Link to slides and Google Colab notebook: http://bit.ly/smuon

MSSM in the context of Dark Matter and $(g\text{-}2)_\mu$

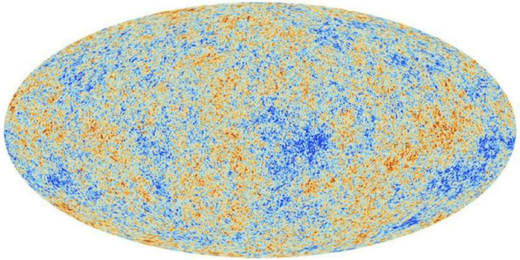
John Tamanas w/ Stefano Profumo

SUSY 2019 - TAMU-CC

arXiv:1906.0xxxx

DM and $(g\text{-}2)_\mu$

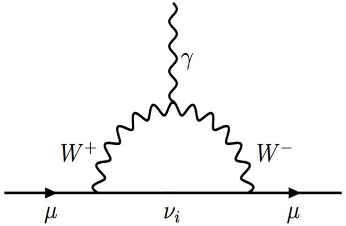
Dark matter exists and can be explained by the neutralino in the MSSM



$$\Omega_c h^2 = 0.12 \pm 0.0012$$

Discrepancy between SM prediction for and experiments measuring $(g\text{-}2)_\mu$

$$\Delta a_\mu = a_\mu^{\mathrm{exp}} - a_\mu^{\mathrm{SM}} = 268(63)(43) \times 10^{-11}$$



arXiv:1807.06209

M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018)

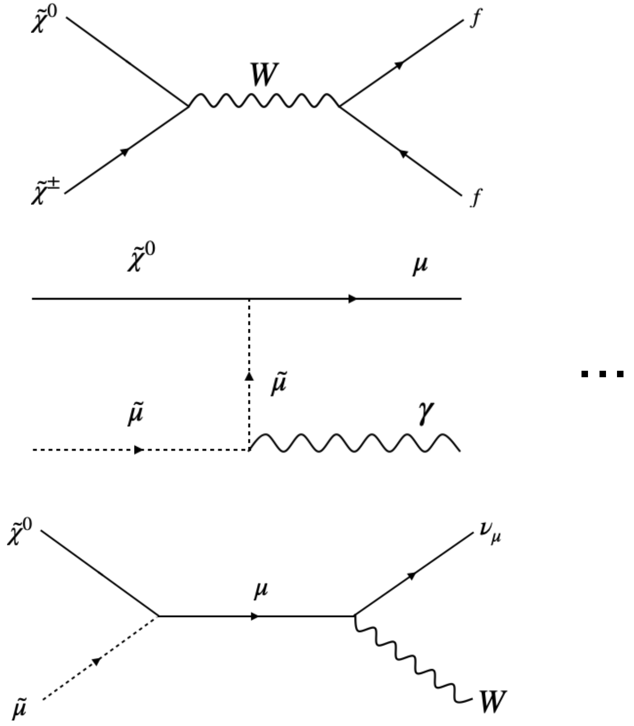
https://bit.ly/smuon

MSSM explanations for dark matter

Bino-like neutralino. $M_{\tilde{\chi}_1^0} \approx M_1 \leq M_2, \mu$

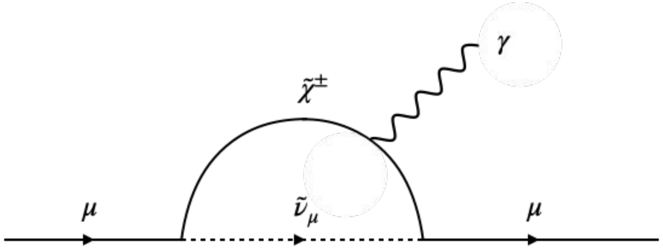
Co-annihilations provide correct relic density for sub-TeV masses.

Assume no R-parity violation so it is absolutely stable.

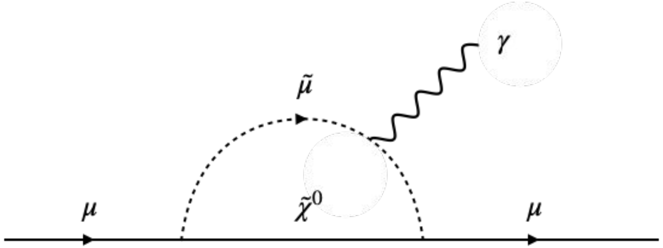


MSSM explanations for (g-2)$_\mu$

Charginos, smuons, and sneutrinos enter loop



$$\begin{aligned}\Delta a_\mu^{\chi^\pm\tilde{\nu}} &= m_\mu \sum_X \Big[m_\mu (C_X^L C_X^L + C_X^R C_X^R) \{ J_4(m^2_{\chi^\pm X}, m^2_{\chi^\pm X}, m^2_{\chi^\pm X}, m^2_{\chi^\pm X}) \\ &+ m^2_{\tilde{\nu}} J_5(m^2_{\chi^\pm X}, m^2_{\chi^\pm X}, m^2_{\chi^\pm X}, m^2_{\chi^\pm X}, m^2_{\tilde{\nu}}) - J_4(m^2_{\chi^\pm X}, m^2_{\chi^\pm X}, m^2_{\chi^\pm X}, m^2_{\tilde{\nu}}) \} \\ &- 2 m_X C_X^L C_X^R J_4(m^2_{\chi^\pm X}, m^2_{\chi^\pm X}, m^2_{\chi^\pm X}, m^2_{\tilde{\nu}}) \Big] \\ &= \frac{1}{16\pi^2} m_\mu \sum_X \left\{ \frac{m_\mu}{3 m^2_{\tilde{\nu}} (1 - x_X)^4} (C_X^L C_X^L + C_X^R C_X^R) \right. \\ &\times \left(1 + \frac{3}{2} x_X - 3 x_X^2 + \frac{1}{2} x_X^3 + 3 x_X \ln x_X \right) \\ &\left. - \frac{3 m_{\chi^\pm X}}{m^2_{\tilde{\nu}} (1 - x_X)^3} C_X^L C_X^R \left(1 - \frac{4}{3} x_X + \frac{1}{3} x_X^2 + \frac{2}{3} \ln x_X \right) \right\}\end{aligned}$$

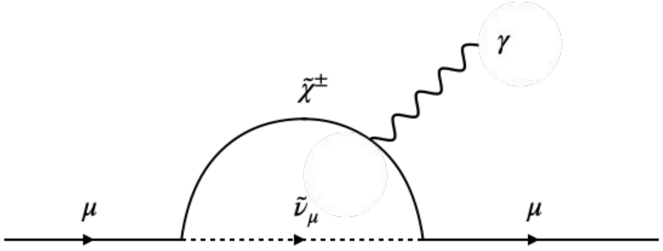


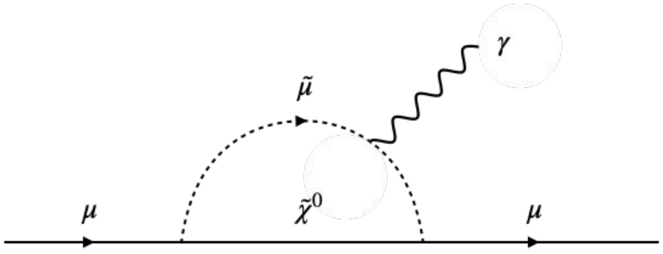
$$\begin{aligned}\Delta a_\mu^{\chi^0\tilde{\mu}} &= m_\mu \sum_{AX} \Big\{ - m_\mu (N_{AX}^L N_{AX}^L + N_{AX}^R N_{AX}^R) m^2_{\tilde{\mu}A} J_5(m^2_{\chi^0 X}, m^2_{\tilde{\mu}A}, m^2_{\tilde{\mu}A}, m^2_{\tilde{\mu}A}, m^2_{\tilde{\mu}A}) \\ &+ m_{\chi^0 X} N_{AX}^L N_{AX}^R J_4(m^2_{\chi^0 X}, m^2_{\chi^0 X}, m^2_{\tilde{\mu}A}, m^2_{\tilde{\mu}A}) \Big\} \\ &= \frac{1}{16\pi^2} m_\mu \sum_{AX} \left\{ - \frac{m_\mu}{6 m^2_{\tilde{\mu}A} (1 - x_{AX})^4} (N_{AX}^L N_{AX}^L + N_{AX}^R N_{AX}^R) \right. \\ &\times (1 - 6 x_{AX} + 3 x_{AX}^2 + 2 x_{AX}^3 - 6 x_{AX}^2 \ln x_{AX}) \\ &\left. - \frac{m_{\chi^0 X}}{m^2_{\tilde{\mu}A} (1 - x_{AX})^3} N_{AX}^L N_{AX}^R (1 - x_{AX}^2 + 2 x_{AX} \ln x_{AX}) \right\}\end{aligned}$$

arXiv:hep-ph/9512396v3

MSSM explanations for $(g\text{-}2)_\mu$

Charginos, smuons, and sneutrinos enter loop





$$\Delta a_\mu^{\chi^\pm\tilde\nu} = m_\mu \sum_X \Big[m_\mu(C^L_X C^L_X + C^R_X C^R_X)\{ \ldots$$

Roughly

$$\Delta a_\mu^{M_1} \propto \frac{1}{M_1} \qquad \Delta a_\mu^{M_{\tilde\mu}} \propto \frac{1}{M_{\tilde\mu}}$$

$$\Delta a_\mu^{M_2} \propto \frac{1}{M_2} \qquad \Delta a_\mu^{\tan\beta} \propto \tan\beta$$

$$\Delta a_\mu^{\mu} \propto \mu + \frac{1}{\mu}$$

$$\ldots, m^2_{\tilde\mu A}, m^2_{\tilde\mu A})$$

$$-\frac{m_{\chi^0 X}}{m^2_{\tilde\mu A}(1-x_{AX})^3} N^L_{AX} N^R_{AX}(1 - x^2_{AX} + 2x_{AX}\ln x_{AX})\Big\}$$

arXiv:hep-ph/9512396v3

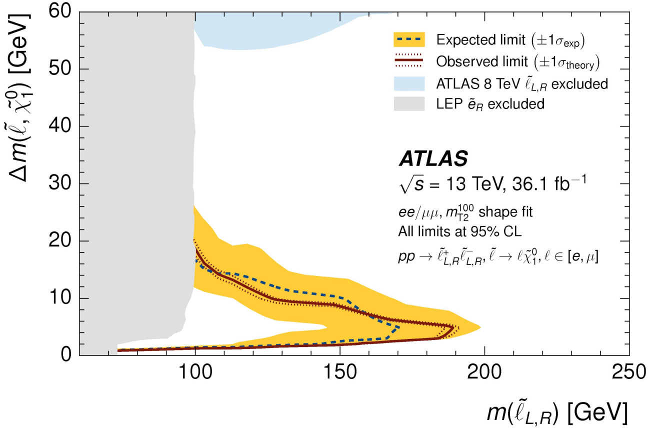
The Goal

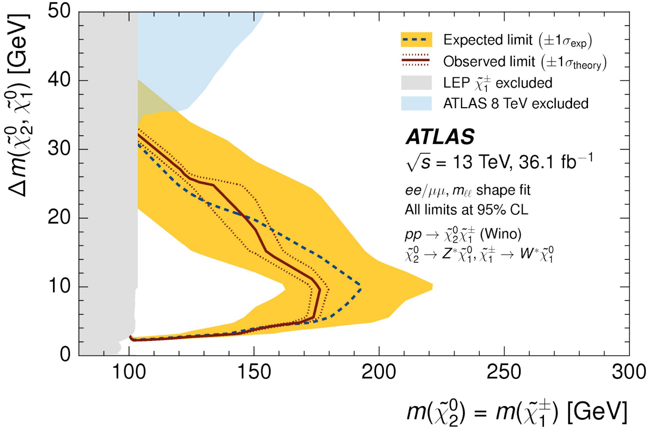
Determine which configurations in parameter space satisfy these experimental limits and how they can be tested.

Parameter Space

For this talk, we'll take the LHC point of view and set all* left and right handed first and second generation slepton masses equal.

Parameters: $(M_1, M_2, \mu, M_{\tilde{\ell}}, \tan\beta)$





* The charged left-handed slepton and sneutrino have a built-in mass splitting coming from interactions with *W*

Sampling a 5-dimensional parameter space is difficult

The Problem

DELFI - Density Estimation Likelihood-free Inference

1. Sample priors and generate a set of points in parameter space, $\boldsymbol{\Theta}$.
2. Evaluate model, *M*, at every element of $\boldsymbol{\Theta}$.
    a. $D = \{M(\theta) | \forall \theta \in \boldsymbol{\Theta}\}$
3. Using how close $d \in D$ is to experimental results, create posteriors.
4. Repeat with geometric mean of priors and posteriors as new priors.

Parametrized Posteriors

Mixture Density Networks

$$p(\mathbf{t}|\boldsymbol{\theta};\mathbf{w}) = \sum_{\text{comp., k}} r_k(\boldsymbol{\theta};\mathbf{w})\,\mathcal{N}\left[\mathbf{t}|\boldsymbol{\mu}_k(\boldsymbol{\theta};\mathbf{w}), \mathbf{C}_k(\boldsymbol{\theta};\mathbf{w}) = \boldsymbol{\Sigma}_k\boldsymbol{\Sigma}_k^T\right]$$

Masked Autoregressive Flows

$$p(\mathbf{t}|\boldsymbol{\theta};\mathbf{w}) = \mathcal{N}\left[\mathbf{u}(\mathbf{t},\boldsymbol{\theta};\mathbf{w})|\mathbf{0},\mathbf{I}\right] \times \prod_{n=1}^{N_{\text{mades}}} \prod_{i=1}^{\dim(\mathbf{t})} \sigma_i^n(\mathbf{t},\boldsymbol{\theta};\mathbf{w})$$

Parametrized Posteriors

Mixture Density Networks

$$p(\mathbf{t}|\boldsymbol{\theta};\mathbf{w}) = \sum_{\text{comp., k}} r_k(\boldsymbol{\theta};\mathbf{w})\,\mathcal{N}\left[\mathbf{t}|\boldsymbol{\mu}_k(\boldsymbol{\theta};\mathbf{w}),\mathbf{C}_k(\boldsymbol{\theta};\mathbf{w}) = \boldsymbol{\Sigma}_k\boldsymbol{\Sigma}_k^T\right]$$

Masked Autoregressive Flows

$$p(\mathbf{t}|\boldsymbol{\theta};\mathbf{w}) = \mathcal{N}\left[\mathbf{u}(\mathbf{t},\boldsymbol{\theta};\mathbf{w})|\mathbf{0},\mathbf{I}\right] \times \prod_{n=1}^{N_{\text{mades}}} \prod_{i=1}^{\dim(\mathbf{t})} \sigma_i^n(\mathbf{t},\boldsymbol{\theta};\mathbf{w})$$

In both cases, individual probability densities are Gaussian.

Neural networks (NNs) take in θ and *d*, and return means, variances, weights, etc.

Parametrized Posteriors

Mixture Density Networks

$$p(\mathbf{t}|\boldsymbol{\theta};\mathbf{w}) = \sum_{\text{comp., k}} r_k(\boldsymbol{\theta};\mathbf{w})\,\mathcal{N}\left[\mathbf{t}|\boldsymbol{\mu}_k(\boldsymbol{\theta};\mathbf{w}),\mathbf{C}_k(\boldsymbol{\theta};\mathbf{w}) = \boldsymbol{\Sigma}_k\boldsymbol{\Sigma}_k^T\right]$$

Masked Autoregressive Flows

$$p(\mathbf{t}|\boldsymbol{\theta};\mathbf{w}) = \mathcal{N}\left[\mathbf{u}(\mathbf{t},\boldsymbol{\theta};\mathbf{w})|\mathbf{0},\mathbf{I}\right] \times \prod_{n=1}^{N_{\text{mades}}} \prod_{i=1}^{\text{dim}(\mathbf{t})} \sigma_i^n(\mathbf{t},\boldsymbol{\theta};\mathbf{w})$$

In both cases, individual probability densities are Gaussian.

Neural networks (NNs) take in θ and *d*, and return means, variances, weights, etc.

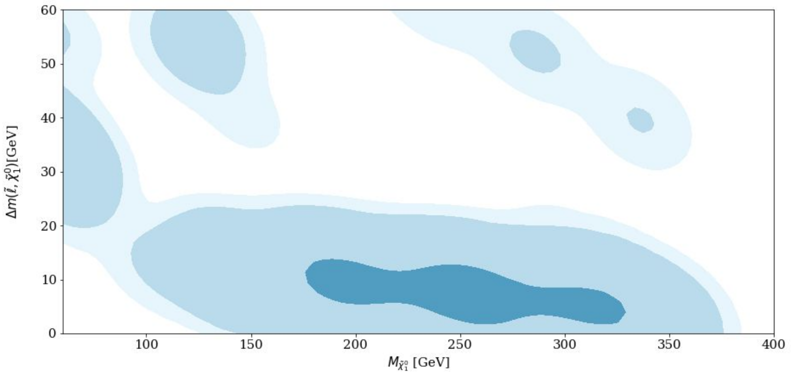
Because probability is estimated at all points parameter space, we can use active learning.

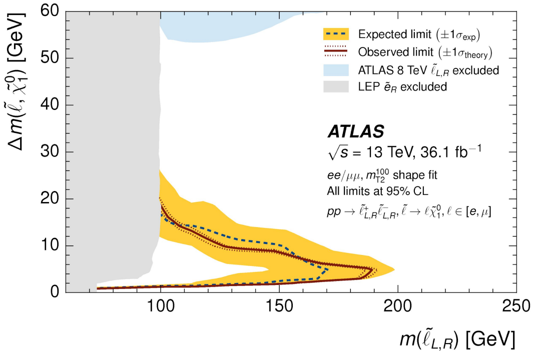
arXiv:1903.00007

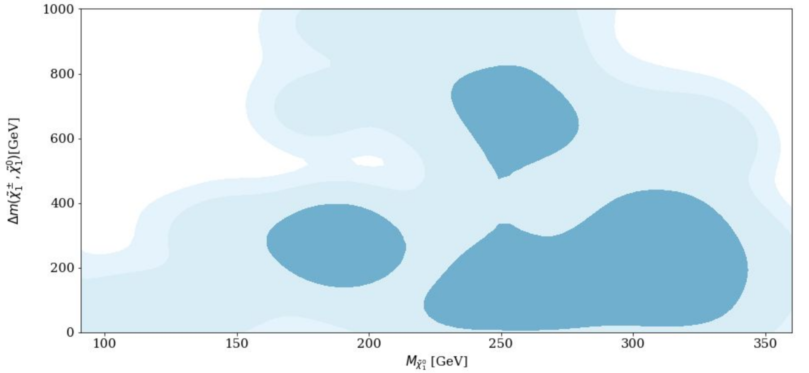
Results

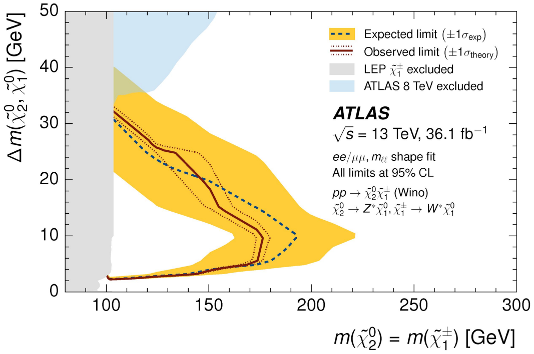
https://bit.ly/smuon

LHC Constraints

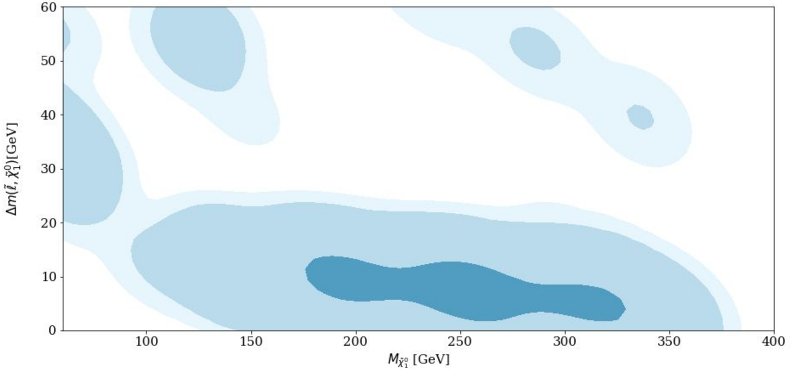


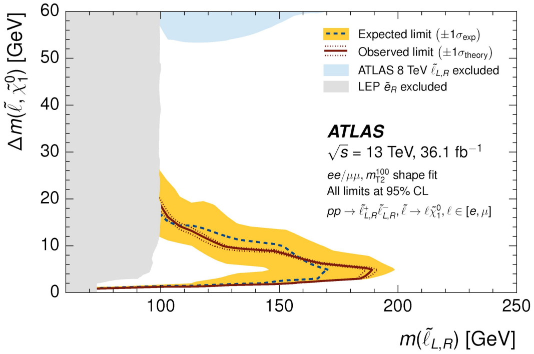


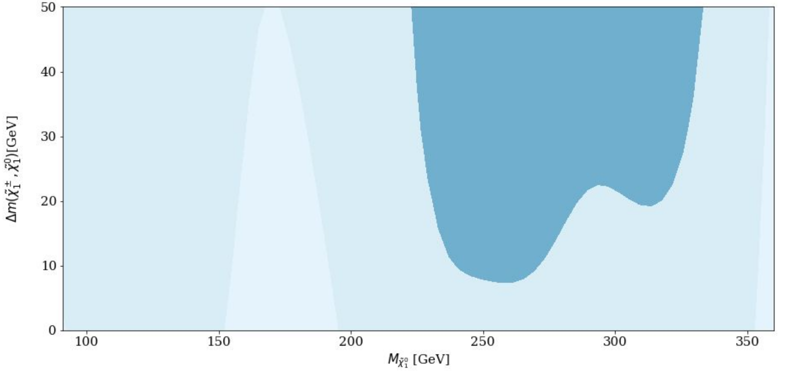


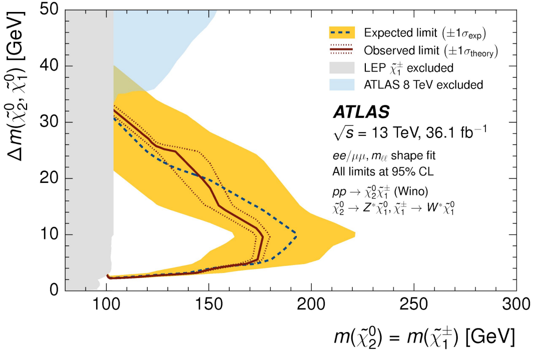


LHC Constraints



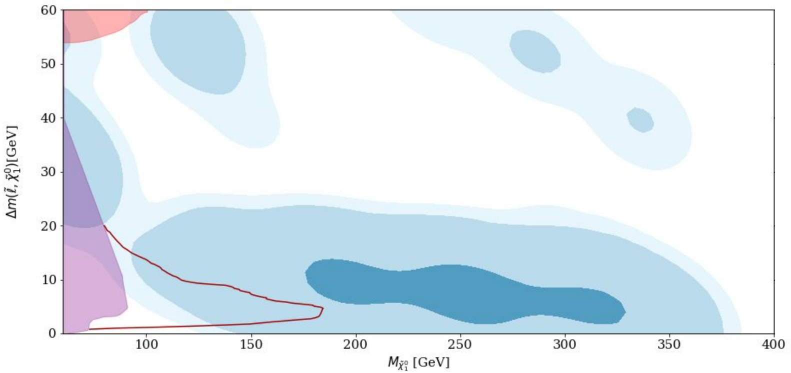


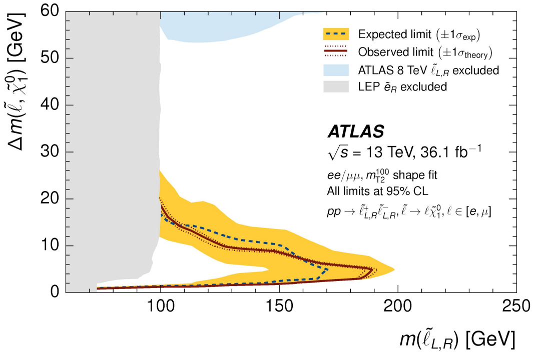


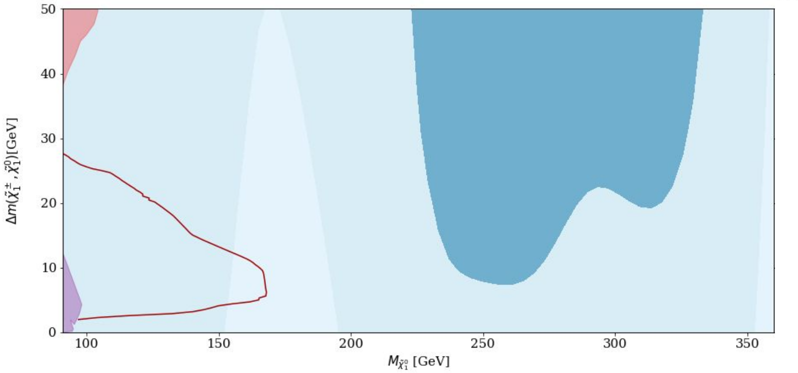


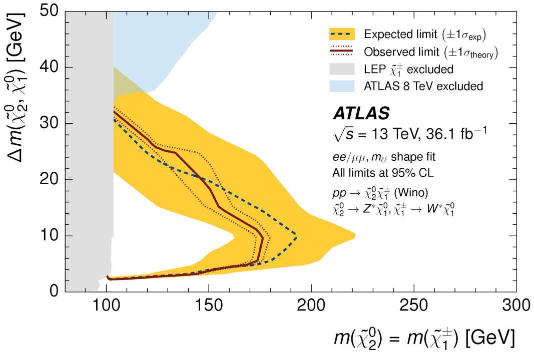
https://bit.ly/smuon

LHC Constraints



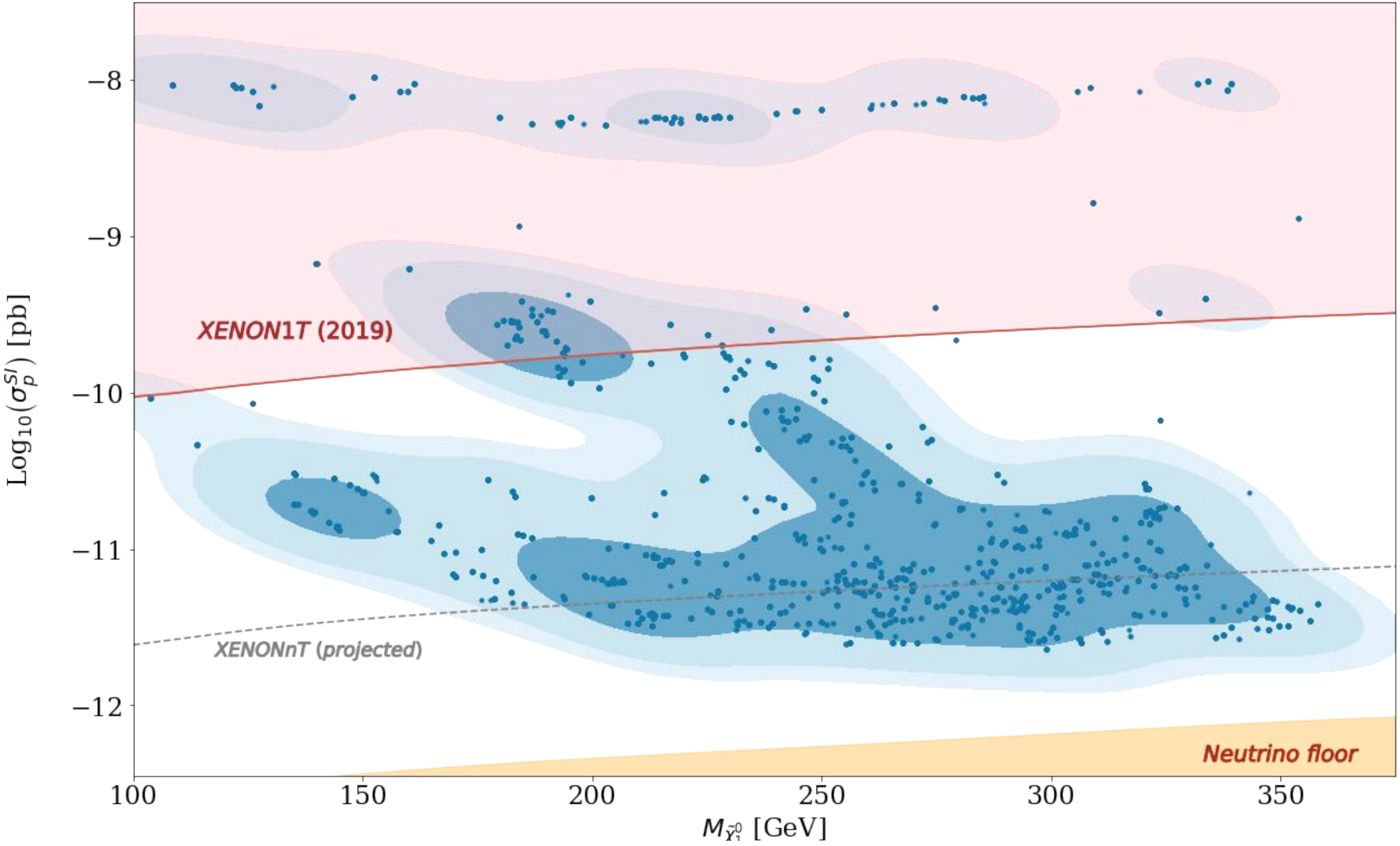






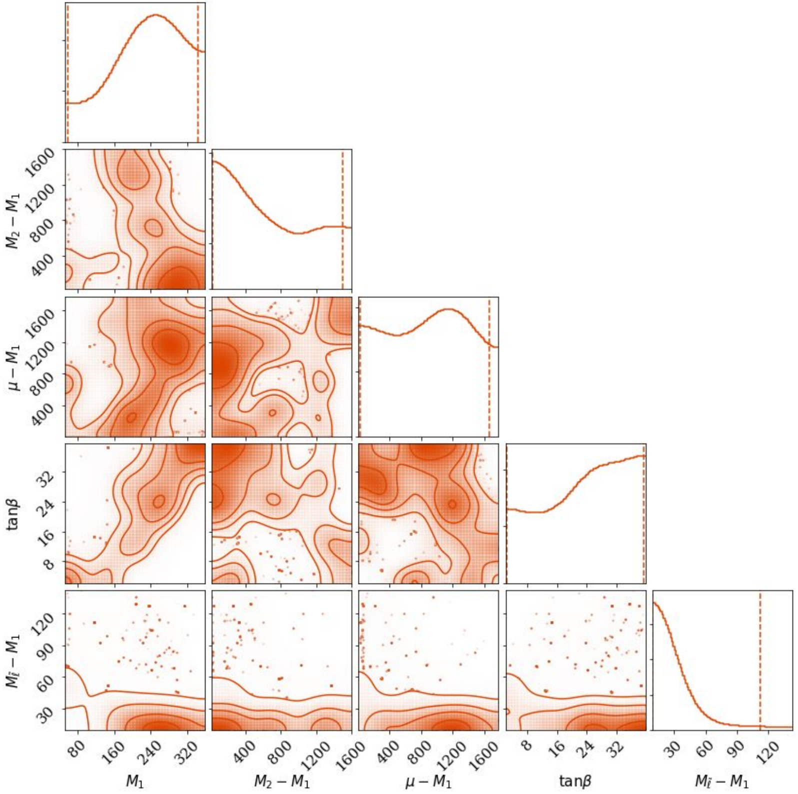
https://bit.ly/smuon

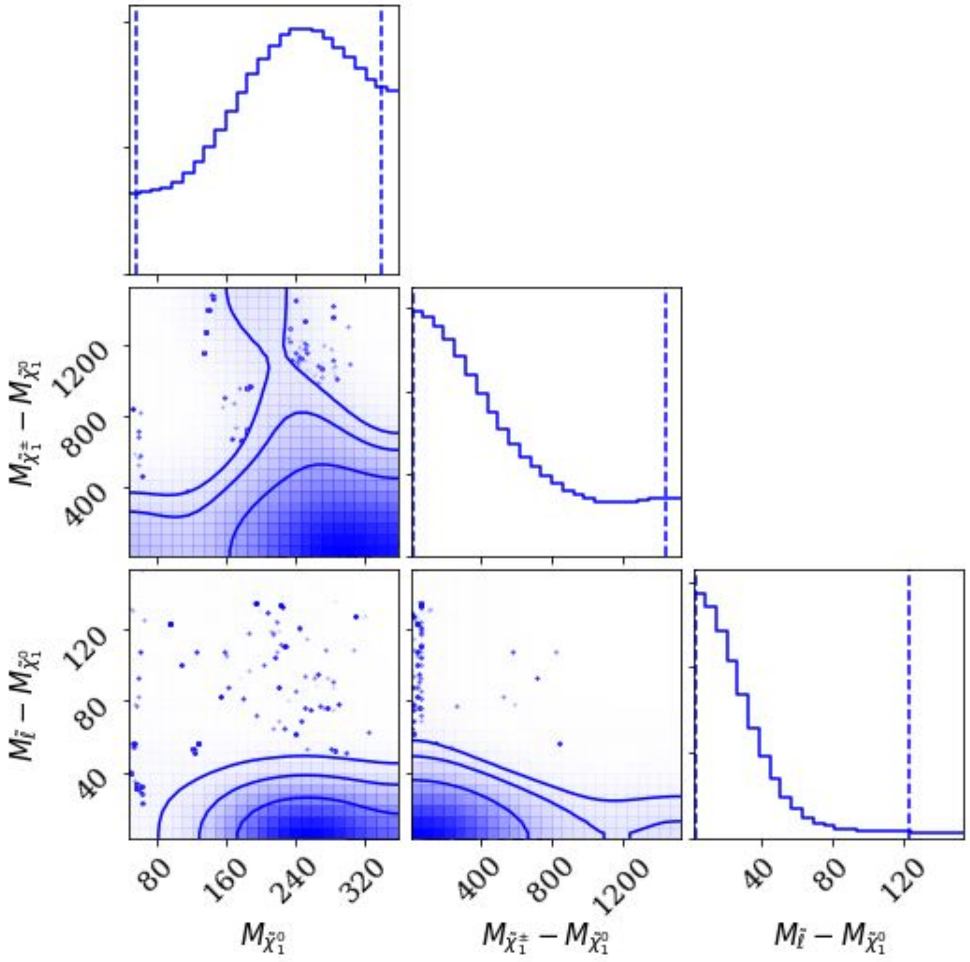
Direct Detection



https://bit.ly/smuon

Density Estimates

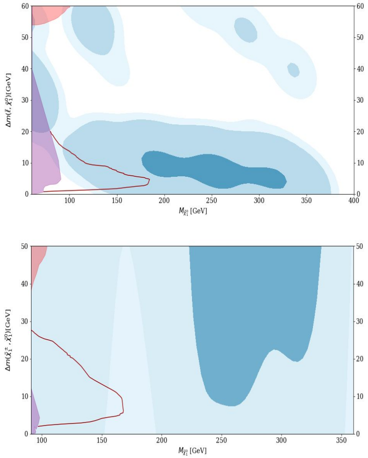


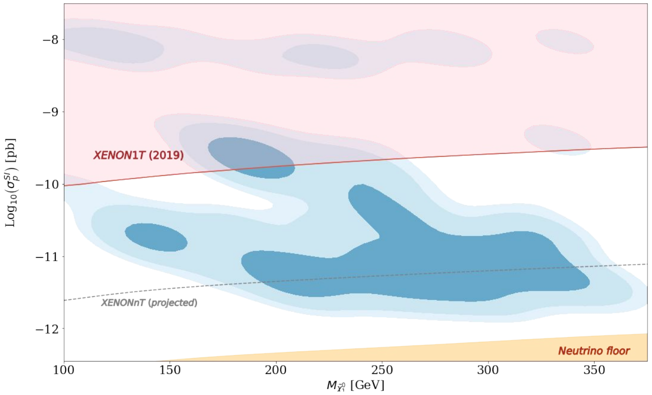


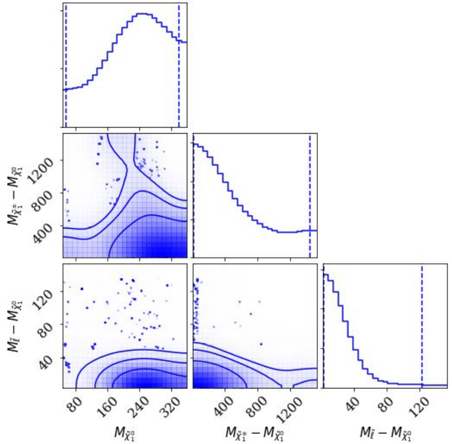
Summary

Using new machine learning tools, we were able to explore and describe areas of parameter space which can explain dark matter relic density and $(g\text{-}2)_\mu$.

Results:



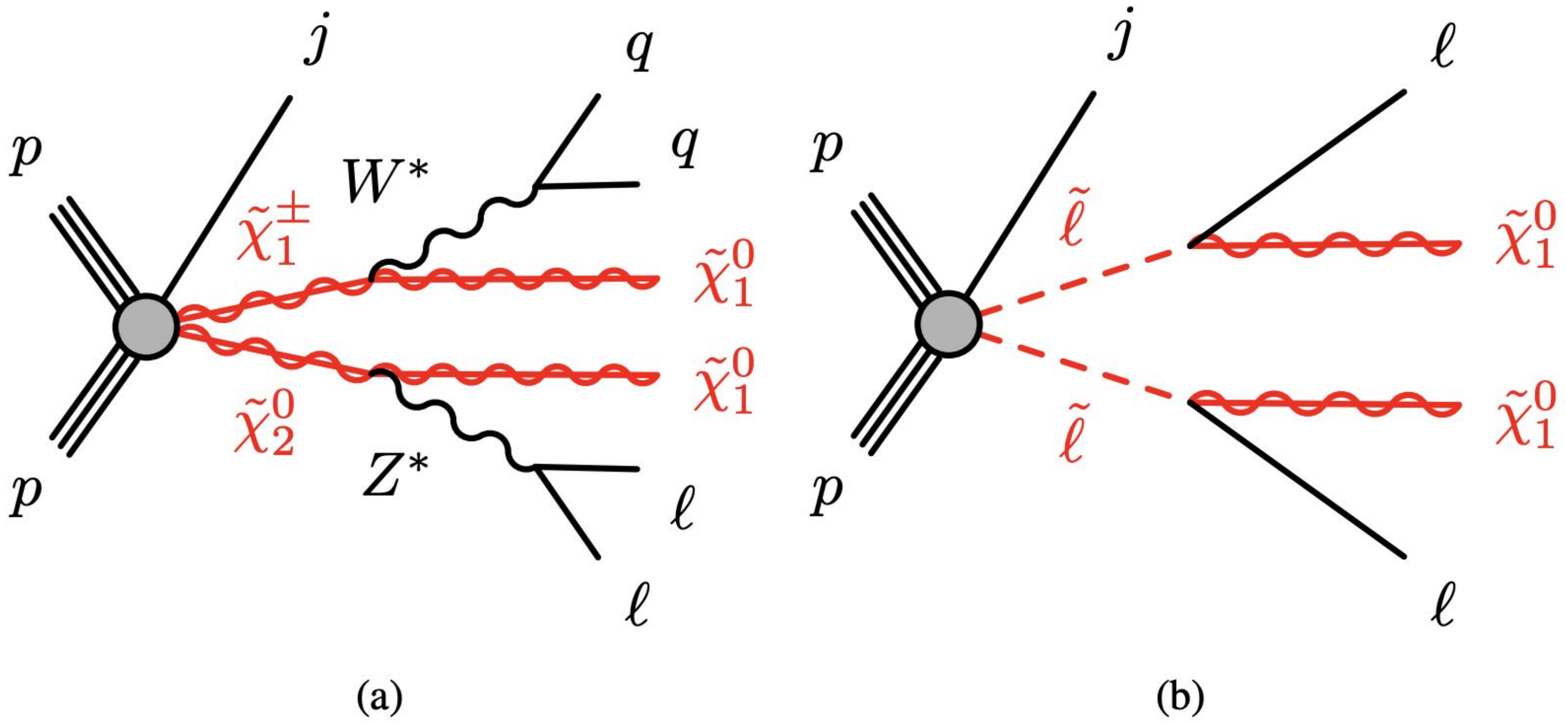




Thank you

Additional Slides

Ewino and slepton pair production



Typical mass spectrum

```
=== MASSES OF HIGGS AND SUSY PARTICLES: ===
Higgs masses and widths
      h   123.99 3.98E-03
      H  6000.01 1.01E+02
     H3  6000.00 1.01E+02
     H+  6000.45 1.01E+02

Masses of odd sector Particles:
~o1      : MNE1    = 145.103 || ~eR      : MSeR    = 242.196 || ~mR      : MSmR    = 242.196
~ne      : MSne    = 243.055 || ~nm      : MSnm    = 243.055 || ~eL      : MSeL    = 255.815
~mL      : MSmL    = 255.815 || ~1+      : MC1     = 406.194 || ~o2      : MNE2    = 406.254
~o3      : MNE3    = 629.535 || ~o4      : MNE4    = 644.603 || ~2+      : MC2     = 645.027
~l1      : MSl1    = 1996.655 || ~nl      : MSnl    = 1998.971 || ~l2      : MSl2    = 2004.367
~t1      : MSt1    = 2338.003 || ~b1      : MSb1    = 2496.438 || ~uL      : MSuL    = 2499.425
~cL      : MScL    = 2499.425 || ~uR      : MSuR    = 2499.752 || ~cR      : MScR    = 2499.752
~dR      : MSdR    = 2500.124 || ~sR      : MSsR    = 2500.124 || ~dL      : MSdL    = 2500.699
~sL      : MSsL    = 2500.699 || ~b2      : MSb2    = 2504.257 || ~t2      : MSt2    = 2658.138
~g       : MSG     = 6000.000 ||
```

Priors

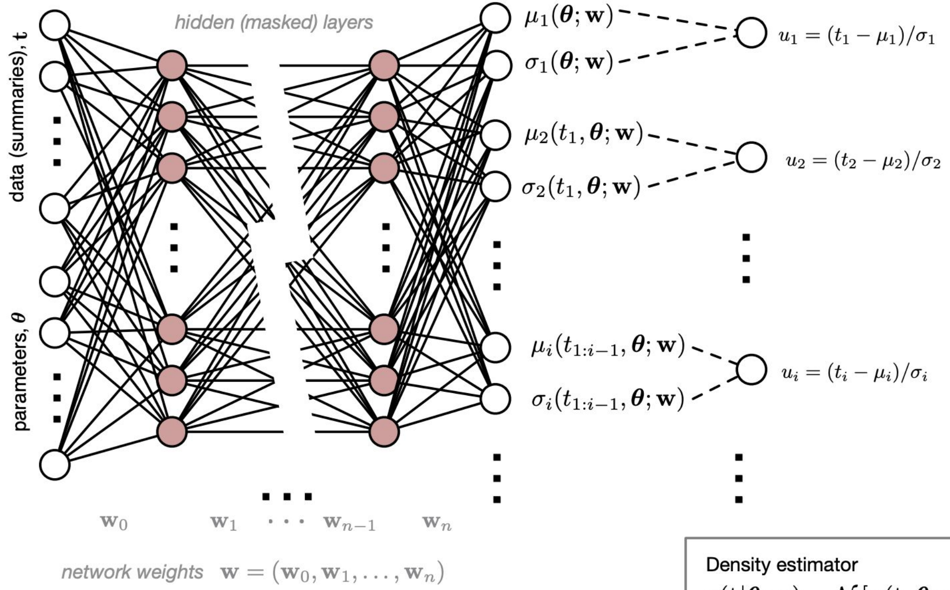
We impose uniform priors with ranges shown on the right

Take differences from M_1 because co-annihilations require small mass splittings with the neutralino.

Search Parameter	Lower	Upper
M_1	50	500
$M_2 - M_1$	0	1750
$M_\mu - M_1$	0	2250
$M_{\tilde{\ell}} - M_1$	0	130
$\tan\beta$	4	40

Masked Autoregressive Flows

Masked Autoencoder for Density Estimation (MADE)



Autoregressive conditionals

$$p(t_1|\boldsymbol{\theta}; \mathbf{w}) = \mathcal{N}\left[\mu_1(\boldsymbol{\theta}; \mathbf{w}), \sigma_1(\boldsymbol{\theta}; \mathbf{w})\right]$$

$$p(t_2|t_1, \boldsymbol{\theta}; \mathbf{w}) = \mathcal{N}\left[\mu_2(t_1, \boldsymbol{\theta}; \mathbf{w}), \sigma_2(t_1, \boldsymbol{\theta}; \mathbf{w})\right]$$

$$\vdots$$

$$p(t_i|t_{1:i-1}, \boldsymbol{\theta}; \mathbf{w}) = \mathcal{N}\left[\mu_i(t_{1:i-1}, \boldsymbol{\theta}; \mathbf{w}), \sigma_i(t_{1:i-1}, \boldsymbol{\theta}; \mathbf{w})\right]$$

$$\vdots$$

Density estimator

$$p(\mathbf{t}|\boldsymbol{\theta}; \mathbf{w}) = \mathcal{N}\left[\mathbf{u}(\mathbf{t}, \boldsymbol{\theta}; \mathbf{w})|0, \mathbf{I}\right] \times \left|\frac{\partial \mathbf{u}(\mathbf{t}, \boldsymbol{\theta}; \mathbf{w})}{\partial \mathbf{t}}\right| = \prod_i p(t_i|t_{1:i-1}, \boldsymbol{\theta}; \mathbf{w})$$

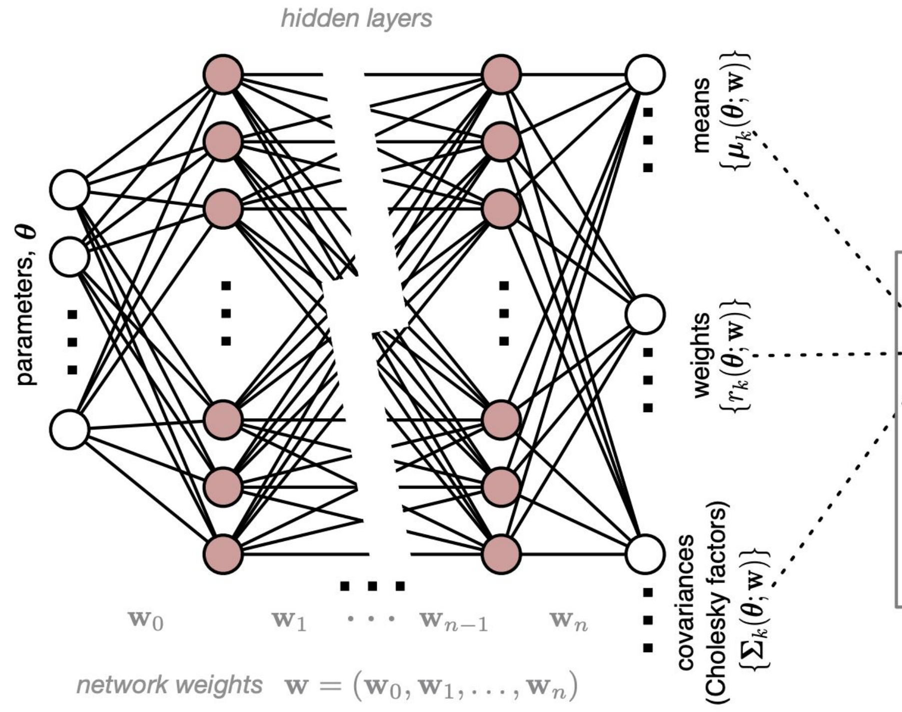
Training data $\{\boldsymbol{\theta}, \mathbf{t}\}$

Loss function $-\ln U = -\sum_i \ln p(\mathbf{t}_i|\boldsymbol{\theta}_i; \mathbf{w})$

pyDELFI: arXiv:1903.00007

Mixture Density Networks

Mixture Density Network (MDN)



Density estimator

$$p(\mathbf{t}|\boldsymbol{\theta};\mathbf{w}) = \sum_{\text{comp., k}} r_k(\boldsymbol{\theta};\mathbf{w})\,\mathcal{N}\left[\mathbf{t}|\boldsymbol{\mu}_k(\boldsymbol{\theta};\mathbf{w}), \mathbf{C}_k(\boldsymbol{\theta};\mathbf{w}) = \boldsymbol{\Sigma}_k\boldsymbol{\Sigma}_k^T\right]$$

Training data $\{\boldsymbol{\theta}, \mathbf{t}\}$

Loss function $-\ln U = -\sum_i \ln p(\mathbf{t}_i|\boldsymbol{\theta}_i;\mathbf{w})$

pyDELFI: arXiv:1903.00007

Active Learning

Sequential Neural Likelihood (SNL)

Sample geometric mean of prior and posterior with MCMC. Take random population of sample to create points in parameter space to search

Bayesian Optimization

Uncertainty is known at every point, so can sample regions where least is known

We used SNL because uncertainty is not necessarily reflective of ignorance, and so optimal choice of parameter space is not trivial (active area of research)

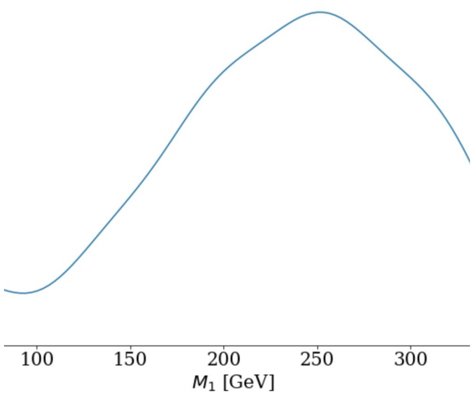
pyDELFI: arXiv:1903.00007

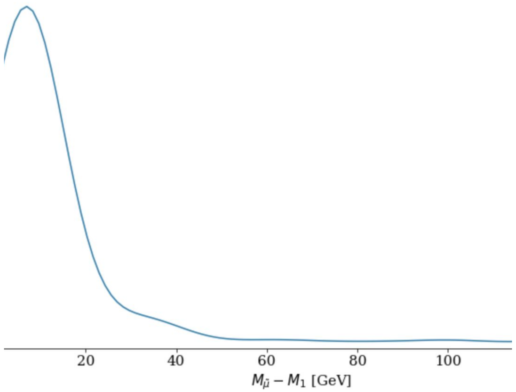
Parameter density estimates

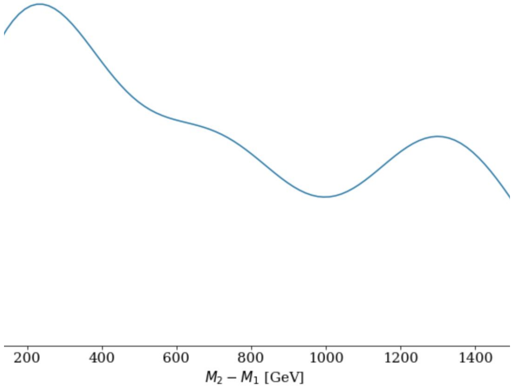
$$\Delta a_\mu^{M_1} \propto \frac{1}{M_1} \qquad \Delta a_\mu^{M_{\tilde{\mu}}} \propto \frac{1}{M_{\tilde{\mu}}}$$

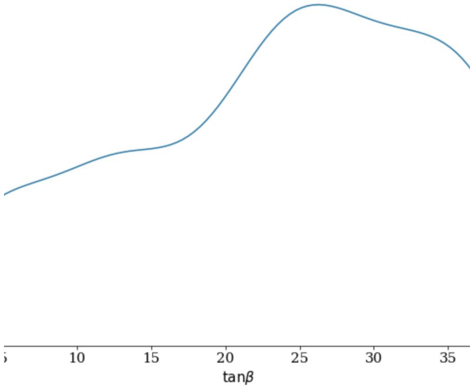
$$\Delta a_\mu^{M_2} \propto \frac{1}{M_2} \qquad \Delta a_\mu^{\tan\beta} \propto \tan\beta$$

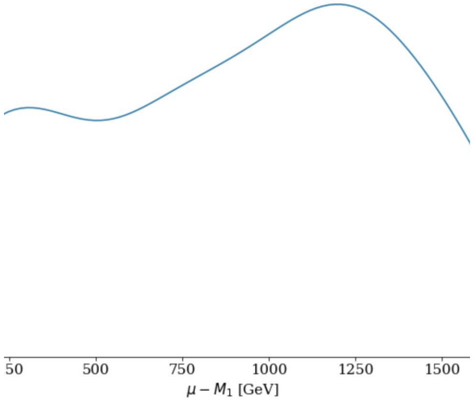
$$\Delta a_\mu^{\mu} \propto \mu + \frac{1}{\mu}$$









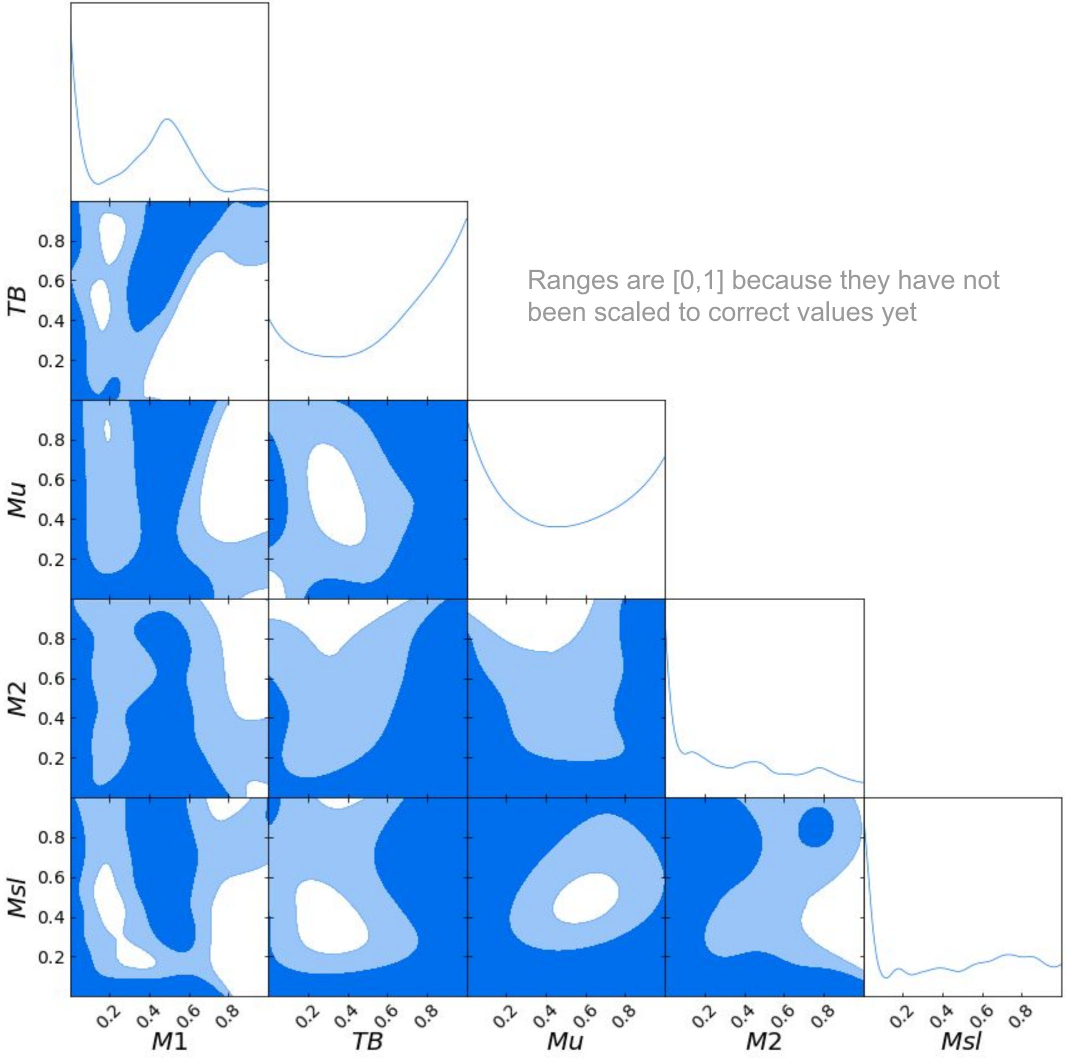


“Polluted” Triangle Plot

After training, we sample the posteriors with MCMC.

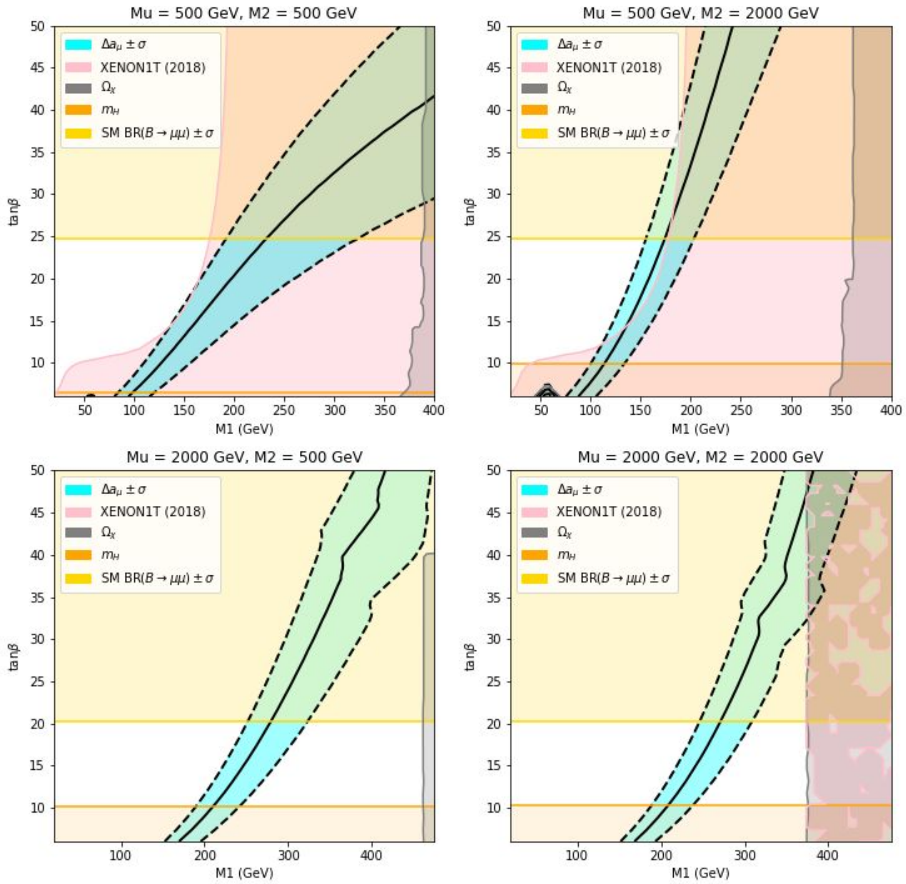
This can lead to (many) points that do not lie within our criteria

This plot is a mix of good and bad points



Ranges are [0,1] because they have not been scaled to correct values yet

Grid search



Mu = 500 GeV, M2 = 500 GeV
Mu = 500 GeV, M2 = 2000 GeV
Mu = 2000 GeV, M2 = 500 GeV
Mu = 2000 GeV, M2 = 2000 GeV
$\Delta a_\mu \pm \sigma$
XENON1T (2018)
Ω_χ
m_H
SM BR($B \to \mu\mu$) $\pm \sigma$
tanβ
M1 (GeV)