

# MACHINE LEARNING THE LIKELIHOODS

Rafał Masełek  
(ラファウ・マセウエク)

in collaboration with

J. Araz, A. Butter, J. Iturizza, S. Kraml, H. Reyes-Gonzalez, and W. Waltenberger

KEK Theory Meeting on Particle Physics Phenomenology, 18-02-2026



# Status of the Particle Physics

## Standard Model of Elementary Particles

three generations of matter (fermions)			interactions / force carriers (bosons)				
	I	II	III				
QUARKS	mass $\approx 2.16 \text{ MeV}/c^2$ charge $\frac{2}{3}$ spin $\frac{1}{2}$ <b>u</b> up	mass $\approx 1.273 \text{ GeV}/c^2$ charge $\frac{2}{3}$ spin $\frac{1}{2}$ <b>c</b> charm	mass $\approx 172.57 \text{ GeV}/c^2$ charge $\frac{2}{3}$ spin $\frac{1}{2}$ <b>t</b> top	0 0 1 <b>g</b> gluon	SCALAR BOSONS	mass $\approx 125.2 \text{ GeV}/c^2$ 0 0 0 <b>H</b> higgs	
	mass $\approx 4.7 \text{ MeV}/c^2$ charge $-\frac{1}{3}$ spin $\frac{1}{2}$ <b>d</b> down	mass $\approx 93.5 \text{ MeV}/c^2$ charge $-\frac{1}{3}$ spin $\frac{1}{2}$ <b>s</b> strange	mass $\approx 4.183 \text{ GeV}/c^2$ charge $-\frac{1}{3}$ spin $\frac{1}{2}$ <b>b</b> bottom	0 0 1 <b><math>\gamma</math></b> photon			
	mass $\approx 0.511 \text{ MeV}/c^2$ charge $-1$ spin $\frac{1}{2}$ <b>e</b> electron	mass $\approx 105.66 \text{ MeV}/c^2$ charge $-1$ spin $\frac{1}{2}$ <b><math>\mu</math></b> muon	mass $\approx 1.77693 \text{ GeV}/c^2$ charge $-1$ spin $\frac{1}{2}$ <b><math>\tau</math></b> tau	mass $\approx 91.188 \text{ GeV}/c^2$ 0 1 <b>Z</b> Z boson			GAUGE BOSONS VECTOR BOSONS
	mass $< 0.8 \text{ eV}/c^2$ charge 0 spin $\frac{1}{2}$ <b><math>\nu_e</math></b> electron neutrino	mass $< 0.17 \text{ MeV}/c^2$ charge 0 spin $\frac{1}{2}$ <b><math>\nu_\mu</math></b> muon neutrino	mass $< 18.2 \text{ MeV}/c^2$ charge 0 spin $\frac{1}{2}$ <b><math>\nu_\tau</math></b> tau neutrino	mass $\approx 80.3692 \text{ GeV}/c^2$ $\pm 1$ 1 <b>W</b> W boson			

! Dark Matter

! matter-antimatter asymmetry

! neutrino masses

! hierarchy problem

! strong CP problem

! quantum gravity?

IMG source: [https://en.wikipedia.org/wiki/Standard\\_Model](https://en.wikipedia.org/wiki/Standard_Model)

# Status of the Particle Physics

## Standard Model of Elementary Particles

	three generations of matter (fermions)			interactions / force carriers (bosons)	
	I	II	III		
mass	$\approx 2.16 \text{ MeV}/c^2$	$\approx 1.273 \text{ GeV}/c^2$	$\approx 172.57 \text{ GeV}/c^2$	0	$\approx 125.2 \text{ GeV}/c^2$

! Dark Matter

! hierarchy problem

! matter-antimatter asymmetry

**! WE NEED BSM PHYSICS**

! strong CP problem

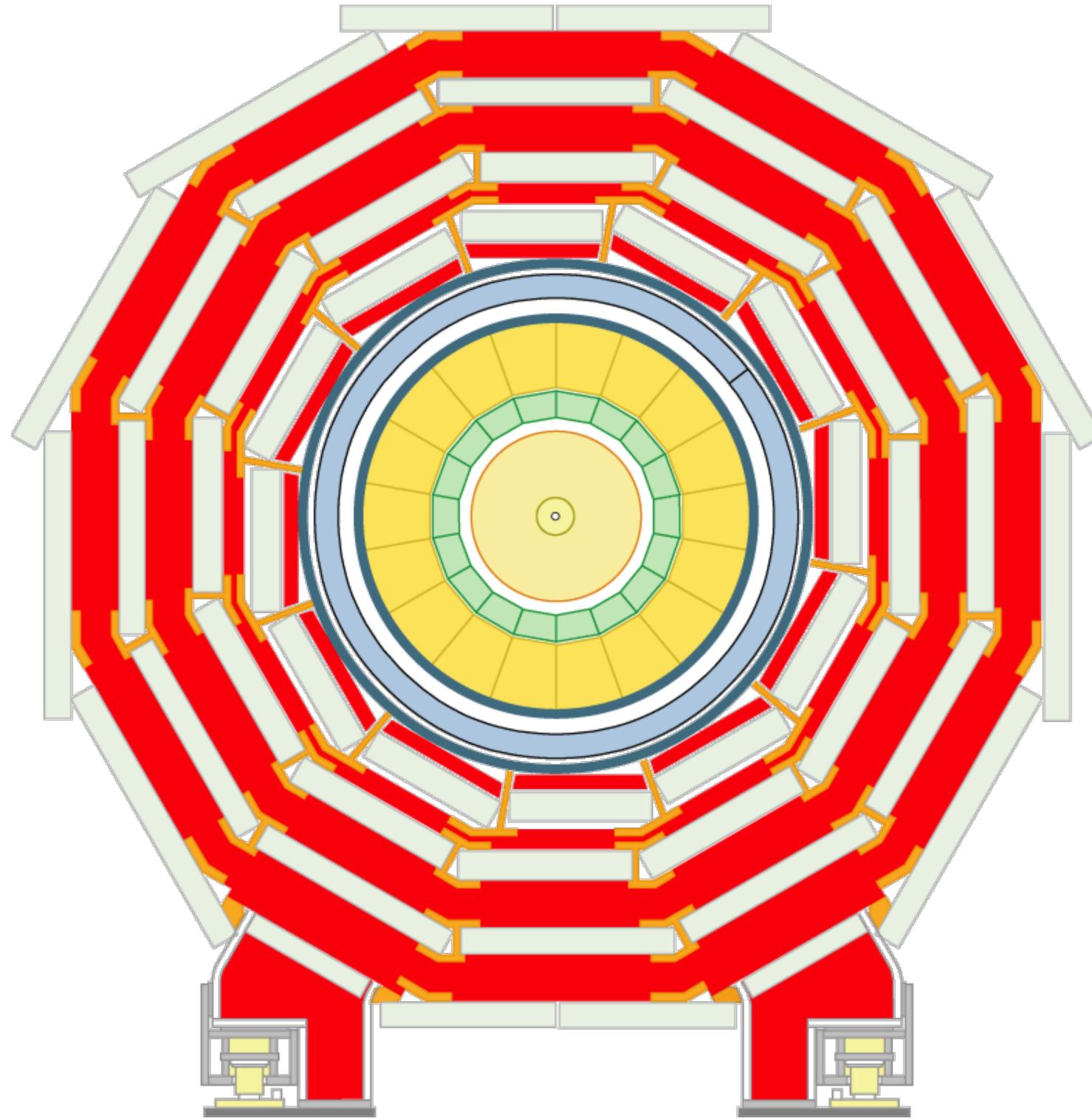
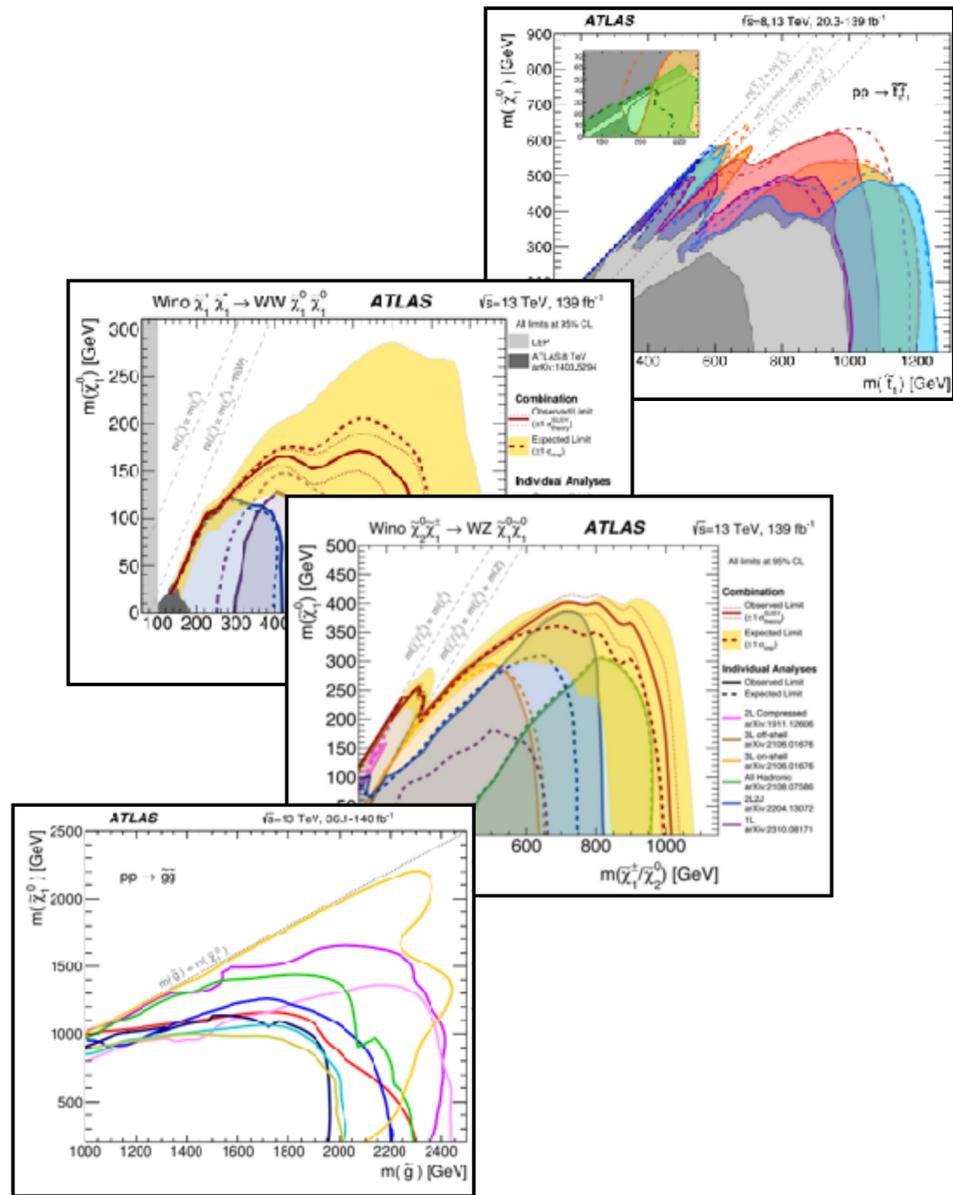
! neutrino masses

LEPTONS	LEPTONS			GAUGE BOSON VECTOR BOSONS	
	$\frac{1}{2}$ $e$ electron	$\frac{1}{2}$ $\mu$ muon	$\frac{1}{2}$ $\tau$ tau	1 $Z$ Z boson	
	$< 0.8 \text{ eV}/c^2$ 0 $\nu_e$ electron neutrino	$< 0.17 \text{ MeV}/c^2$ 0 $\nu_\mu$ muon neutrino	$< 18.2 \text{ MeV}/c^2$ 0 $\nu_\tau$ tau neutrino	$\approx 80.3692 \text{ GeV}/c^2$ $\pm 1$ 1 $W$ W boson	

! quantum gravity?

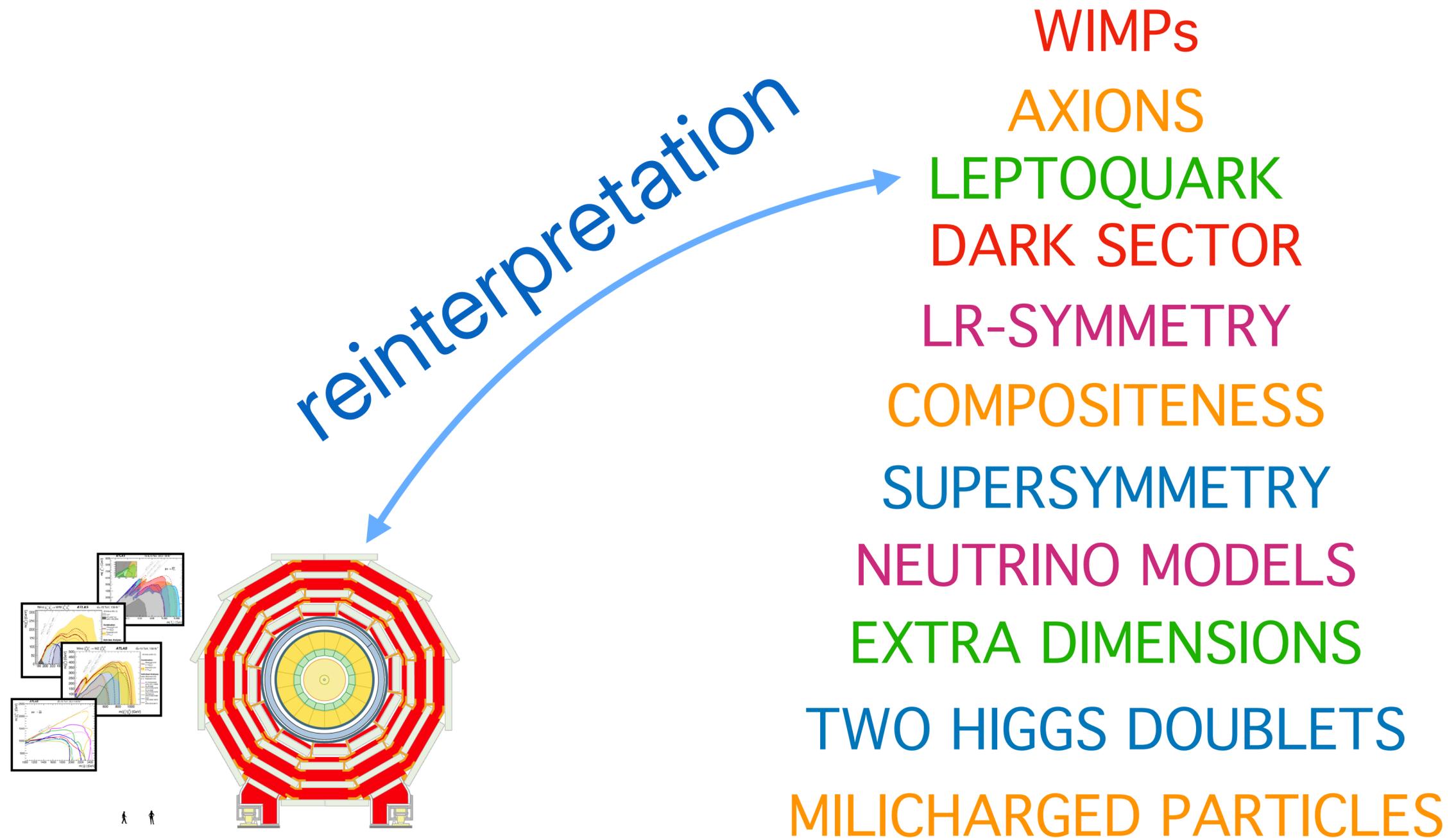
IMG source: [https://en.wikipedia.org/wiki/Standard\\_Model](https://en.wikipedia.org/wiki/Standard_Model)

# Searches for BSM at the LHC



plots from: <https://doi.org/10.1016/j.physrep.2024.09.010>

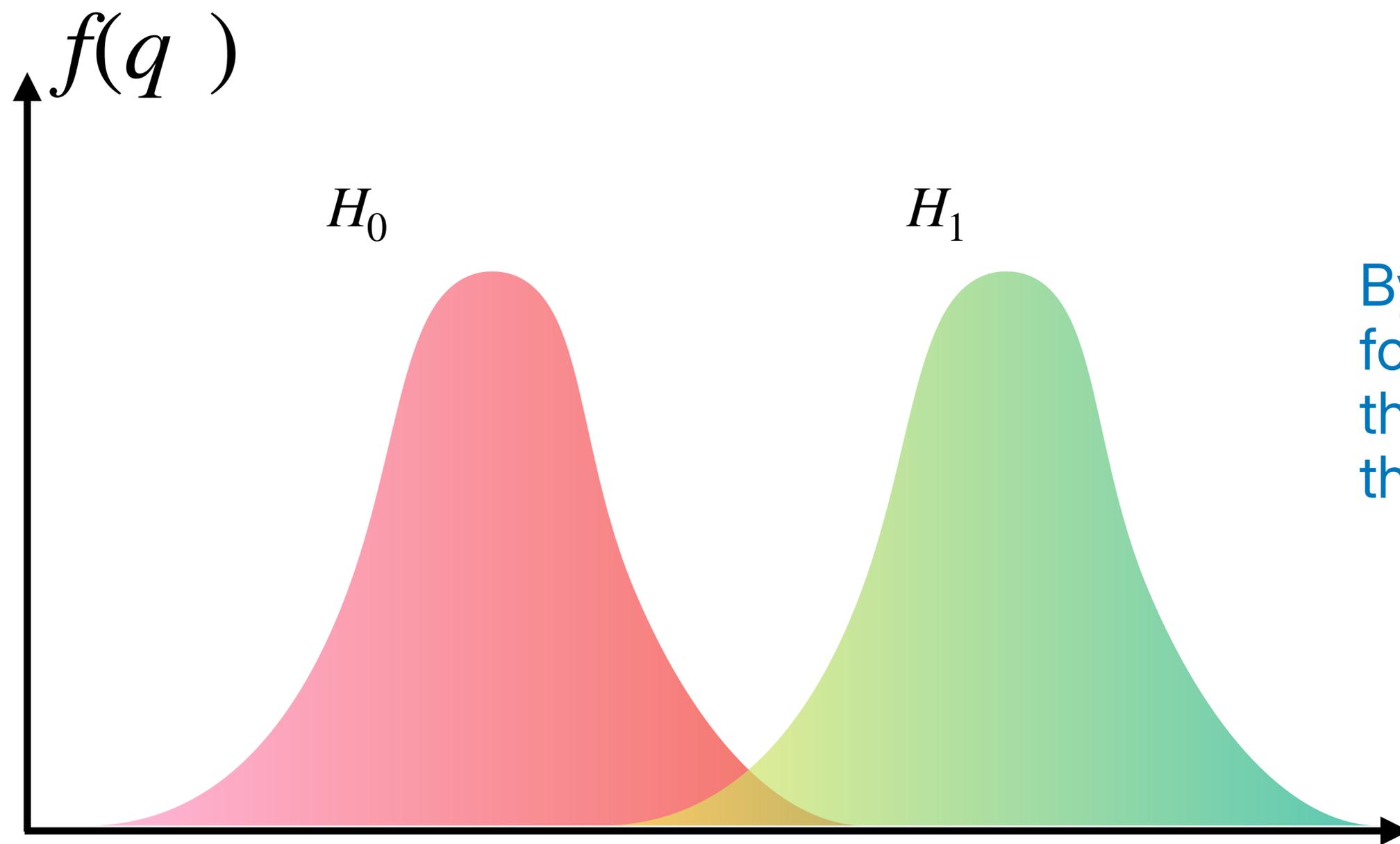
# Searches for BSM at the LHC



# Reinterpretation



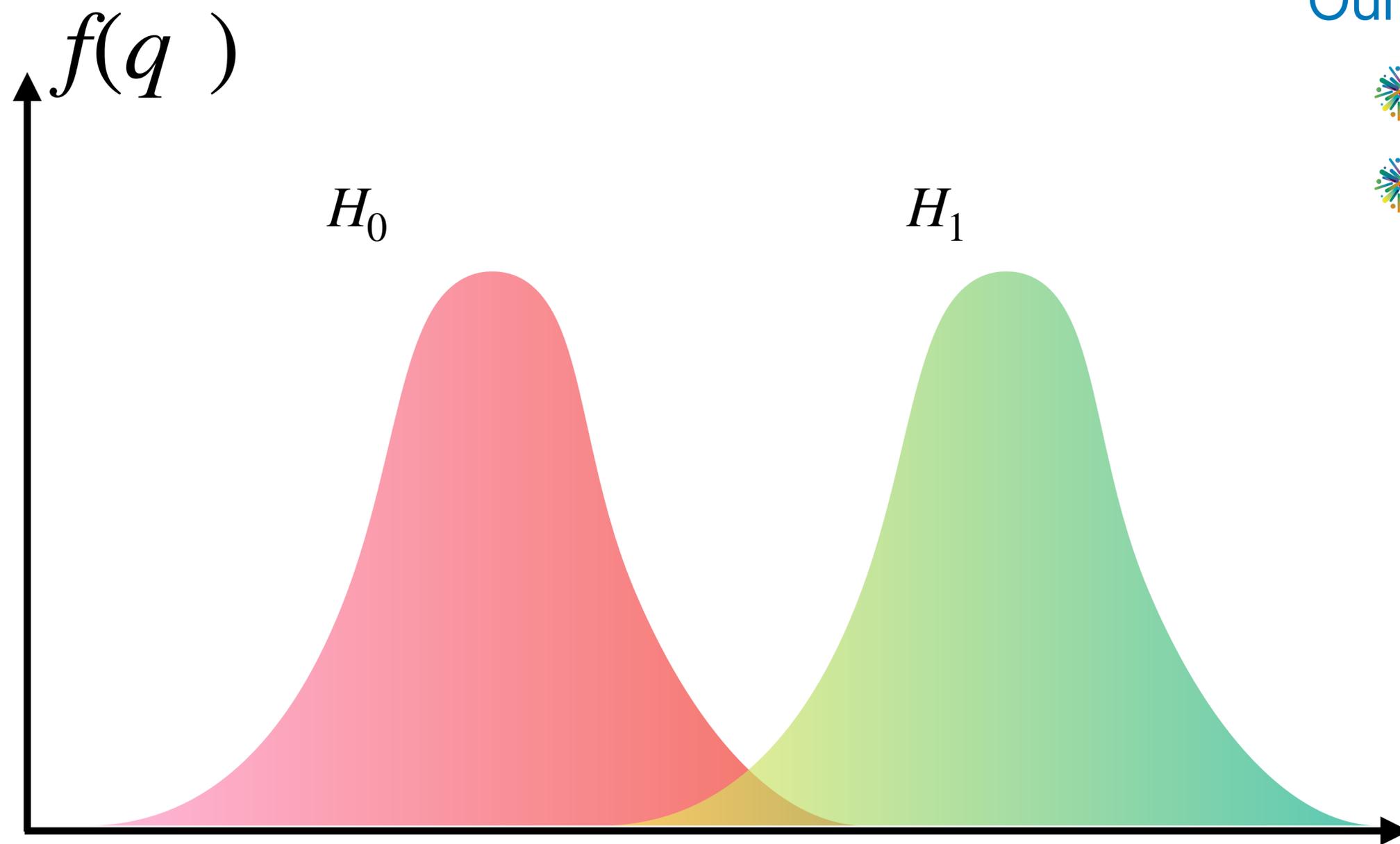
# Hypothesis testing



By the Neyman-Pearson lemma, for testing simple hypotheses the most powerful test statistic is the likelihood ratio

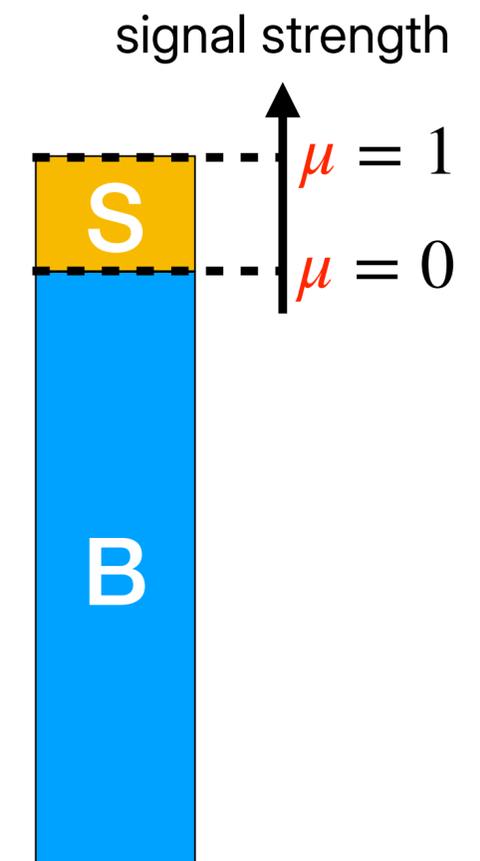
$$q = -2 \ln \frac{L(\text{data} | H_0)}{L(\text{data} | H_1)}$$

# Hypothesis testing

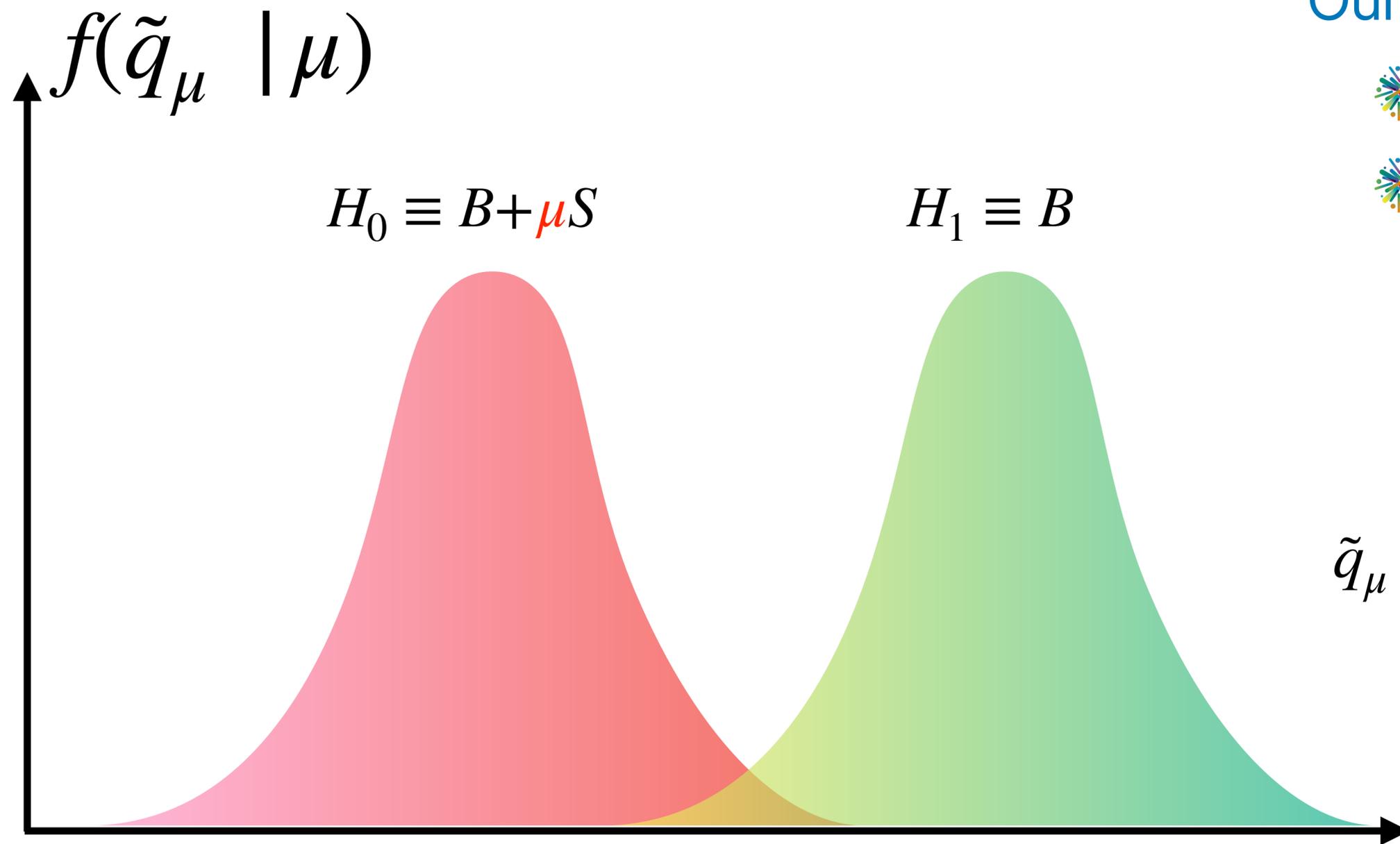


Our hypotheses are composite:

- ✿ We have nuisance parameters  $\theta$
- ✿ We have a parameter of interest  $\mu$



# Hypothesis testing



Our hypotheses are composite:

- ✿ We have nuisance parameters  $\theta$
- ✿ We have a parameter of interest  $\mu$

$$\tilde{q}_\mu = \begin{cases} -2 \ln \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(0, \hat{\hat{\theta}}(0))} & \hat{\mu} < 0, \\ -2 \ln \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\theta})} & 0 \leq \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu \end{cases}$$

$\hat{\square}$  unconditional MLE

$\hat{\hat{\square}}(\mu)$  conditional MLE

# Full likelihood models by ATLAS (Hist Factory)

$$L(n, a | \mu, \theta) = \prod_c^{\text{channels}} \prod_b^{\text{bins}_c} \text{Pois}(n_{cb} | \nu_{cb}(\mu, \theta)) \prod_{\theta} p_{\theta}(a_{\theta} | \theta)$$

# Full likelihood models by ATLAS (Hist Factory)

channel data

auxiliary data

free parameters

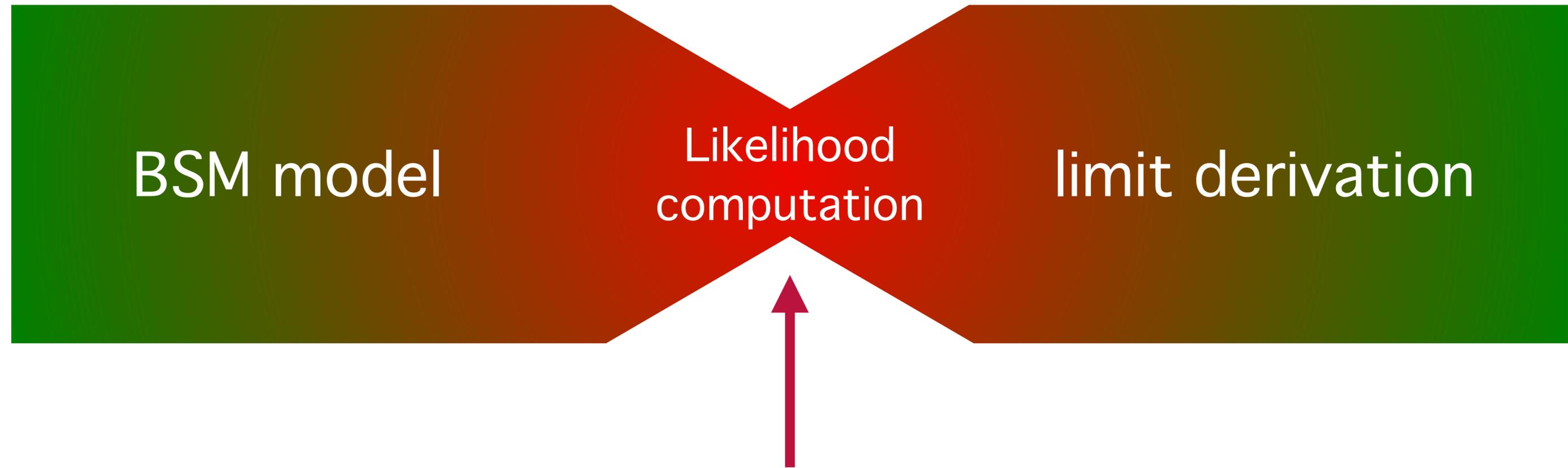
constrained parameters

$$L(n, a | \mu, \theta) = \prod_c \prod_{b \in \text{bins}_c} \text{Pois}(n_{cb} | \nu_{cb}(\mu, \theta)) \prod_{\theta} p_{\theta}(a_{\theta} | \theta)$$

simultaneous measurement of multiple channels

constraint terms for "auxiliary measurements"

# Computational bottleneck



Full statistical model calculations enter here

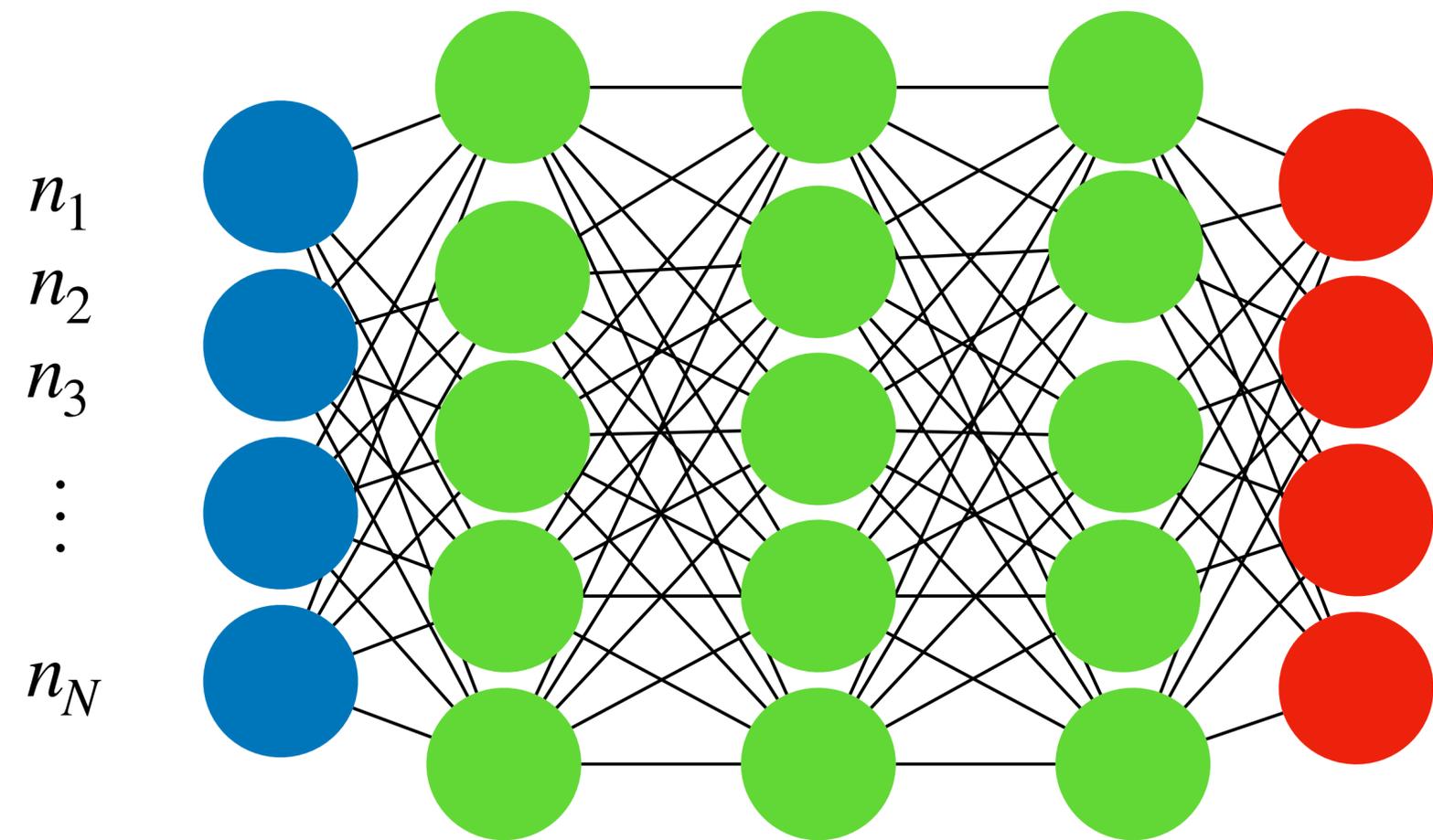
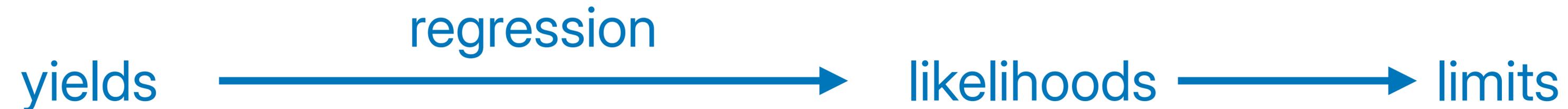
# Computational bottleneck



Machine learning enters here

Use Machine Learning to create surrogates to Full Statistical Models that would be **accurate** but **much faster** to use.

# The approach



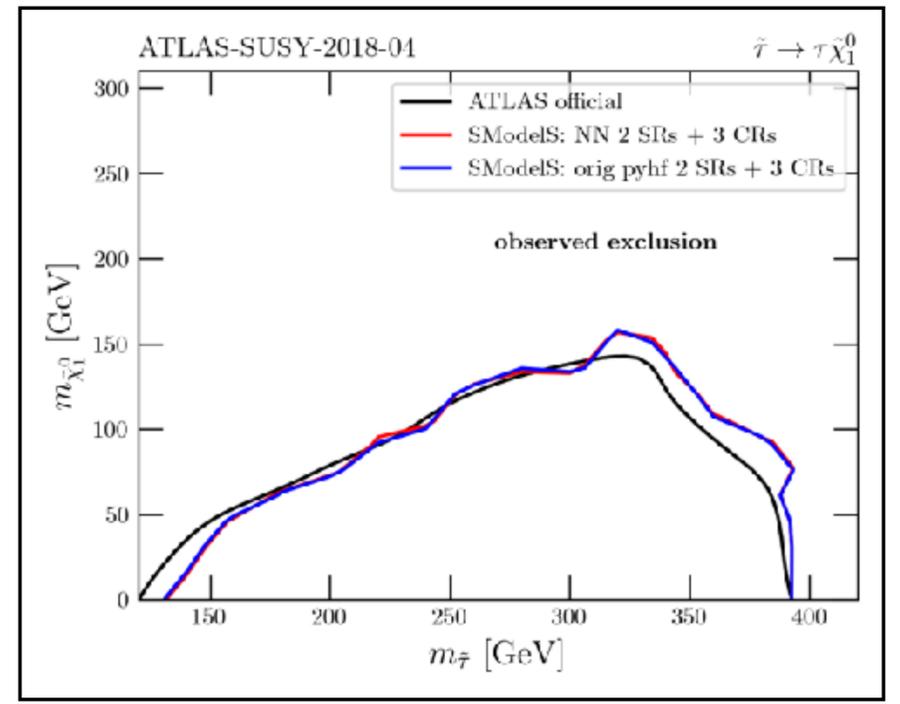
Fully connected network

$$-2 \log \frac{L^{\text{exp}}(\mu_1, \dots, \mu_N)}{L^{\text{exp}}(0, \dots, 0)}$$

$$-2 \log \frac{L_{\text{Asimov}}^{\text{exp}}(\mu_1, \dots, \mu_N)}{L_{\text{Asimov}}^{\text{exp}}(0, \dots, 0)}$$

$$-2 \log \frac{L^{\text{obs}}(\mu_1, \dots, \mu_N)}{L^{\text{obs}}(0, \dots, 0)}$$

$$-2 \log \frac{L_{\text{Asimov}}^{\text{obs}}(\mu_1, \dots, \mu_N)}{L_{\text{Asimov}}^{\text{obs}}(0, \dots, 0)}$$



A wide-angle photograph of Moszna Castle in Poland, a large Gothic Revival castle with multiple towers and a red-tiled roof. The castle is reflected in a large, rectangular pond in the foreground. The sky is clear blue, and the surrounding area is lush with green trees.

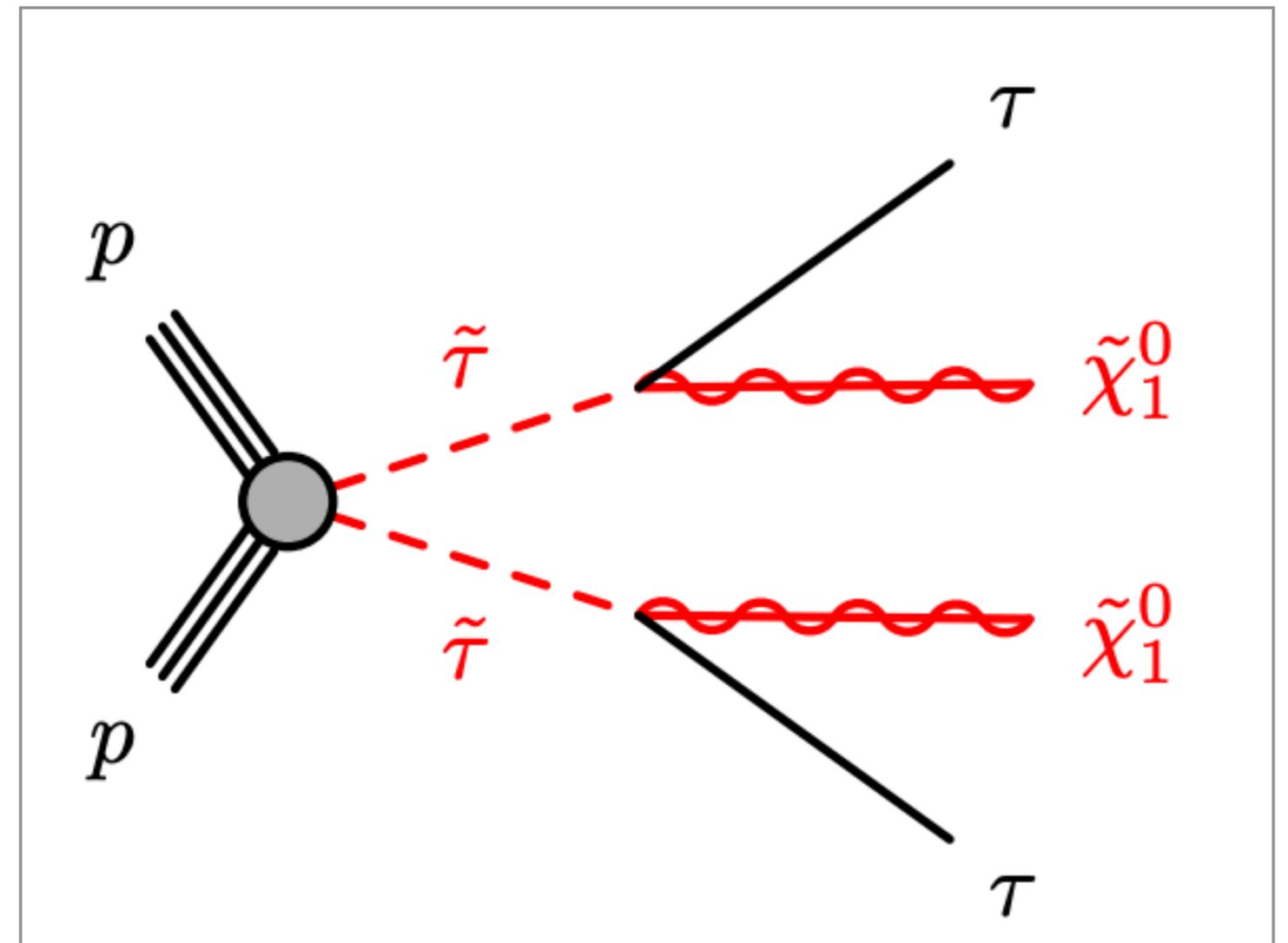
# PRELIMINARY RESULTS

Moszna Castle, Poland

Search for direct stau production in events with two hadronic  $\tau$ -leptons in  $\sqrt{s}=13$  TeV pp collisions with the ATLAS detector

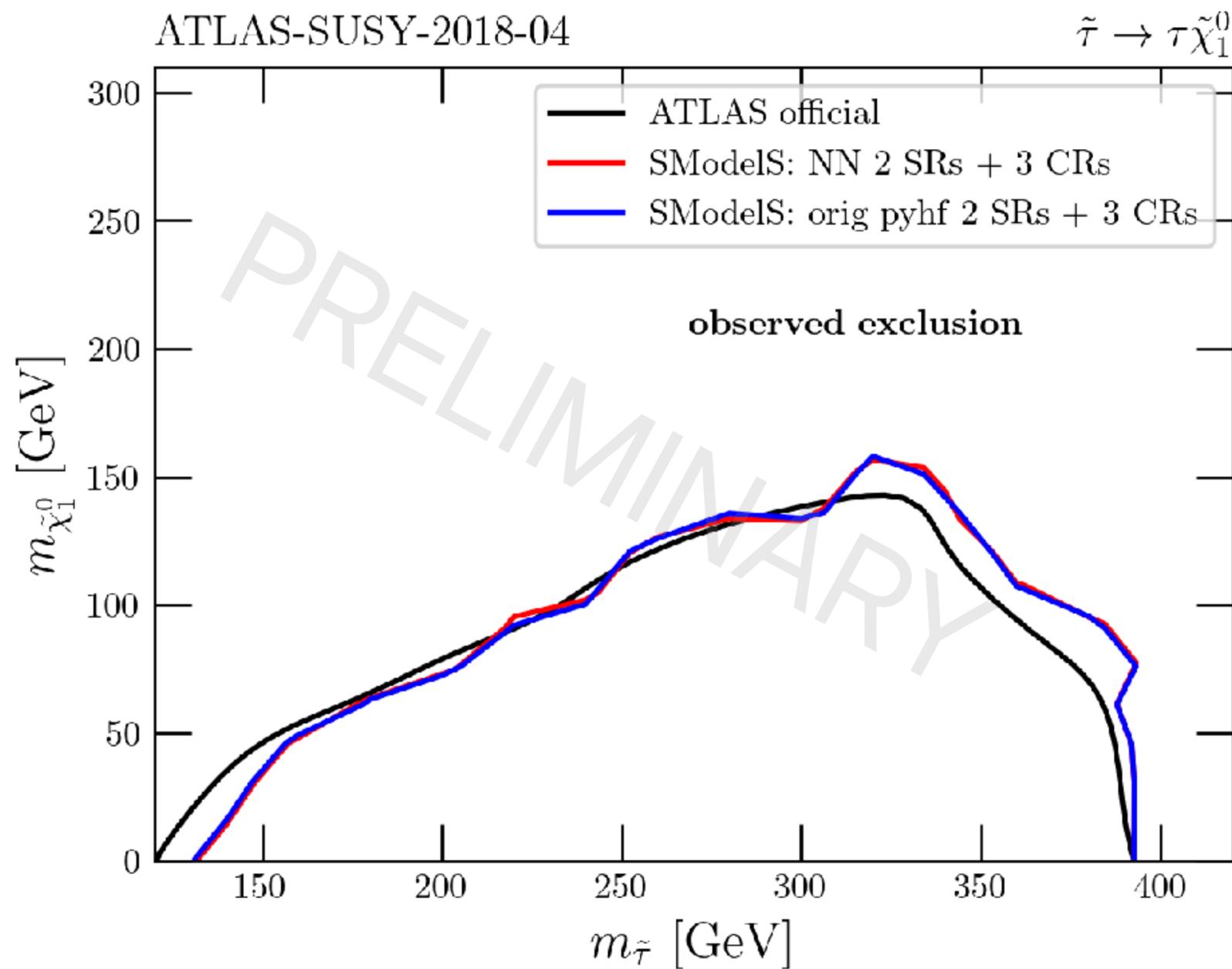
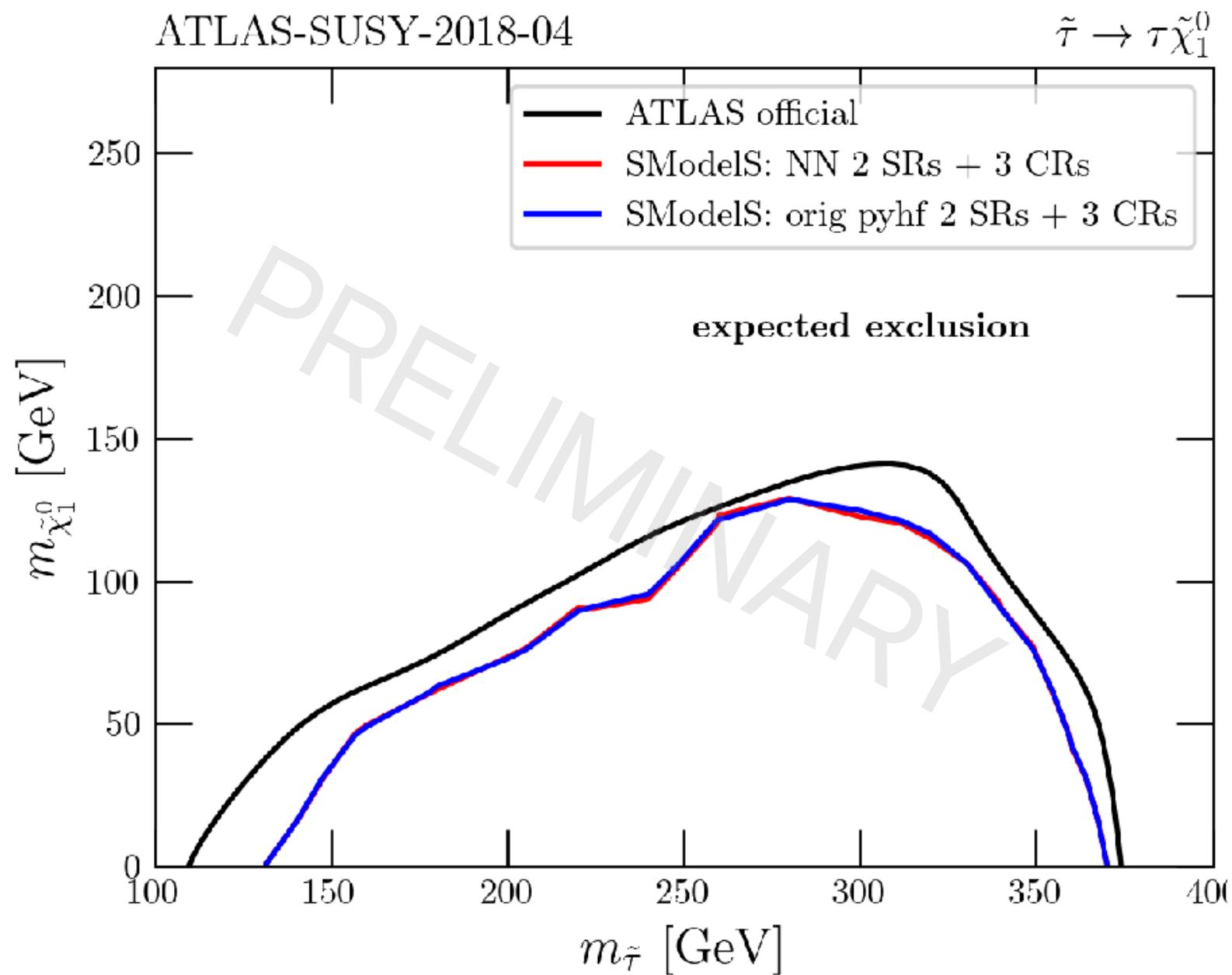
2 signal bins, 3 control bins

fit **without** the CRs



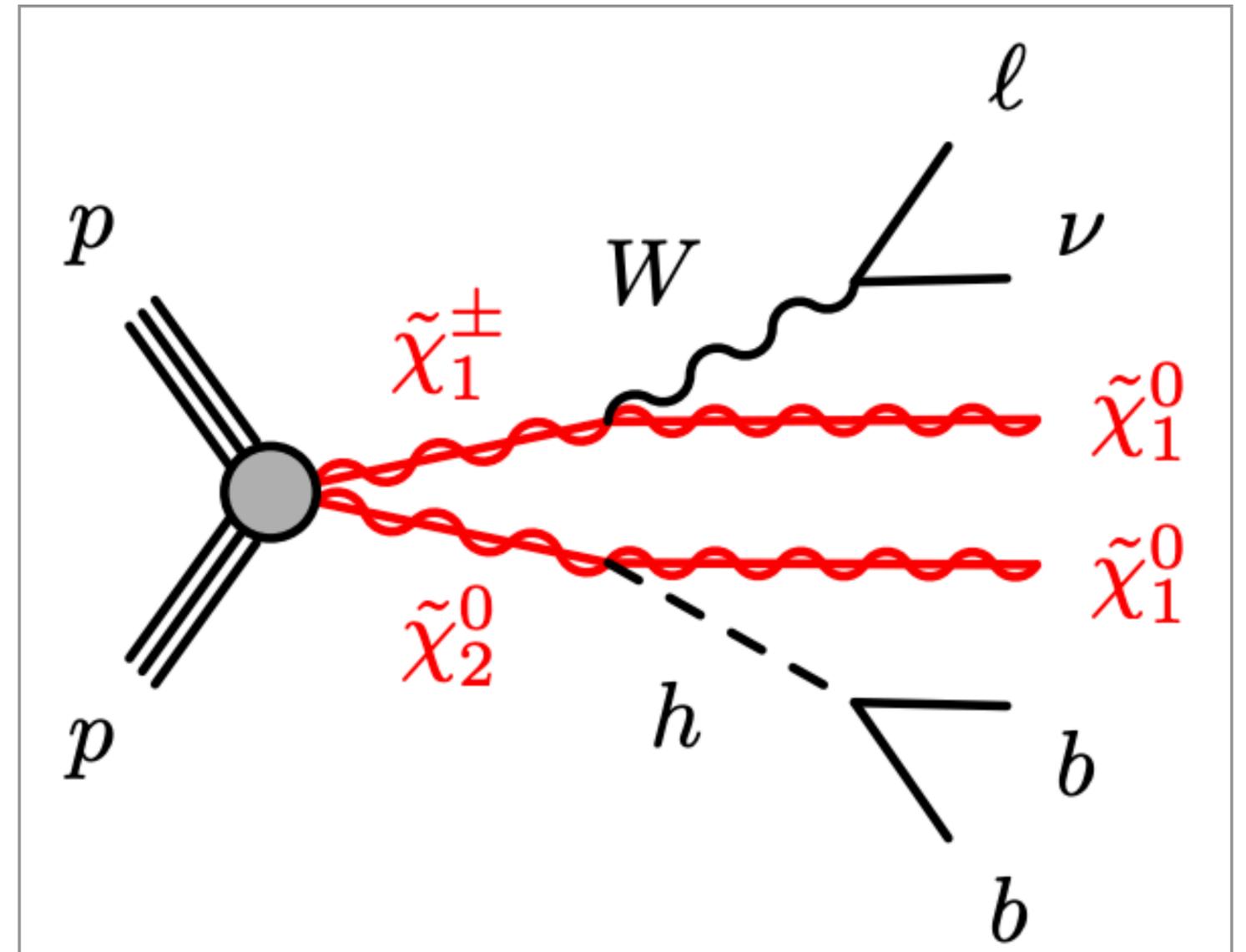
expected

observed



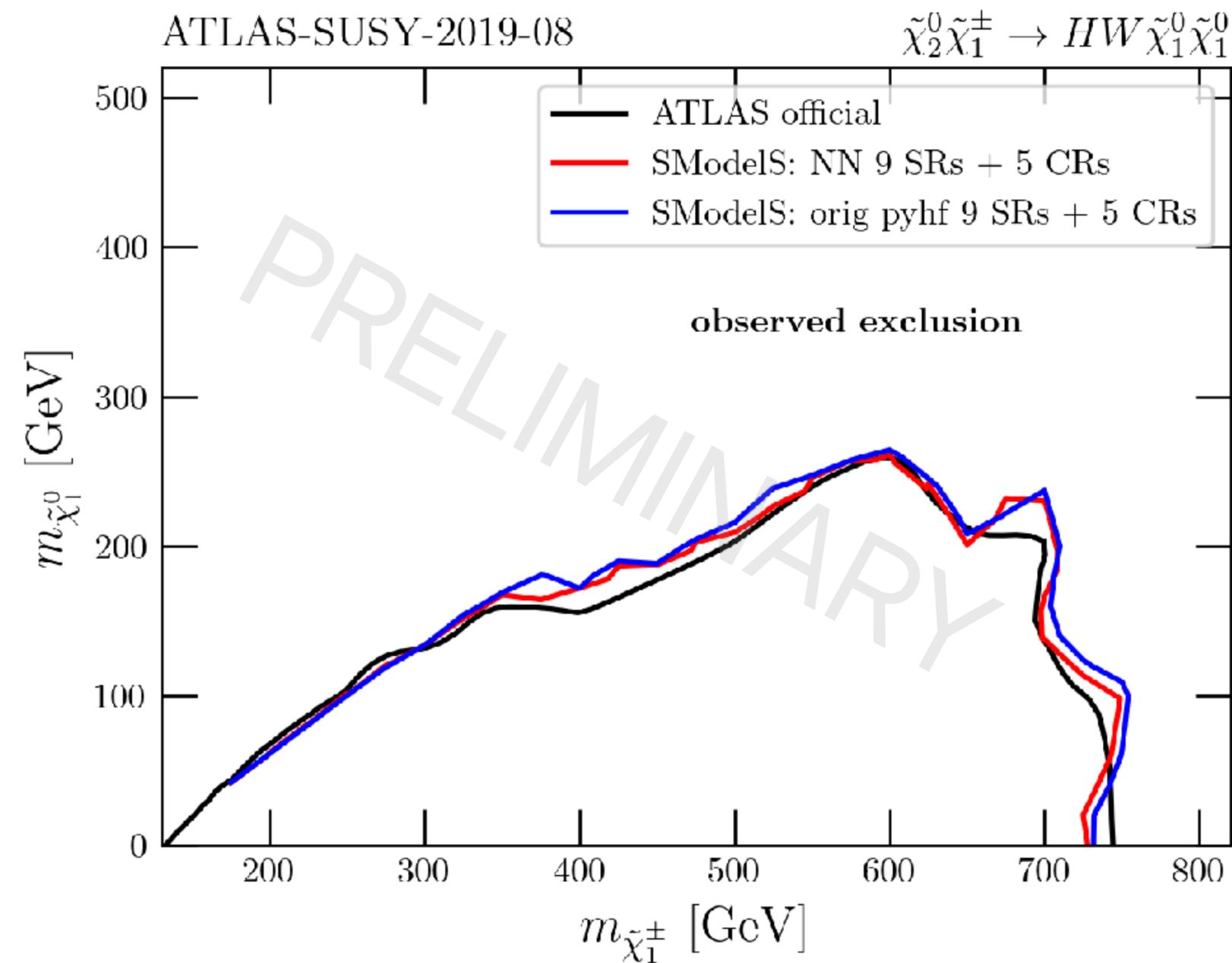
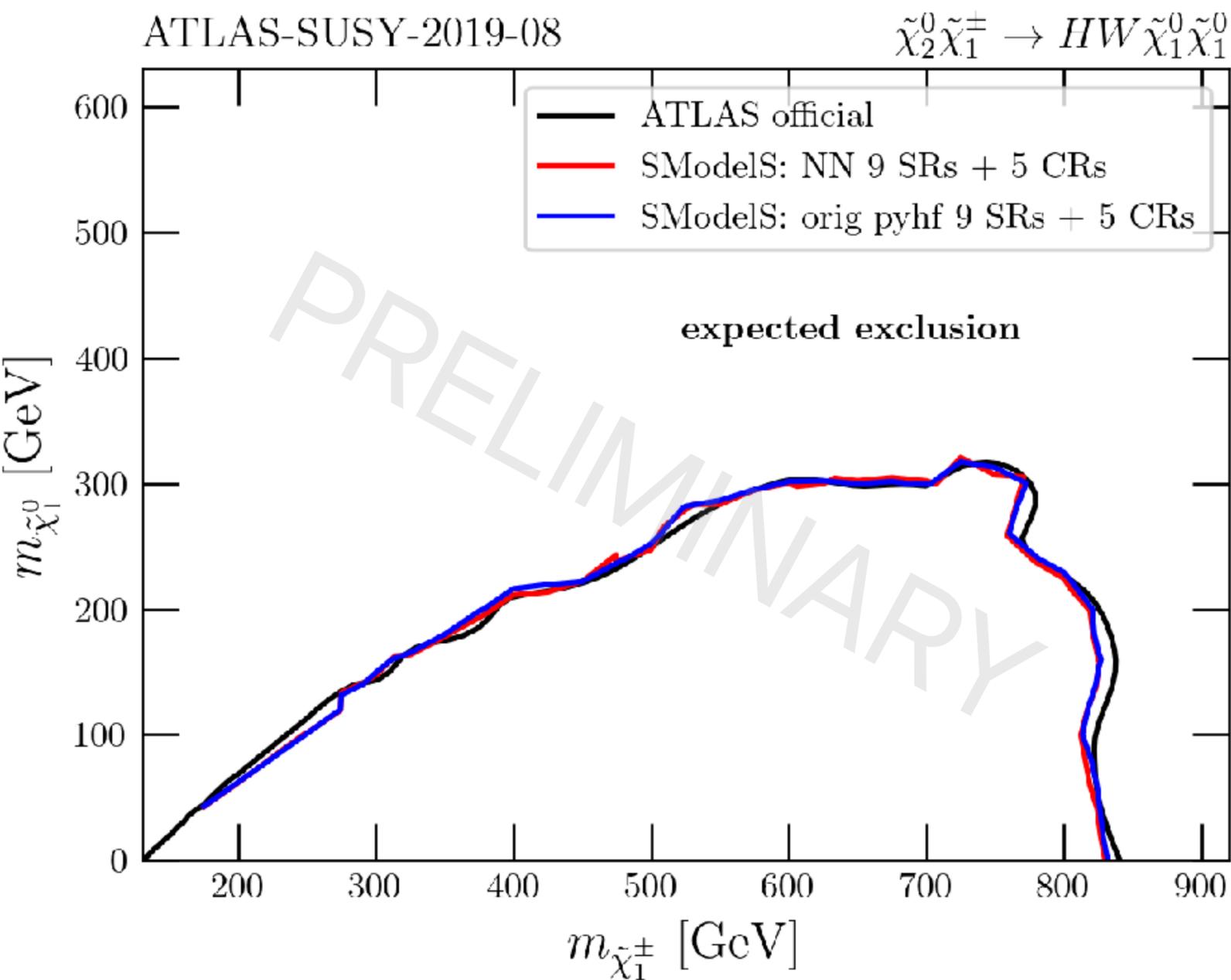
Search for direct production of electroweakinos in final states with one lepton, missing transverse momentum and a Higgs boson decaying into two b-jets in  $pp$  collisions at  $\sqrt{s}=13$  TeV with the ATLAS detector

9 signal bins, 5 control bins



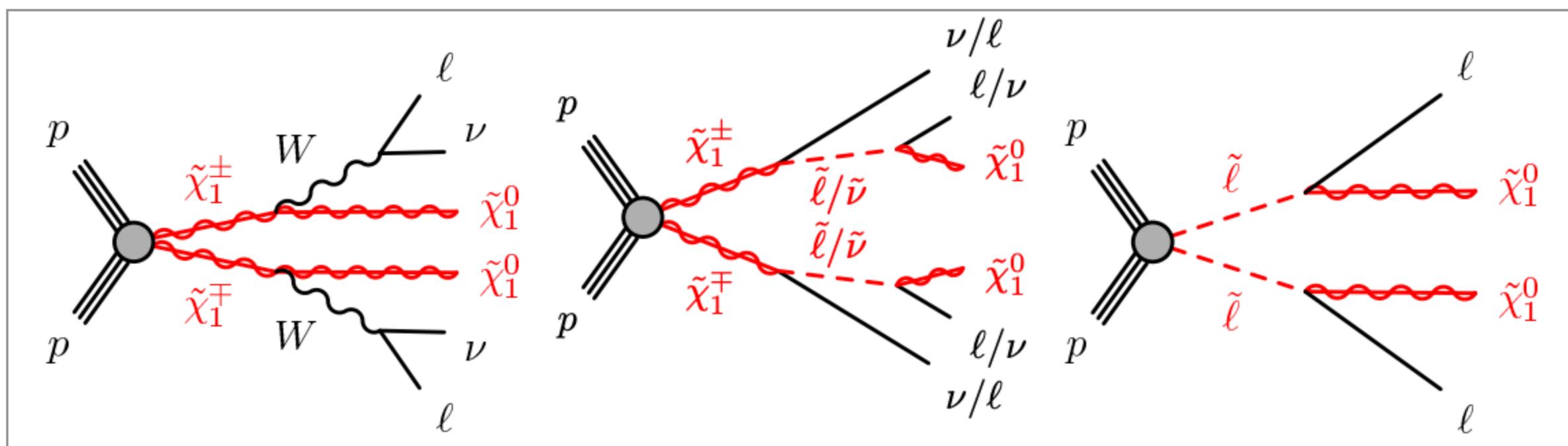
expected

observed



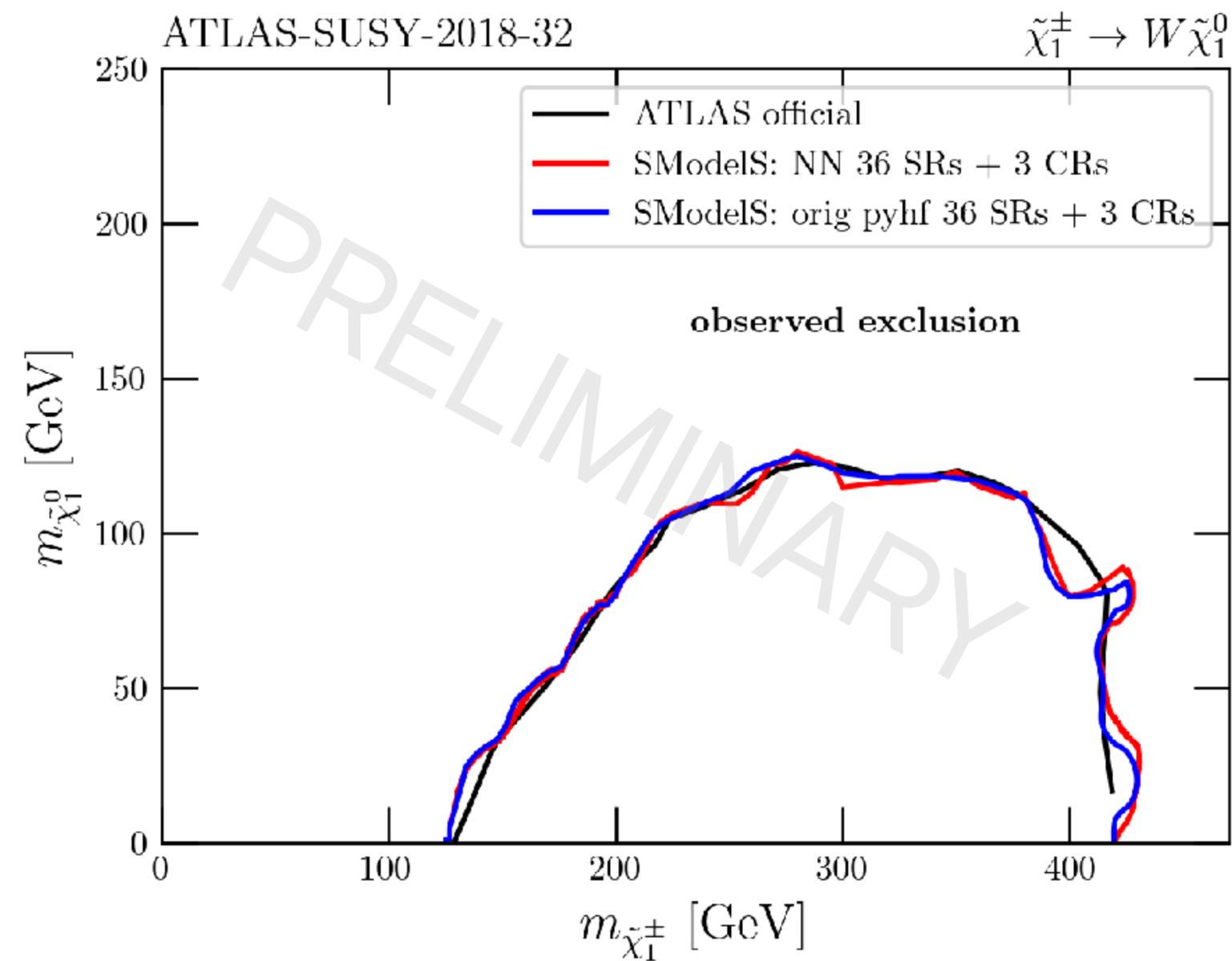
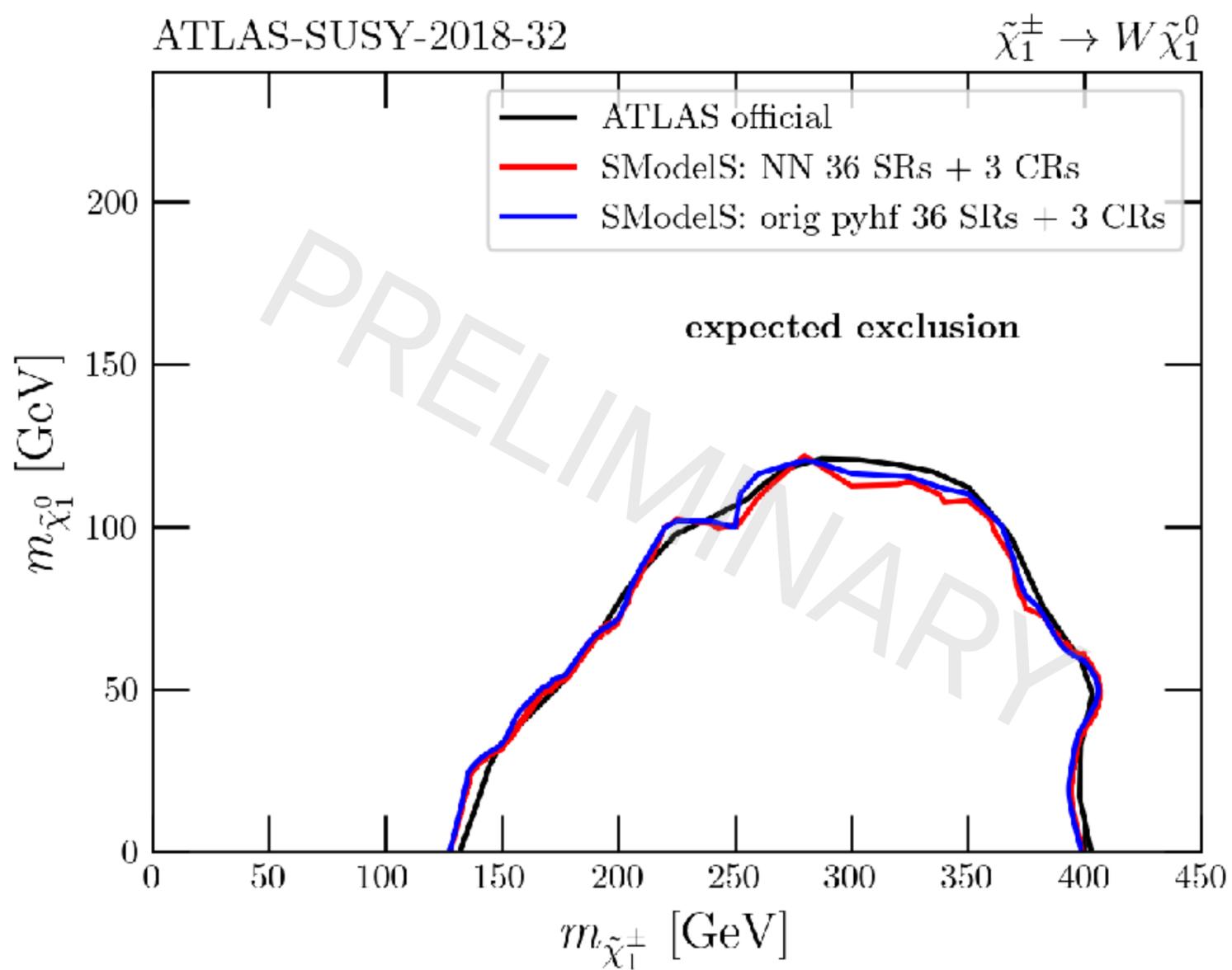
Search for electroweak production of charginos and sleptons decaying into final states with two leptons and missing transverse momentum in  $\sqrt{s}=13$  TeV pp collisions using the ATLAS detector

36 SRs and 3 CRs

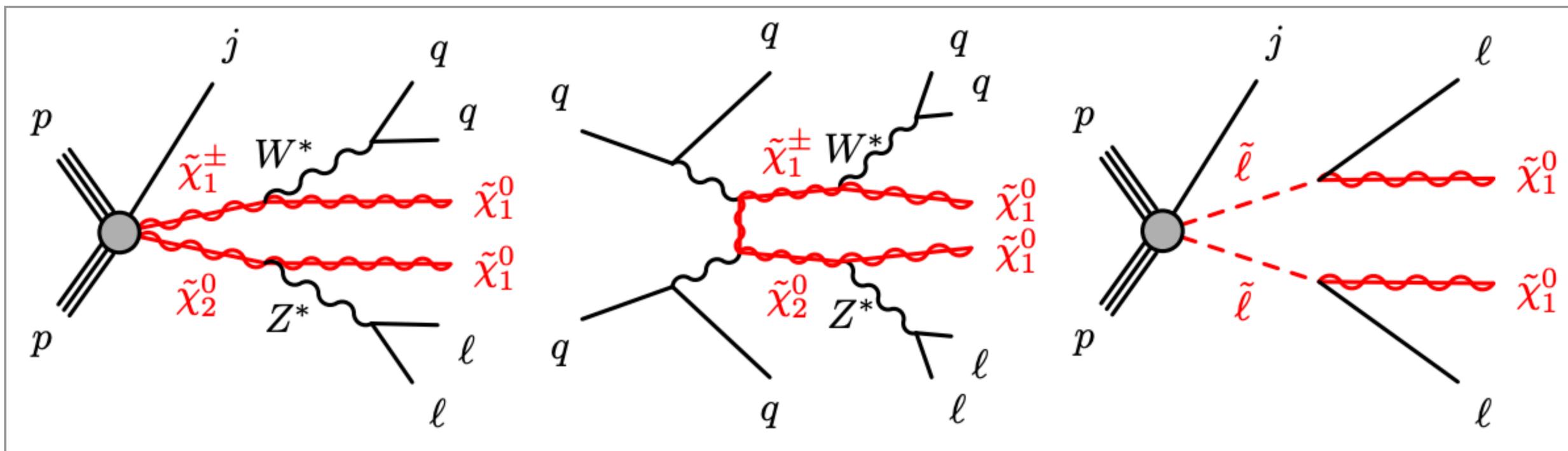


expected

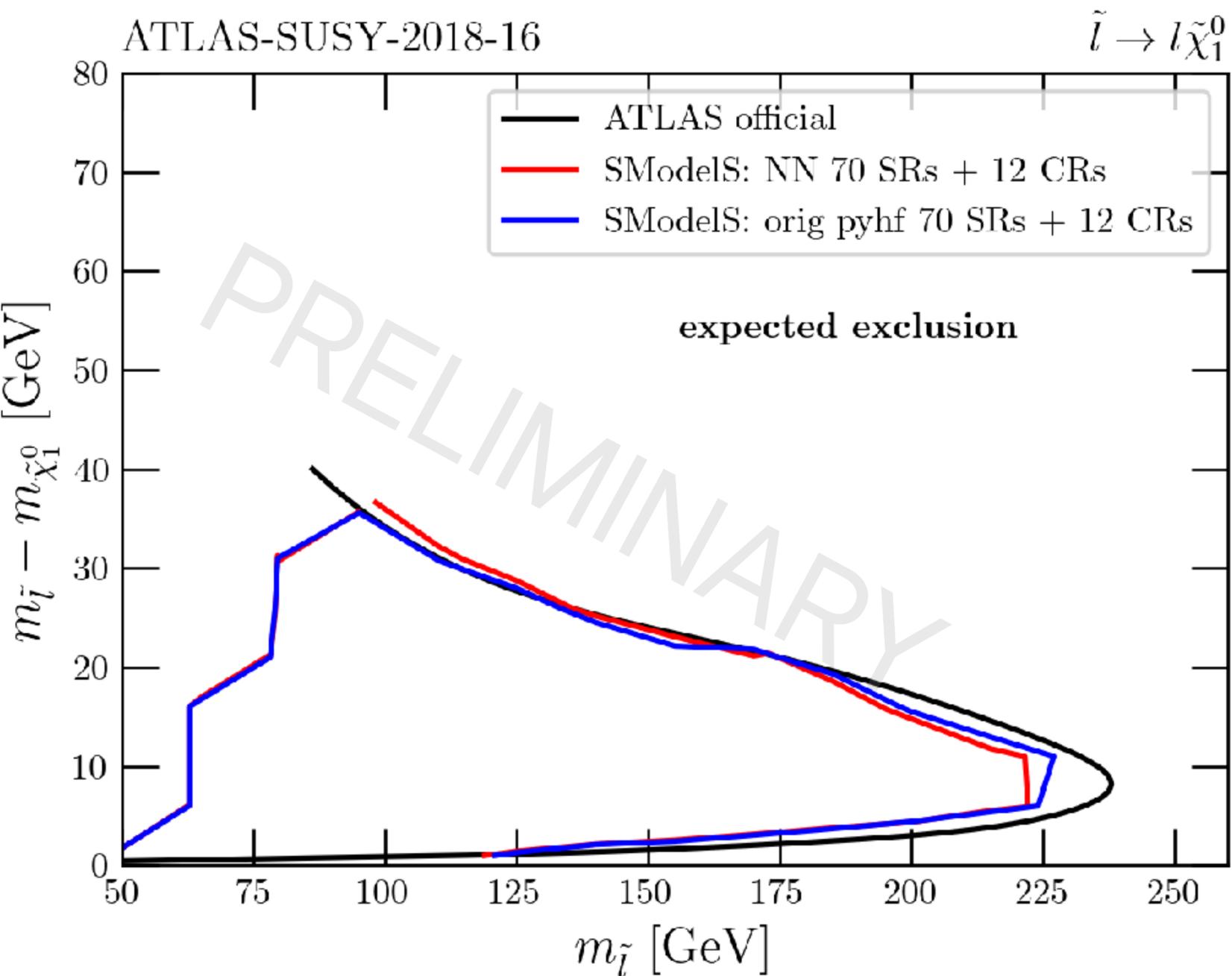
observed



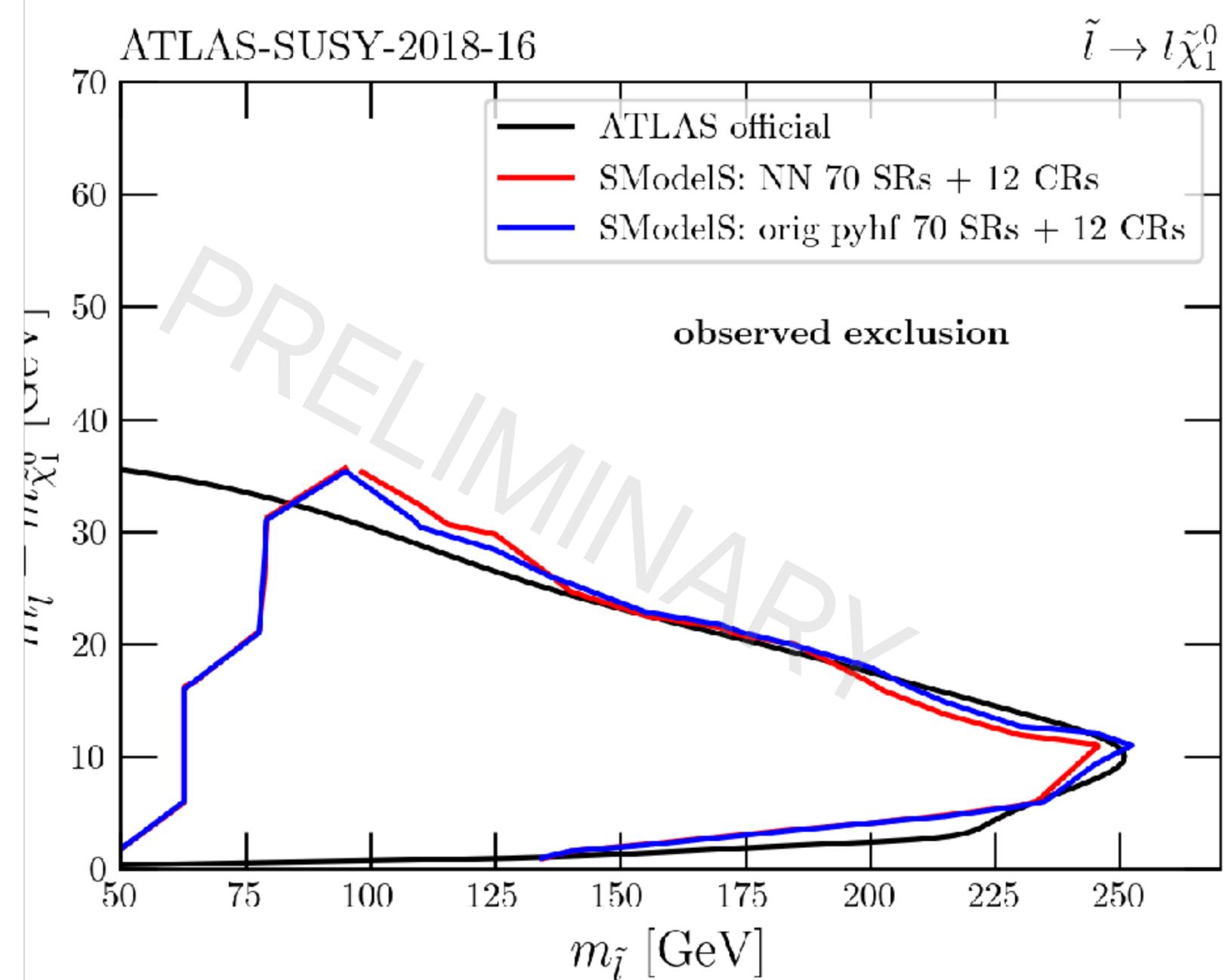
- Searches for electroweak production of supersymmetric particles with compressed mass spectra in  $\sqrt{s}=13$  TeV pp collisions with the ATLAS detector
- 2 subanalyses: sleptons (32+6) and EWKinos (38+6)



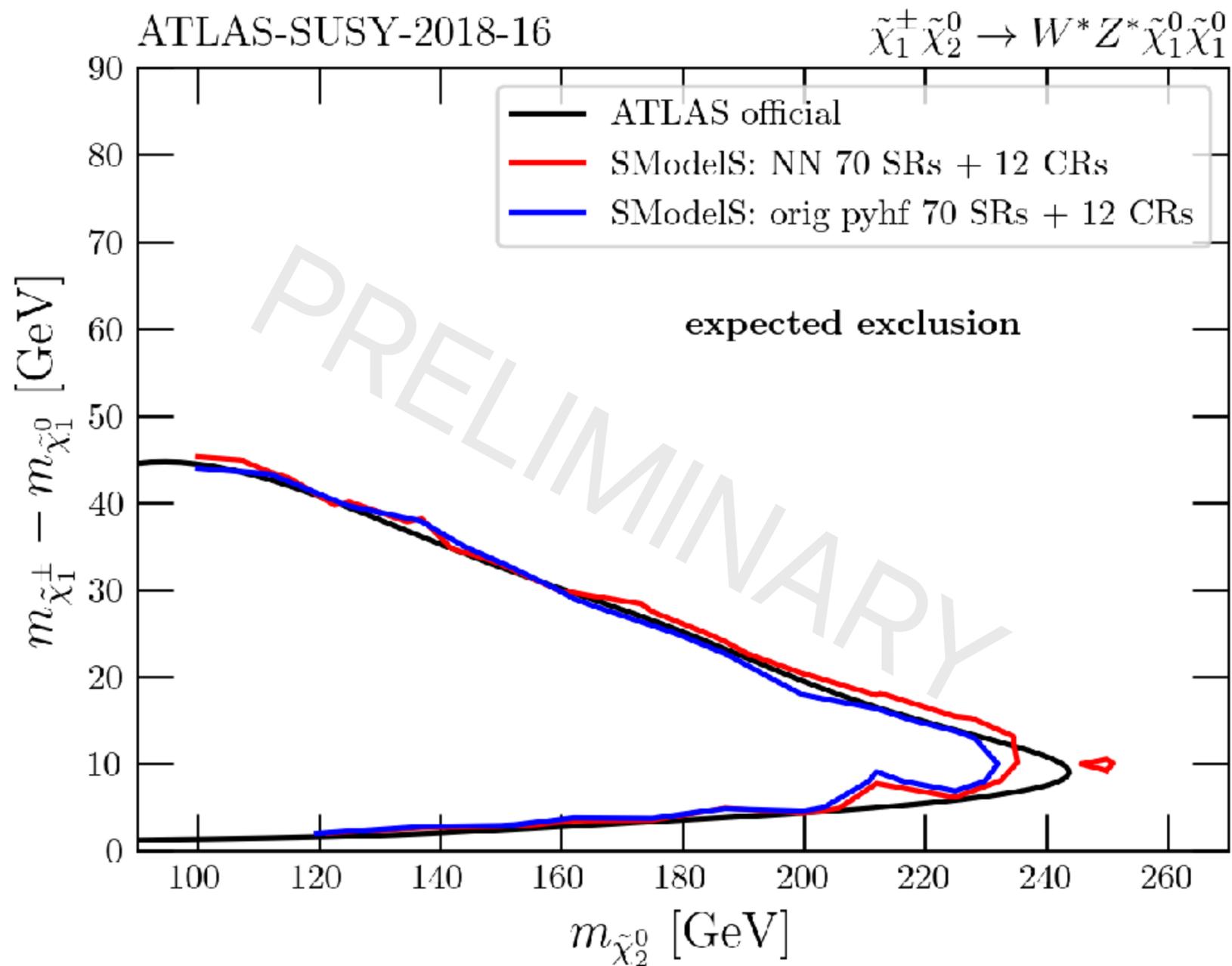
expected (sleptons)



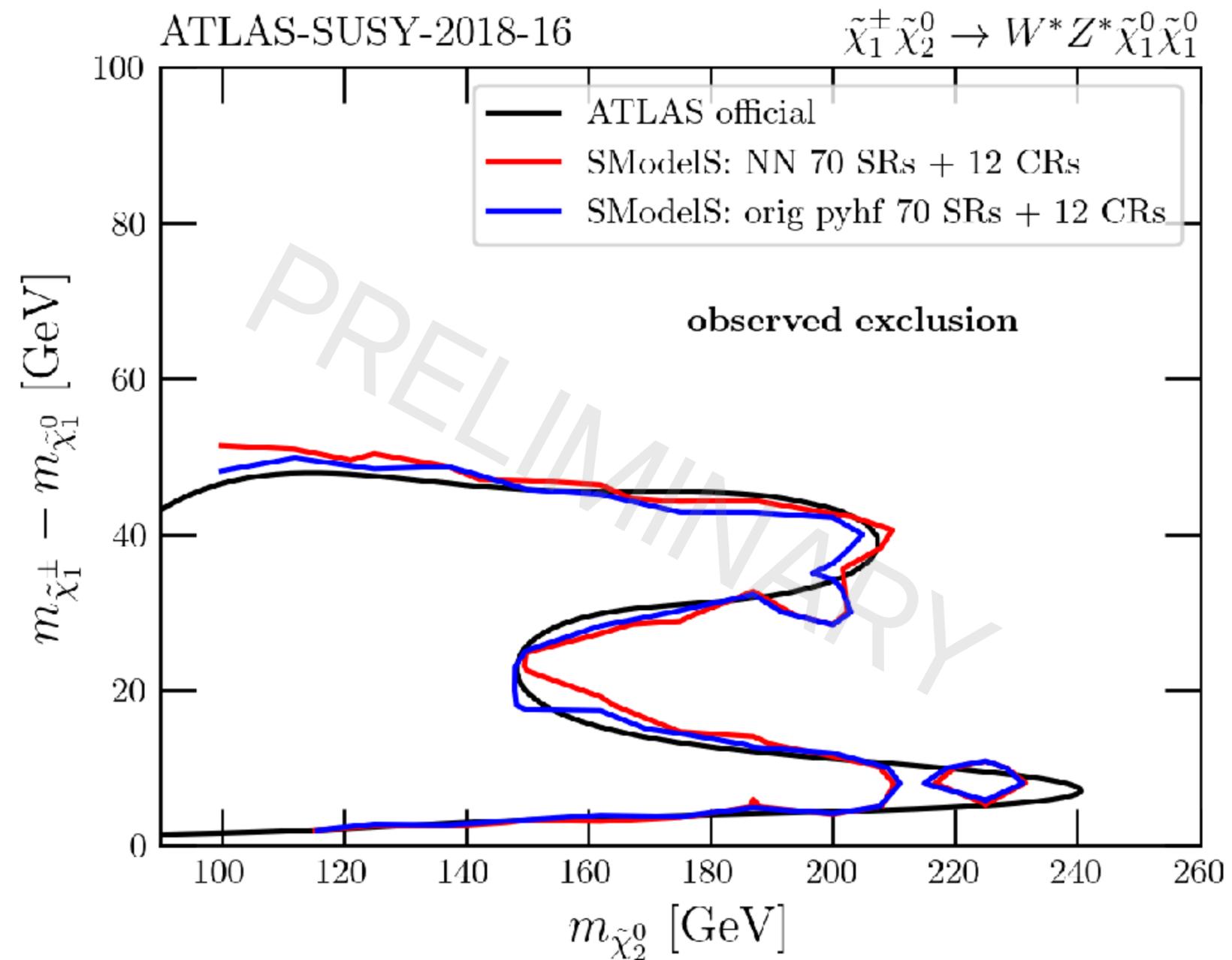
observed (sleptons)



expected (EWKinosh)



observed (EWKinosh)



Search for chargino-neutralino pair production in final states with three leptons and missing transverse momentum in  $\sqrt{s}=13$  TeV pp collisions with the ATLAS detector

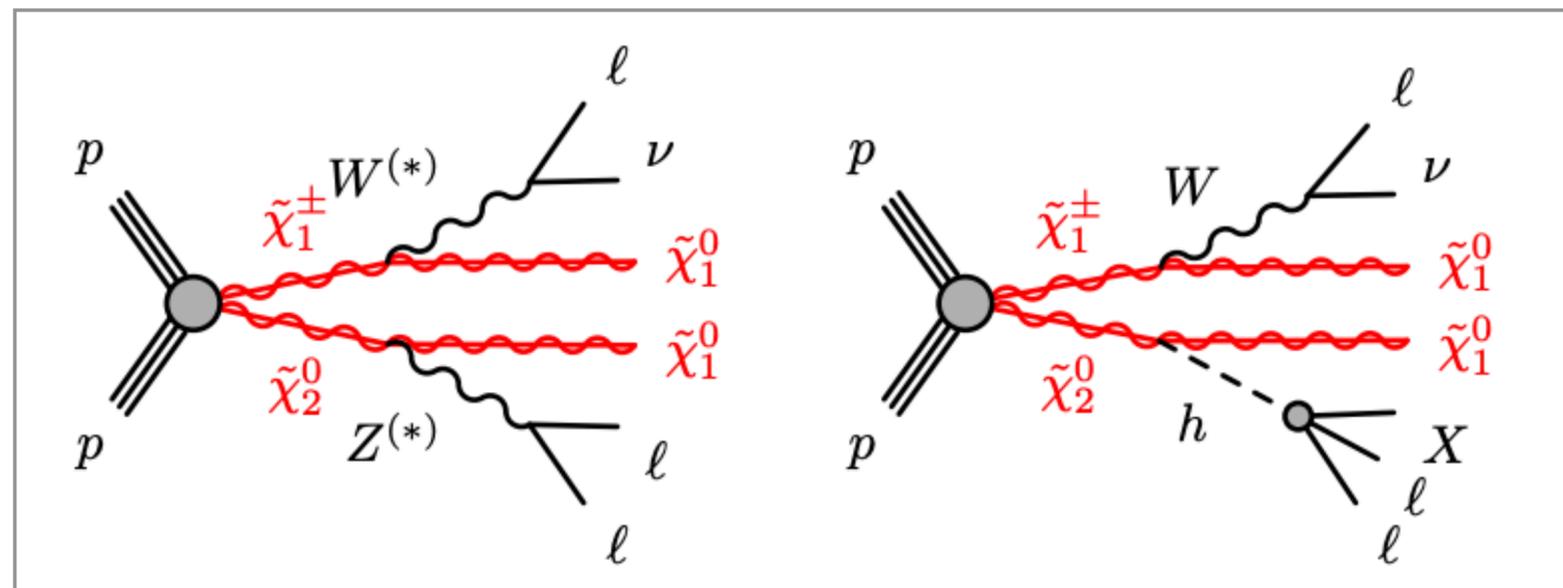
4 subanalyses:

onshell (W/Z when decaying) **wino/bino** (22+2) ❌

offshell **wino/bino minus** (sign of the product of two neutralino mass eigenstates) (32+2) ✔️

offshell **wino/bino plus** (32+2) ✔️

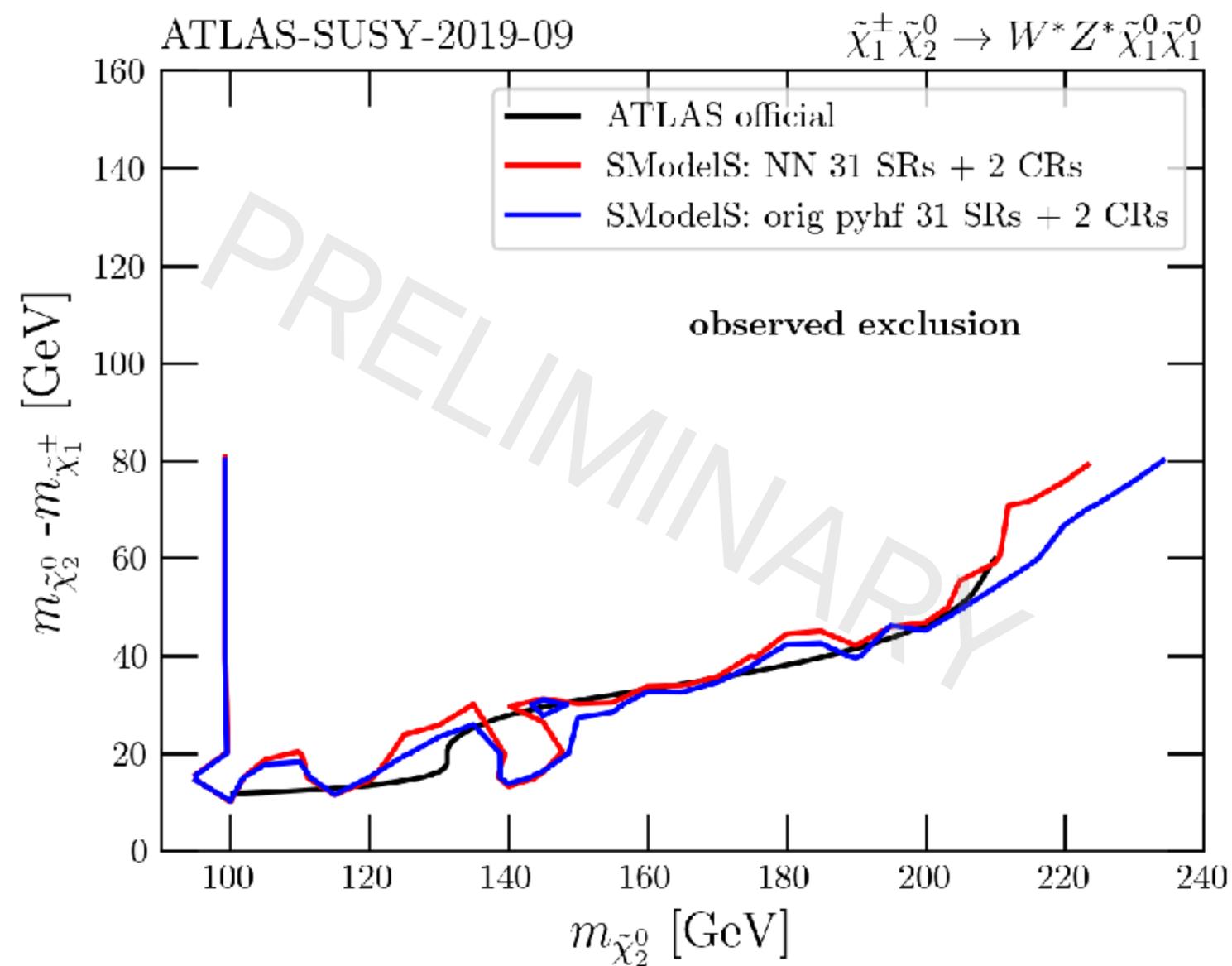
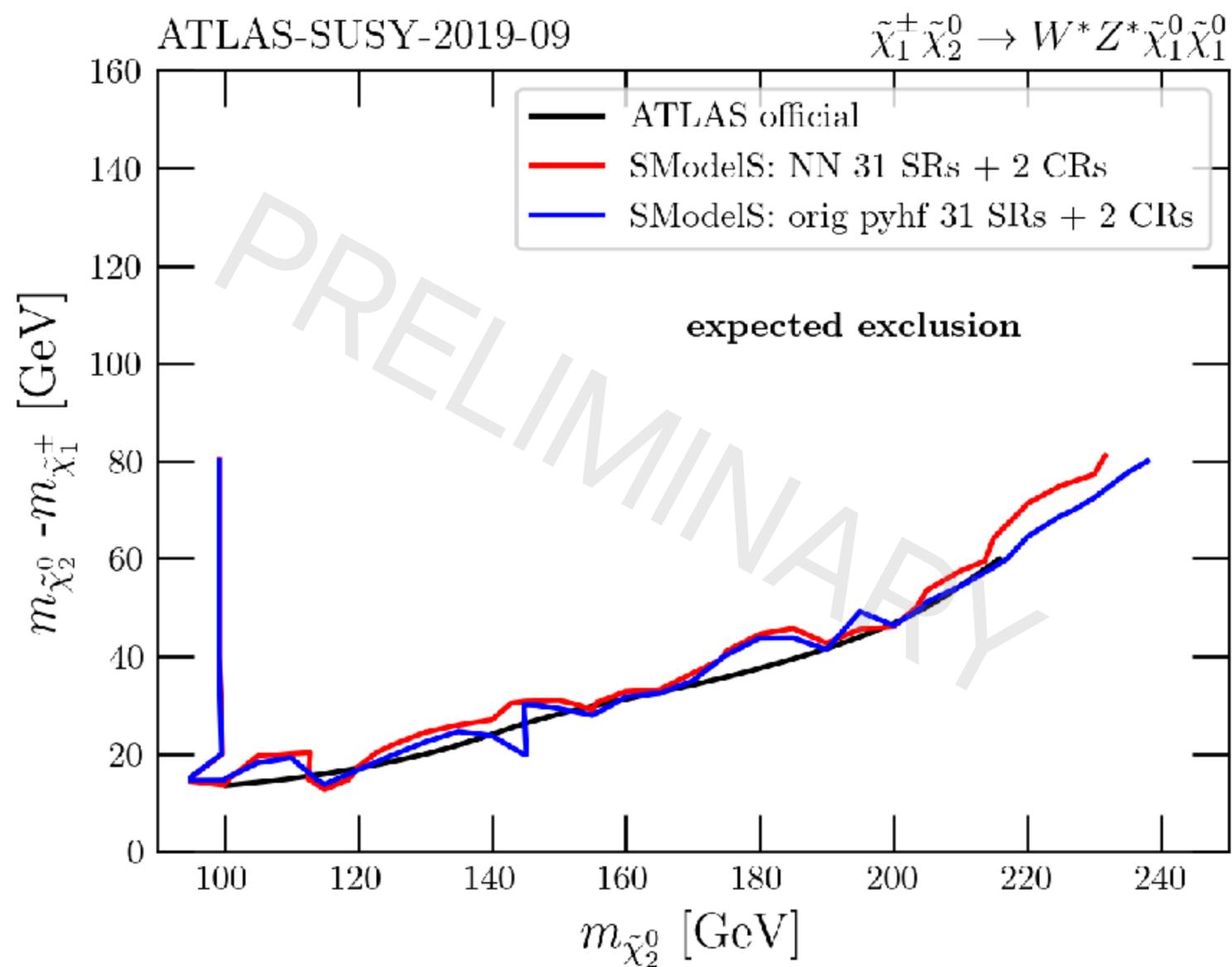
offshell **higgsino** (32+2) ✔️



- **On-shell WZ selection:**  $\tilde{\chi}_2^0 \rightarrow Z\tilde{\chi}_1^0$  with 100% branching ratio, where  $\Delta m(\tilde{\chi}_2^0, \tilde{\chi}_1^0) \geq m_Z$ , for the wino/bino (+) scenario.
- **Off-shell WZ selection:**  $\tilde{\chi}_2^0 \rightarrow Z^{(*)}\tilde{\chi}_1^0$  with 100% branching ratio, where  $\Delta m(\tilde{\chi}_2^0, \tilde{\chi}_1^0) < m_Z$ , for the wino/bino (+), the wino/bino (-), and the higgsino scenarios.
- **Wh selection:**  $\tilde{\chi}_2^0 \rightarrow h\tilde{\chi}_1^0$  with 100% branching ratio, where  $\Delta m(\tilde{\chi}_2^0, \tilde{\chi}_1^0) > m_h$ , for the wino/bino (+) scenario.

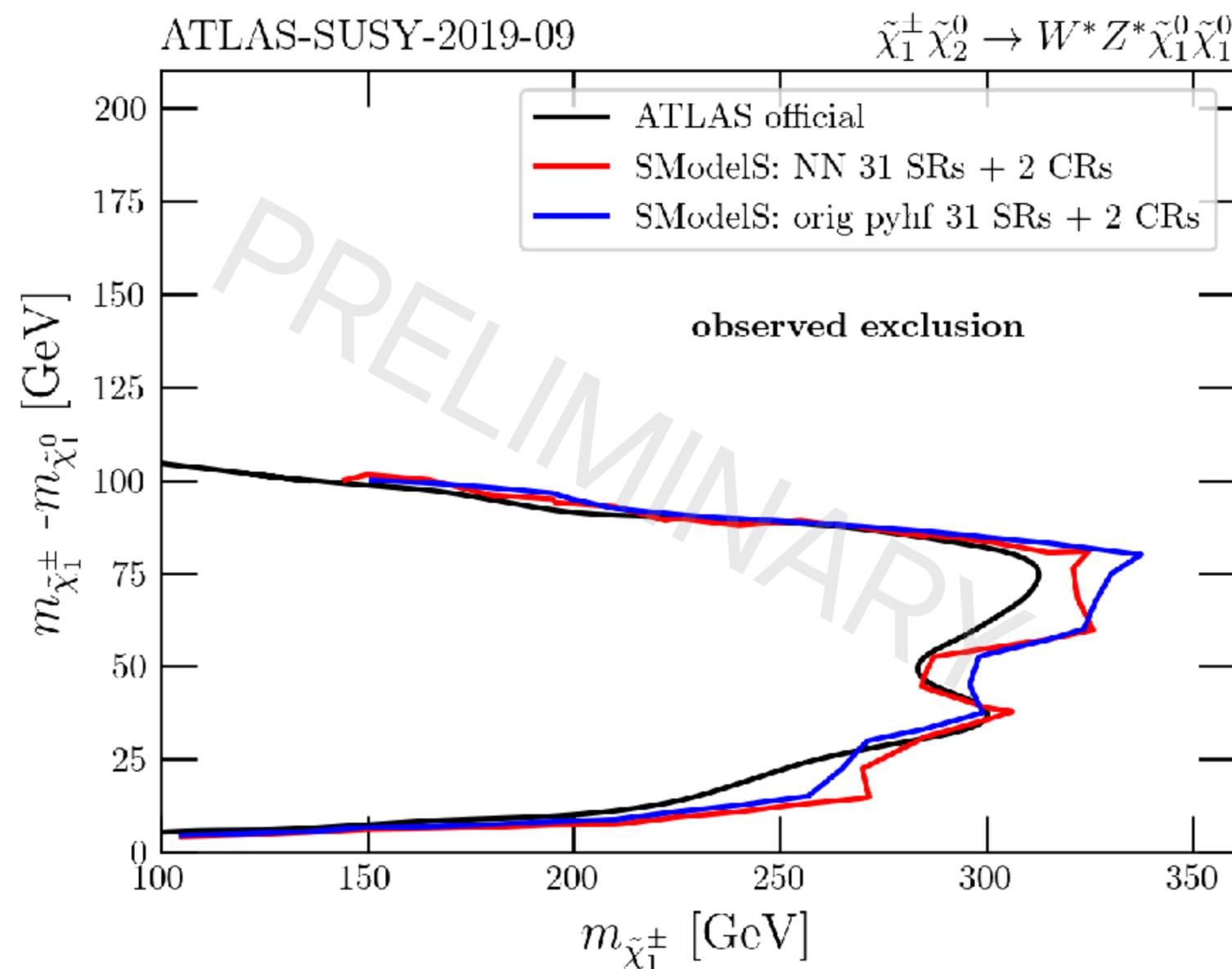
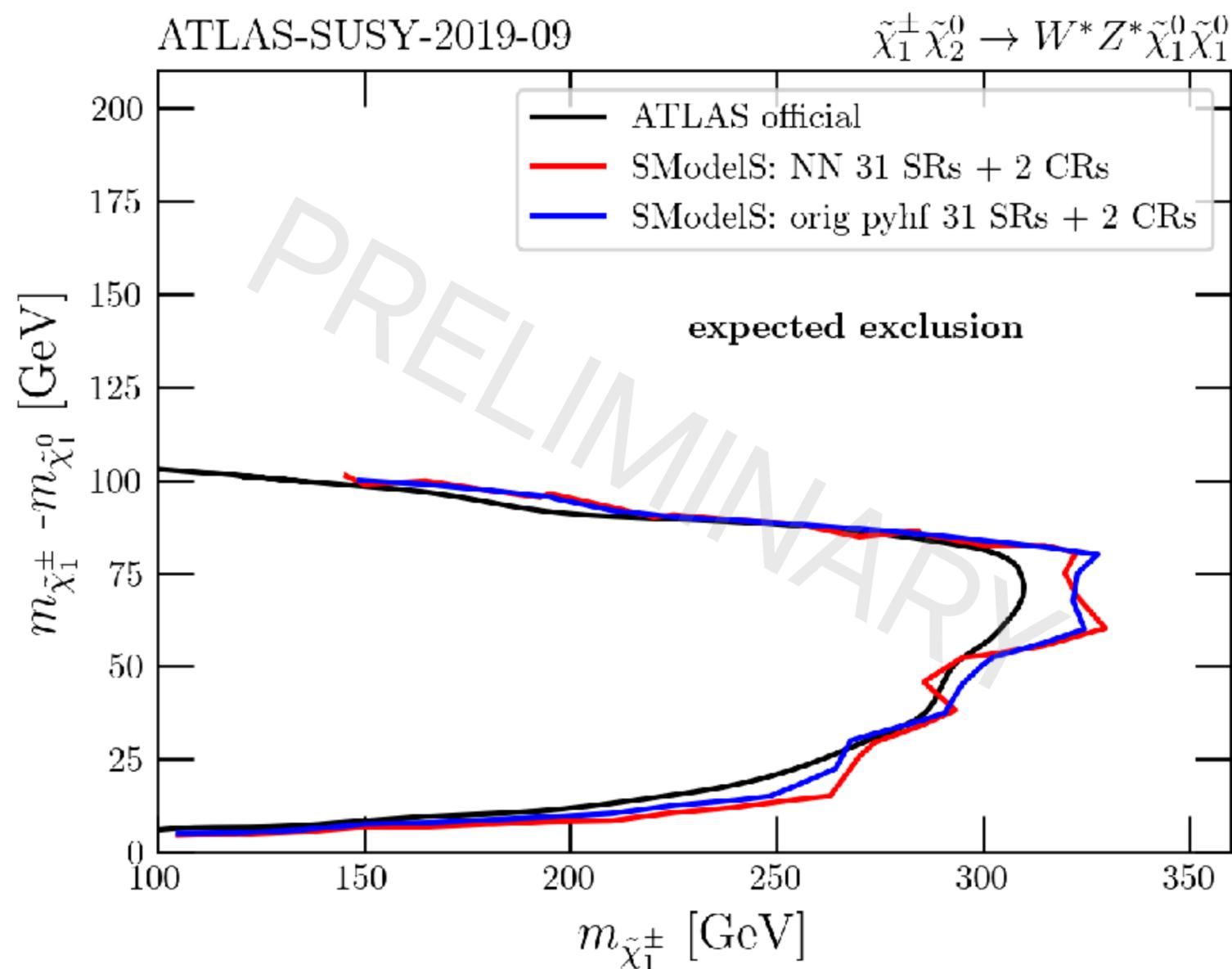
expected (offshell Higgsino)

observed (offshell Higgsino)



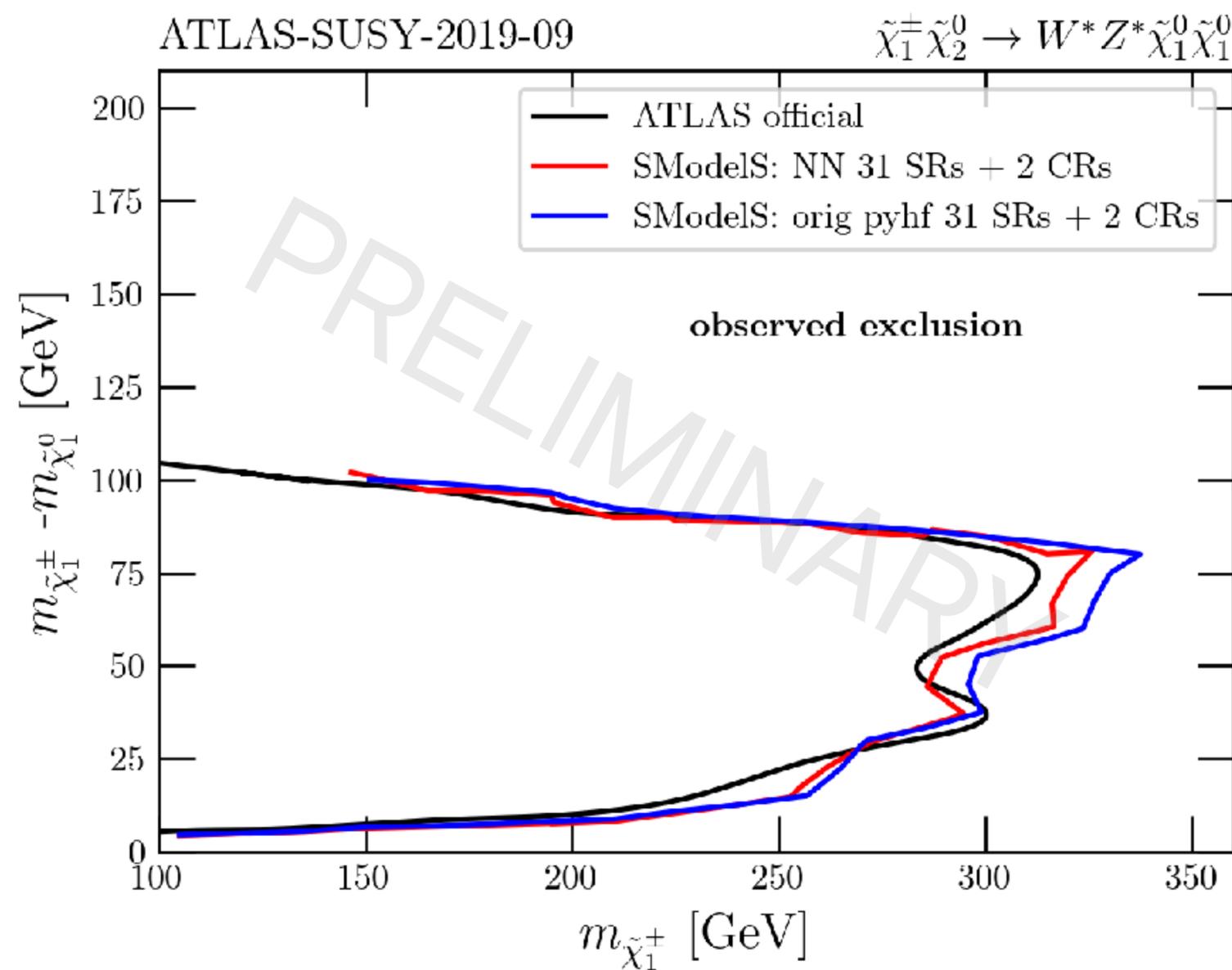
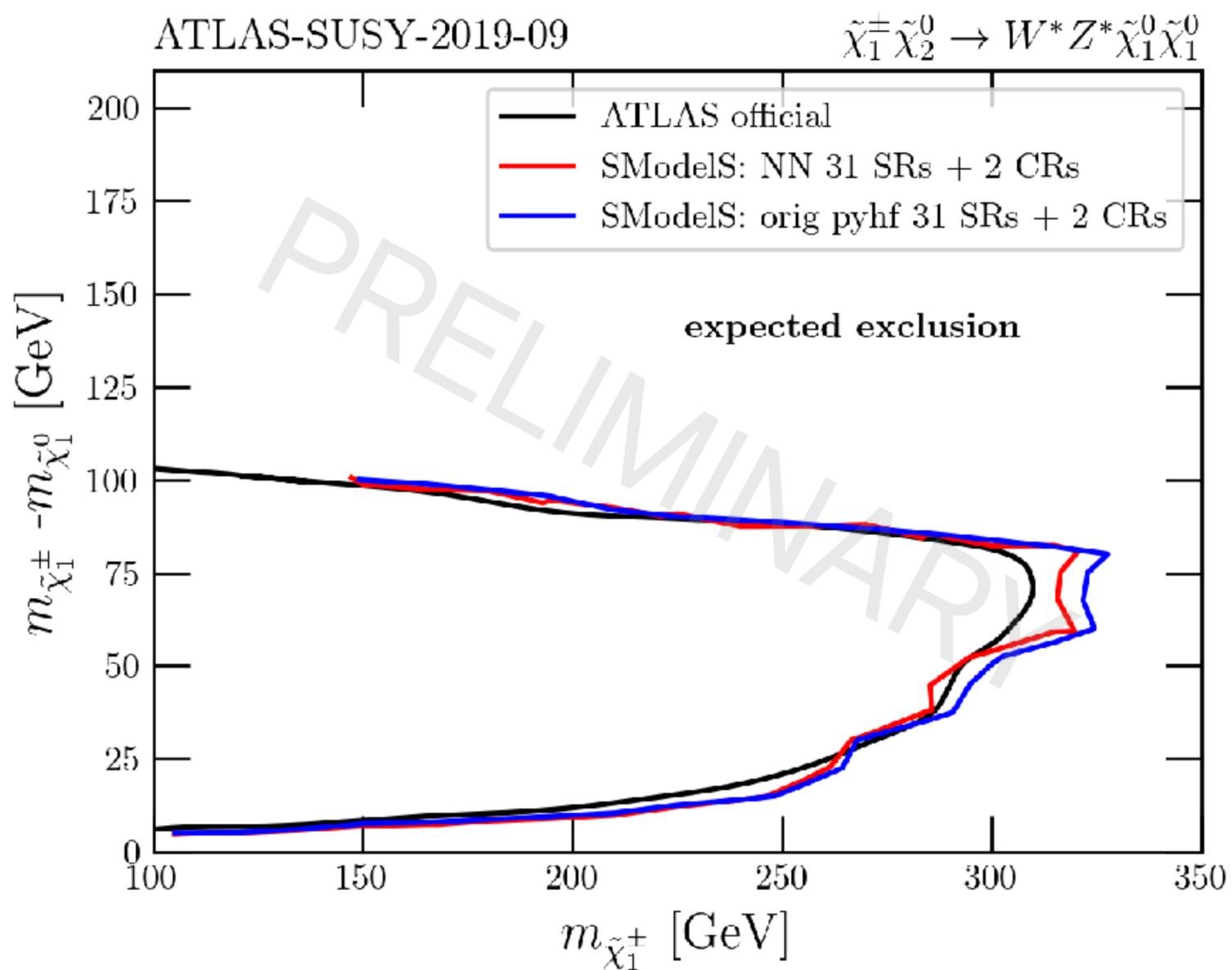
expected (offshell wino/bino -)

observed (offshell wino/bino -)



expected (offshell wino/bino +)

observed (offshell wino/bino +)



# Summary

- ✿ Selecting the right BSM model enhances chances of new physics discovery
- ✿ Reinterpretation allows to constrain untested models
- ✿ Proper hypotheses testing is possible with published Full Statistical Models
- ✿ However, using these models constitute a **computational bottleneck**
- ✿ Machine Learning surrogates offer **good accuracy with much faster computations**
- ✿ Preliminary results with MLPs are promising
- ✿ Trained models will be published to be used by the community



Thank you for attention!

[rafal.maselek@ijs.si](mailto:rafal.maselek@ijs.si)