Efficient Search for New Physics Using Active Learning in the ATLAS Experiment

Zubair Bhatti, Kyle Cranmer, Irina Espejo, Lukas Heinrich, Phillip Gadow, Patrick Rieck and Janik von Ahne on behalf of the ATLAS Collaboration

May 9, 2023 CHEP

ATL-PHYS-PUB-2022-045

The Team

Kyle Cranmer
(UW Madison,
former NYU)

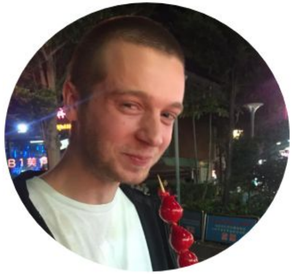
Lukas Heinrich
(TMU, former
NYU)

Patrick Rieck
(NYU)

Irina Espejo
(NYU)

Phillip Gadow
(DESY)

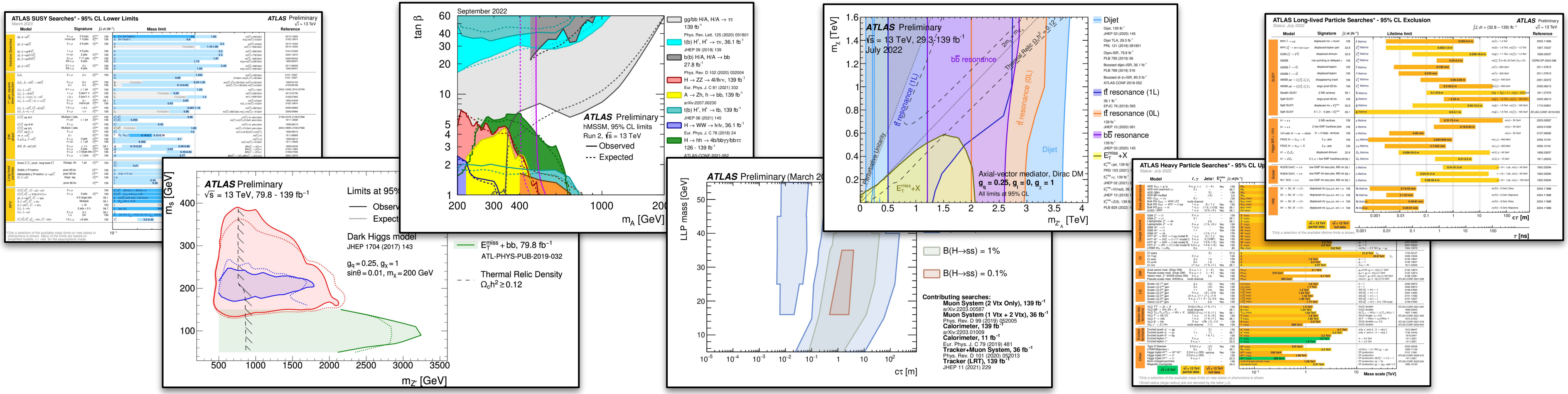
Eleni Skorda
(Lund University)



Janik von Ahnen
(DESY)

Ruth Pöttgen
(Lund University)

1000's of New Physics Papers



- Proliferation of new physics models with a multitude of parameters
  - Need to raise efficiency in limit setting

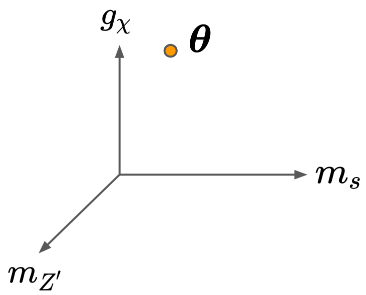
- How do we increase efficiency in Beyond Standard Model parameters limit setting?
  - Active learning of exclusion contours

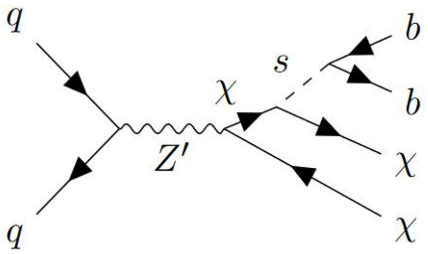
[I. Espejo et al. RIF 2022](#)
[P. Rieck et al. ACAT 2022](#)
https://atlaspo.cern.ch/public/summary_plots/

Limit Setting at ATLAS

Full Pipeline

$\mathcal{L}_{SM} + \mathcal{L}_{BSM}(\boldsymbol{\theta})$ → $\sigma(pp \to bg, sig)$ → MC Event Simulation

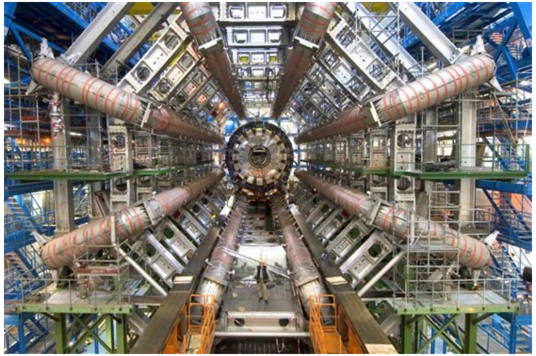


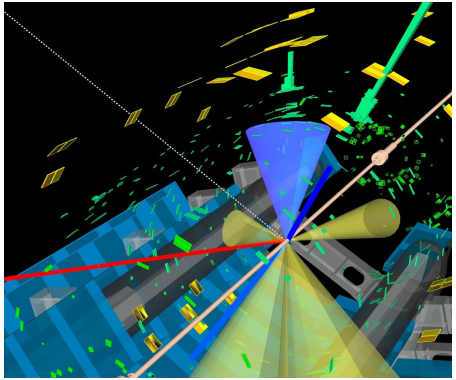


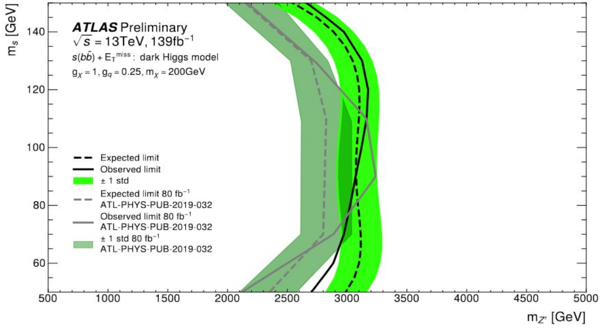
1. Event Generation
2. Showering/Hadronization
3. Detector Simulation
4. Digitization

MadGraph, Pythia, Herwig, Geant4, …

+ ATLAS Data → Event Reconstruction → Analysis and Inference







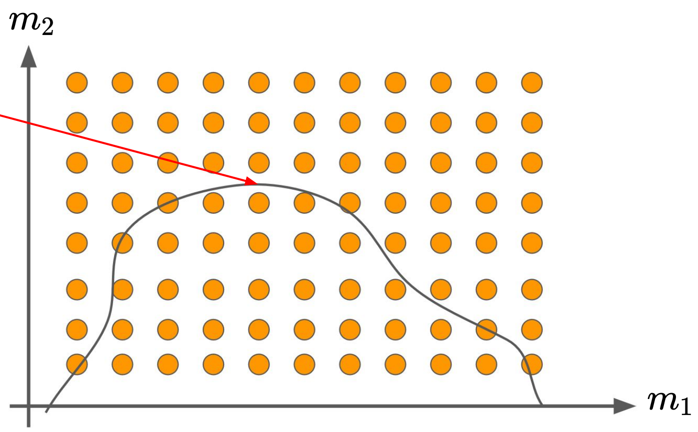
Determining Exclusion Contours

f, the full pipeline is expensive to evaluate:

$$f : \boldsymbol{\theta} \rightarrow \boldsymbol{y} := (\log \mu^{UL}_{exp}(\boldsymbol{\theta}), \log \mu^{UL}_{+1\sigma}(\boldsymbol{\theta}), \log \mu^{UL}_{-1\sigma}(\boldsymbol{\theta}), \log \mu^{UL}_{obs}(\boldsymbol{\theta}))$$

And afterwards, we perform a regression across our grid on the signal strength upper limit:

$$\mu = \frac{\sigma^{excluded}}{\sigma^{theory}}$$



The exclusion contour is what we are interested in:

$$y(\boldsymbol{\theta}) = 0 \Rightarrow \mu^{UL}_{exp} = 1 \Rightarrow \sigma(\boldsymbol{\theta}) \approx \sigma_{BSM}(\boldsymbol{\theta})$$

L. Heinrich et al. ACAT 2019

Indiscriminate

"Arbitrary" 4D cartesian grid, **e.g.:**

11 m_1, 8 m_2, 10 m_3, 4 g_1: 3500 points

1 shot parallelized

Compromise on density?

Compromise on dimensionality?

Informed

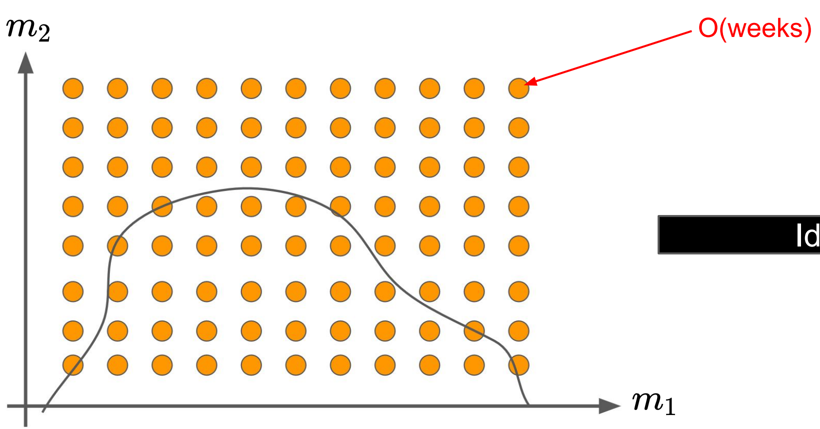
Start with prior knowledge:

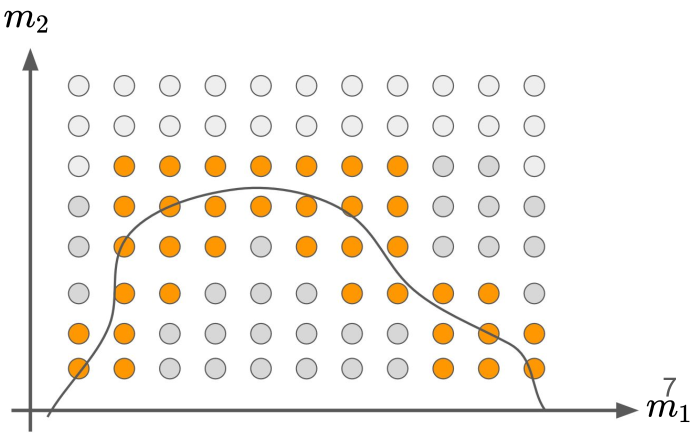
Reduce total number of points

Iterative and parallelized

High density near contour & vice versa

Maintain full theory dimensionality





Dark Matter Search

ATLAS CONF Note

ATLAS-CONF-2018-039

25th July 2018

Search for Dark Matter Produced in Association with a Higgs Boson decaying to $b\bar{b}$ at $\sqrt{s}$ = 13 TeV with the ATLAS Detector using 79.8 fb^{-1} of proton-proton collision data

Reinterpretation

ATLAS PUB Note

ATL-PHYS-PUB-2019-032

11th August 2019

RECAST framework reinterpretation of an ATLAS Dark Matter Search constraining a model of a dark Higgs boson decaying to two b-quarks

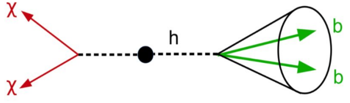
4D Model and Full Run 2 data

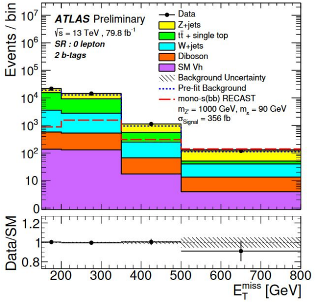
ATLAS PUB Note

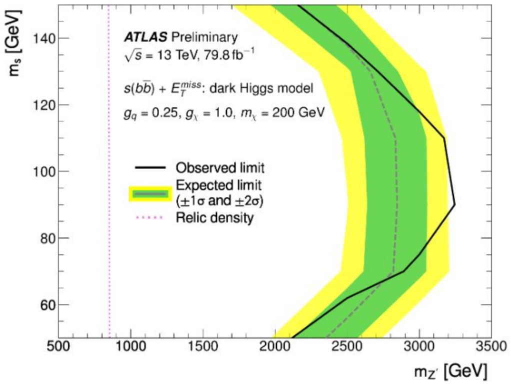
ATL-PHYS-PUB-2022-045

November 3, 2022

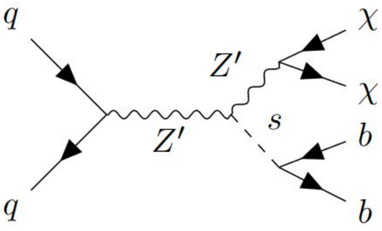
Active Learning reinterpretation of an ATLAS Dark Matter search constraining a model of a dark Higgs boson decaying to two b-quarks



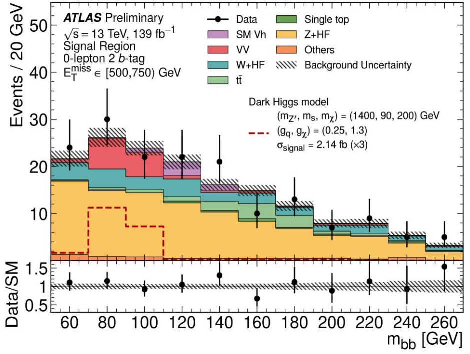




- # From 2D Reinterpretation
  - Sophisticated reconstruction techniques: Large-Radius jets with b-tagged subjects as Higgs boson decay candidates
  - Numerous signal and control regions



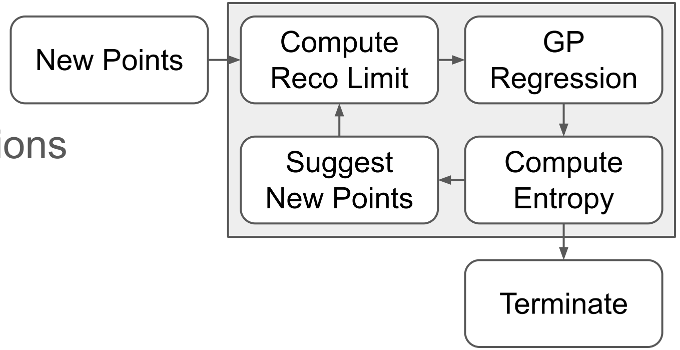
$$\theta = \left(m_{Z'}, m_s, m_\chi, g_\chi\right)$$



- # To a Complete 4D Search
  - Efficient parameter space sampling with Active Learning
  - Full Run 2 139 fb^{-1} data

The Active Learning Approach

Overview of Active Learning



➔ Follow up on ACAT 2017, ACAT 2019 & ACAT 2022 presentations

for each new batch of BSM parameter points $\{\boldsymbol{\theta_i}\}$:

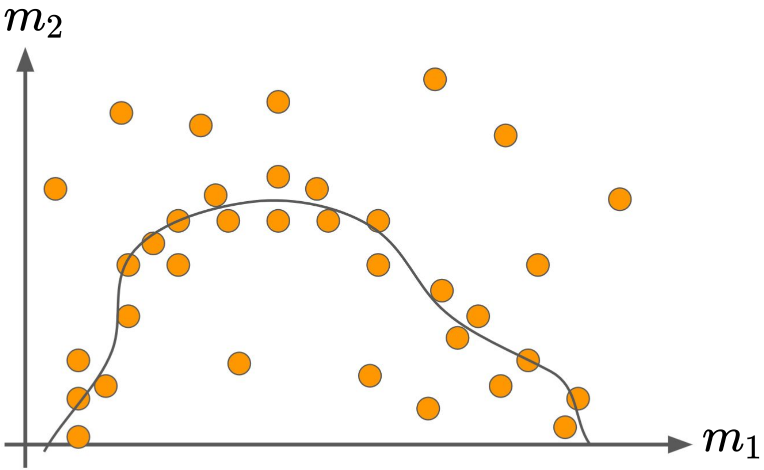
    compute the reco limits $\mu_i = \mu(\boldsymbol{\theta_i})$

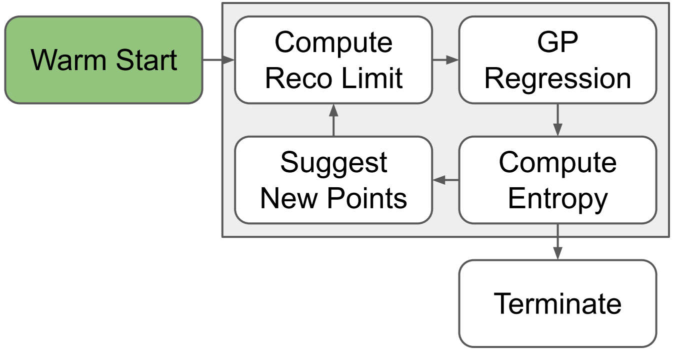
    estimate the function $\mu(\boldsymbol{\theta})$ using GP regression
    given all $\{\mu_k\}$

    compute exclusion entropy

    if the surface $\mu == 1$ is not yet sharp:
        suggest new points $\{\boldsymbol{\theta}_j\}$ where exclusion
        is most uncertain

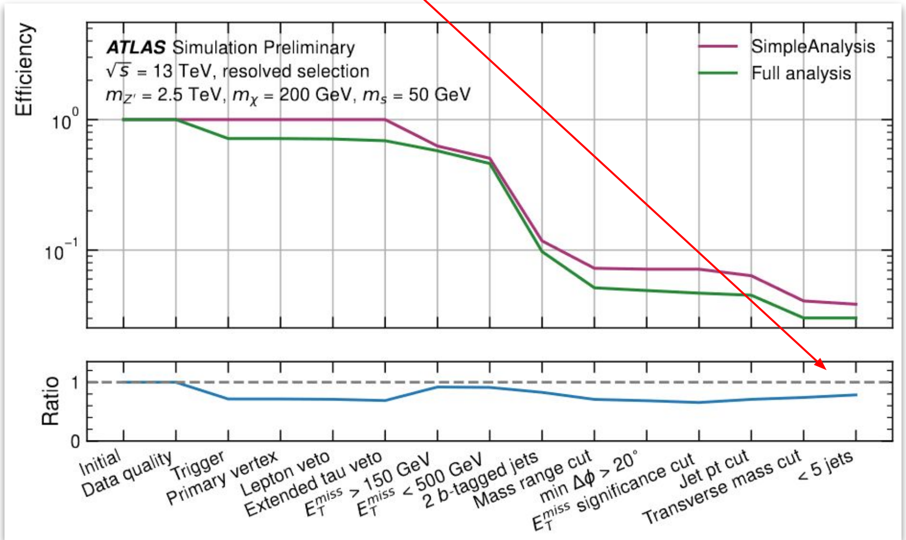
    else:
        break





- Simplest measure: total cross-section
  - Exclusion limits O(1 fb) for Run 2 Mono-H(bb)
  - Calculated during first steps in MC simulation

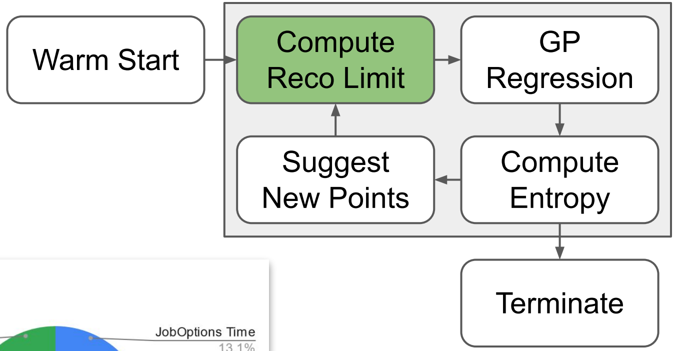
~20% discrepancy in selection efficiency

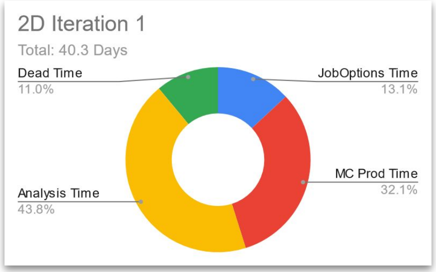


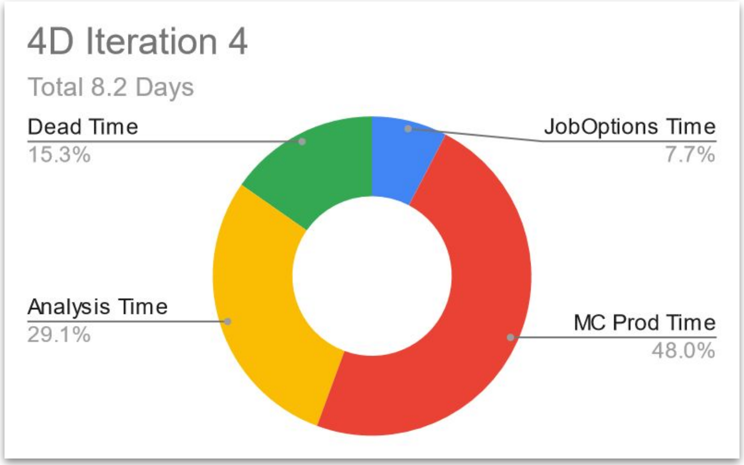
- Next level of accuracy: [SimpleAnalysis]
  - Approximate efficiencies and smearing functions applied on generator-level MC events
  - Grid of 5k points
  - Discrepancies compared to full reinterpretation:
    - Resolved regime : E_T^{miss} significance cut
    - Merged regime : large-R jet substructure

- Full accuracy achieved with RECAST efficiently, running the captured original workflow in a Docker container on CERN's reusable analysis platform (reana)

- Improvements automating MC production:
  - Auto approval of JobOptions through GitLab
  - API based clone MC Production submission
  - First Iteration: ~40 Days for 15 points
  - Last Iteration: ~8 Days for 200 points







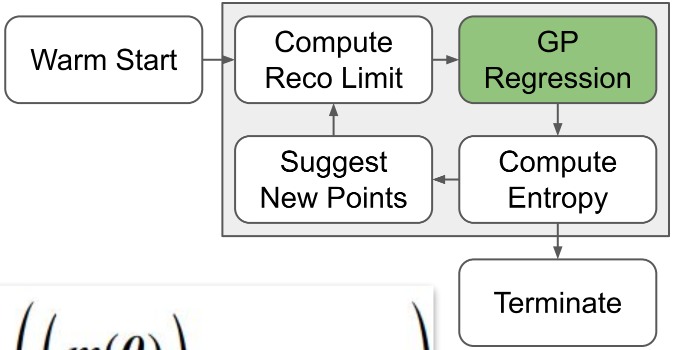
2-Task Gaussian Process

- SimpleAnalysis derived limits on a regular, fine grid support the limit determination
- SimpleAnalysis limit y_s and RECAST limit y_r

Training: 10 parameters determined from SimpleAnalysis and RECAST data (maximum likelihood)

Inference: Gaussian Process prediction of limits and their uncertainties

Python GP library



$$\begin{pmatrix} y_s(\theta) \\ y_r(\theta') \end{pmatrix} \sim \mathcal{GP}\left(\begin{pmatrix} m(\theta) \\ m(\theta') \end{pmatrix}, \Sigma_{sr}(\theta, \theta')\right)$$

$$m(\theta) = w^T\theta + b$$

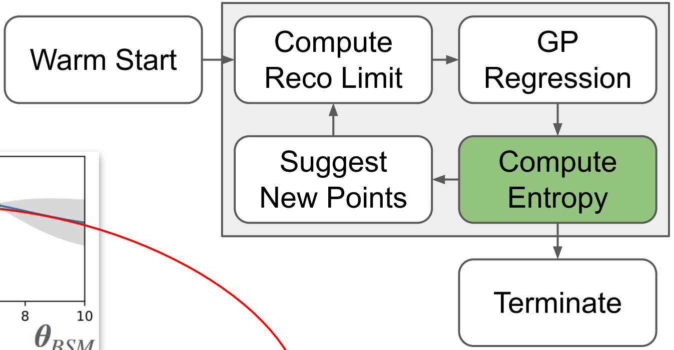
$$\Sigma_{sr}(\theta, \theta') = \begin{pmatrix} k_{ss}(\theta, \theta) & k_{sr}(\theta, \theta') \\ k_{sr}(\theta, \theta') & k_{rr}(\theta', \theta') \end{pmatrix}$$

$$k_{ij}(\theta, \theta') = k(\theta, \theta')\kappa_{ij} + \epsilon^2\delta(\theta, \theta')$$

$$k(\theta, \theta') = \exp\left(-\frac{||\theta - \theta'||^2}{2\,l^2}\right)$$

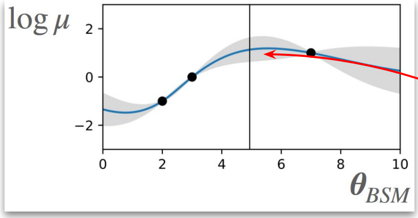
$$\kappa_{ij} = \begin{cases} \sigma_s \text{ if } i = j = s \\ \sigma_r \text{ if } i = j = r \\ \sigma_{sr} \text{ if } i \neq j \end{cases}$$

$$\delta(\theta, \theta') = \begin{cases} 1, & \text{if } \theta = \theta' \\ 0, & \text{else} \end{cases}$$



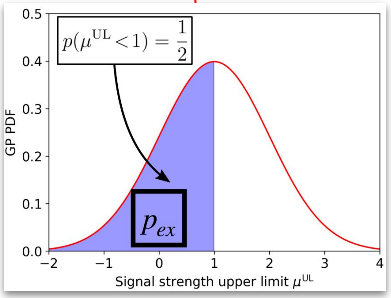
Exclusion probability / Uncertainty Estimate

$$p_{\mathrm{excl}}(\boldsymbol{\theta}) = \int_{-\infty}^{0} g\left(y \mid \mu(\boldsymbol{\theta}), \sigma(\boldsymbol{\theta})\right) dy$$



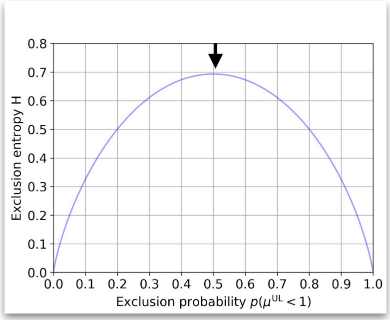
Exclusion entropy for each new point

$$H_{\mathrm{excl}}(\boldsymbol{\theta}) = -p_{\mathrm{excl}}(\boldsymbol{\theta}) \log p_{\mathrm{excl}}(\boldsymbol{\theta}) - \left(1 - p_{\mathrm{excl}}(\boldsymbol{\theta})\right) \log \left(1 - p_{\mathrm{excl}}(\boldsymbol{\theta})\right)$$



Is exclusion entropy across parameter space low?

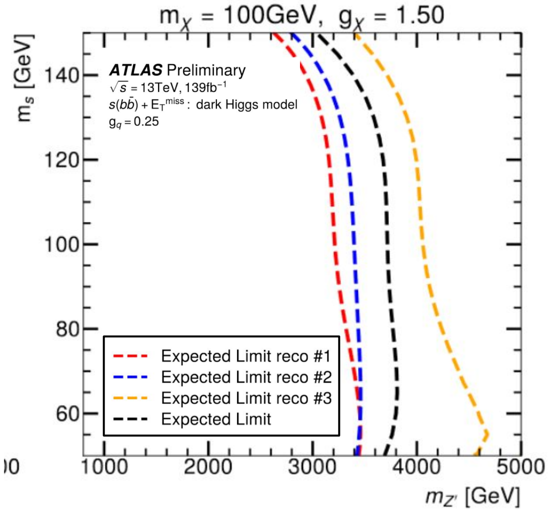
$p_{excl} = 0$ $\qquad$ $p_{excl} = 1$



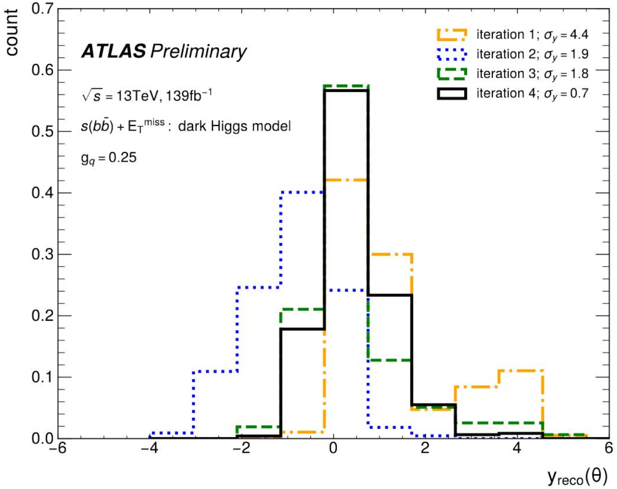
Acquire new points such that H_{excl} is reduced

GP Convergence and Performance

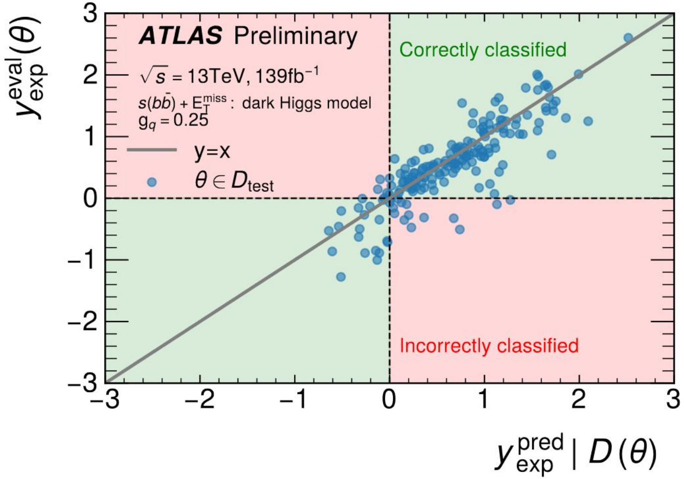
- 4 iterations of RECAST, approx. 200 new points **θ** each
  - Selecting 50% of new points according to max exclusion entropy and 50% Poisson discs around the expected exclusion contour



Stabilization of Limits



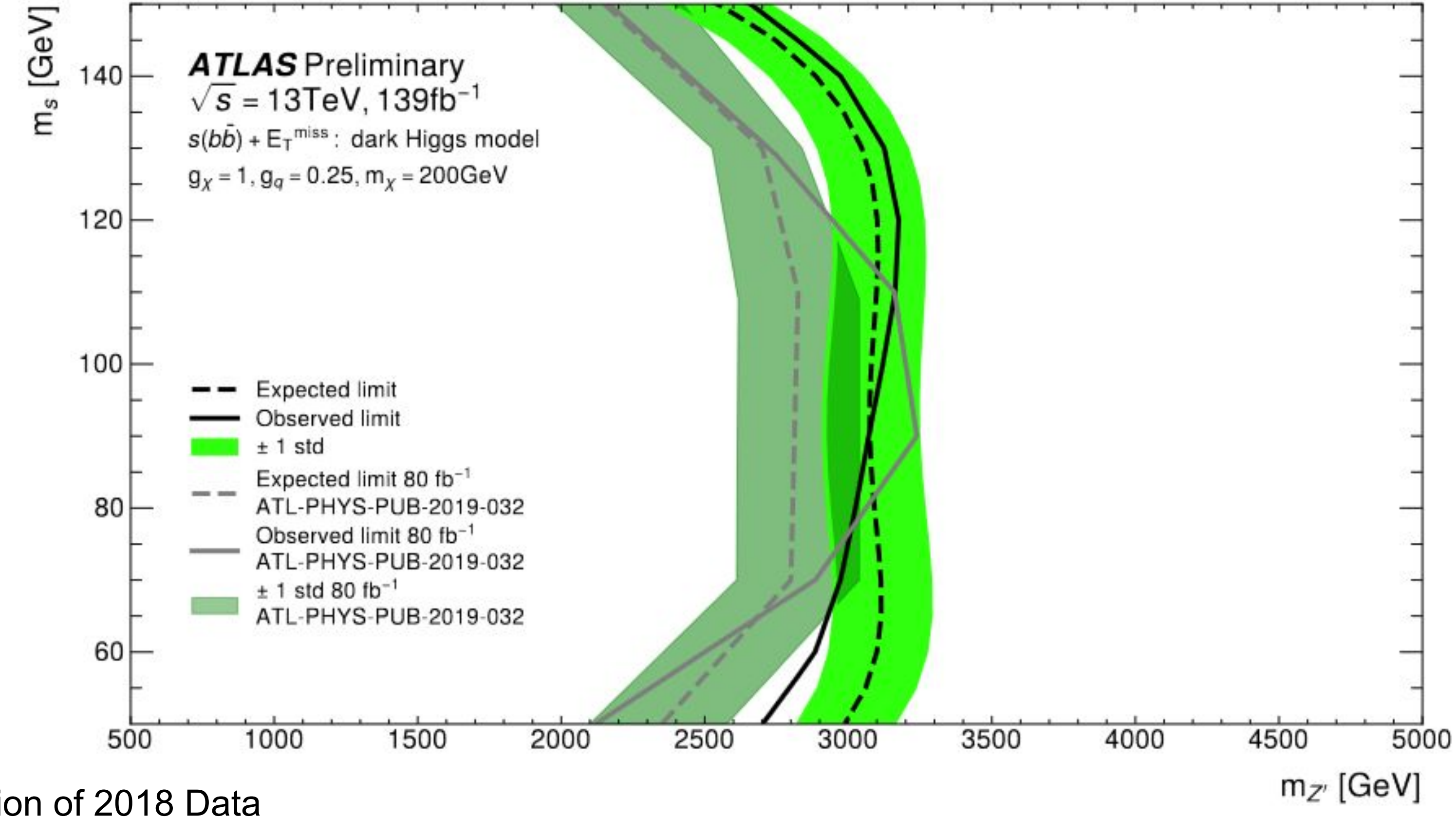
Convergence of Limits



Classification Accuracy 88%

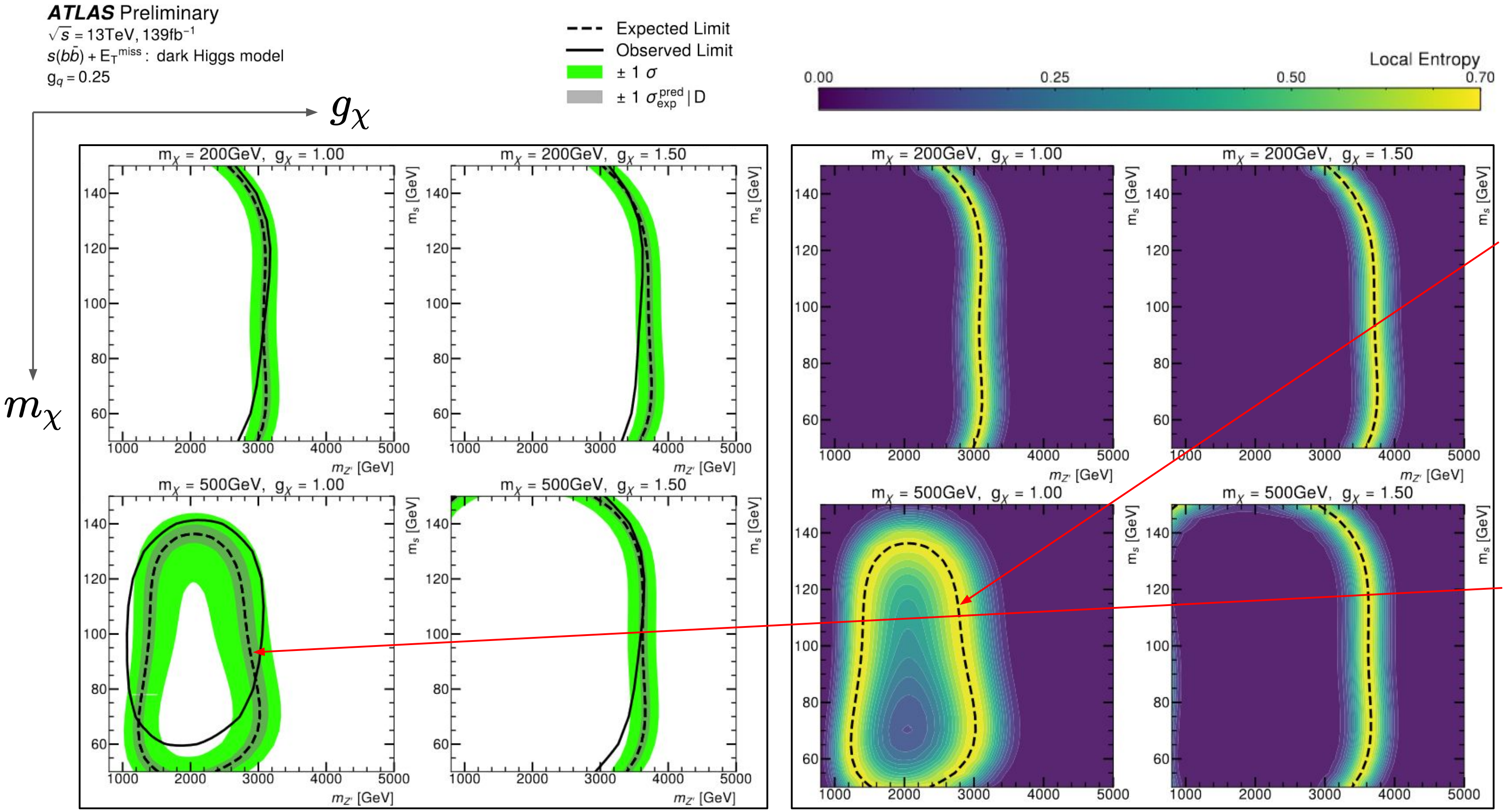
Exclusion Limit Results

Improved Limits Over Previous Search



- Addition of 2018 Data
- Addition of m_χ and g_χ dimensions
- Gaussian Process determination of limits

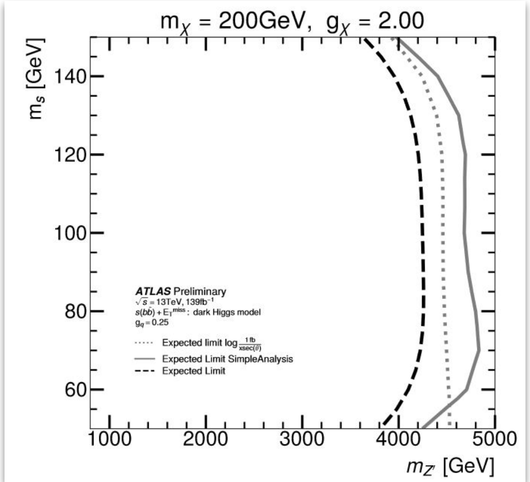
Dark Higgs Mass vs Z’ Mass Exclusion Limits and Entropy in 4D

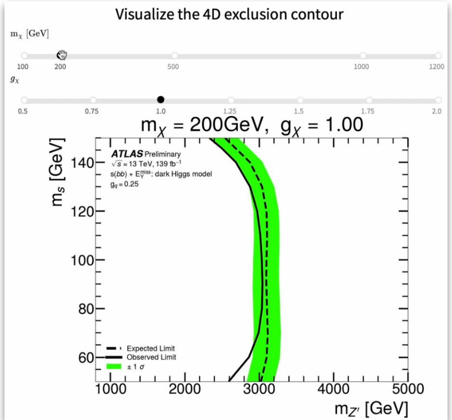


Small exclusion entropy except for exclusion contour region

Negligible intrinsic GP uncertainty

Further Consideration





- Cross-sections provide valuable approximation
  - Warm start with total cross-sections is a quick, simple and useful approximation. SimpleAnalysis may only lead to small gains
- Need for full RECAST in case of this complex physics analysis
  - This finding is true not just for a particular slice but for the whole parameter space, as determined efficiently by means of Active Learning

- 4D Exclusion Contour Visualization
- Trained GP published for internal ATLAS use
- Check out Christian's talk from earlier!

Closing Remarks

- First application of Active Learning for an efficient, accurate and comprehensive interpretation of a search for new physics, using the Run 2 Mono-H(bb) Dark Matter search from ATLAS

- Gaussian Process for exclusion contour determination, making use of its uncertainty estimate in order to acquire new limits efficiently

- Improved dark Higgs boson limits

- Complexity of this analysis ⇒ need for full limit accuracy obtained with RECAST