12606

⚠ **Ties in with g-2 on previous slide!**

Example: reinterpretation

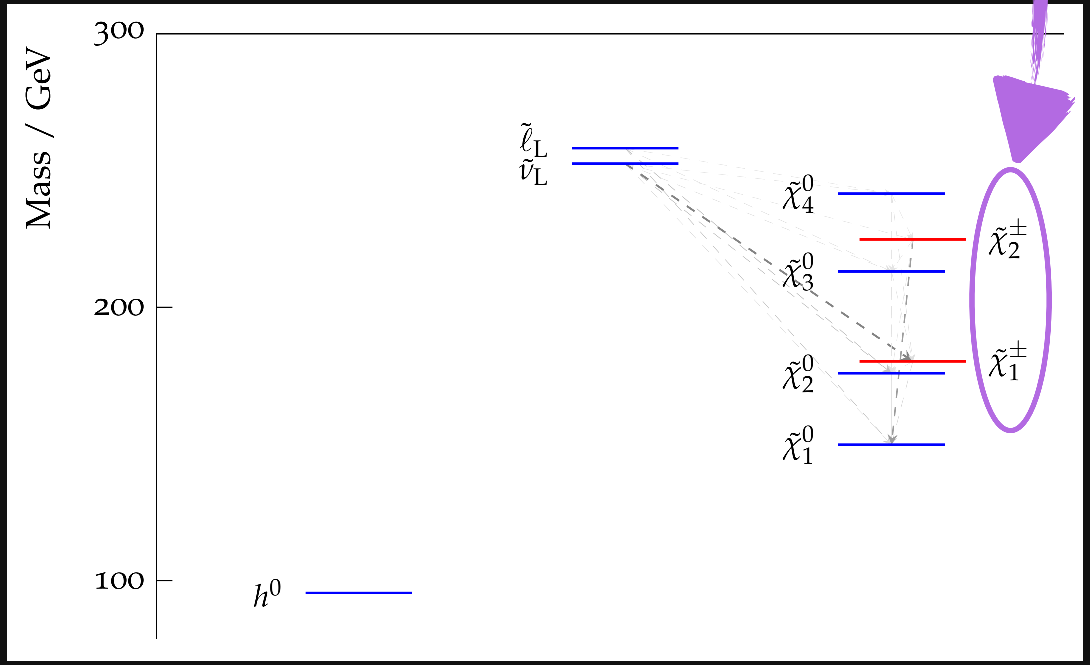
mapyde





$\tilde{\mu}^{\pm}/\tilde{\chi}^0$ or $\tilde{\chi}^{\pm}/\tilde{\nu}_{\mu}$

- Including wino increases phase-space, perhaps contributing more to g-2?



See my tutorial at RiF2022

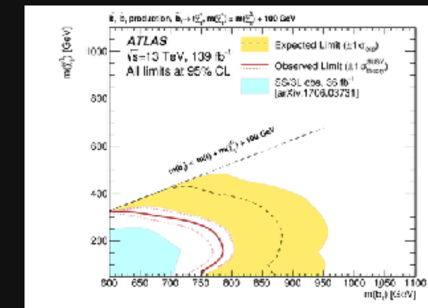
approachable pheno tooling for students(!)

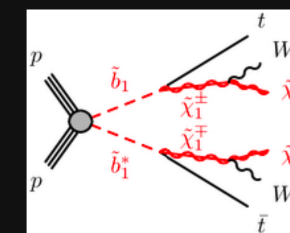
https://twiki.cern.ch/twiki/bin/view/AtlasPublic/SupersymmetryPublicResults

More public models!

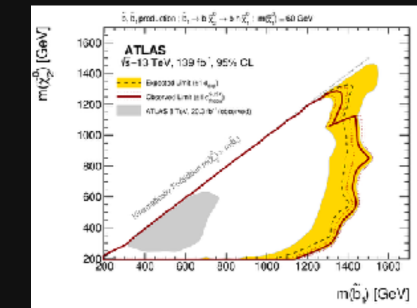
(since 2021!)

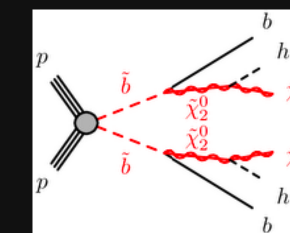
SUSY-2018-09



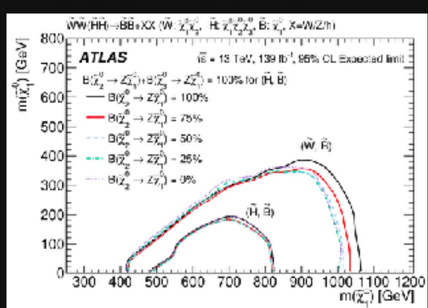


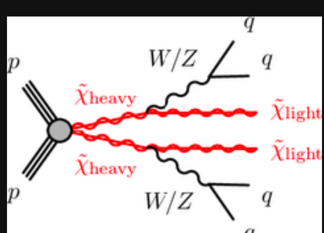
SUSY-2018-31



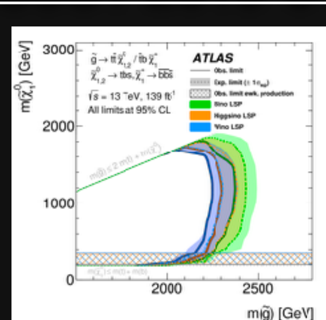


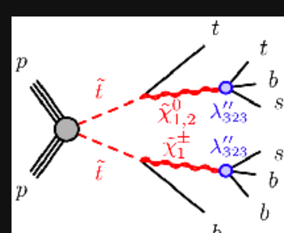
SUSY-2018-41



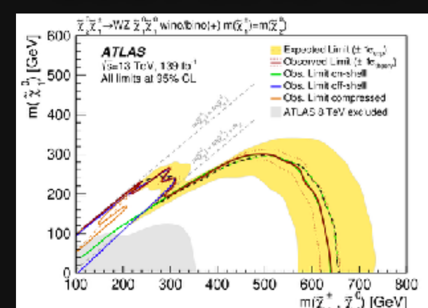


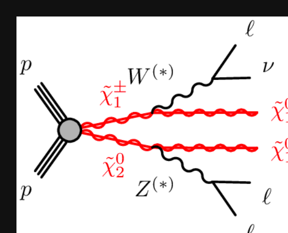
SUSY-2019-04



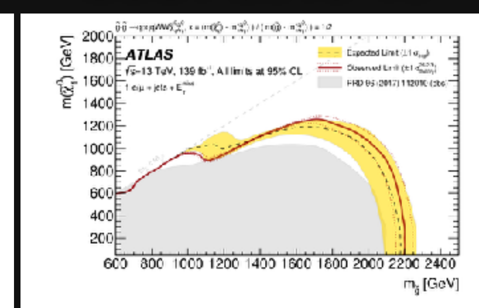


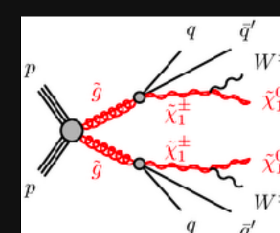
SUSY-2019-09



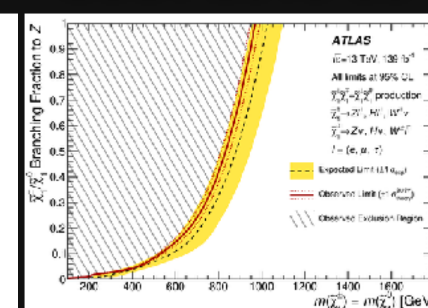


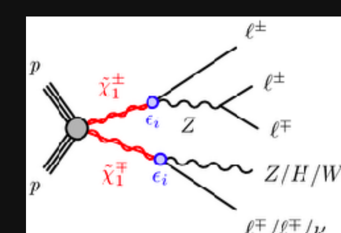
SUSY-2018-10



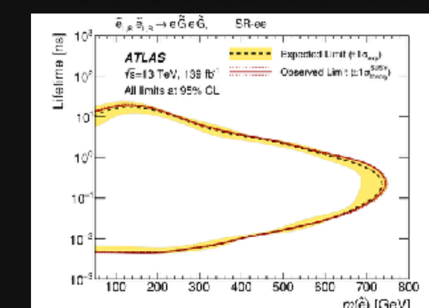


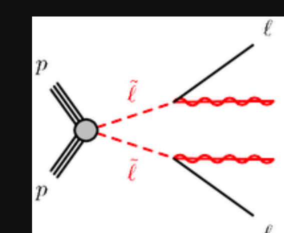
SUSY-2018-36



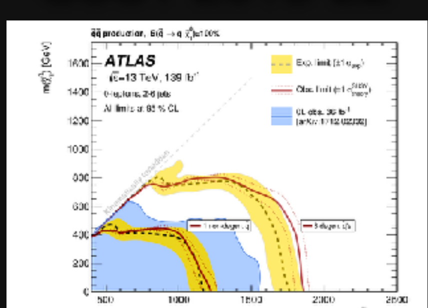


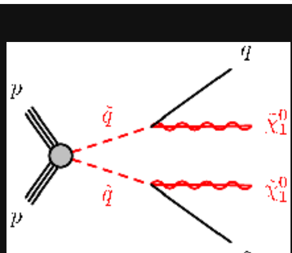
SUSY-2018-14



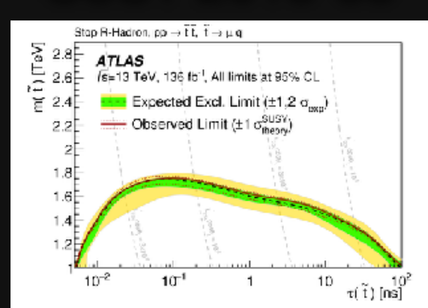


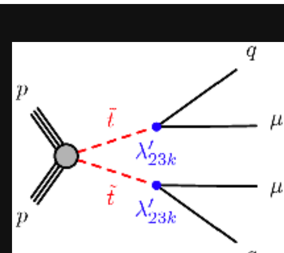
SUSY-2018-22



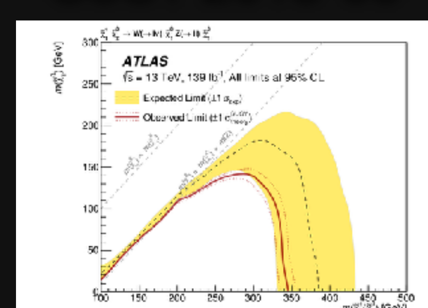


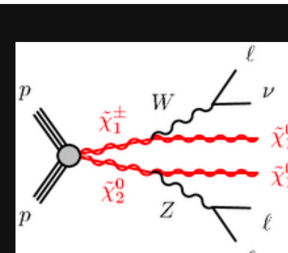
SUSY-2018-33



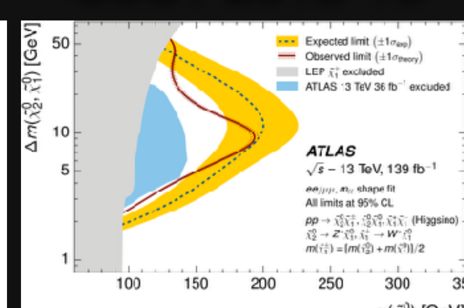


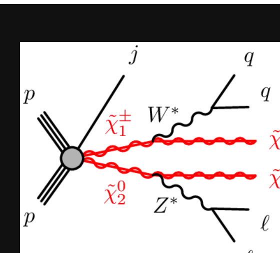
SUSY-2018-06



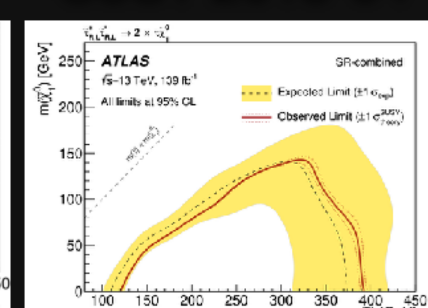


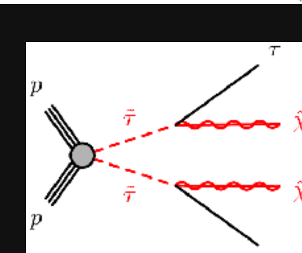
SUSY-2018-16



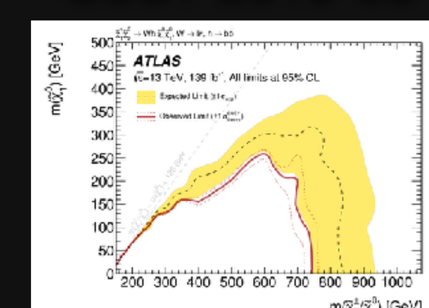


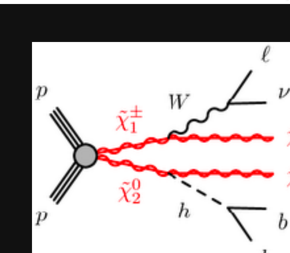
SUSY-2018-04





SUSY-2019-08





❗ Builds on top of my work with pyhf

Integration into theory tools

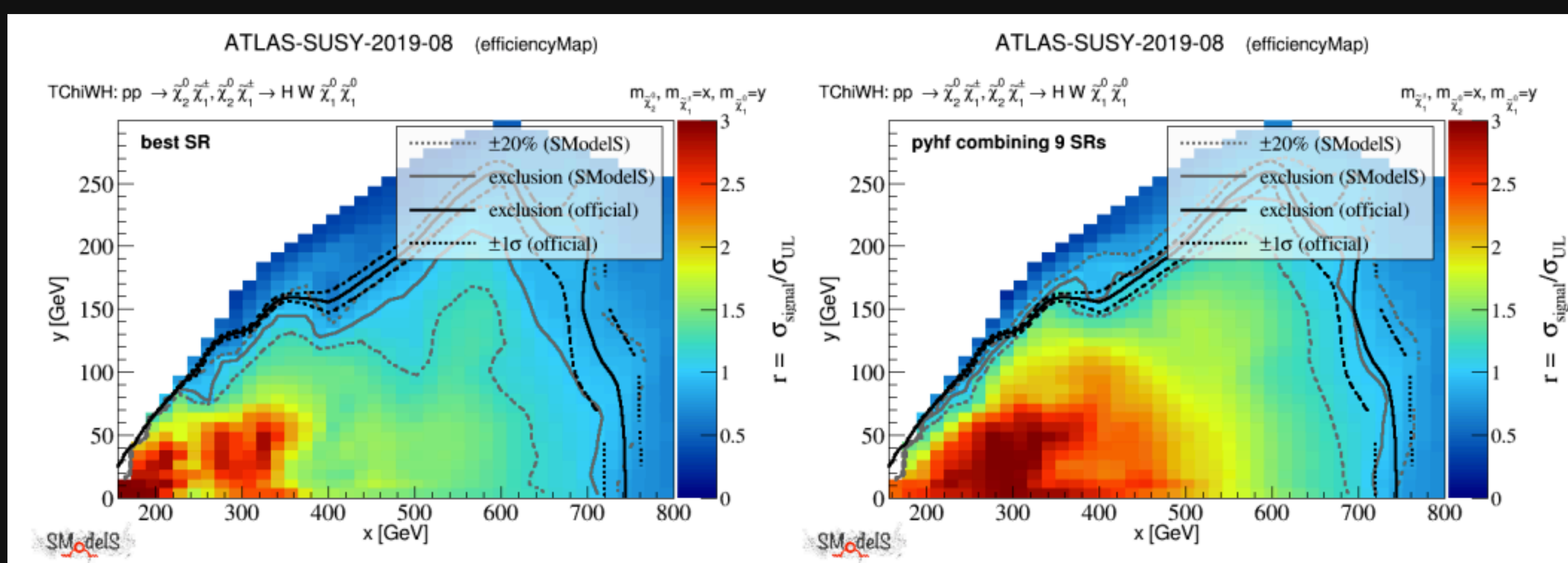
“if you build it, they will come”

[arXiv:2009.01809]

A SModelS interface for pyhf likelihoods

Gaël Alguero[a], Sabine Kraml[a], Wolfgang Waltenberger[b,c]

[a] *Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, 53 Avenue des Martyrs, F-38026 Grenoble, France*
[b] *Institut für Hochenergiephysik, Österreichische Akademie der Wissenschaften, Nikolsdorfer Gasse 18, 1050 Wien, Austria*
[c] *University of Vienna, Faculty of Physics, Boltzmanngasse 5, A-1090 Wien, Austria*



https://scikit-hep.org/pyhf/citations.html

…and other experiments

LHC QCD: arXiv:2012.09120

How to discover QCD Instantons at the LHC[1]

Simone Amoroso[a] Deepak Kar[b] Matthias Schott[2c]
[a] DESY, Hemburg, Germany
[b] University of Witwatersrand, South Africa
[c] Johannes Gutenberg-University, Mainz, Germany
E-mail: matthias.schott@cern.ch

Sensitivity of Future Hadron Colliders to Leptoquark Pair Production in the Di-Muon Di-Jets Channel

B. C. Allanach[1], Tyler Corbett[2], Maeve Madigan[a,1]
[1] DAMTP, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, United Kingdom
[2] The Niels Bohr International Academy, Blegdamsvej 17, University of Copenhagen, DK-2100 Copenhagen, Denmark

FCC: arXiv:1911.04455

EIC: arXiv:2102.06176

Charged Lepton Flavor Violation at the EIC

Vincenzo Cirigliano, Kaori Fuyuto, Christopher Lee, Emanuele Mereghetti, and Bin Yan
Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, U.S.A.
E-mail: cirigliano@lanl.gov, kfuyuto@lanl.gov, clee@lanl.gov, emereghetti@lanl.gov, binyan@lanl.gov

Search for new phenomena in events with two opposite-charge leptons, jets and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS Collaboration

ATLAS: arXiv:2102.01444

SEARCH FOR $B^+ \to K^+ \nu\bar{\nu}$ DECAYS WITH AN INCLUSIVE TAGGING METHOD AT THE BELLE II EXPERIMENT

On the single leptoquark solutions to the B-physics anomalies

Andrei Angelescu,[1,*] Damir Bečirević,[2,†] Darius A. Faroughy,[3,‡] Florentin Jaffredo,[2,§] and Olcyr Sumensari[2,¶]

Belle II: arXiv:2103.12504, arXiv:2105.05754

Searching for dark tridents with the MicroBooNE detector - 05/04/2023

µBooNE

Expected li

- Final BDT/CNN distributions are passed to ***Pyhf*** (

µBooNE: indico:e1261135

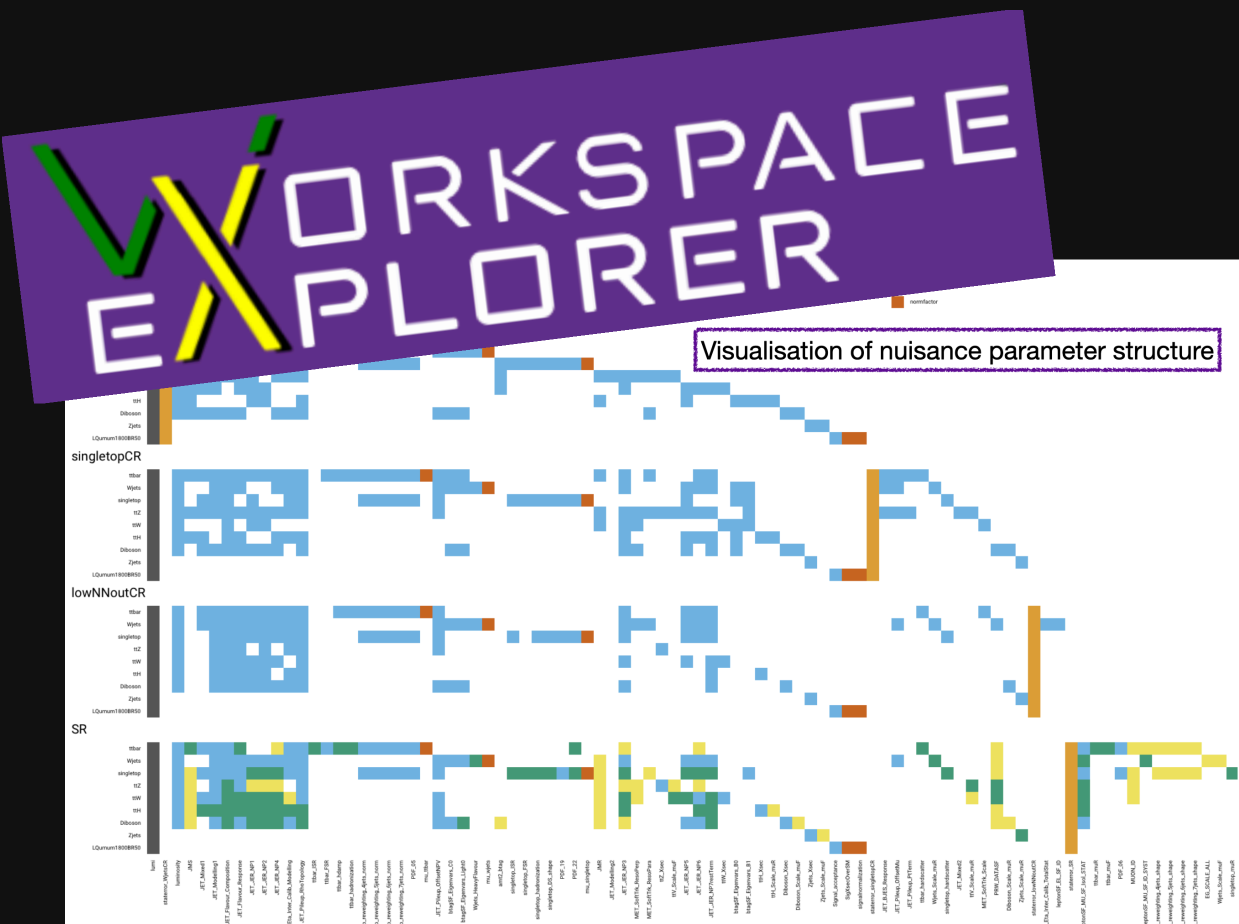
Hunting wino and higgsino dark matter at the muon collider with disappearing tracks

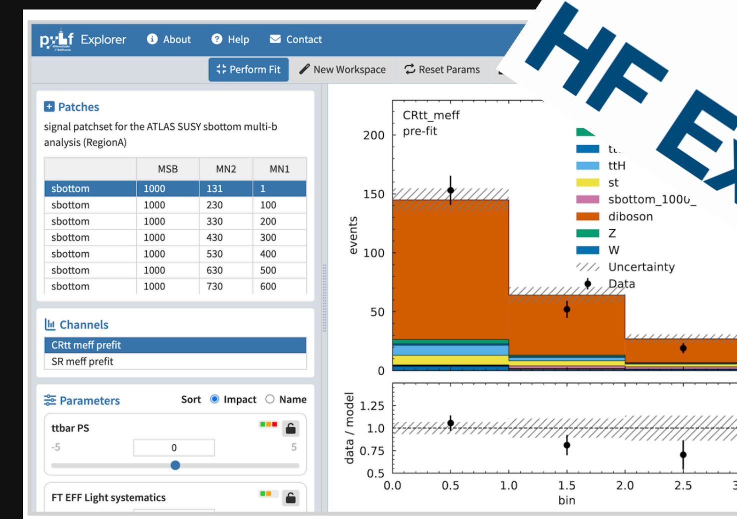
Rodolfo Capdevilla,[a,b] Federico Meloni,[c] Rosa Simoniello,[d] Jose Zurita[e]

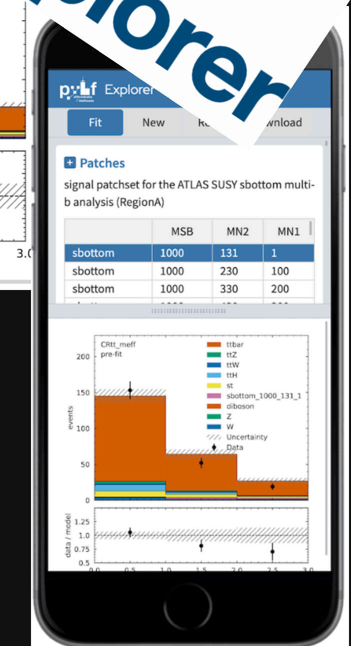
µ-collider: arXiv:2102.11292

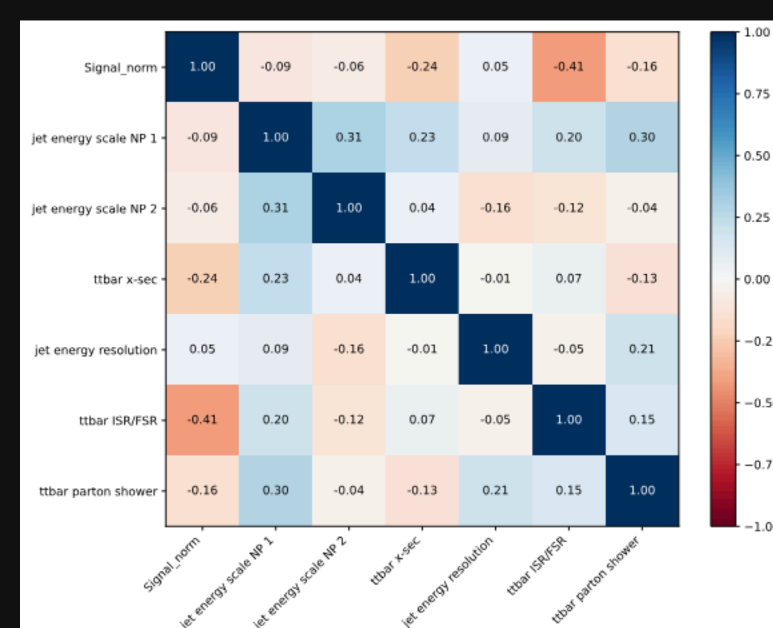
❗ Theory/experiment adoption across the field

… and other tools



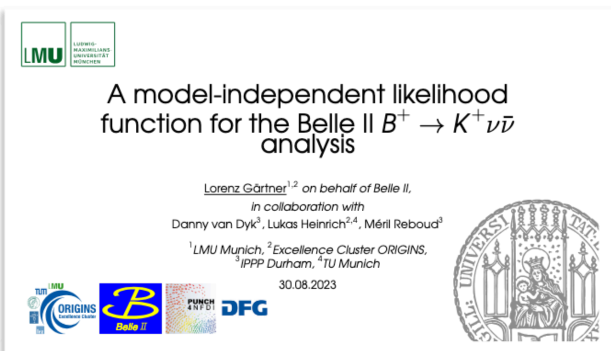


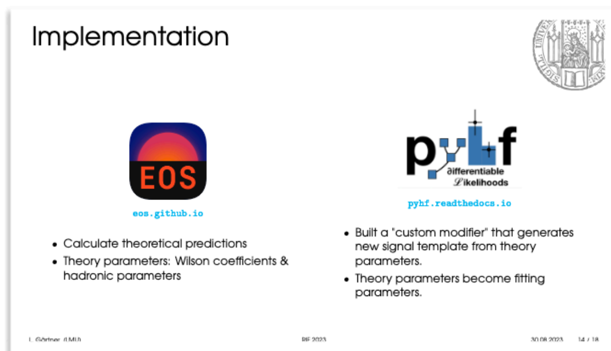


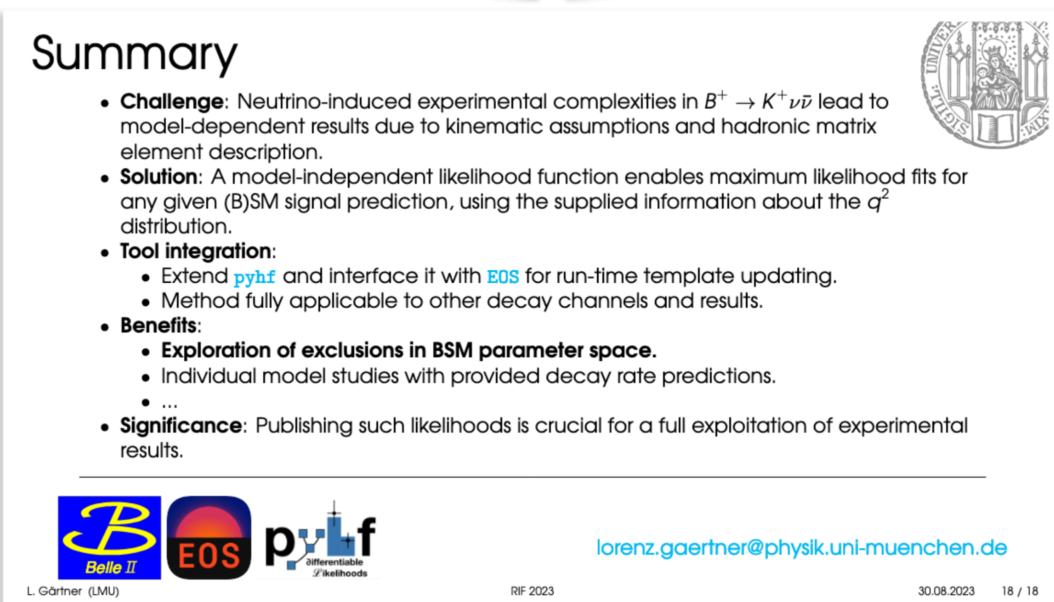


(brief flash) EFT fits

Extending pyhf for EFT fits







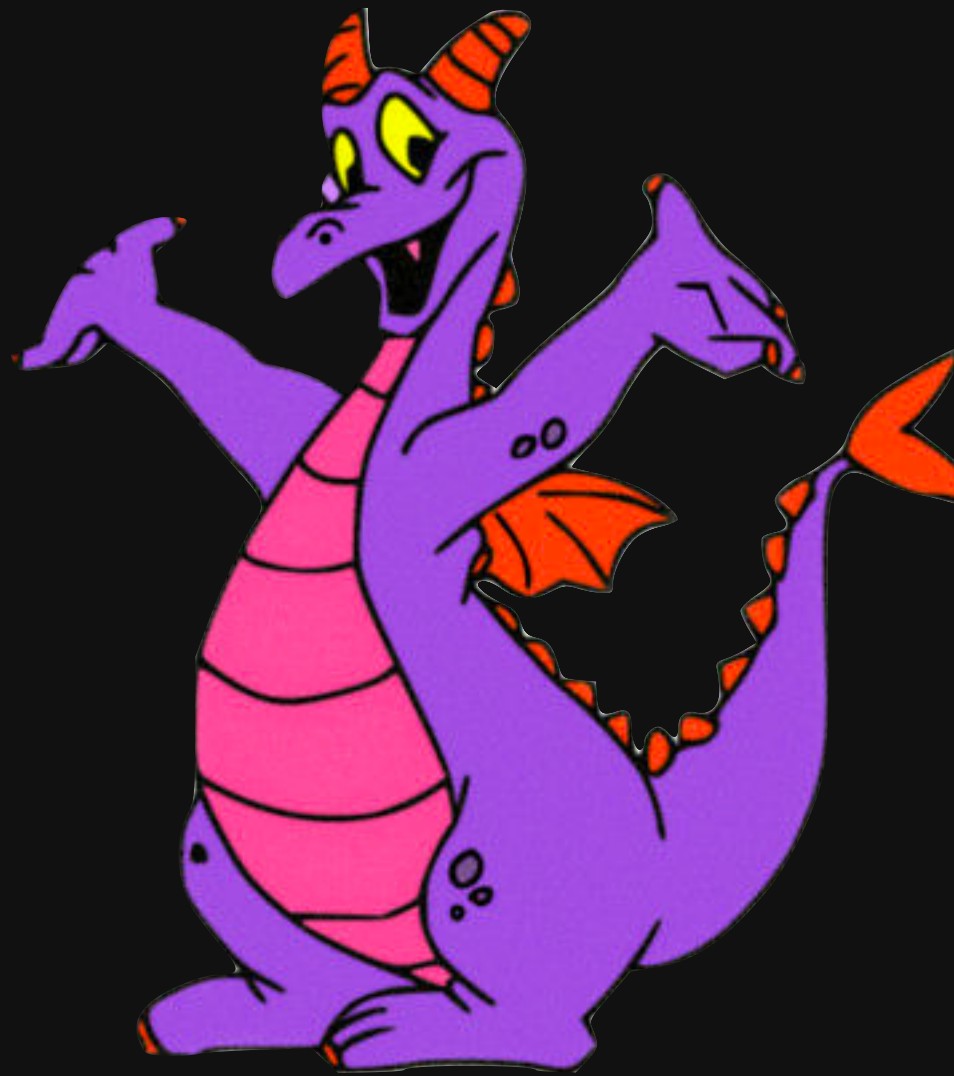
Happy to see pyhf getting attention beyond the LHC Reinterpretation Working Group

- Specifically the EFT working group
- Nice follow-up of BelleII result

https://conference.ippp.dur.ac.uk/event/1178/contributions/6443/
https://indico.cern.ch/event/1296757/timetable/#3-area-4-a-practical-framework

Published Models

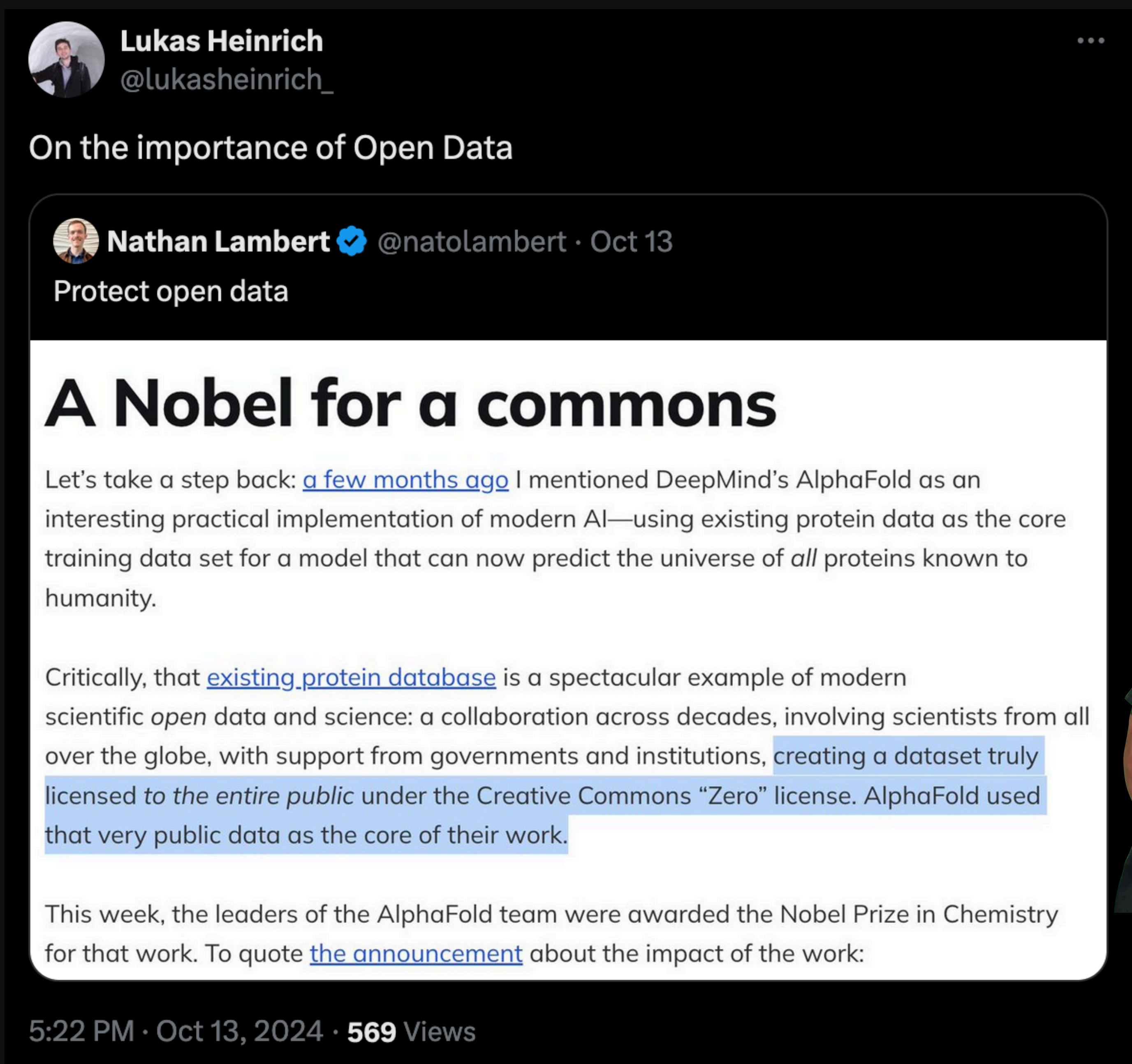
- Because it's plain-text, likelihoods are easy to manipulate, and easy to play with. You can do anything your imagination comes up with.
    - Making new statistical models from old ones (e.g. RECAST/reinterpretation/global fits)
    - Propagating updated theory predictions (and uncertainties) to published statistical models and re-running inference
    - Statistical combinations between different statistical models targeting a similar region of phase-space



arXiv:2109.04981

HEP-wide call to action!

What's the point?


Lukas Heinrich @lukasheinrich_

On the importance of Open Data

Nathan Lambert @natolambert · Oct 13

Protect open data

A Nobel for a commons

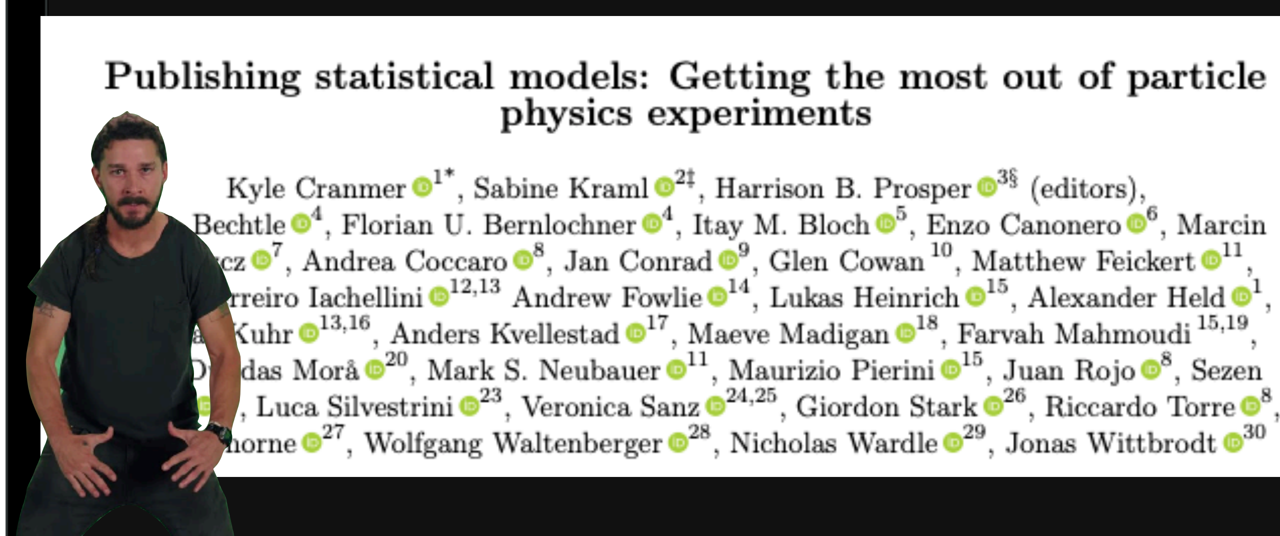
Let's take a step back: a few months ago I mentioned DeepMind's AlphaFold as an interesting practical implementation of modern AI—using existing protein data as the core training data set for a model that can now predict the universe of *all* proteins known to humanity.

Critically, that existing protein database is a spectacular example of modern scientific *open* data and science: a collaboration across decades, involving scientists from all over the globe, with support from governments and institutions, creating a dataset truly licensed *to the entire public* under the Creative Commons "Zero" license. AlphaFold used that very public data as the core of their work.

This week, the leaders of the AlphaFold team were awarded the Nobel Prize in Chemistry for that work. To quote the announcement about the impact of the work:

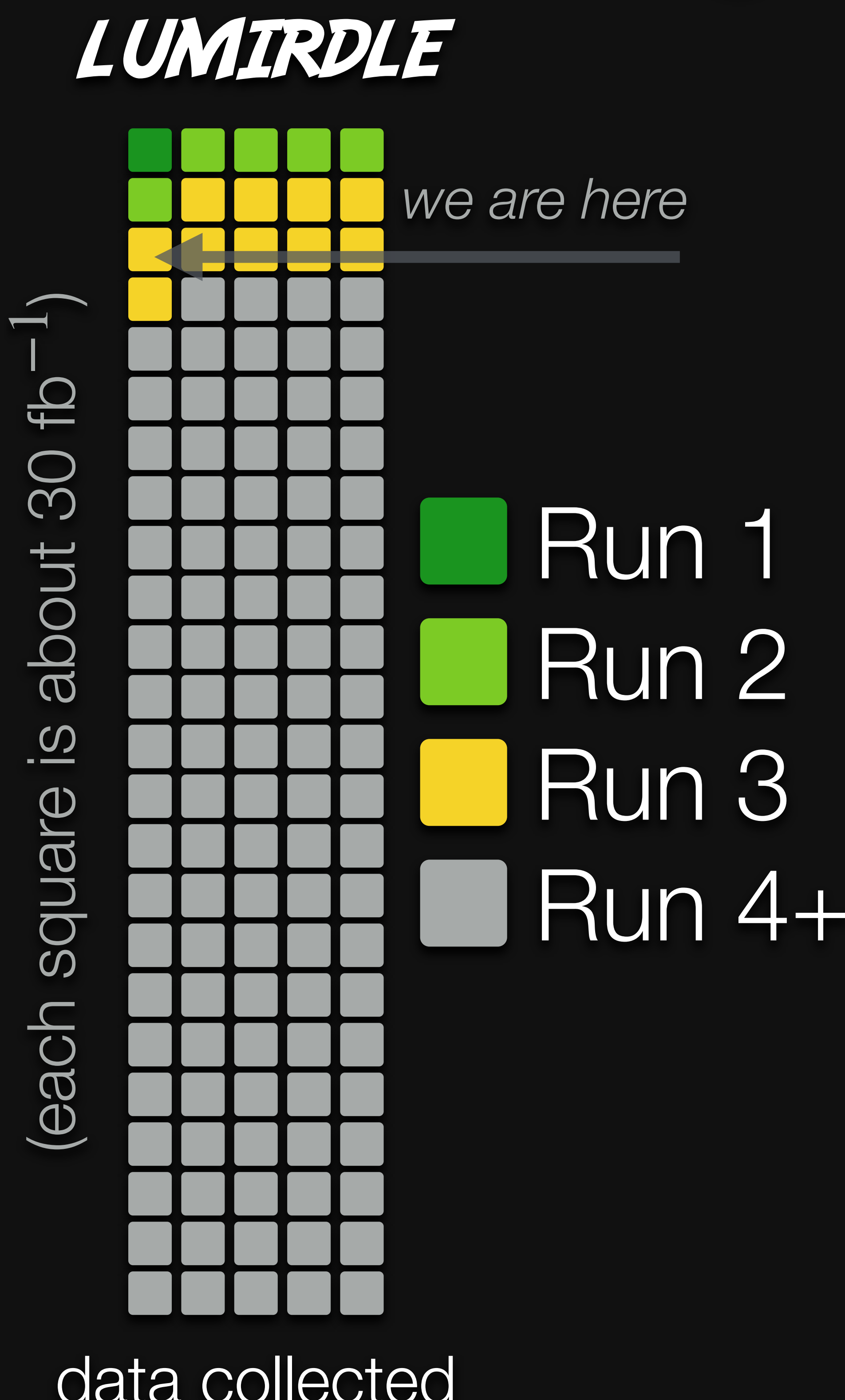
5:22 PM · Oct 13, 2024 · **569** Views

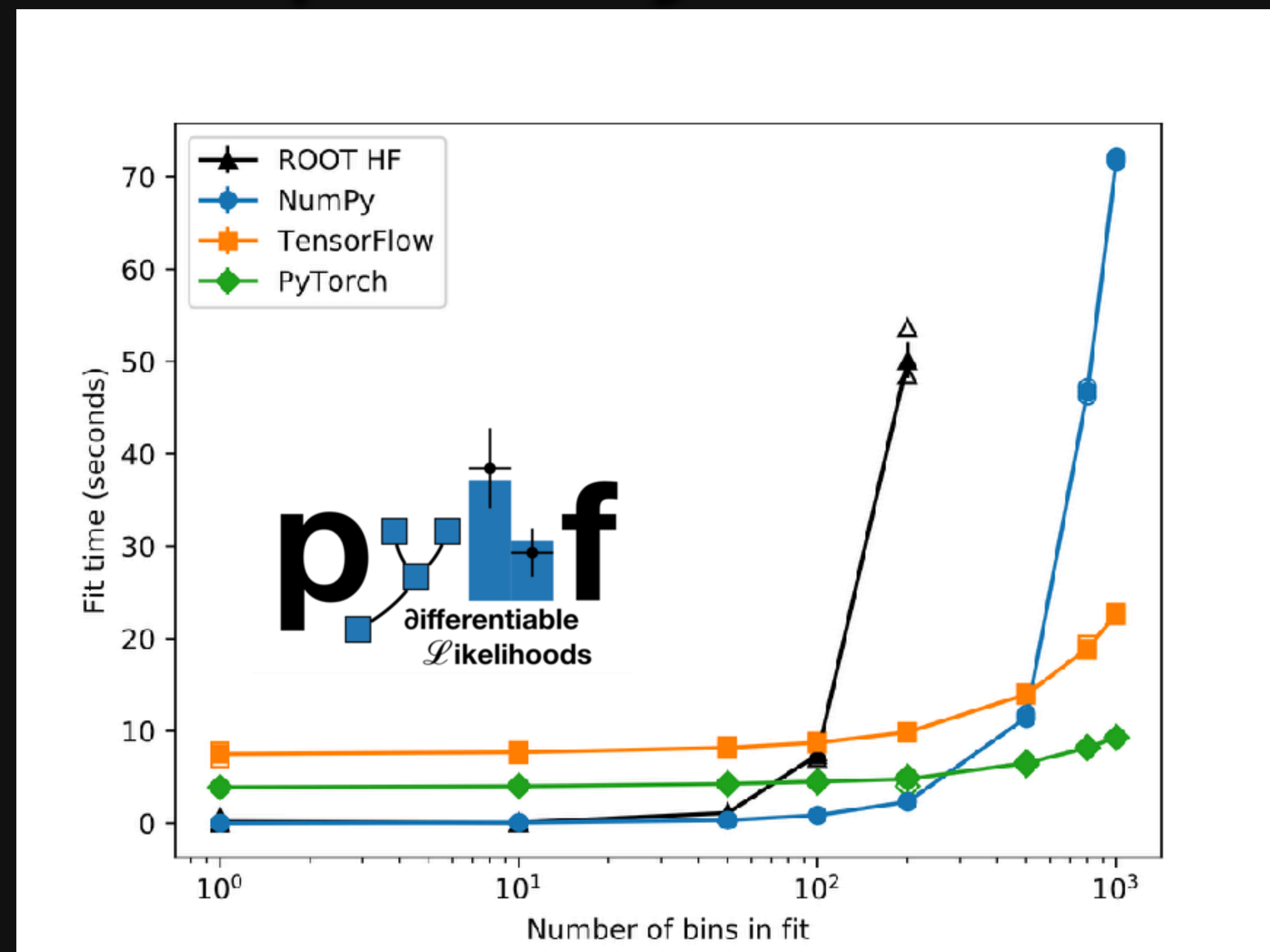
arXiv:2109.04981


Publishing statistical models: Getting the most out of particle physics experiments

Kyle Cranmer[1*], Sabine Kraml[2†], Harrison B. Prosper[3§] (editors), Bechtle[4], Florian U. Bernlochner[4], Itay M. Bloch[5], Enzo Canonero[6], Marcin ...cz[7], Andrea Coccaro[8], Jan Conrad[9], Glen Cowan[10], Matthew Feickert[11], ...rreiro Iachellini[12,13] Andrew Fowlie[14], Lukas Heinrich[15], Alexander Held[1], ...Kuhr[13,16], Anders Kvellestad[17], Maeve Madigan[18], Farvah Mahmoudi[15,19], ...das Morå[20], Mark S. Neubauer[11], Maurizio Pierini[15], Juan Rojo[8], Sezen ..., Luca Silvestrini[23], Veronica Sanz[24,25], Giordon Stark[26], Riccardo Torre[8], ...horne[27], Wolfgang Waltenberger[28], Nicholas Wardle[29], Jonas Wittbrodt[30]

Increasing Complexity

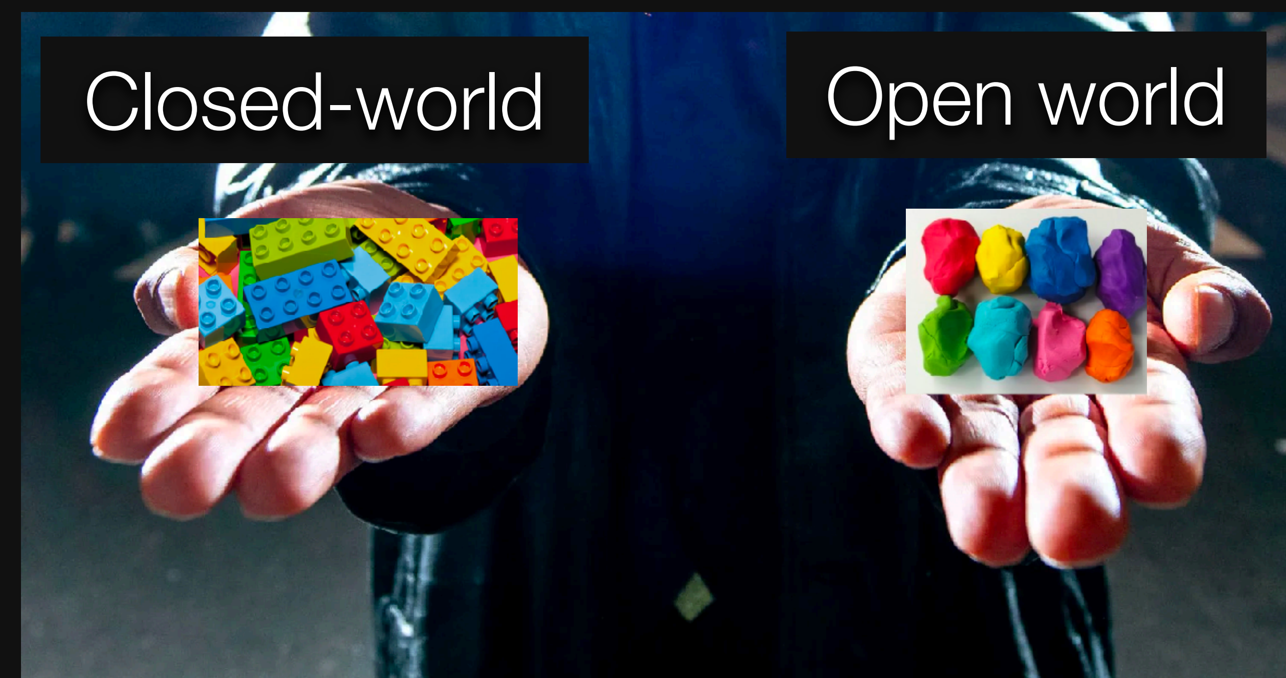




❗ Analyses will evolve in complexity. HL-LHC demands more computing resources.

Diversity of statistical models

- So, now we're at the point of the story where we've got **all the important pieces in place**… except
  - what analysis do we want to publish? **All of them?**
  - arbitrary likelihood functions? Need common tools across experiments!
- This gave us two paths forward:
  - "**closed-world**" implementation: build statistical models from a finite set of building blocks (e.g. linear combinations of Gaussian and Poisson distributions)
  - "**open-world**" implementation: anything goes! This needs work.



pyhf ∂ifferentiable ℒikelihoods

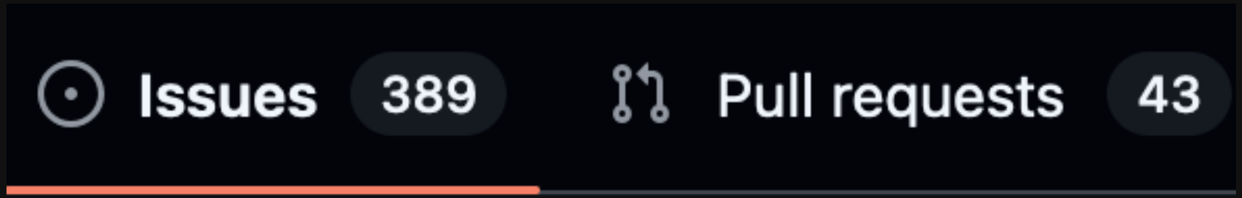
the next step

Serializing the “open world”

- HS3 spec would support RooFit, pyhf, BAT (Julia), and others
- pyhf has work-in-progress to adopt (need more person power)
- **still need an implementation for non-HistFactory models — I know how to do it, but I need help (students, postdocs, anyone who has spare time)!**

Other wishlists/incoming

- HS3 support [!1978, !2397]
- Custom modifiers (arbitrary functional modifiers but within HistFactory) [!1991, !2398]
- Custom test statistics (errors-on-errors, r_u): BIRS Systematics 2023 [E. Canonero, G. Cowan, et. al]
- Parallelized toy calculations [!1158]
- Multi-POI fits, systematic pruning, and handling of non-closure
- Improve documentation, add more examples and tutorials
- …and many more issues, PRs — **we need more personpower**



Tutorial tomorrow!

3:30 PM → 5:00 PM **pyhf-tutorial**

Repository: https://github.com/lorenzennio/pyhf-tutorial/tree/belle2-physics-week

Format: allow everyone to clone the repo and execute as we go

Conveners: Slavomira Stefkova (KIT), Torben Ferber (Karlsruhe Institute of Technology)

3:30 PM	**Pyhf. (setup, simple model, fit)** **Speaker**: Giordon Stark	30m
4:00 PM	**Frequentist inference: CLs limit setting with pyhf** **Speaker**: Slavomira Stefkova (KIT)	30m
4:30 PM	**Bayesian inference with pyhf** **Speaker**: Lorenz Ennio Gaertner (BELLE (BELLE II Experiment))	30m

- Join the hands-on tutorial tomorrow with Sally, Lorenz, and yours truly
- You do not need a computer to follow along the jupyter notebook tutorial if you're just curious how easy it is!

Summary

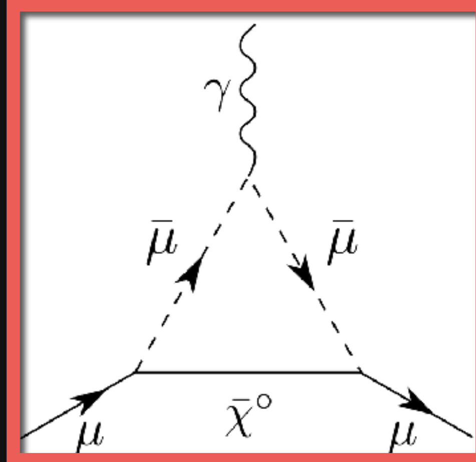
- Accelerated fitting library
  - reducing time to inference/insight
  - Hardware acceleration on CPU and GPU
- Fully-open sourced and developed on GitHub
  - user guides, tutorials, and contributions welcome!
- Key things to move forward on:
  - standardize naming of channels, regions, systematics
  - integrate HS3 more widely
  - improve tools for visualization, debugging, inspection, and validation

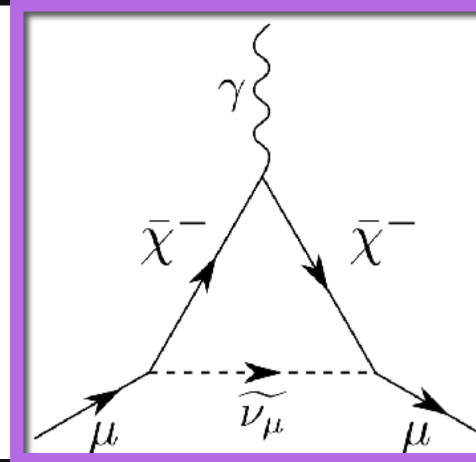
Backup

want $a_\mu \neq 0$ to match experimental evidence

Example: global fit (g-2) (II)

$\tilde{\mu}^\pm/\tilde{\chi}^0$ or $\tilde{\chi}^\pm/\tilde{\nu}_\mu$





$$a_\mu^{\tilde{\chi}^0-\tilde{\mu}} \approx \frac{\alpha m_\mu^2 M_1(\mu \tan\beta - A_\mu)}{4\pi \cos^2\theta_W (m_{\tilde{\mu}_R}^2 - m_{\tilde{\mu}_L}^2)} \left[\frac{f_{\chi^0}\left(M_1^2/m_{\tilde{\mu}_R}^2\right)}{m_{\tilde{\mu}_R}^2} - \frac{f_{\chi^0}\left(M_1^2/m_{\tilde{\mu}_L}^2\right)}{m_{\tilde{\mu}_L}^2} \right]$$

$$a_\mu^{\tilde{\chi}^\pm-\tilde{\nu}_\mu} \approx \frac{\alpha m_\mu^2 M_2 \tan\beta}{4\pi \sin^2\theta_W m_{\tilde{\nu}_\mu}^2} \left[\frac{f_{\chi^\pm}\left(M_2^2/m_{\tilde{\nu}_\mu}^2\right) - f_{\chi^\pm}\left(\mu^2/m_{\tilde{\nu}_\mu}^2\right)}{M_2^2 - \mu^2} \right]$$

! Light smuon/light neutralino needed!

⭐ compressed is crucial ⁈ — publish Wave-1 likelihoods?

Summary: g-2 EWK pMSSM

- EWK sector of MSSM can account for the g-2 discrepancy — assuming N1 is the DM candidate (relic density constraint applied as upper limit)

- Due to lower xsec in EWK, lighter scenarios are preferred
  - [a] $\mu < M_1, M_2$ — higgsino DM
  - [b] $M_2 < M_1, \mu$ — wino DM
  - [c] $M_1 < M_2, \mu$ — mixed bino/wino DM in two ways
    - [c_1] C_1 coannihilation ($M_1 < M_2$, see LianTao's talk 🔗)
    - [c_2] slepton coannihilation (bino DM, slepton mass close to LSP)

- Using relic density, upper limits on m(LSP) are: [a] 500 GeV, [b] 600 GeV, [c_1] 600 GeV, [c_2] ~500 GeV

- Additionally constraining using the assumption that FNAL concurs with BNL, upper limits are decreased: [a] 480 GeV, [b] 500 GeV, [c_1] 500 GeV, [c_2] 380-450 GeV