SUSY-AI - Generalizing LHC limits on Supersymmetry with Machine Learning

Bob Stienen

November 4th, 2016

The BSM-AI project: SUSY-AI - Generalizing LHC limits on Supersymmetry with Machine Learning

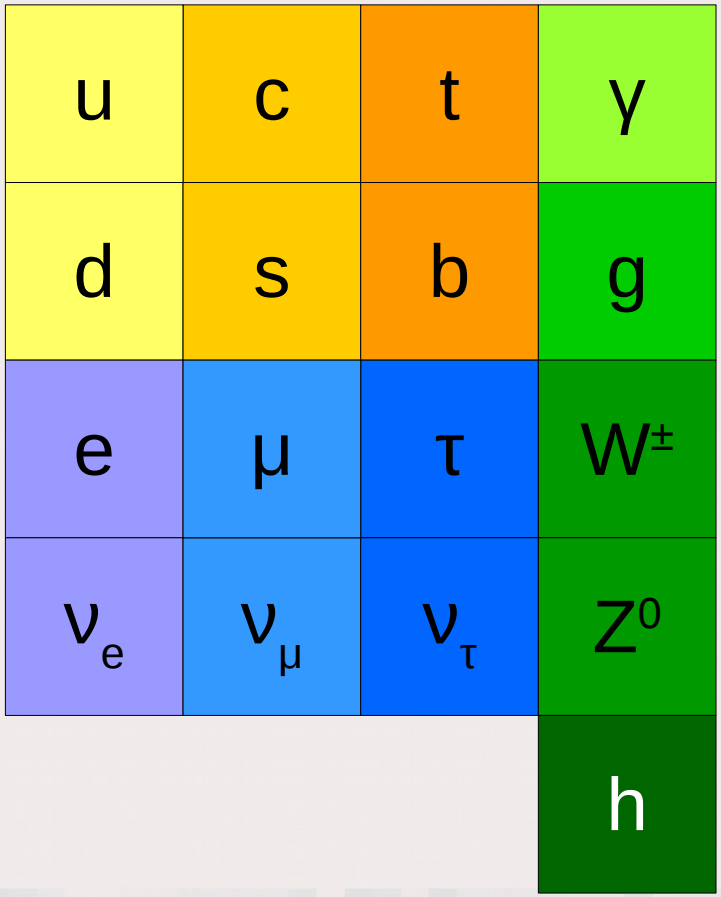
Sascha Caron, Jong Soo Kim, Krzysztof Rolbiecki, Roberto Ruiz de Austri, Bob Stienen

https://arxiv.org/abs/1605.02797

Radboud University

Particle Physics
The Standard Model

- The Standard Model describes all particles and forces in nature
- Best theory we have at the moment
- But... it is incomplete:
  - No gravity
  - No dark matter
  - Serious theoretical concerns



Supersymmetry

General idea All fermions and bosons have a partner particle of the other kind that only differs in spin from themselves

- All problems on previous slide are solved (except gravity)
- **pMSSM** (19 free variables) is the simplest model containing supersymmetry while still obeying experimental constraints.

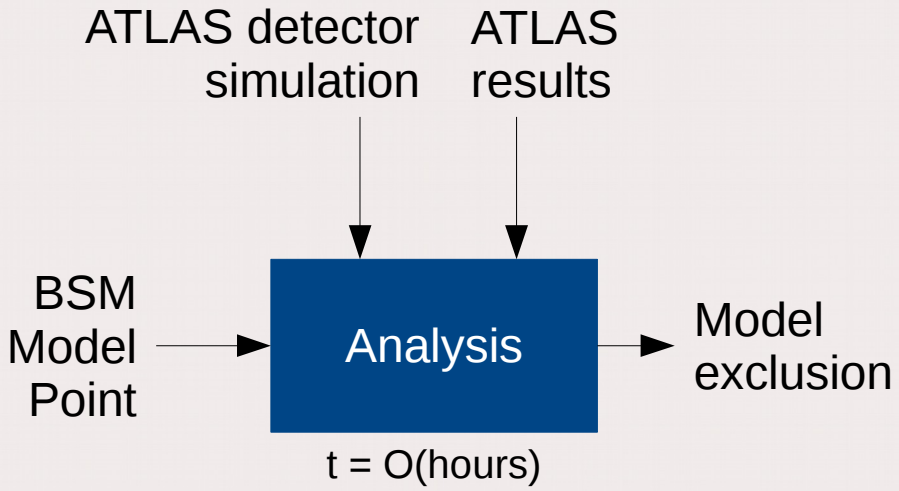
Model points

- A model point is defined by:
  - A theoretical model with X free parameters
  - X values for the free parameters
- For the pMSSM, this means 19 values define a model point
  → point in 19 dimensional space
- Each model point defines unique physics!

- Data from ATLAS: https://arxiv.org/abs/1508.06608
  → *Summary of the ATLAS experiment's sensitivity to supersymmetry after LHC Run 1 - interpreted in the phenomenological MSSM*

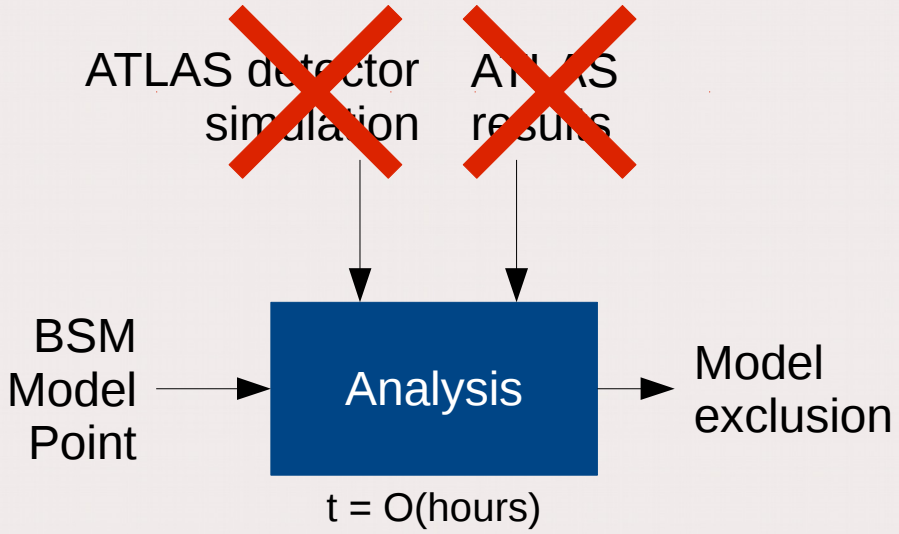
Exclusion analysis

- Create model point
- Simulate events (Monte Carlo)
- Simulate detector
- Simulate detector response
- Compare detector response to actual measurements
- Check for confidence on simulation



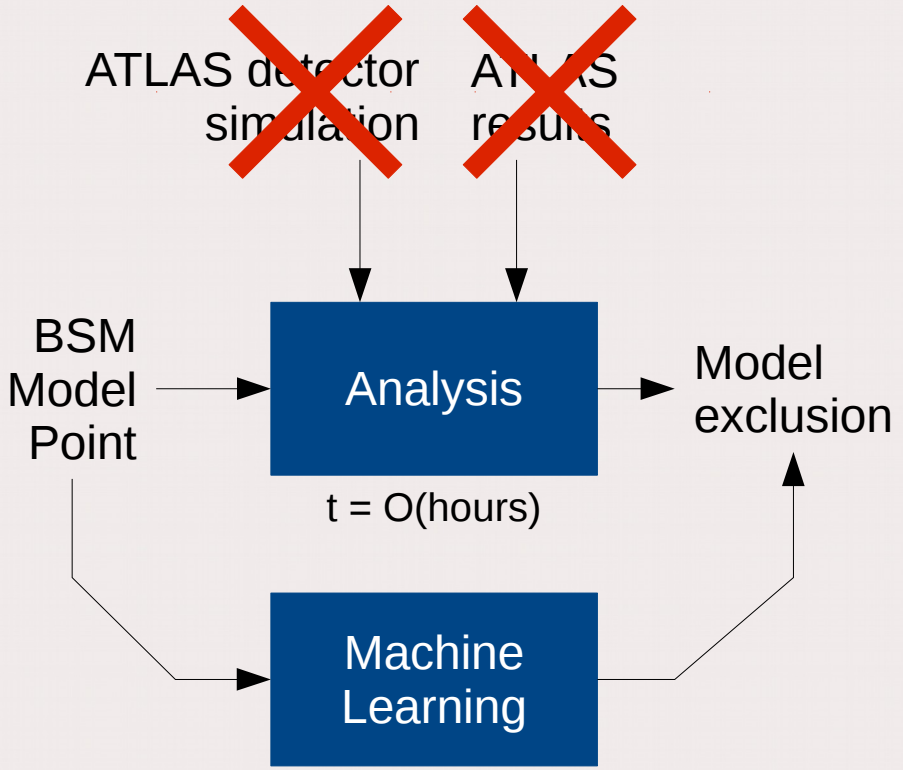
Exclusion analysis

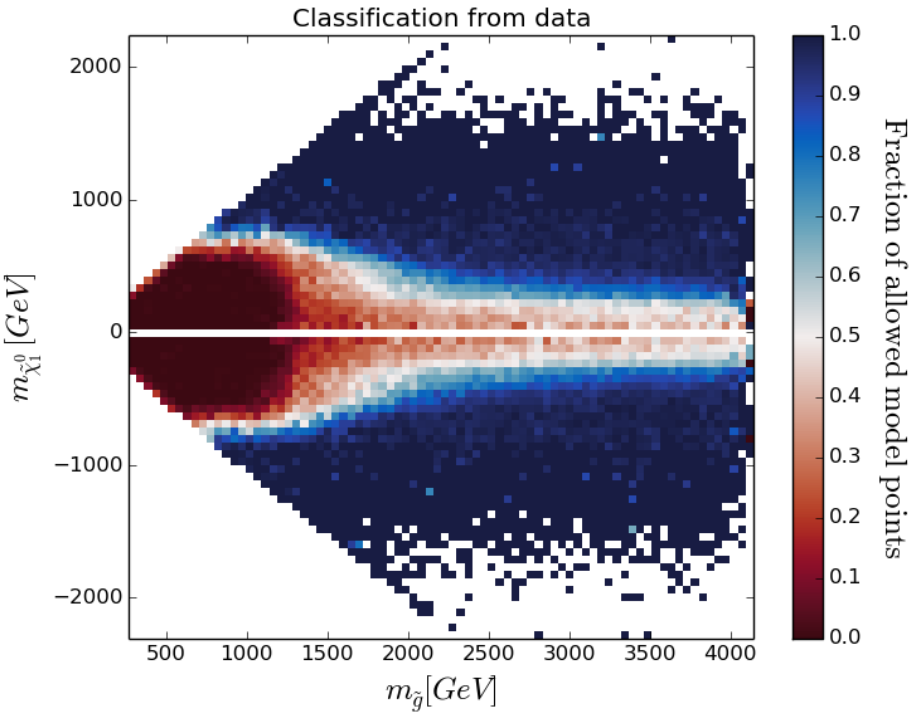
- Create model point
- Simulate events (Monte Carlo)
- Simulate detector
- Simulate detector response
- Compare detector response to actual measurements
- Check for confidence on simulation

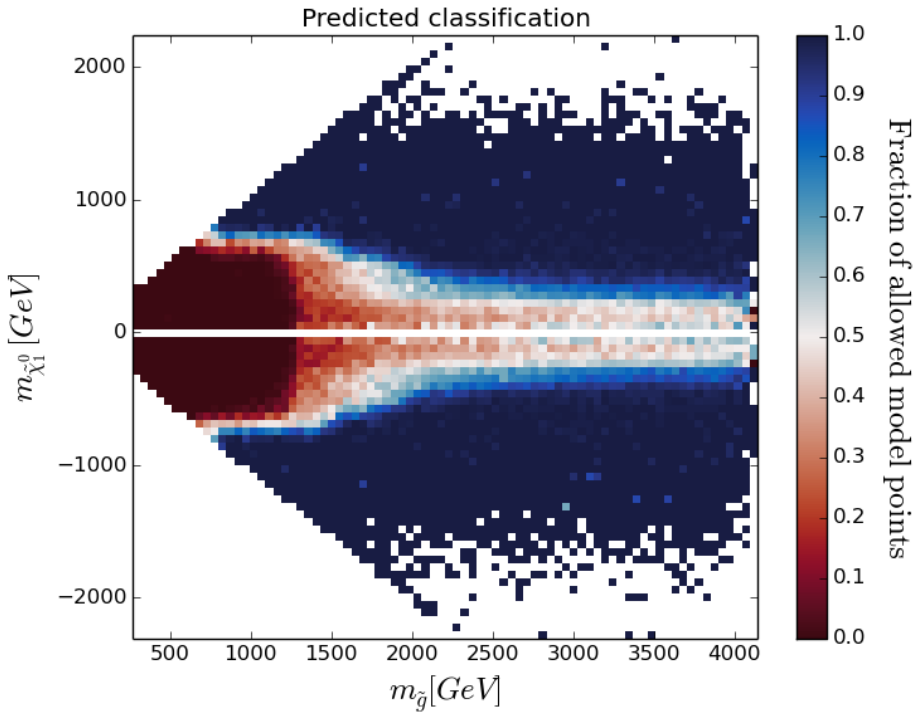


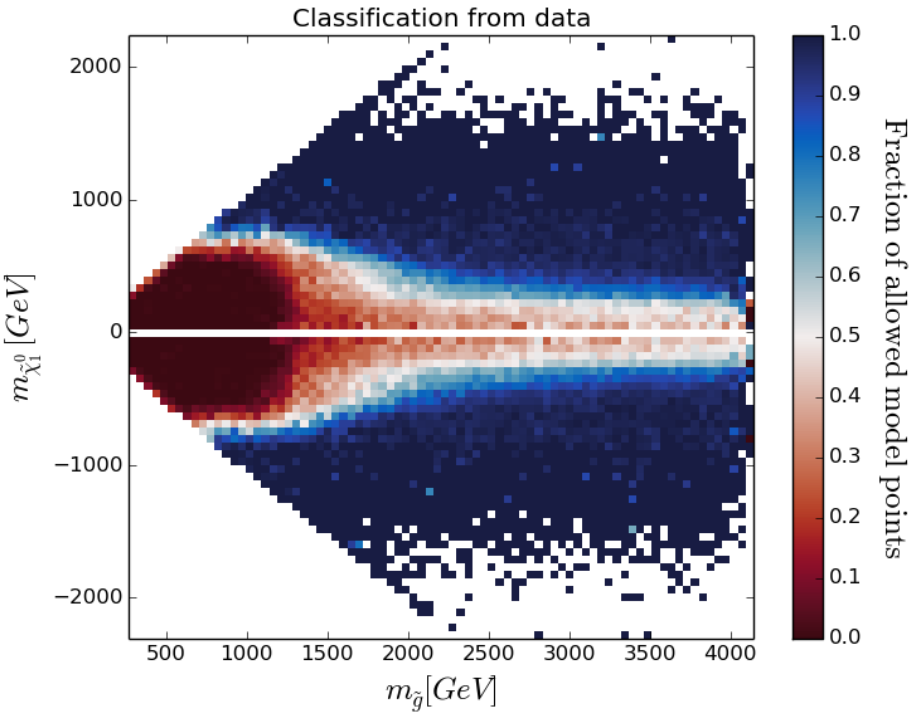
Exclusion analysis

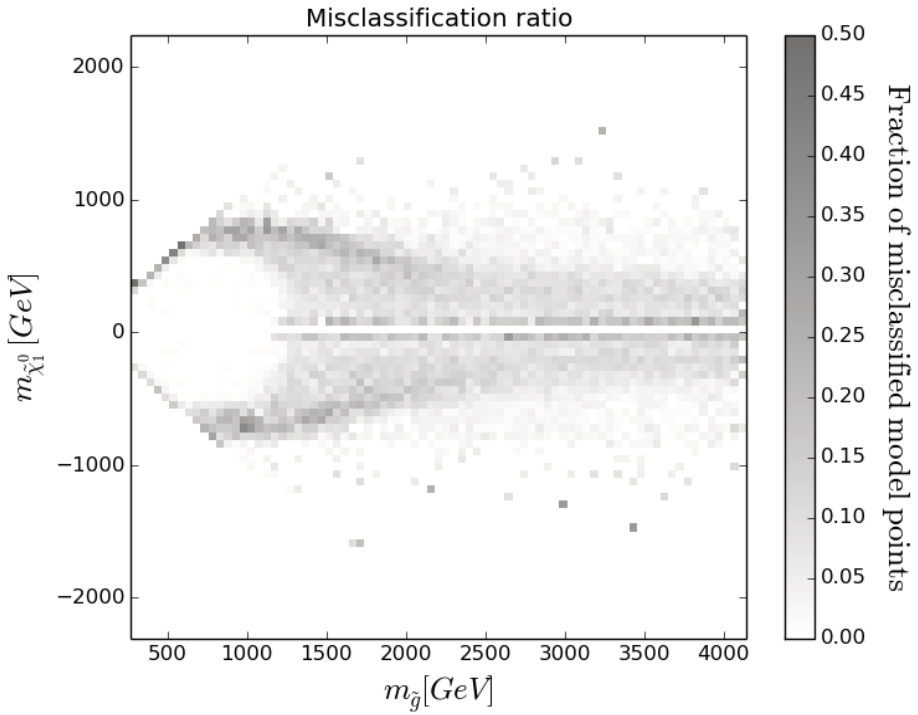
- Create model point
- Simulate events (Monte Carlo)
- Simulate detector
- Simulate detector response
- Compare detector response to actual measurements
- Check for confidence on simulation

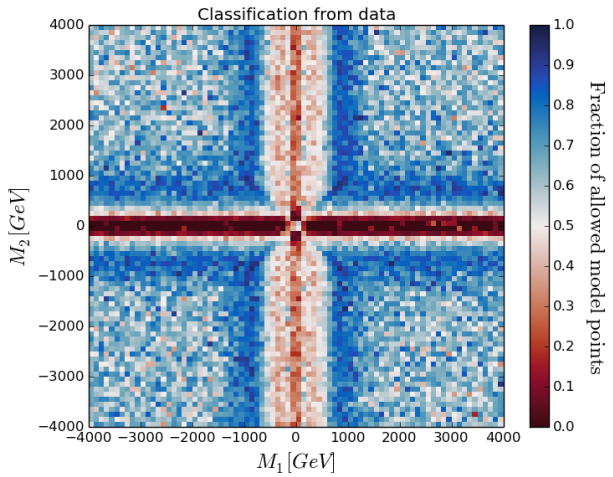


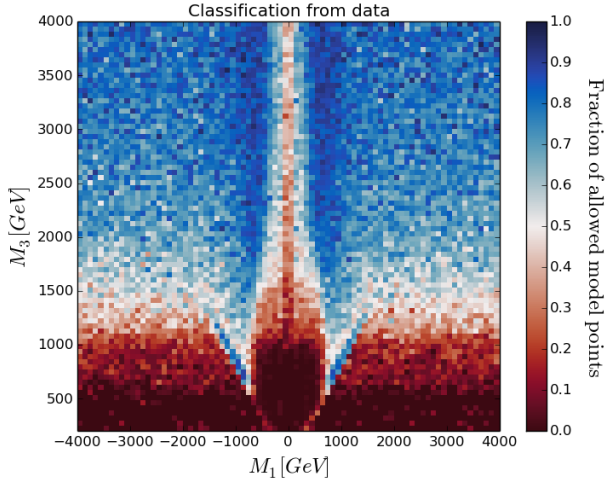

Classification from data
2000
1000
0
−1000
−2000
500
1000
1500
2000
2500
3000
3500
4000
$m_{\tilde{\chi}_1^0}[GeV]$
$m_{\tilde{g}}[GeV]$
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Fraction of allowed model points

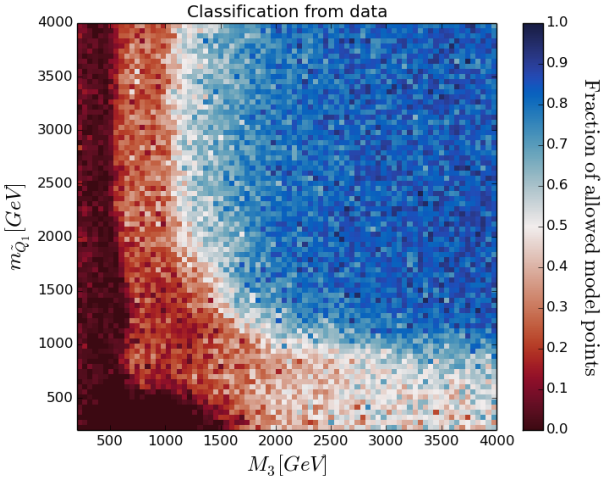

Predicted classification
2000
1000
0
−1000
−2000
500
1000
1500
2000
2500
3000
3500
4000
$m_{\tilde{\chi}_1^0}[GeV]$
$m_{\tilde{g}}[GeV]$
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Fraction of allowed model points

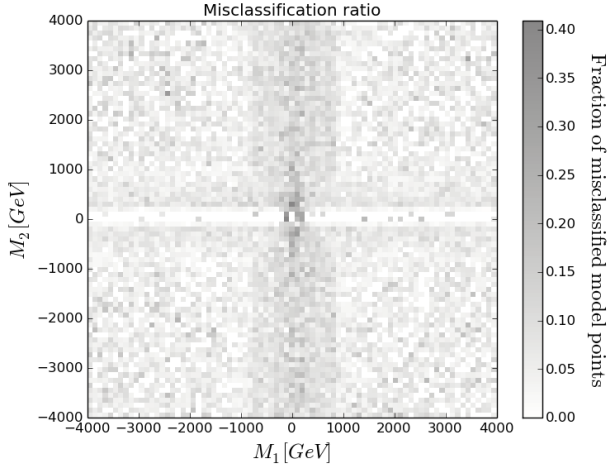

Classification from data
$m_{\tilde{\chi}_1^0}[GeV]$
$m_{\tilde{g}}[GeV]$
Fraction of allowed model points

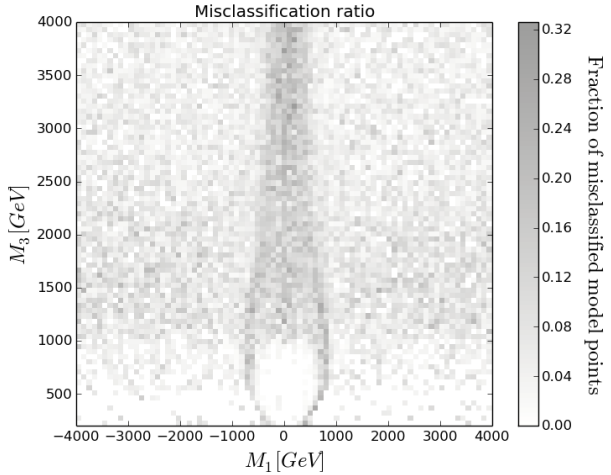

Misclassification ratio
$m_{\tilde{\chi}_1^0}[GeV]$
$m_{\tilde{g}}[GeV]$
Fraction of misclassified model points

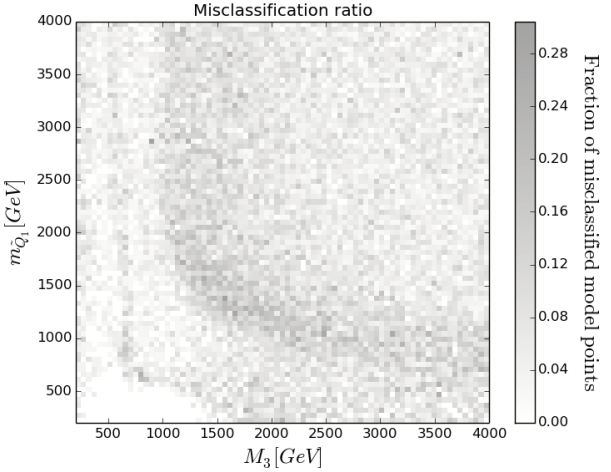

Classification from data
$M_2\,[GeV]$
$M_1\,[GeV]$
Fraction of allowed model points

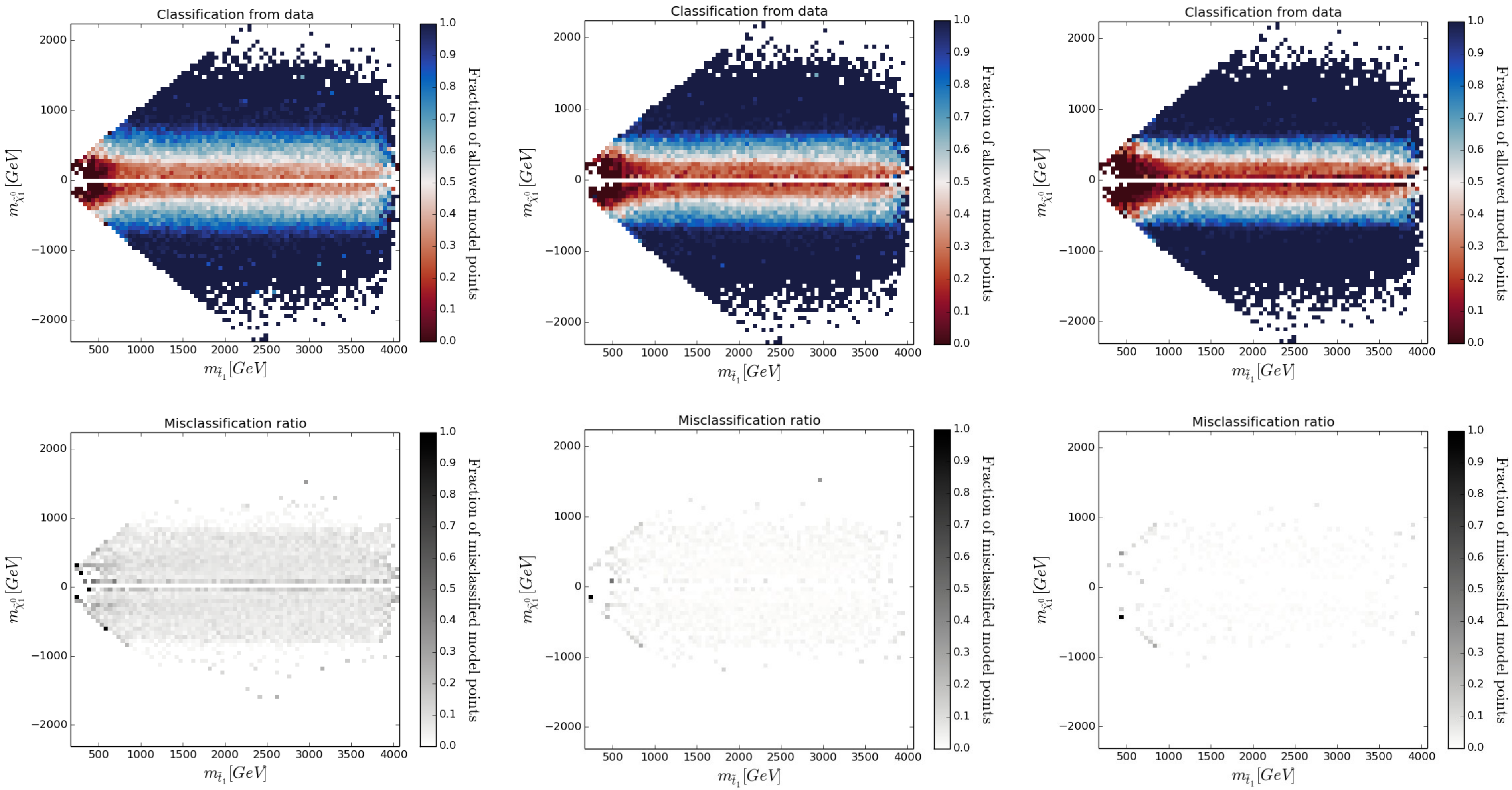

Classification from data
$M_3\,[GeV]$
$M_1\,[GeV]$
Fraction of allowed model points


Classification from data
$m_{\tilde{Q}_1}\,[GeV]$
$M_3\,[GeV]$
Fraction of allowed model points


Misclassification ratio
$M_2\,[GeV]$
$M_1\,[GeV]$
Fraction of misclassified model points

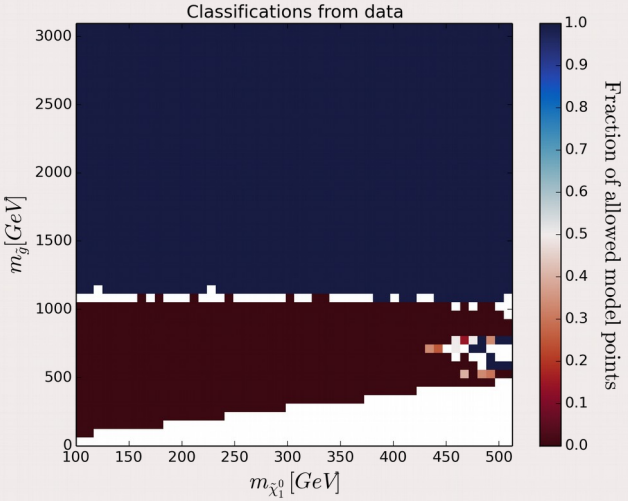

Misclassification ratio
$M_3\,[GeV]$
$M_1\,[GeV]$
Fraction of misclassified model points

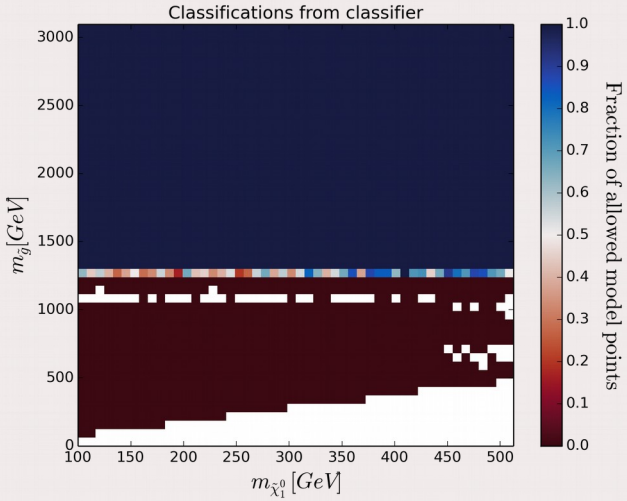

Misclassification ratio
$m_{\tilde{Q}_1}\,[GeV]$
$M_3\,[GeV]$
Fraction of misclassified model points

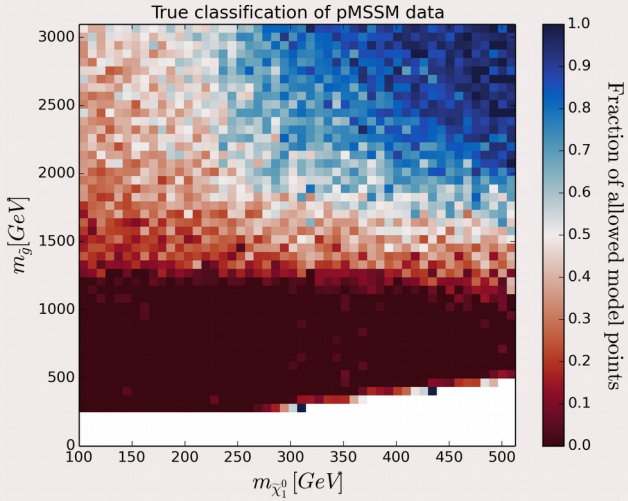


Classification from data
Classification from data
Classification from data
Fraction of allowed model points
Misclassification ratio
Misclassification ratio
Misclassification ratio
Fraction of misclassified model points
$m_{\tilde{t}_1}$ [GeV]
$m_{\tilde{\chi}_1^0}$ [GeV]

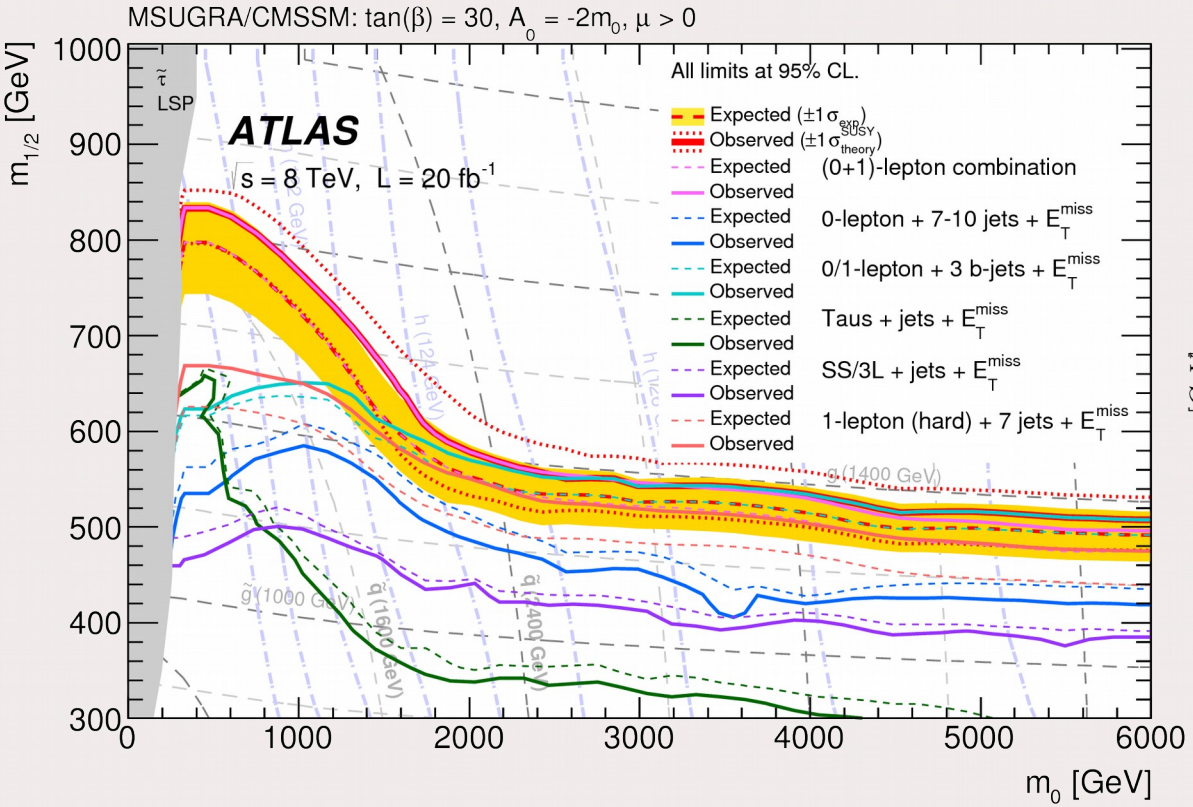
Natural SUSY

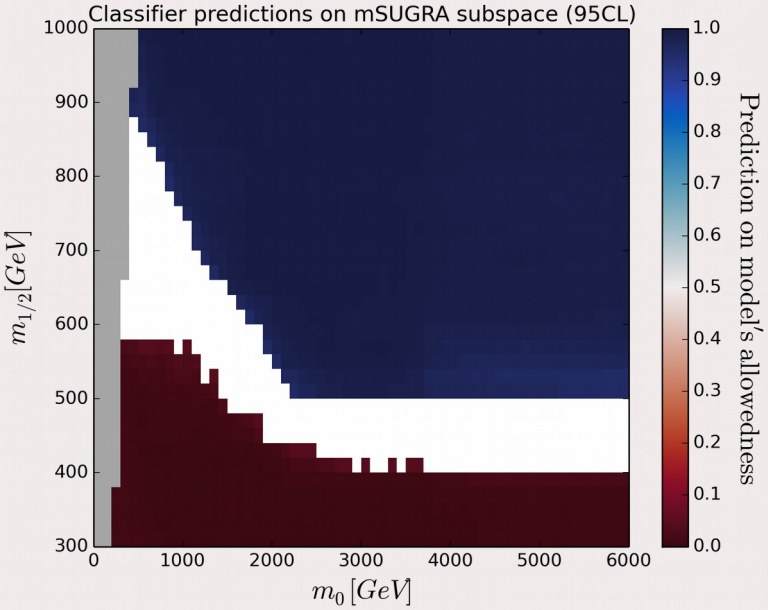






cMSSM


MSUGRA/CMSSM: tan(β) = 30, A0 = -2m0, μ > 0
All limits at 95% CL.
ATLAS
√s = 8 TeV, L = 20 fb-1
LSP
Expected (±1σexp)
Observed (±1σSUSY theory)
Expected (0+1)-lepton combination
Observed
Expected 0-lepton + 7-10 jets + E_T^miss
Observed
Expected 0/1-lepton + 3 b-jets + E_T^miss
Observed
Expected Taus + jets + E_T^miss
Observed
Expected SS/3L + jets + E_T^miss
Observed
Expected 1-lepton (hard) + 7 jets + E_T^miss
Observed
g̃ (1400 GeV)
g̃ (1000 GeV)
q̃ (1400 GeV)
q̃ (1600 GeV)
m1/2 [GeV]
m0 [GeV]
300
400
500
600
700
800
900
1000
0
1000
2000
3000
4000
5000
6000


Classifier predictions on mSUGRA subspace (95CL)
m1/2[GeV]
m0 [GeV]
Prediction on model's allowedness
300
400
500
600
700
800
900
1000
0
1000
2000
3000
4000
5000
6000
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

How to use SUSY-AI / Machine Learning

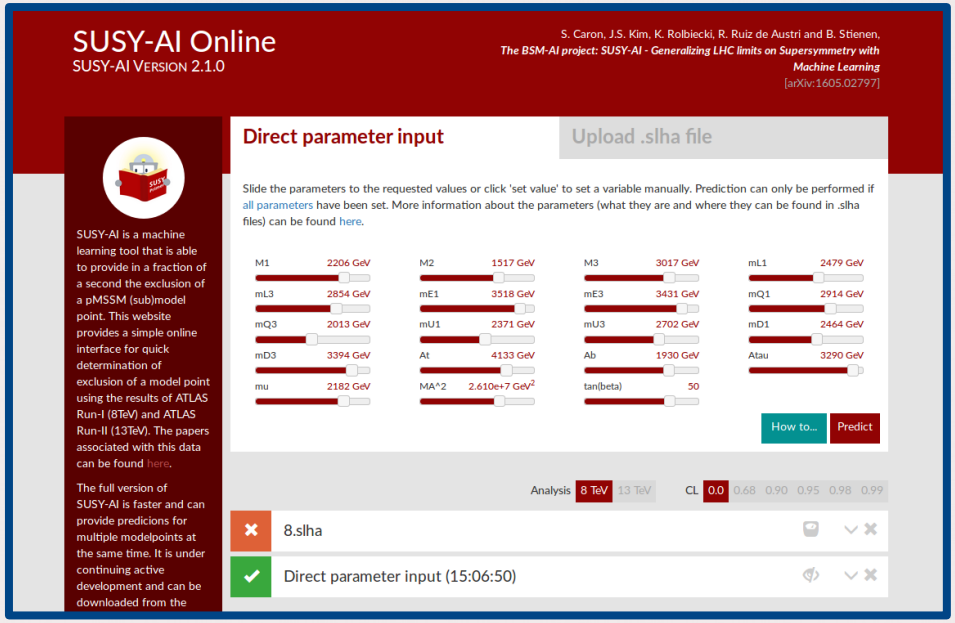
- Publishing of multivariate data
- Generating projection plots not present in the accompanying paper
- Fast check on ATLAS analysis

Benchmark test (10.000 model points)

	Time (s)	*Duration (s)*
Reading files	0	97.177
Loading pickle	97.177	1.856
Mapping coordinates	99.033	0.006
Predicting	**99.039**	**1.048**
Calculating results	100.087	0.073
Apply min CL 0.95	100.160	0.019
Total run time	*100.179*	

SUSY-AI (Online)

- Tool has been published https://susyai.hepforge.org/
  - Python interface to classifier
  - Sklearn package for ML implementation
- Online interface http://susy-ai.org/
  - All functionalities except batch predictions
  - Predictions in < 2 seconds



Summary / Conclusions

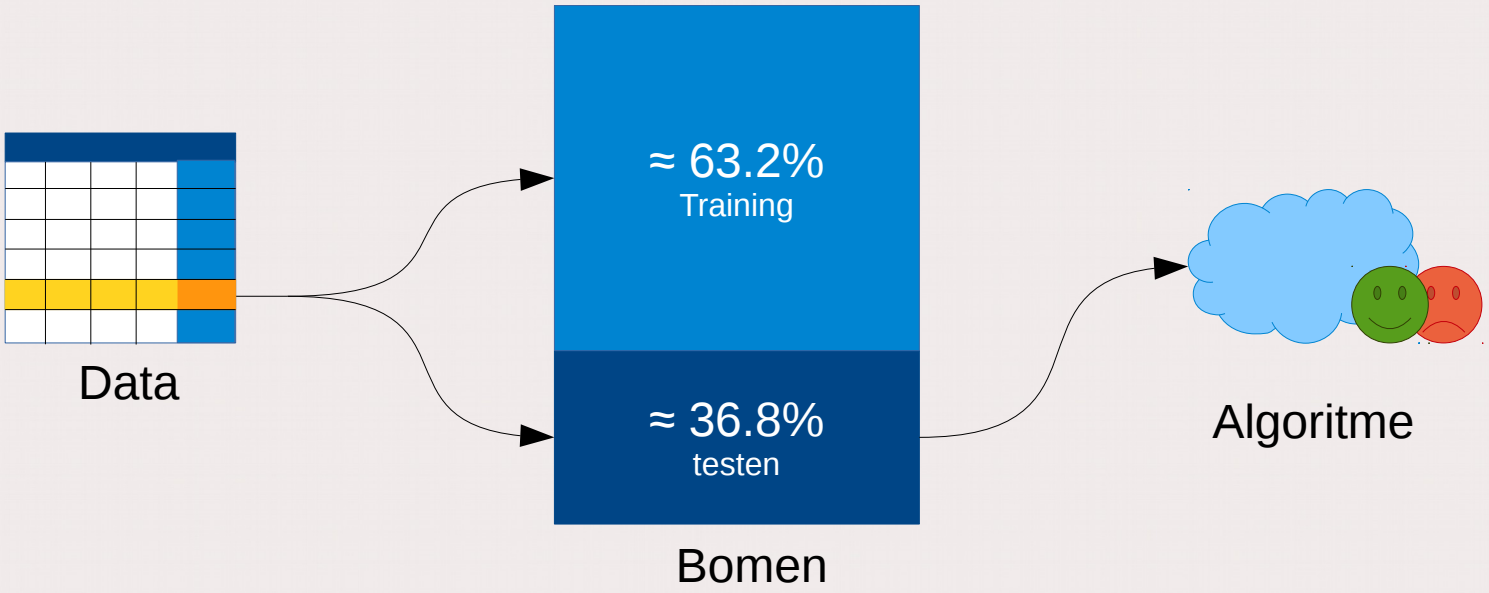
- ATLAS exclusion limits are predicted with 93.2% accuracy
- Applicable to pMSSM and submodels
- The use of machine learning is just being discovered

- Programmatic interface (SUSY-AI: https://susyai.hepforge.org/)
- Online interface (http://susy-ai.org)

- More public data needed to broaden this use of Machine Learning

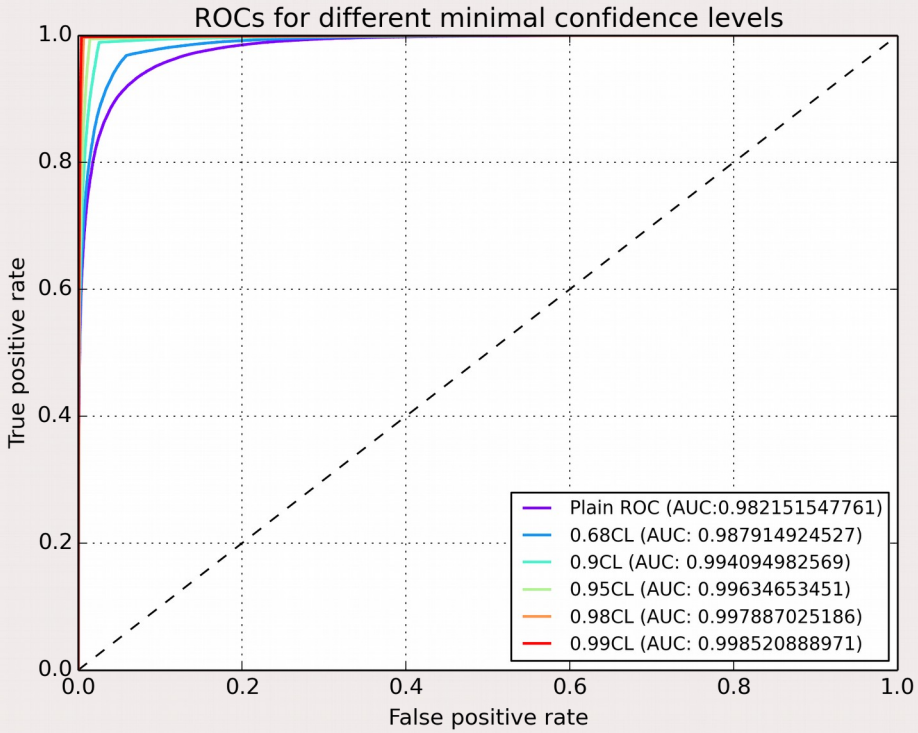
Thank you for your attention!

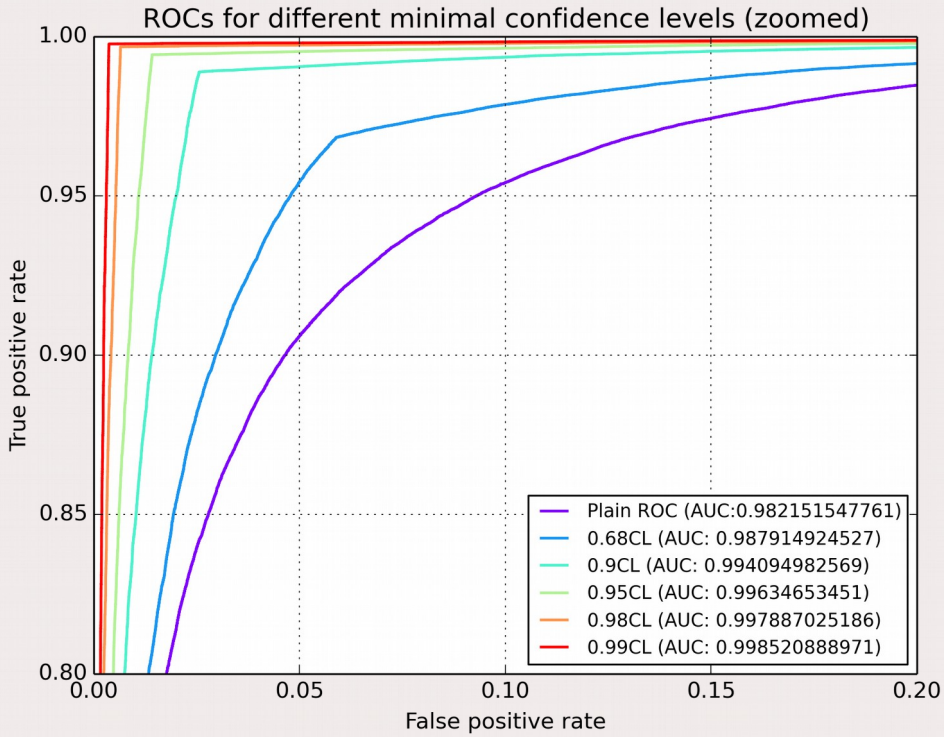
Out-of-bag schatting



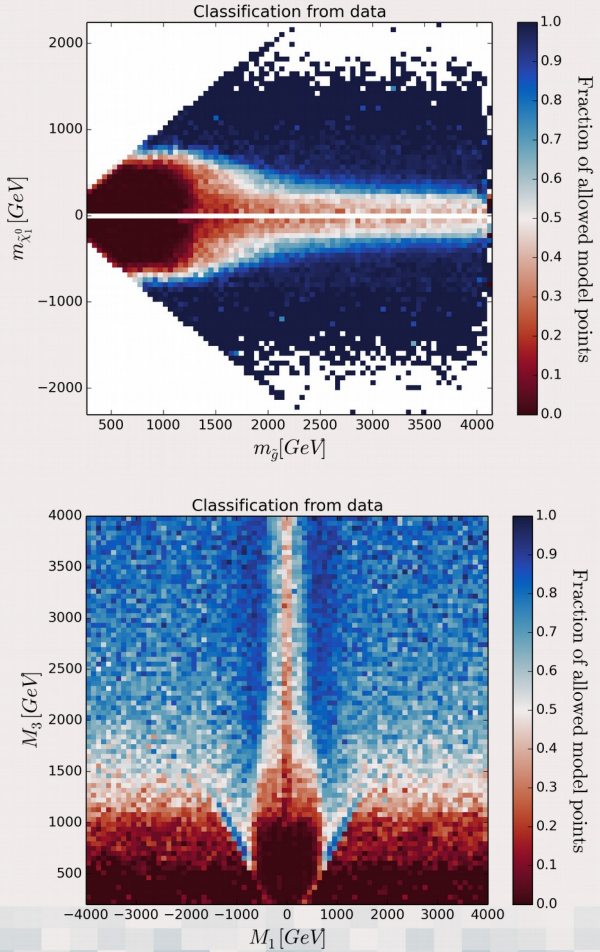
- ≈ 36.8% van de bomen hebben niet op attributeset i getraind
- Vraagt om een groot aantal bomen
- Geen testdataset nodig!

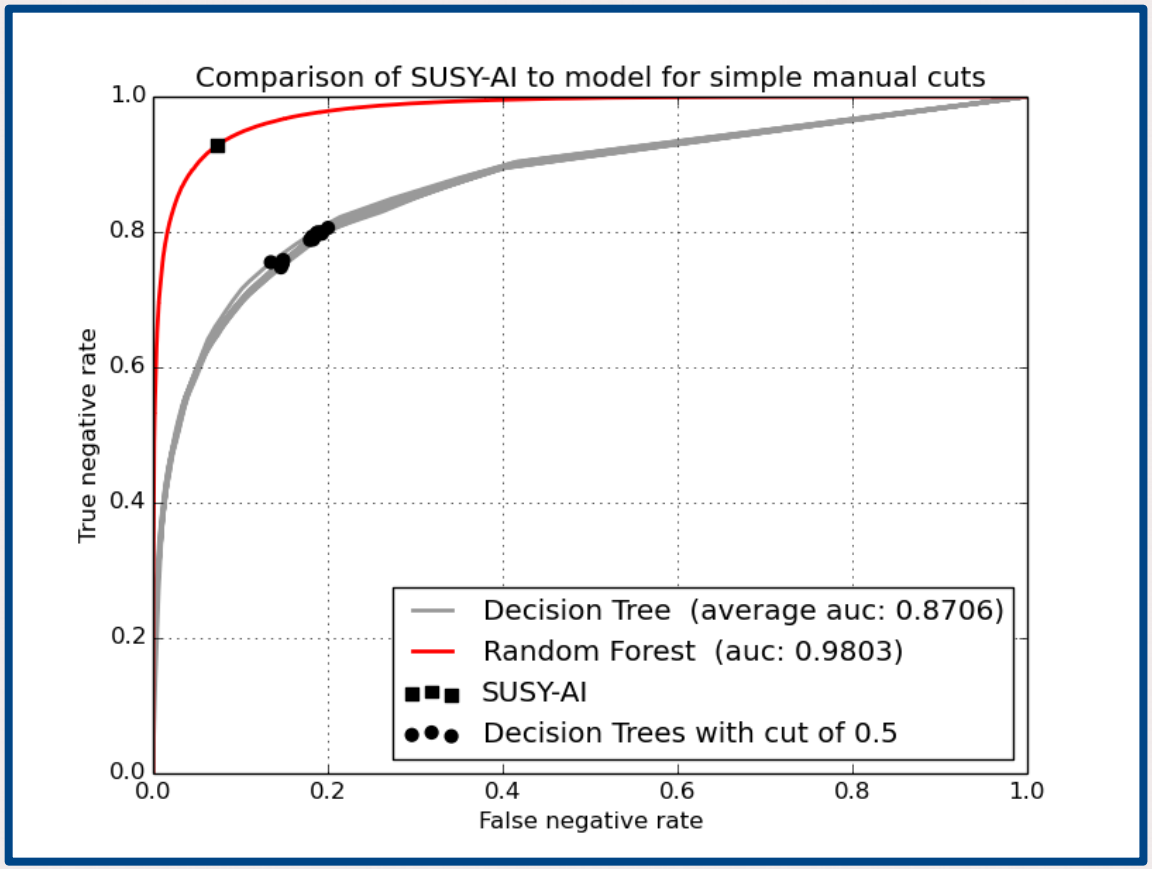
ROC curve


ROCs for different minimal confidence levels
True positive rate
False positive rate
0.0
0.2
0.4
0.6
0.8
1.0
Plain ROC (AUC:0.982151547761)
0.68CL (AUC: 0.987914924527)
0.9CL (AUC: 0.994094982569)
0.95CL (AUC: 0.99634653451)
0.98CL (AUC: 0.997887025186)
0.99CL (AUC: 0.998520888971)


ROCs for different minimal confidence levels (zoomed)
True positive rate
False positive rate
0.80
0.85
0.90
0.95
1.00
0.00
0.05
0.10
0.15
0.20
Plain ROC (AUC:0.982151547761)
0.68CL (AUC: 0.987914924527)
0.9CL (AUC: 0.994094982569)
0.95CL (AUC: 0.99634653451)
0.98CL (AUC: 0.997887025186)
0.99CL (AUC: 0.998520888971)

Comparison to user





Performance maten

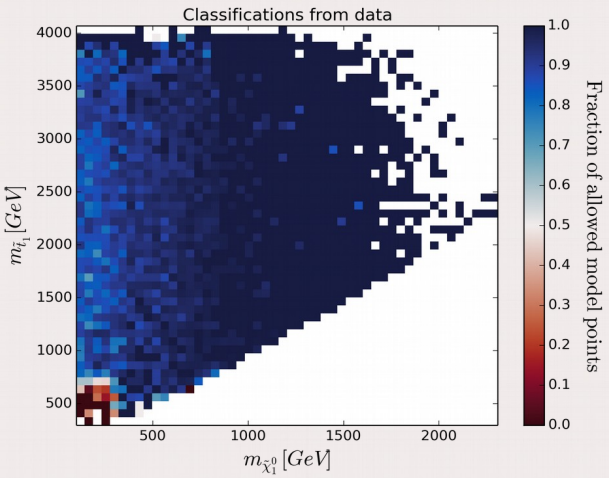
$$\frac{TP}{TP + FP}$$

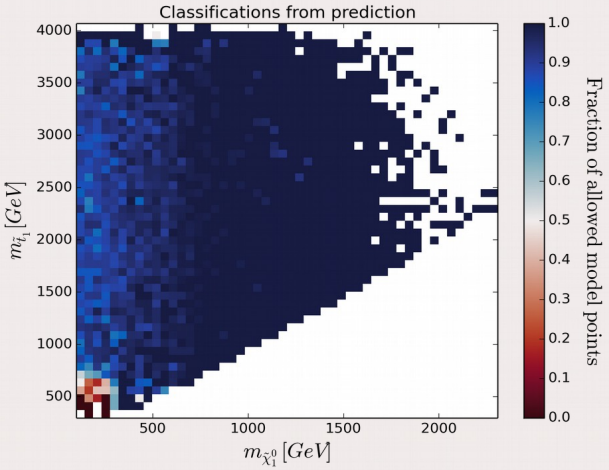
$$\frac{1}{N}\sum_{i=1}^{N}(p_i - t_i)^2$$

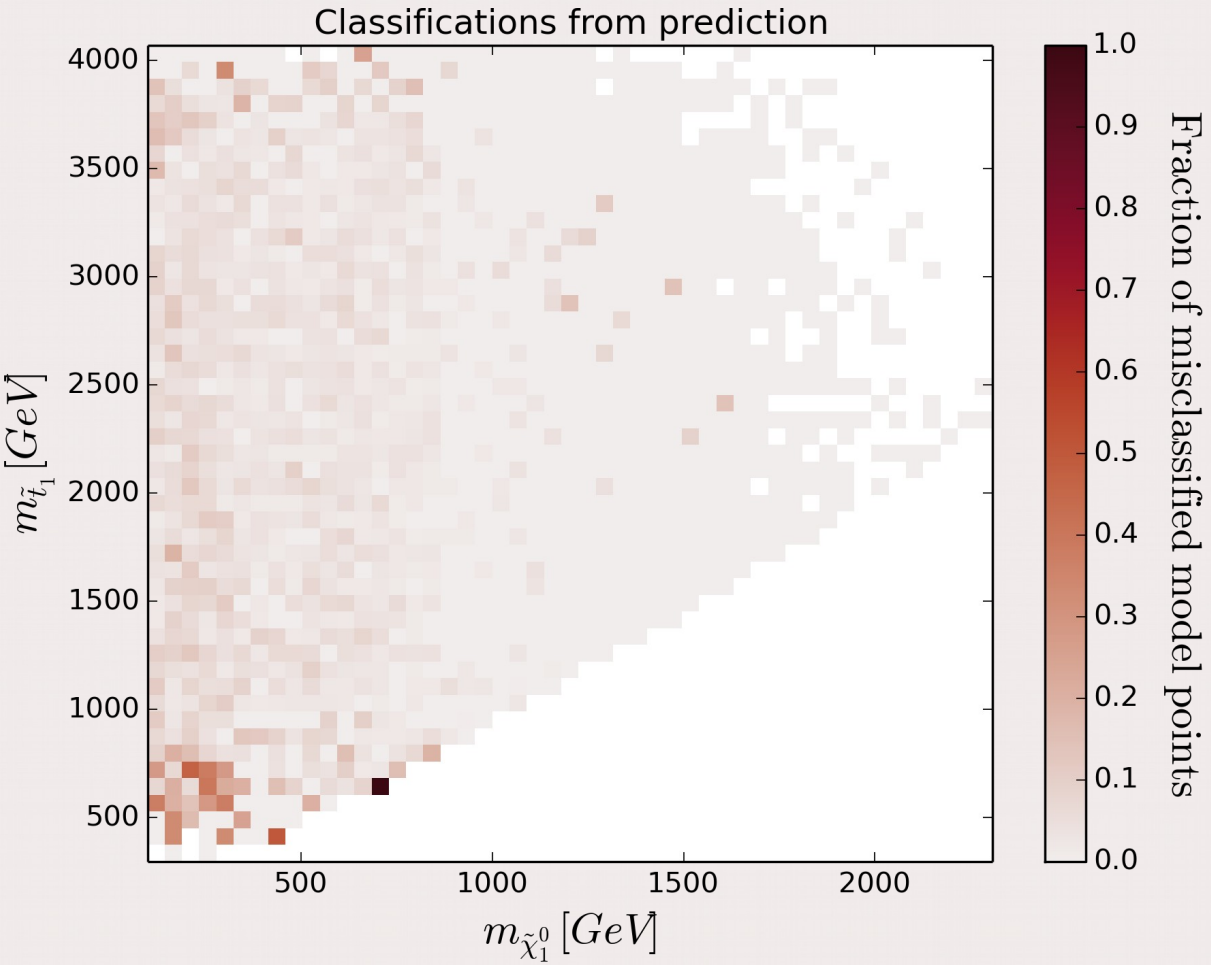
CL	#	Fraction	Accuracy	Precision	Sensitivity	Brier	AUC
0.0	310324	1.0	0.93226	0.93951	0.94665	0.04951	0.98209
0.68	289371	0.93248	0.95735	0.96072	0.96835	0.03573	0.98783
0.95	219233	0.70646	0.99094	0.99092	0.99426	0.00885	0.99618
0.98	184230	0.59367	0.99543	0.99573	0.99672	0.00452	0.99767
0.99	160034	0.5157	0.99708	0.99747	0.99764	0.00291	0.99825

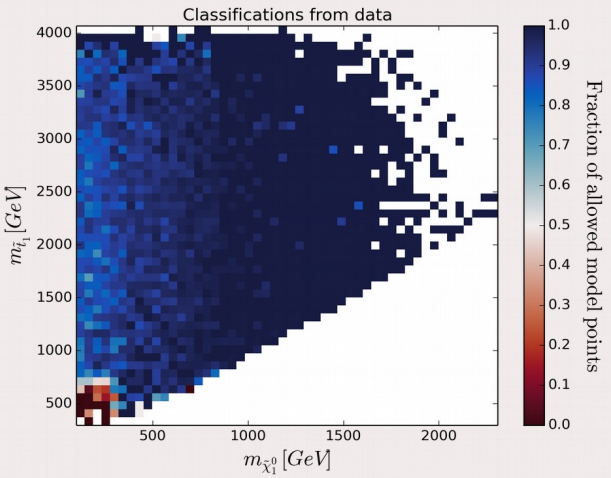
$$\frac{TP + TN}{TP + FP + FN + TN}$$

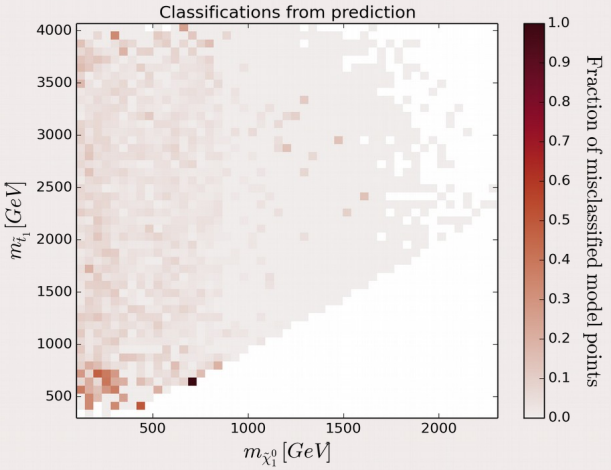
$$\frac{TP}{TP + FN}$$

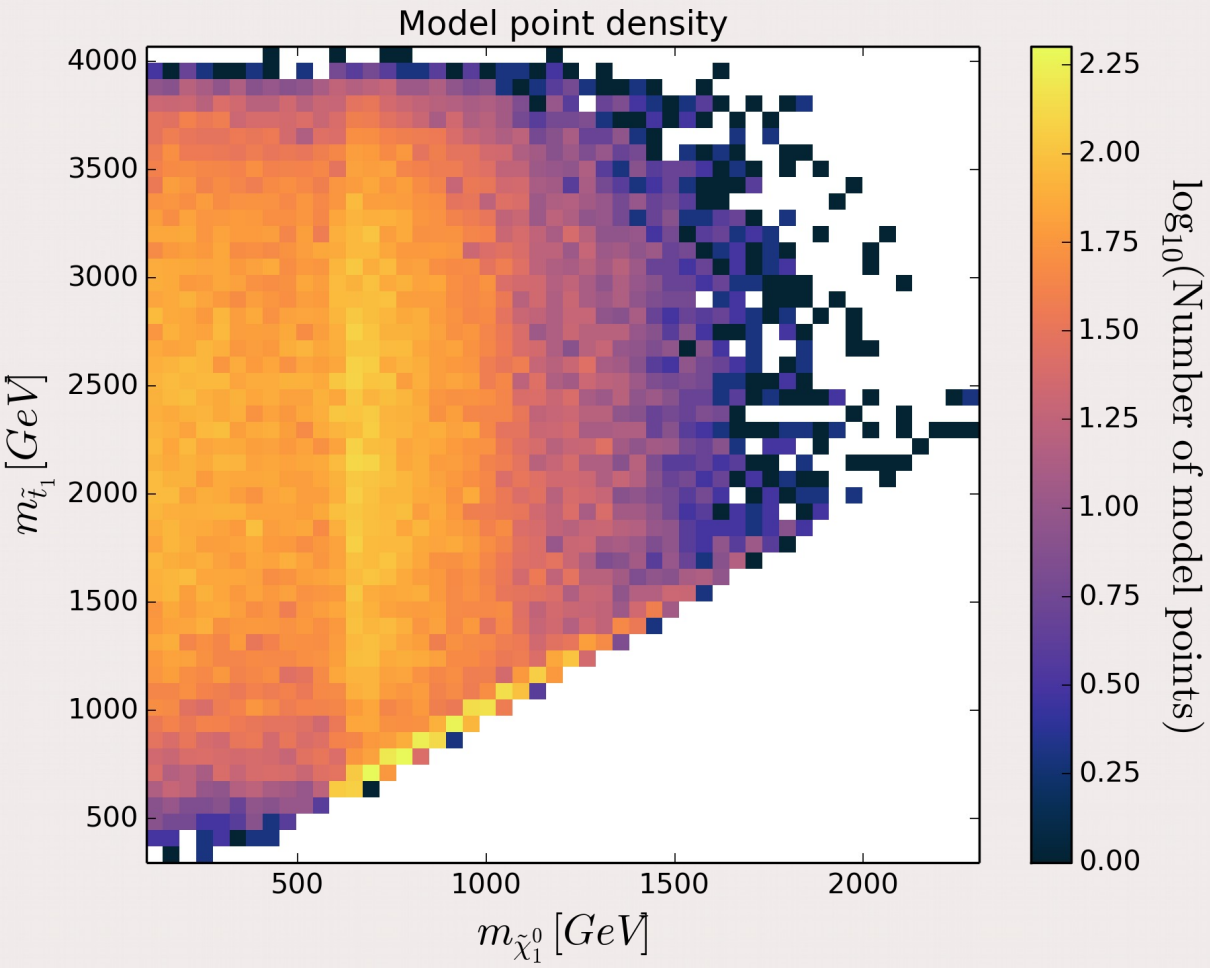

Classifications from data
$m_{\tilde{t}_1}$ [GeV]
$m_{\tilde{\chi}_1^0}$ [GeV]
Fraction of allowed model points

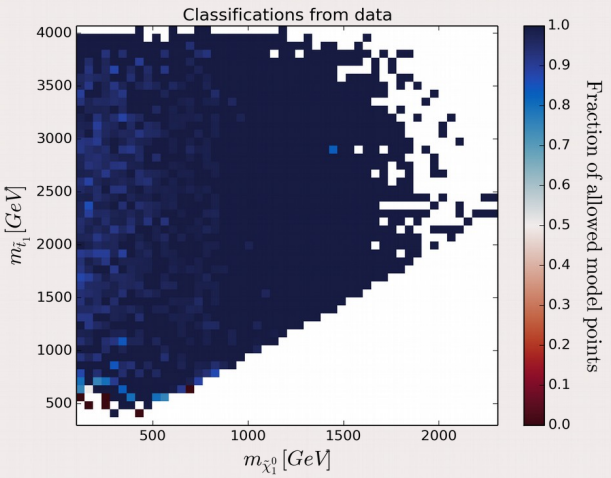

Classifications from prediction
$m_{\tilde{t}_1}$ [GeV]
$m_{\tilde{\chi}_1^0}$ [GeV]
Fraction of allowed model points

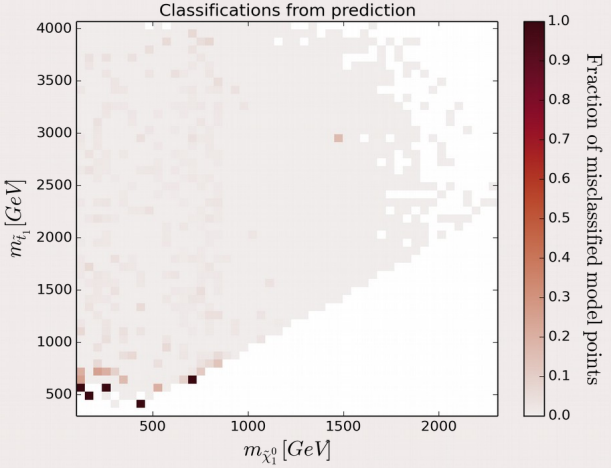

Classifications from prediction
$m_{\tilde{t}_1}$ [GeV]
$m_{\tilde{\chi}_1^0}$ [GeV]
Fraction of misclassified model points

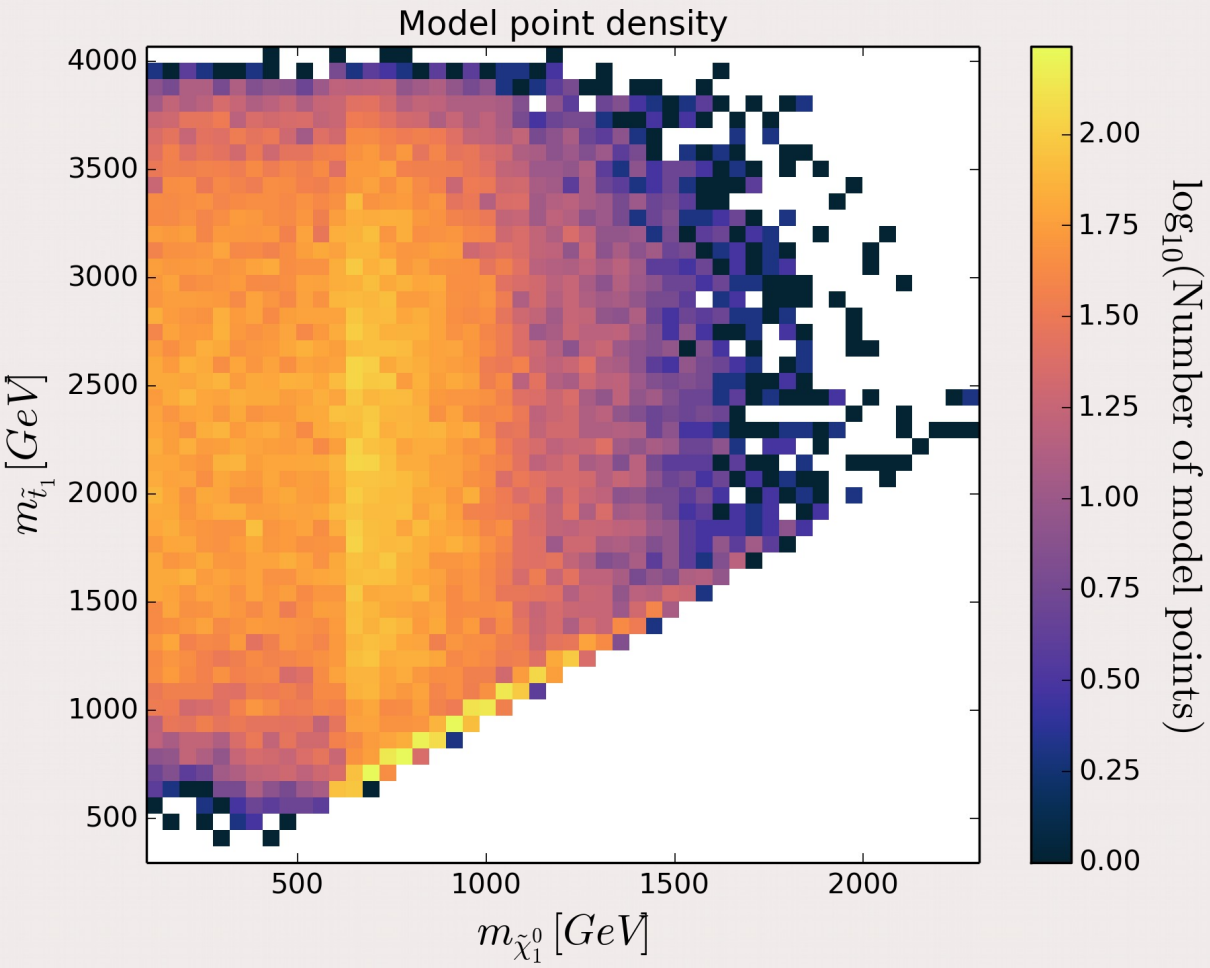

Classifications from data
$m_{\tilde{t}_1}\,[GeV]$
$m_{\tilde{\chi}_1^0}\,[GeV]$
Fraction of allowed model points


Classifications from prediction
$m_{\tilde{t}_1}\,[GeV]$
$m_{\tilde{\chi}_1^0}\,[GeV]$
Fraction of misclassified model points


Model point density
$m_{\tilde{t}_1}\,[GeV]$
$m_{\tilde{\chi}_1^0}\,[GeV]$
$\log_{10}$(Number of model points)


Classifications from data
$m_{\tilde{t}_1}\,[GeV]$
$m_{\tilde{\chi}_1^0}\,[GeV]$
Fraction of allowed model points


Classifications from prediction
$m_{\tilde{t}_1}\,[GeV]$
$m_{\tilde{\chi}_1^0}\,[GeV]$
Fraction of misclassified model points


Model point density
$m_{\tilde{t}_1}\,[GeV]$
$m_{\tilde{\chi}_1^0}\,[GeV]$
$\log_{10}$(Number of model points)