

# Generalising LHC Limits on Supersymmetry with Machine Learning

Bob Stienen  
[b.stienen@science.ru.nl](mailto:b.stienen@science.ru.nl)



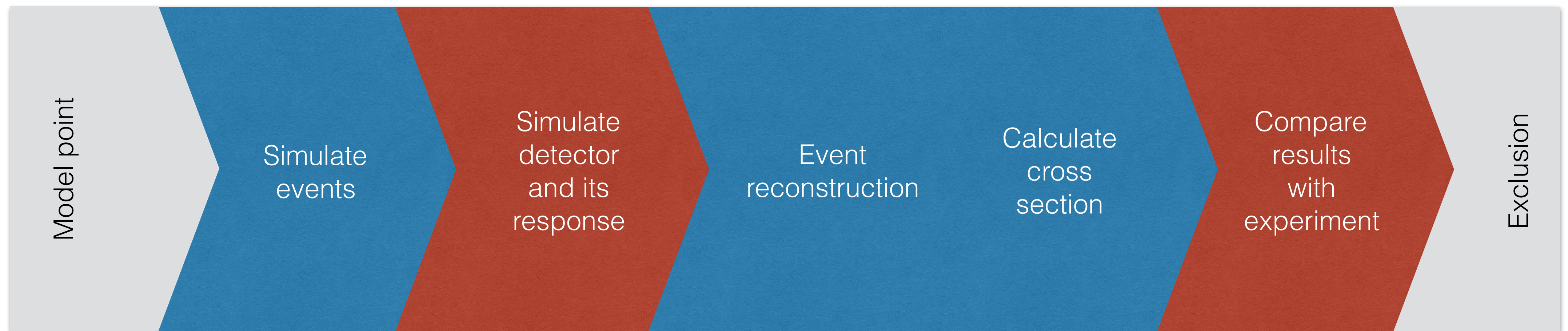
# Contents

- Exclusion Determination
- SUSY-AI Current status
- „ Future

# Exclusion Determination

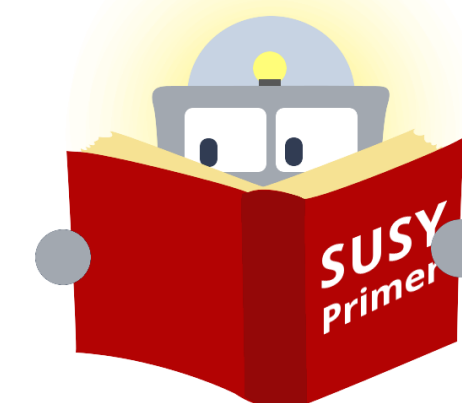
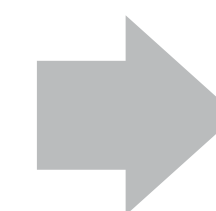
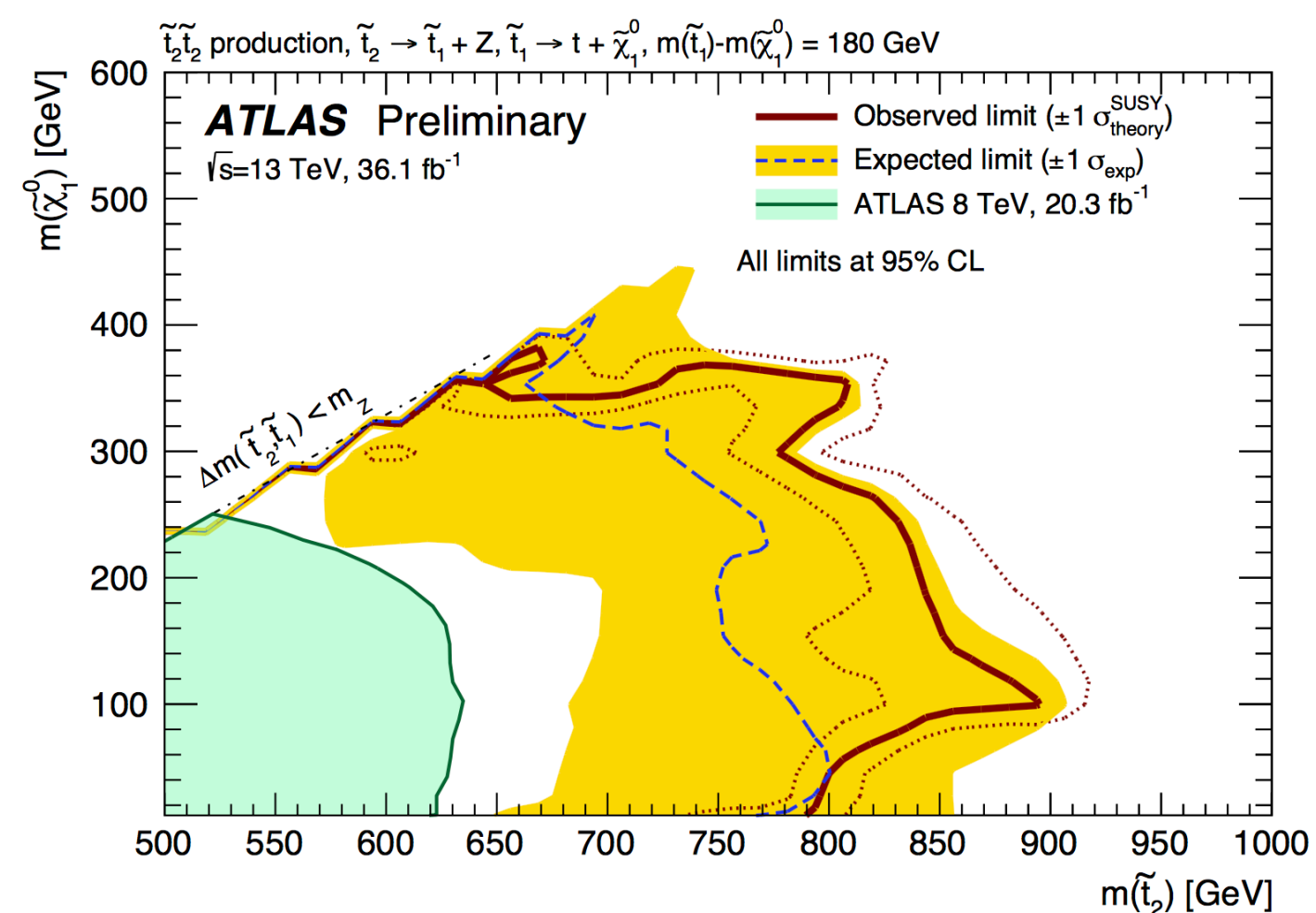
- In search for BSM physics, no signs were found
- Second best thing: setting limits on models
- Exclusion of parameter space takes a long time

Time =  $O(\text{hours})$



# SUSY-AI

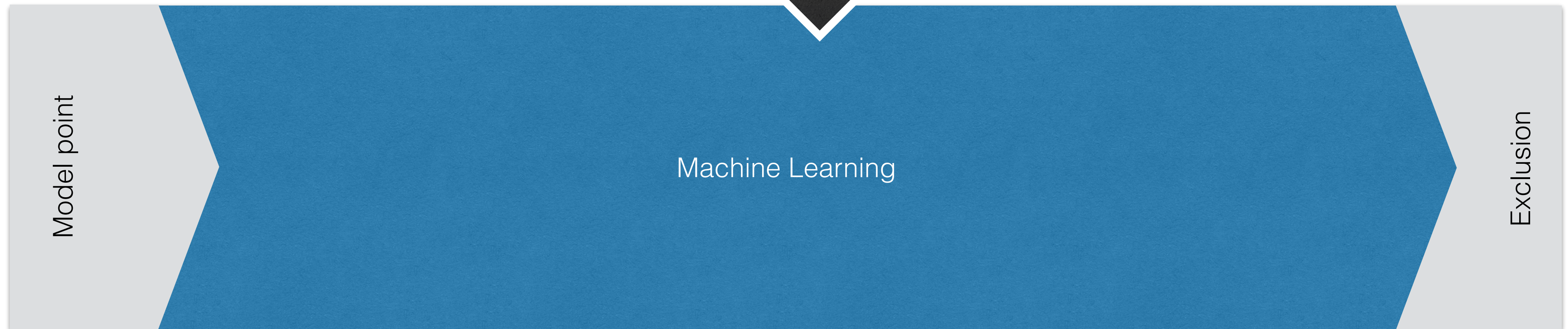
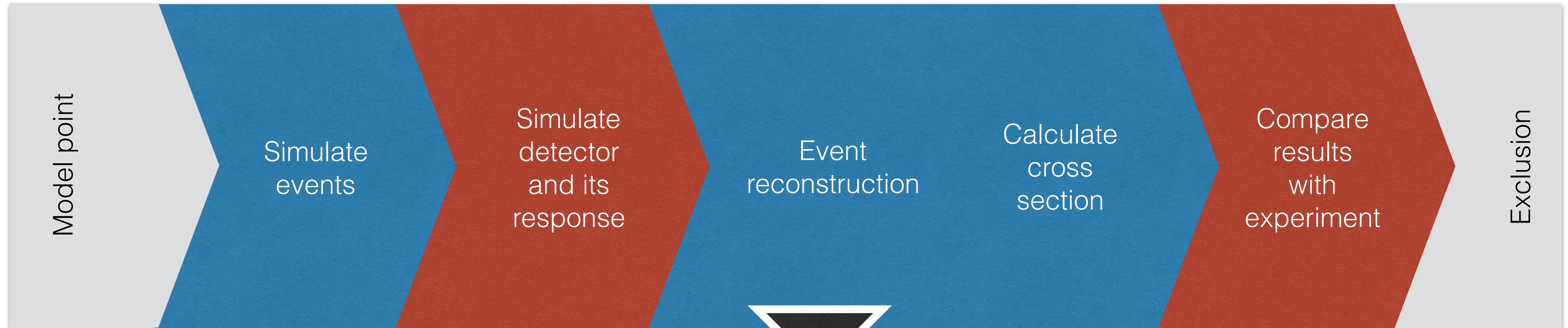
- Exclusion determination in 19d pMSSM
- 310,324 model points with known exclusion as data input
- Algorithm: Random Forest in scikit-learn
- **Idea: going from 2d slices to N-dim representations**





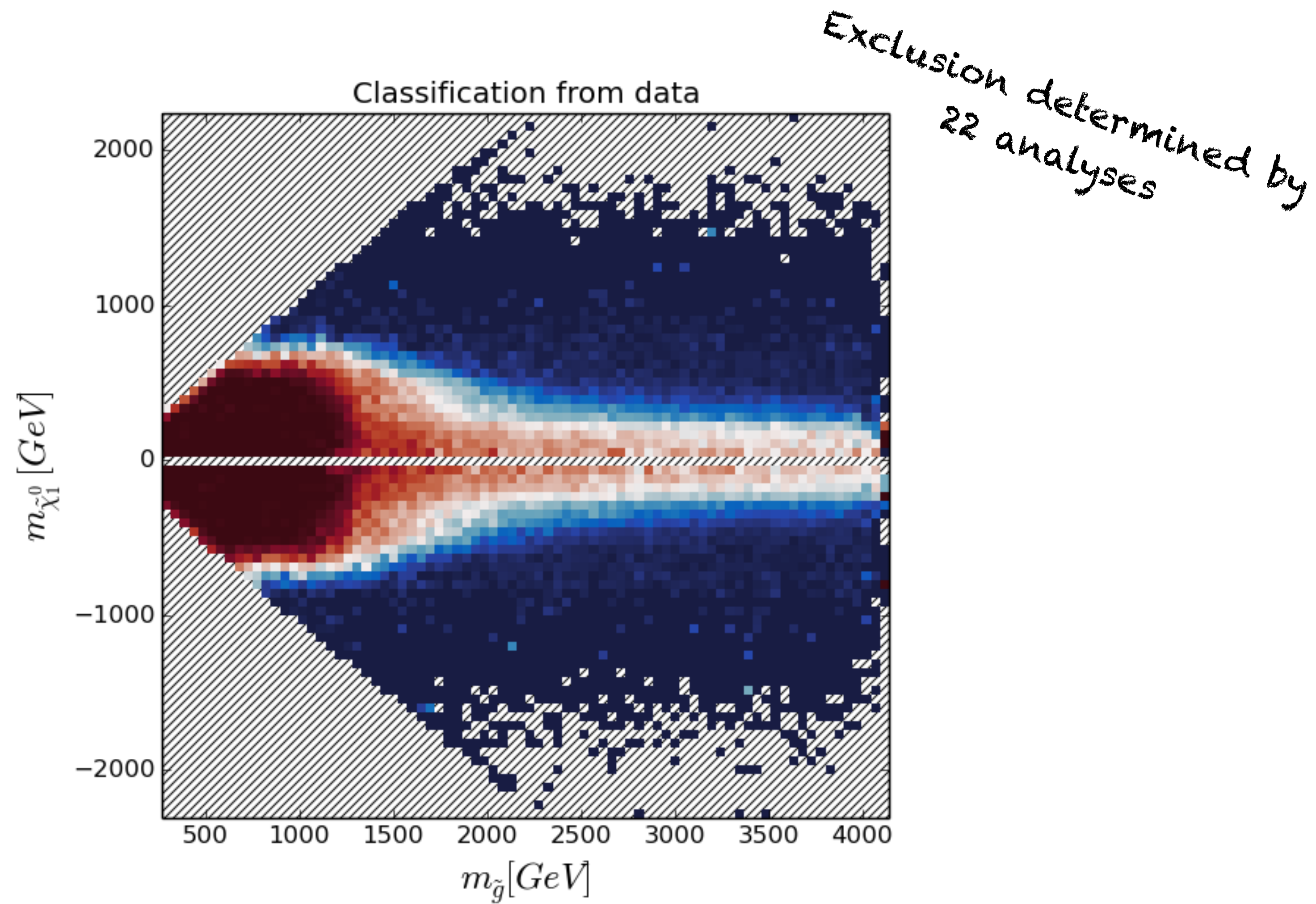
# Exclusion Determination

Time = O(hours)



Time = O(ms)

# SUSY-AI in the pMSSM



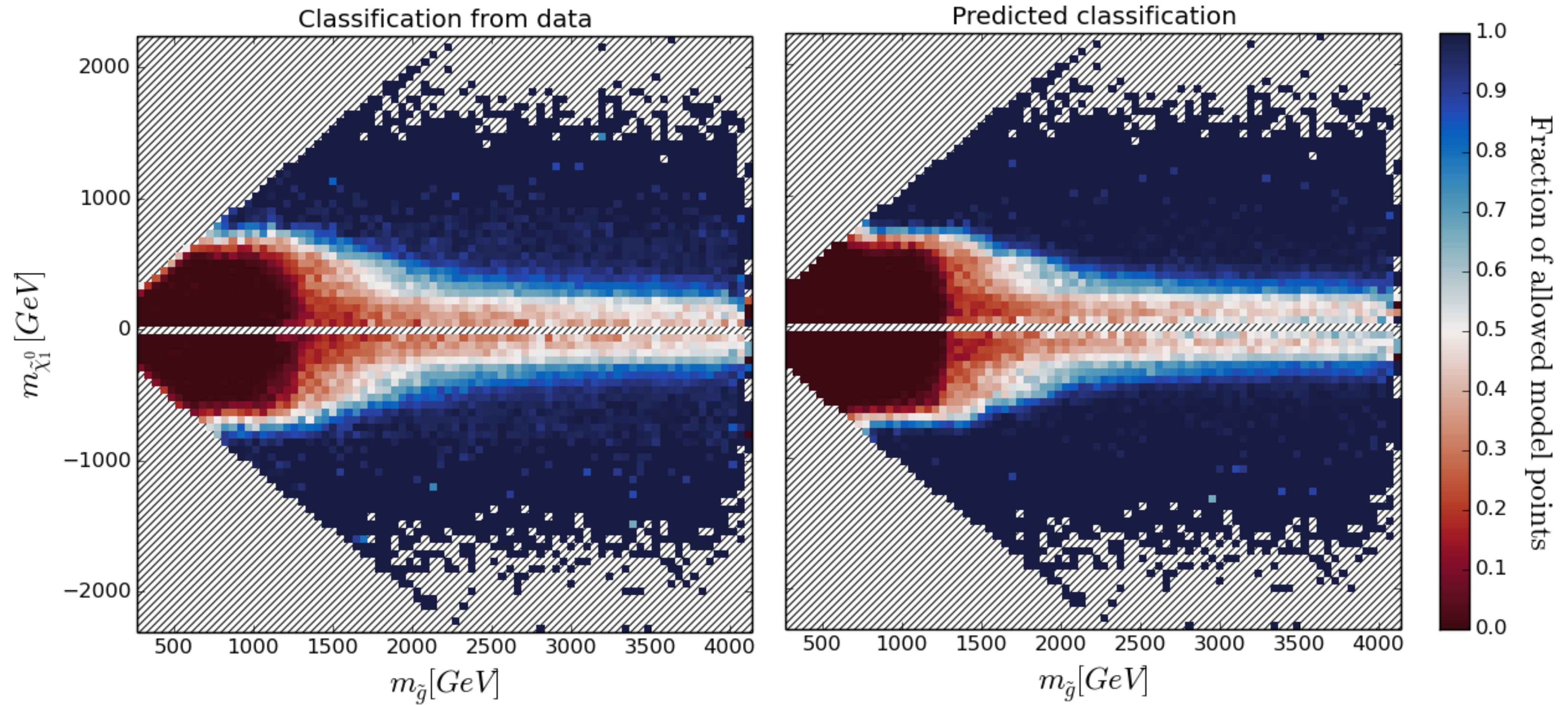


# SUSY-AI in the pMSSM

93.2% accuracy @ 8TeV

92.7% accuracy @ 13 TeV

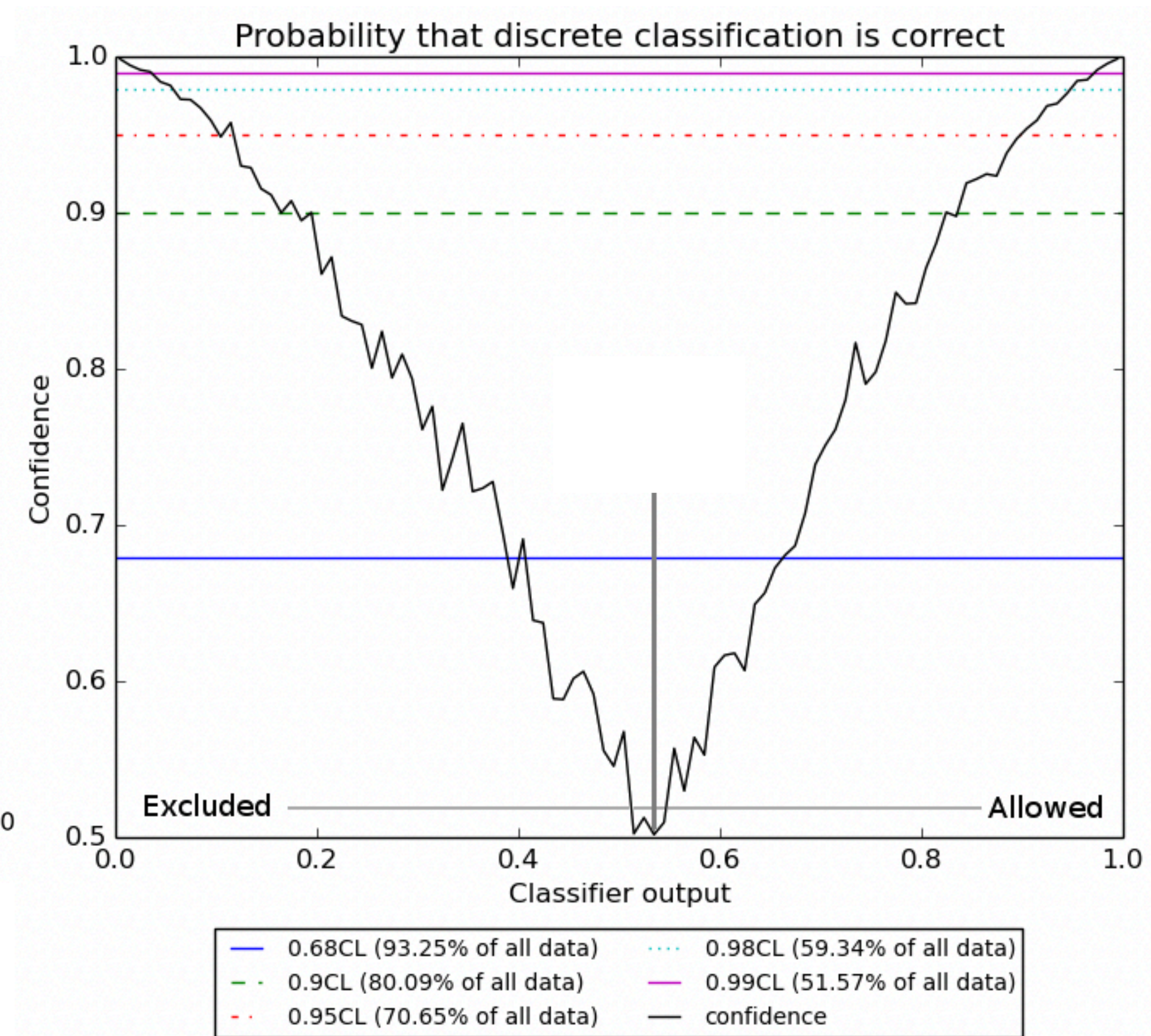
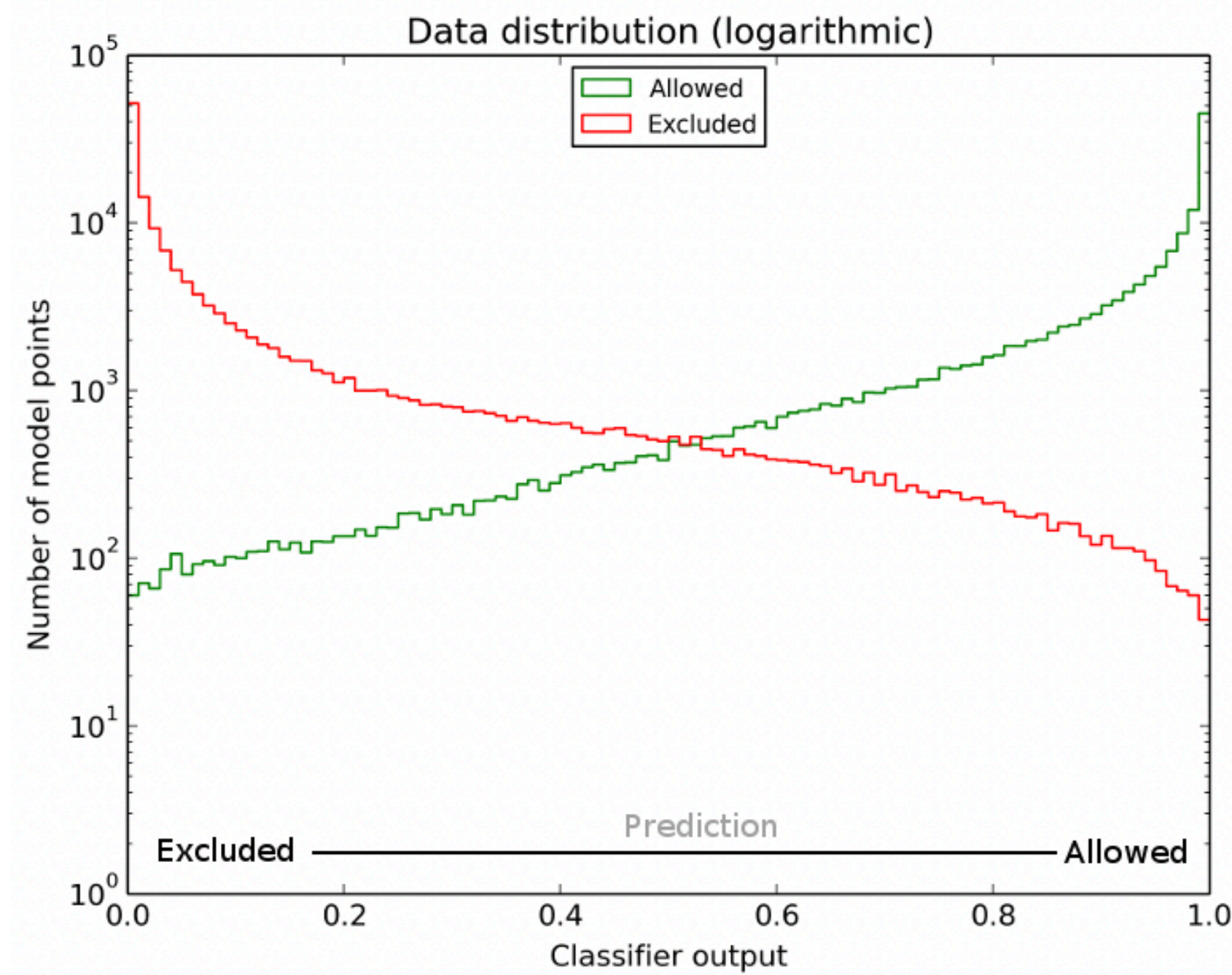
<http://www.susy-ai.org>  
for online demo!





# Confidence measure

Ratio of majority class per bin



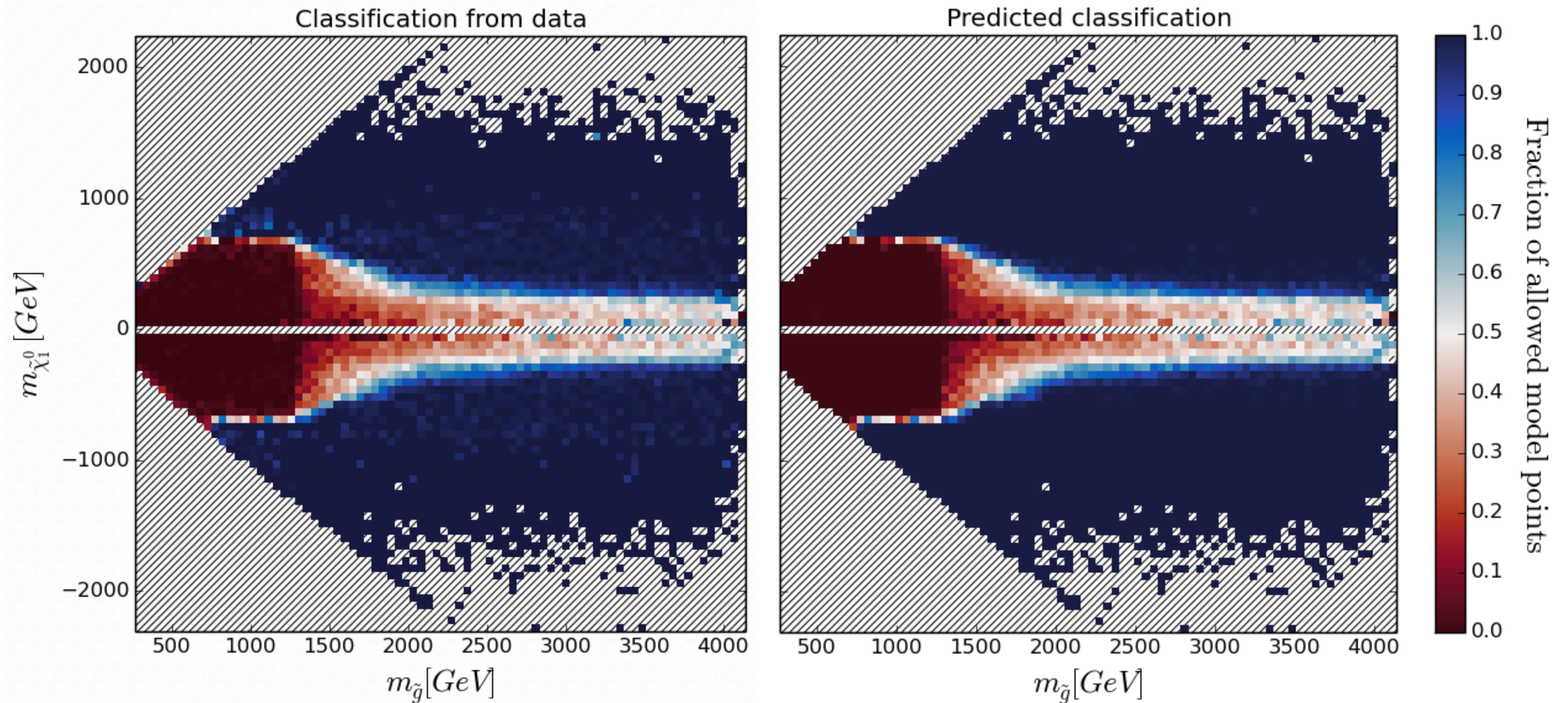


# SUSY-AI in the pMSSM (95CL)

99.1% accuracy on 70.6% of total data @ 8TeV

99.0% accuracy on 68.0% of total data @ 13 TeV

<http://www.susy-ai.org>  
for online demo!



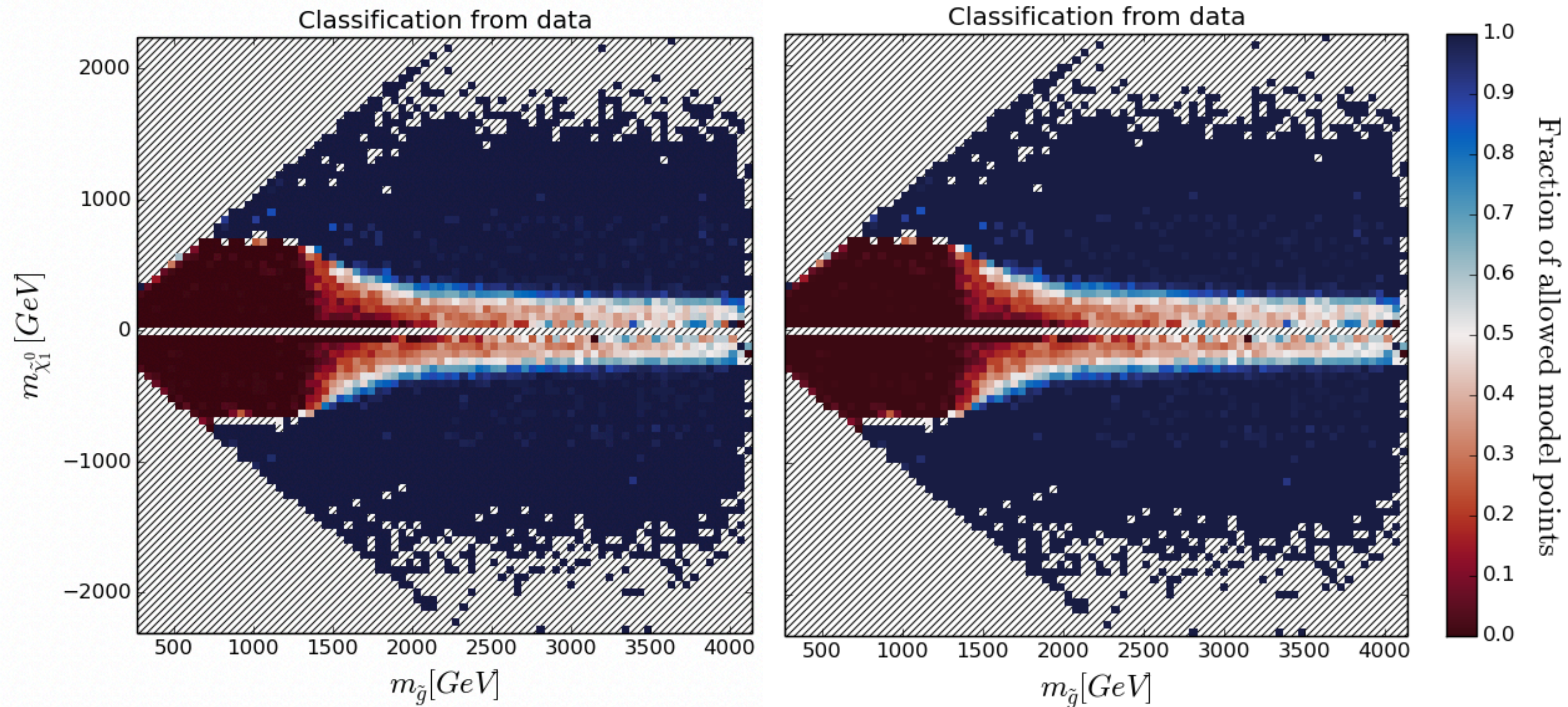


# SUSY-AI in the pMSSM (99CL)

99.7% accuracy on 51.6% of total data @ 8TeV

99.7% accuracy on 47.6% of total data @ 13 TeV

<http://www.susy-ai.org>  
for online demo!





# Metrics

CL	#	# / total	Accuracy	Precision	Sensitivity	NPV	Specificity
0.0	310 324	1.0000	0.93226	0.93951	0.94665	0.92152	0.91133
0.68	289 371	0.93248	0.95735	0.96072	0.96835	0.95222	0.94094
0.95	219 233	0.70646	0.99094	0.99092	0.99426	0.99096	0.98573
0.98	184 230	0.59367	0.99543	0.99573	0.99672	0.99496	0.99346
0.99	160 034	0.51570	0.99708	0.99747	0.99764	0.99649	0.99624

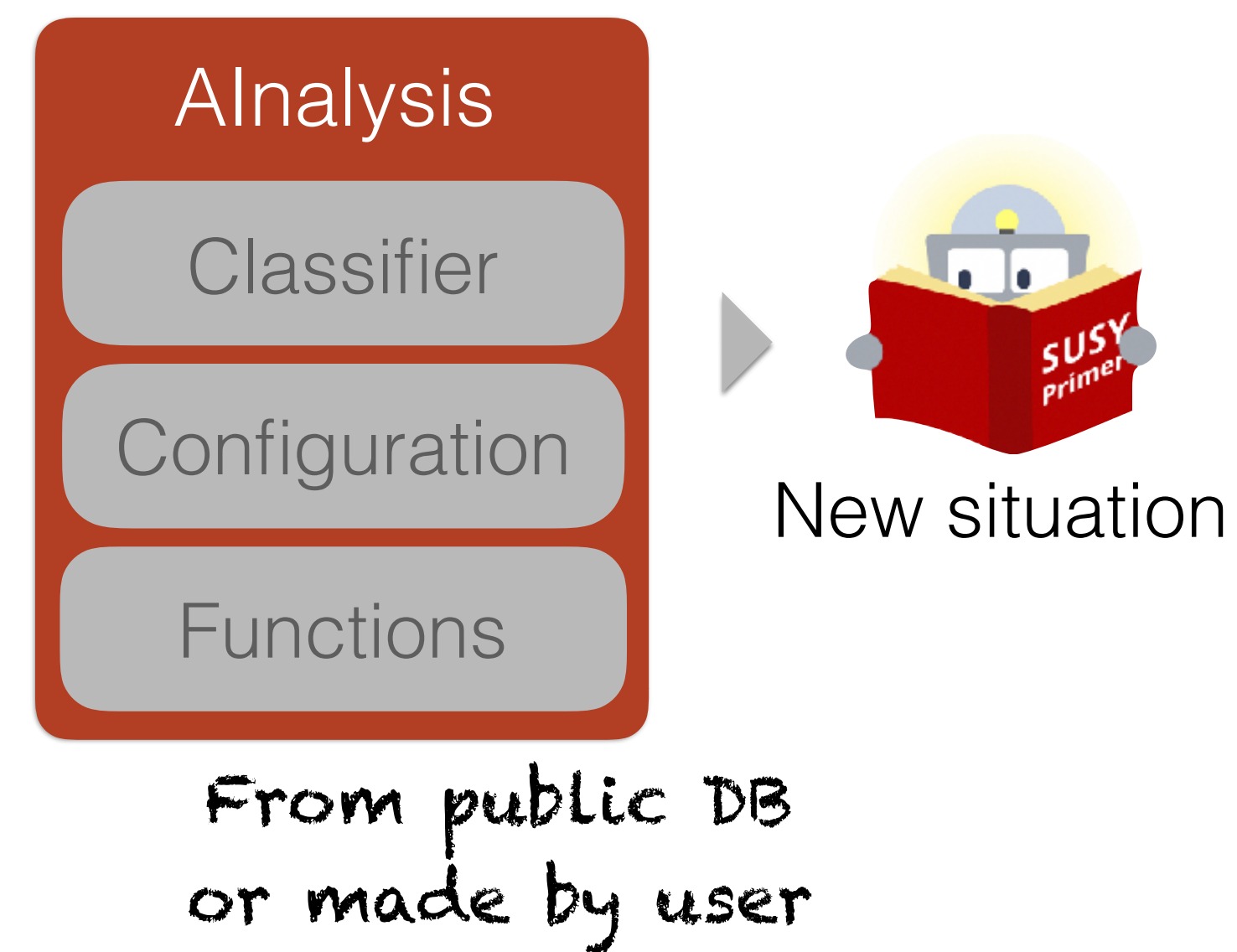
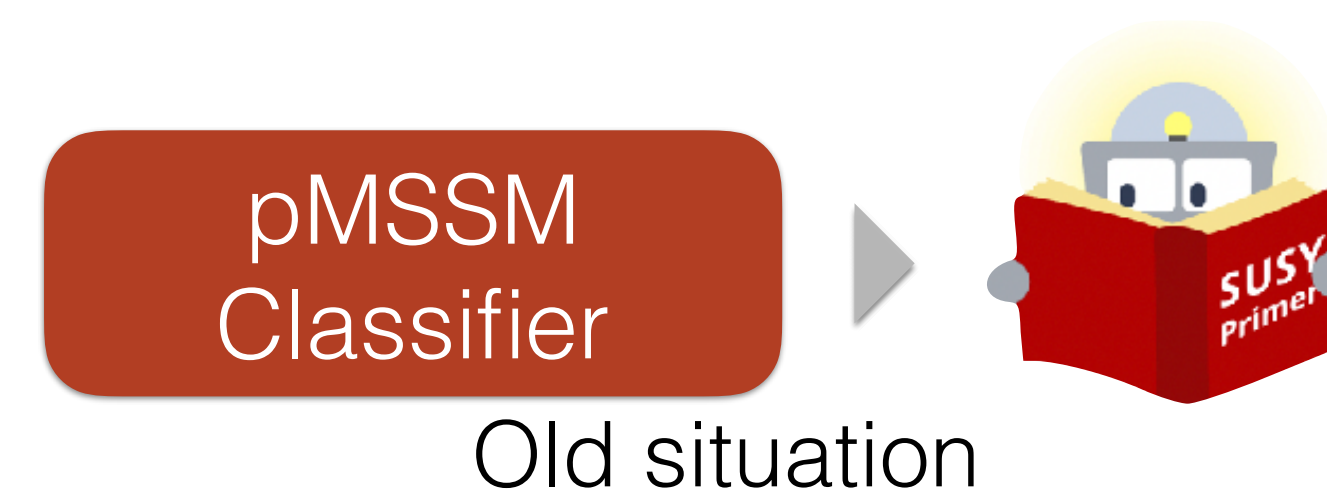
# SUSY-AI Future

- Generalisation
- Built-in server-client structure
- Multiple classifiers (stacking)

Next release  
in coming weeks

Analysis	1	2	3	4	5	6	7	Result
Excluded	No	No	Yes	Yes	No	No	No	Yes

- Boundary and parameter space exploration

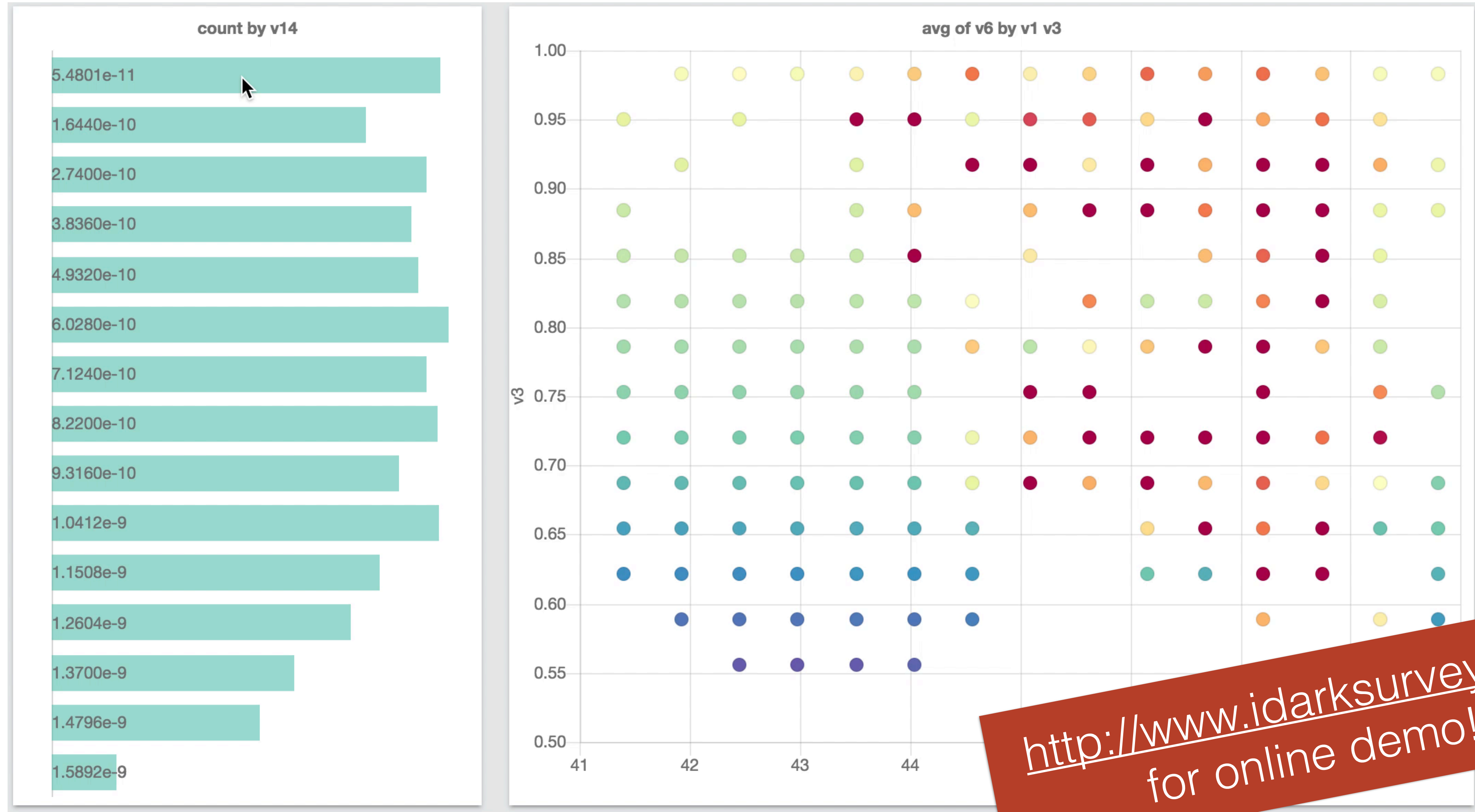




# Obtaining data is still a problem

- Phenomenology data is hard and time consuming to generate
- Much of data is made, but not easy to make public in accessible way
- iDark will host a public database and plotting interface: [idarksurvey.com](http://idarksurvey.com)
  - Collecting and visualising world's phenomenology data
  - Let (your) data live to its fullest potential

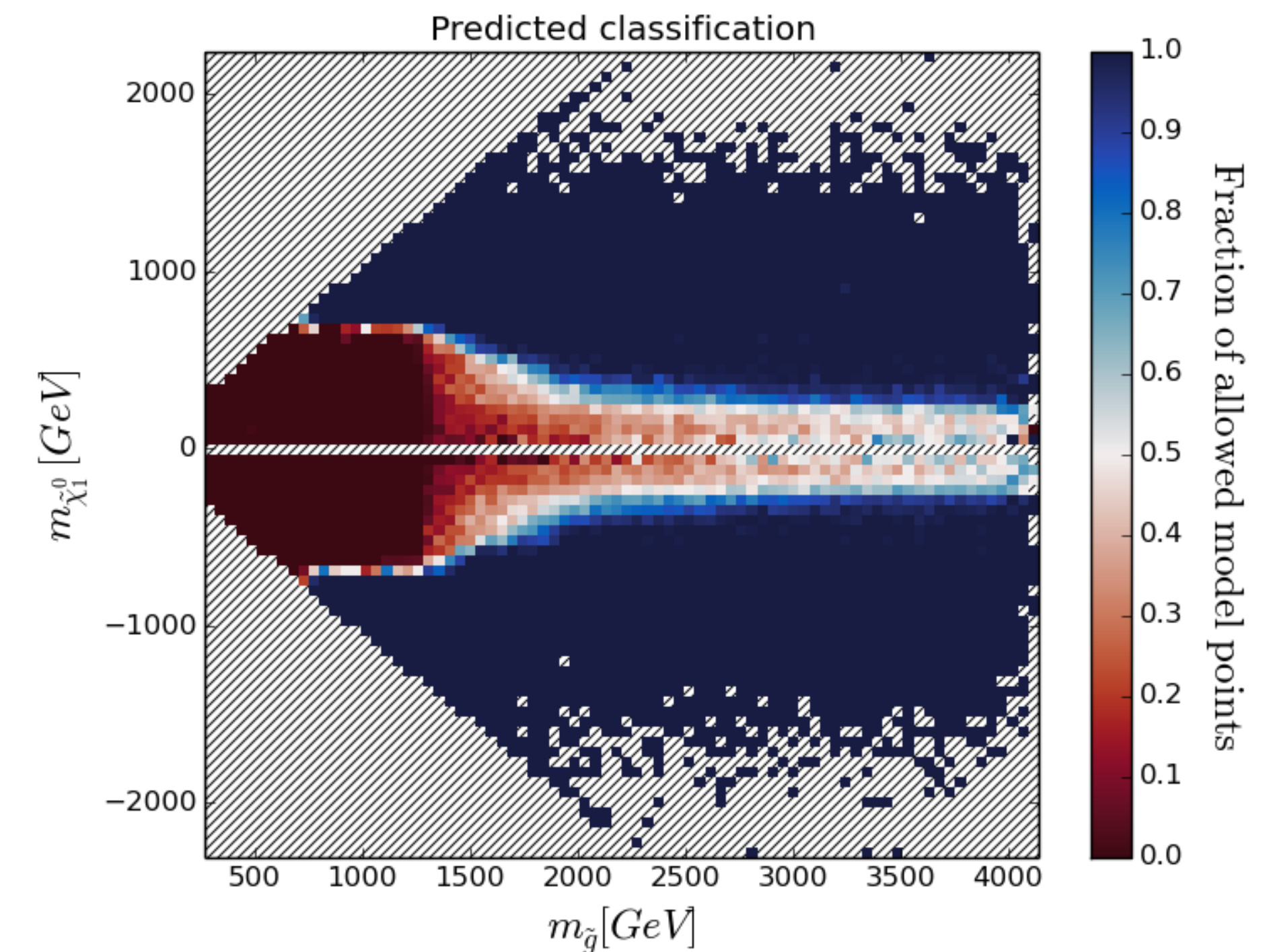
# iDark's idea: Database and interface





# Take-home messages

- SUSY-AI is already fast and reliable, but will be further improved, specifically on:
  - Generalisation
  - Usability (server-client, stacking, etc.)
- Next version will be public in only a few weeks
- Lack of data will be addressed by iDarkSurvey: a database that will host world's phenomenology data and provide tools to view and compare multiple data sets
- **Just the start:** Machine Learning will open up the multi-dimensional character of decision boundaries in order to aid searches and amplify recasting efforts



<http://www.idarksurvey.com>  
<http://www.susy-ai.org>