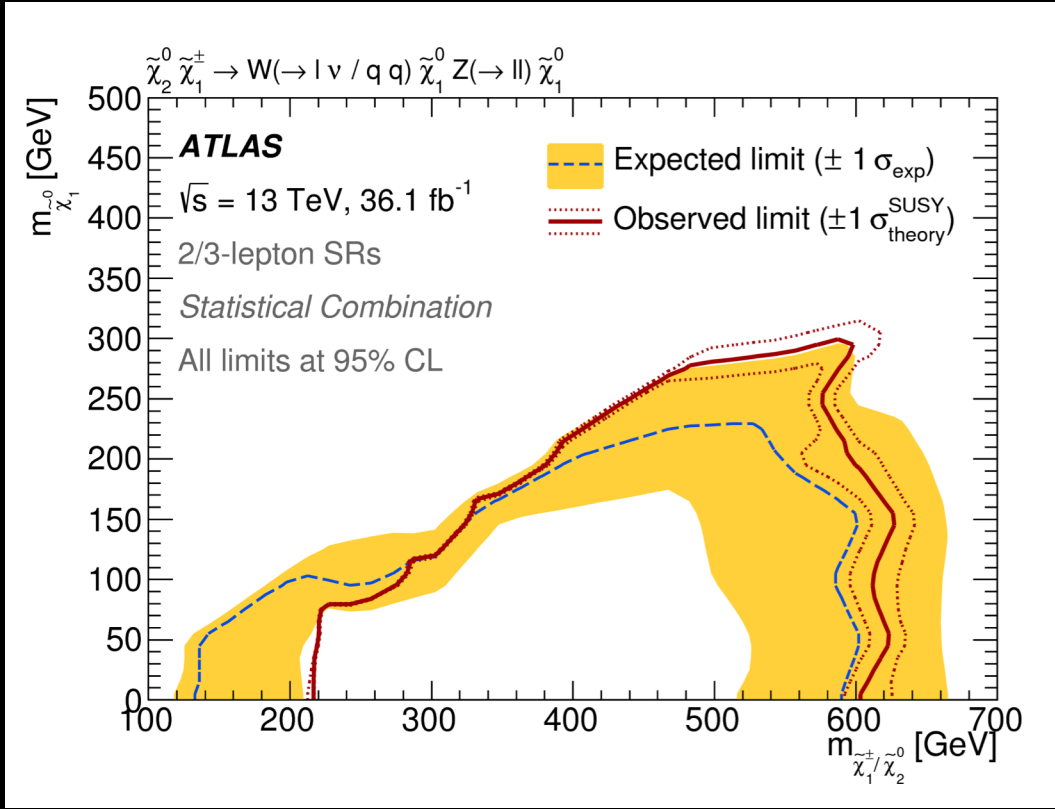
Constraining the Parameters of High-Dimensional Models with Active Learning

Sascha Caron Tom Heskes Sydney Otten Bob Stienen

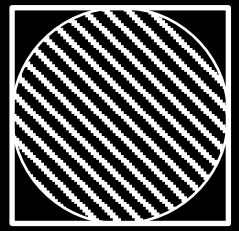
netherlands eScience center

Radboud Universiteit

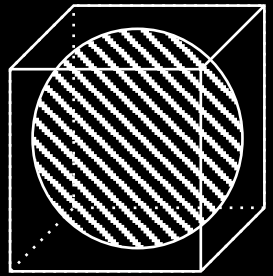
Parameter	Description
$m_{\tilde{L}_1}$	$1^{\text{st}}/2^{\text{nd}}$ gen. $SU(2)$ doublet soft breaking slepton mass
$m_{\tilde{E}_1}$	$1^{\text{st}}/2^{\text{nd}}$ gen. $SU(2)$ singlet soft breaking slepton mass
$m_{\tilde{L}_3}$	3^{rd} gen. $SU(2)$ doublet soft breaking slepton mass
$m_{\tilde{E}_3}$	3^{rd} gen. $SU(2)$ singlet soft breaking slepton mass
$m_{\tilde{Q}_1}$	$1^{\text{st}}/2^{\text{nd}}$ gen. $SU(2)$ doublet soft breaking squark mass
$m_{\tilde{U}_1}$	$1^{\text{st}}/2^{\text{nd}}$ gen. $SU(2)$ singlet soft breaking squark mass
$m_{\tilde{D}_1}$	$1^{\text{st}}/2^{\text{nd}}$ gen. $SU(2)$ singlet soft breaking squark mass
$m_{\tilde{Q}_3}$	3^{rd} gen. $SU(2)$ doublet soft breaking squark mass
$m_{\tilde{U}_3}$	3^{rd} gen. $SU(2)$ singlet soft breaking squark mass
$m_{\tilde{D}_3}$	3^{rd} gen. $SU(2)$ singlet soft breaking squark mass
A_t	Stop trilinear coupling
A_b	Sbottom trilinear coupling
A_τ	Stau trilinear coupling
$\lvert\mu\rvert$	Higgsino mass parameter
$\lvert M_1\rvert$	Bino mass parameter
$\lvert M_2\rvert$	Wino mass parameter
M_3	Gluino mass parameter
M_A	Pseudoscalar Higgs mass
$\tan\beta$	Ratio of vacuum expectation values


$\tilde{\chi}_2^0 \tilde{\chi}_1^\pm \rightarrow W(\rightarrow l\,\nu\,/\,q\,q)\,\tilde{\chi}_1^0\,Z(\rightarrow ll)\,\tilde{\chi}_1^0$
ATLAS
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2/3-lepton SRs
Statistical Combination
All limits at 95% CL
Expected limit ($\pm$ 1 σ_{exp})
Observed limit ($\pm$1 σ_{theory}^{SUSY})
$m_{\tilde{\chi}_1^0}$ [GeV]
$m_{\tilde{\chi}_1^\pm/\tilde{\chi}_2^0}$ [GeV]
0
50
100
150
200
250
300
350
400
450
500
100
200
300
400
500
600
700

2 dimensions

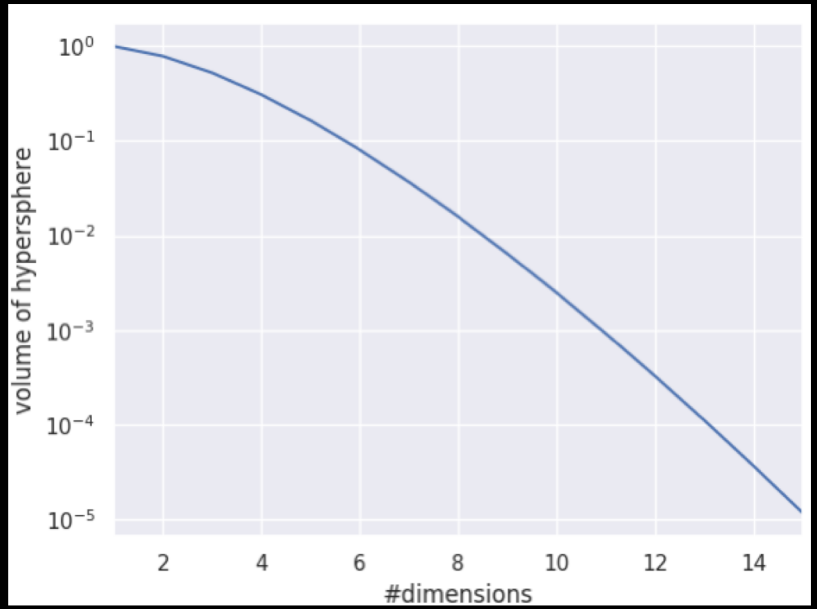


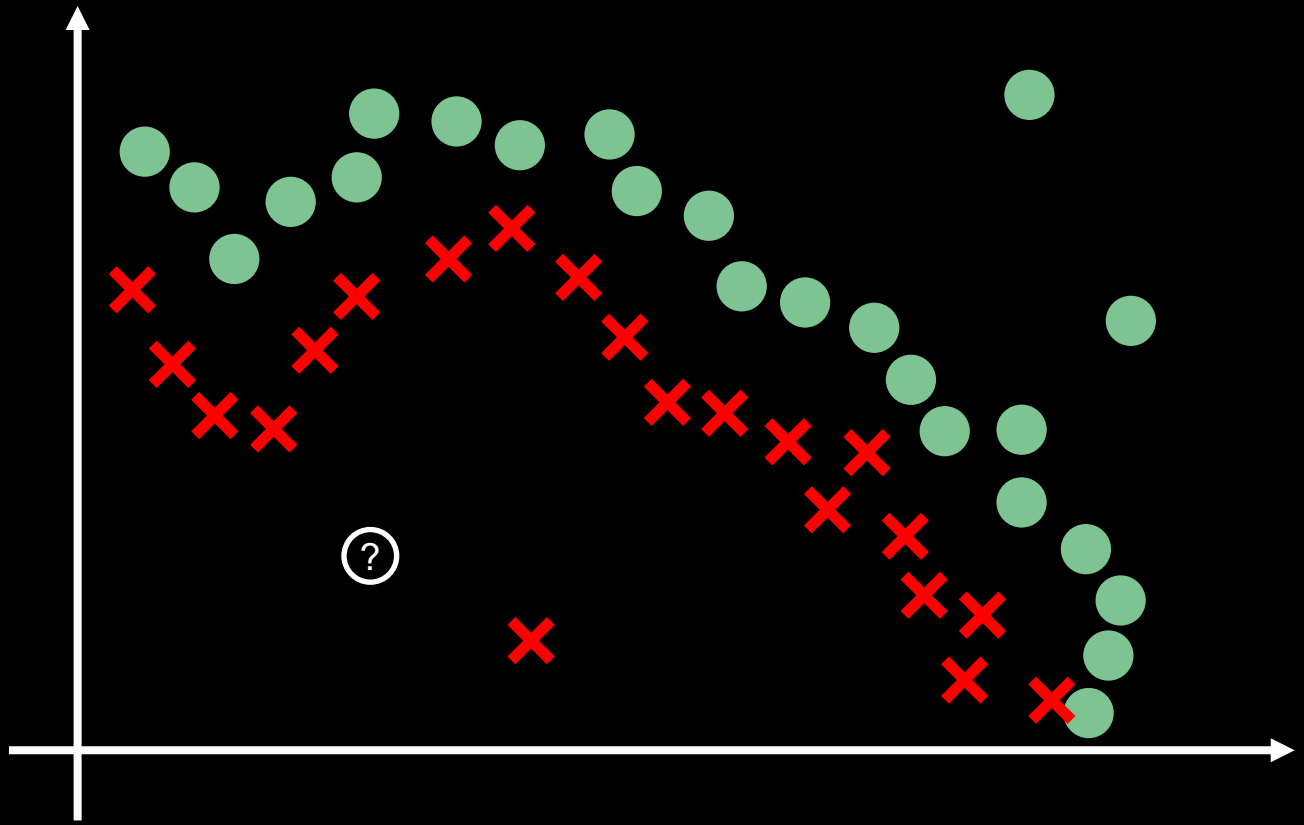
3 dimensions

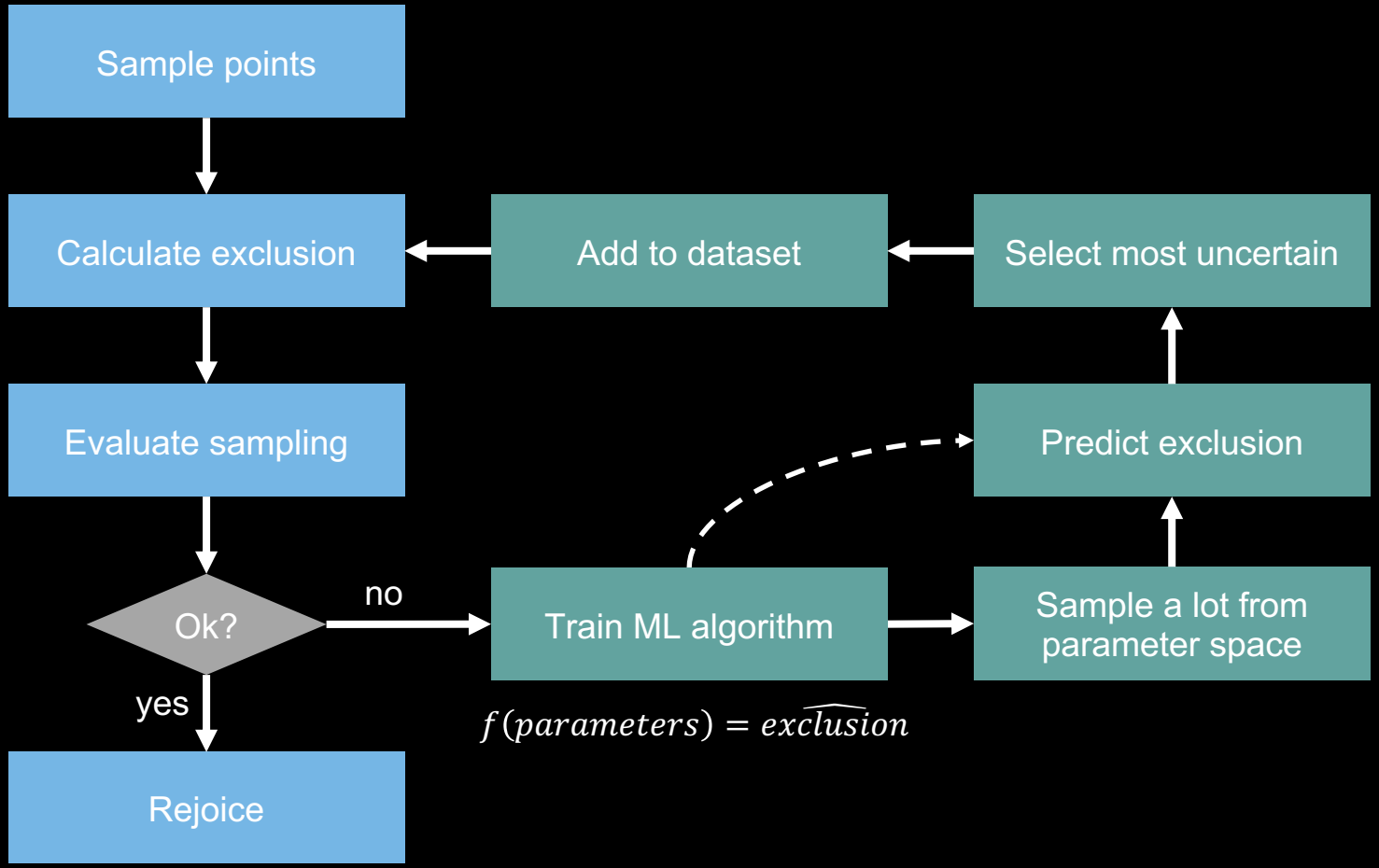


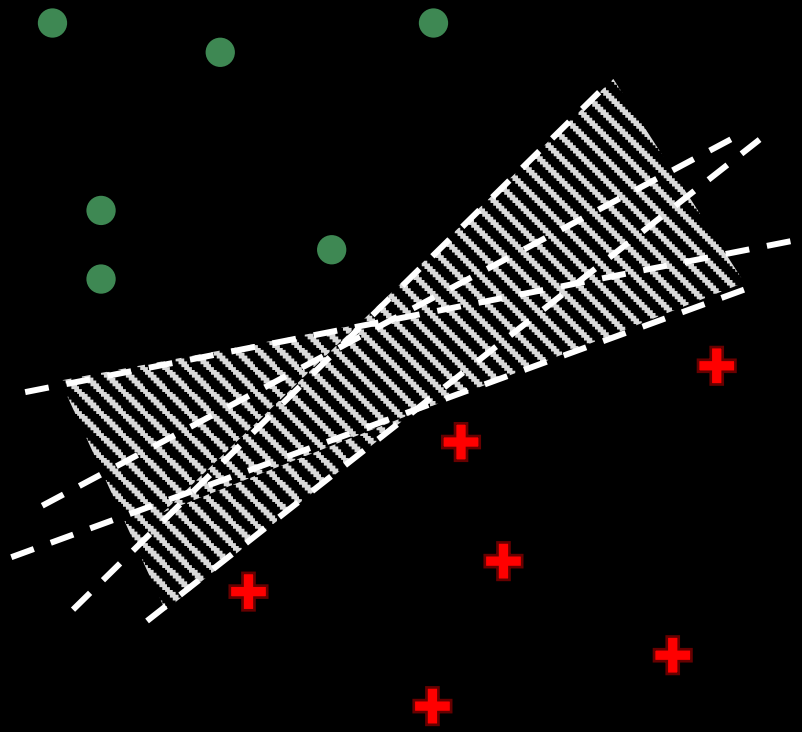
…

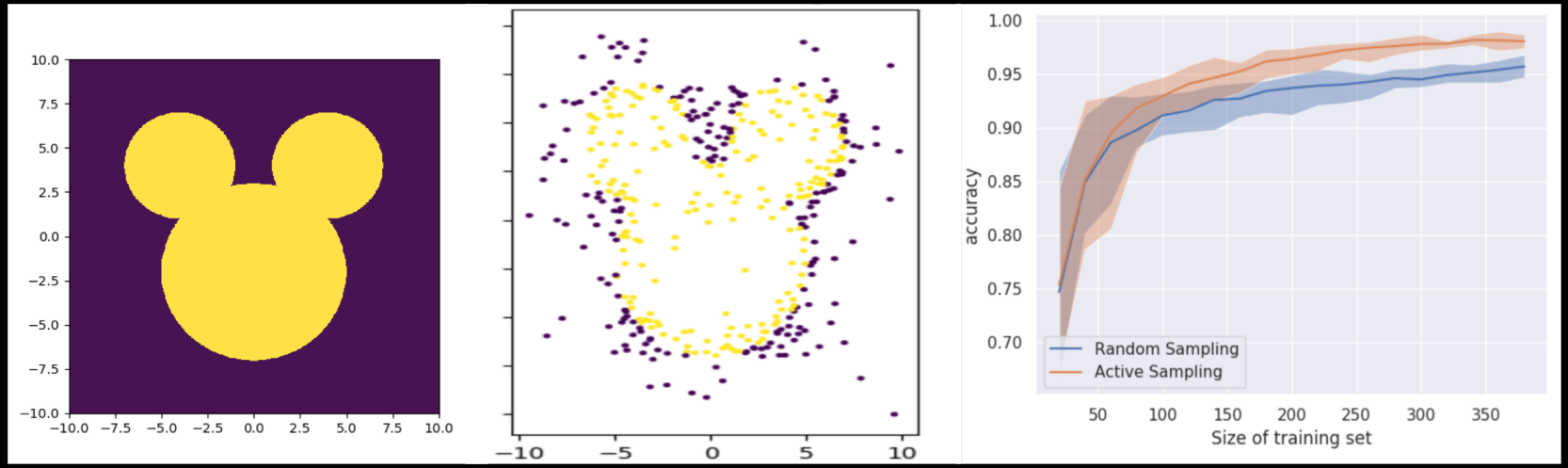
…

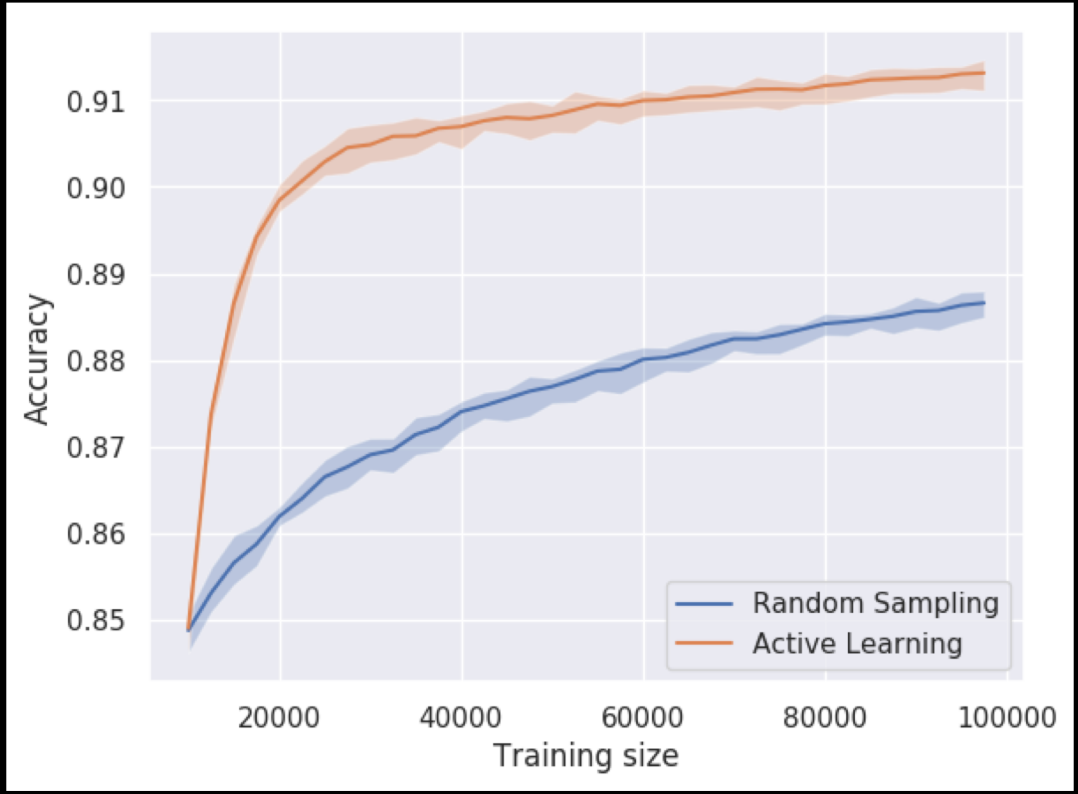





Sample points
Calculate exclusion
Add to dataset
Select most uncertain
Evaluate sampling
Predict exclusion
Ok?
no
Train ML algorithm
Sample a lot from parameter space
yes
$f(parameters) = \widehat{exclusion}$
Rejoice




10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
−10.0 −7.5 −5.0 −2.5 0.0 2.5 5.0 7.5 10.0
−10 −5 0 5 10
1.00
0.95
0.90
0.85
0.80
0.75
0.70
accuracy
Random Sampling
Active Sampling
50 100 150 200 250 300 350
Size of training set

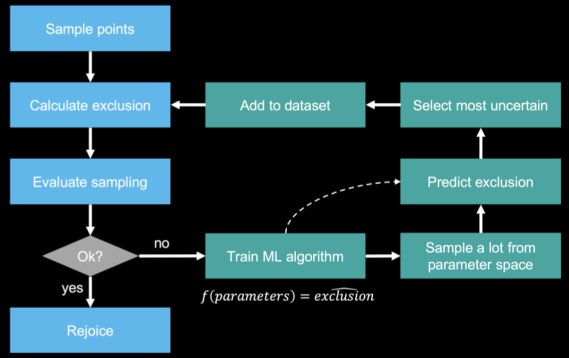

0.91
0.90
0.89
0.88
0.87
0.86
0.85
Accuracy
20000
40000
60000
80000
100000
Training size
Random Sampling
Active Learning

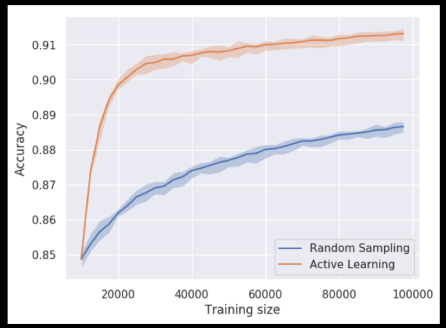
Constraining the Parameters of High-Dimensional Models with Active Learning

Active learning is a machine learning technique that can **reduce the amount of time** used **to calculate likelihoods** and exclusions

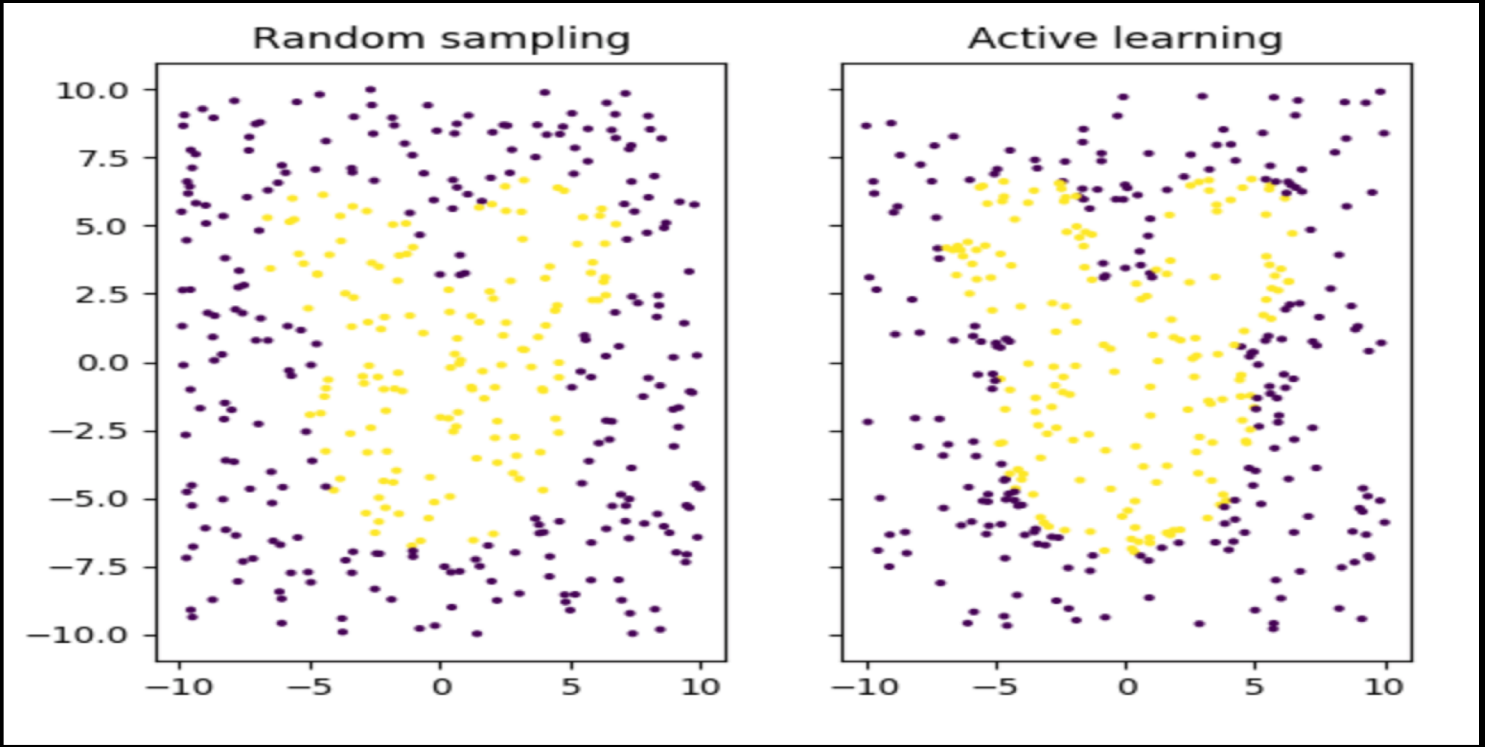
This creates time to sample more points, **allowing the exploration of higher dimensional parameter spaces**..

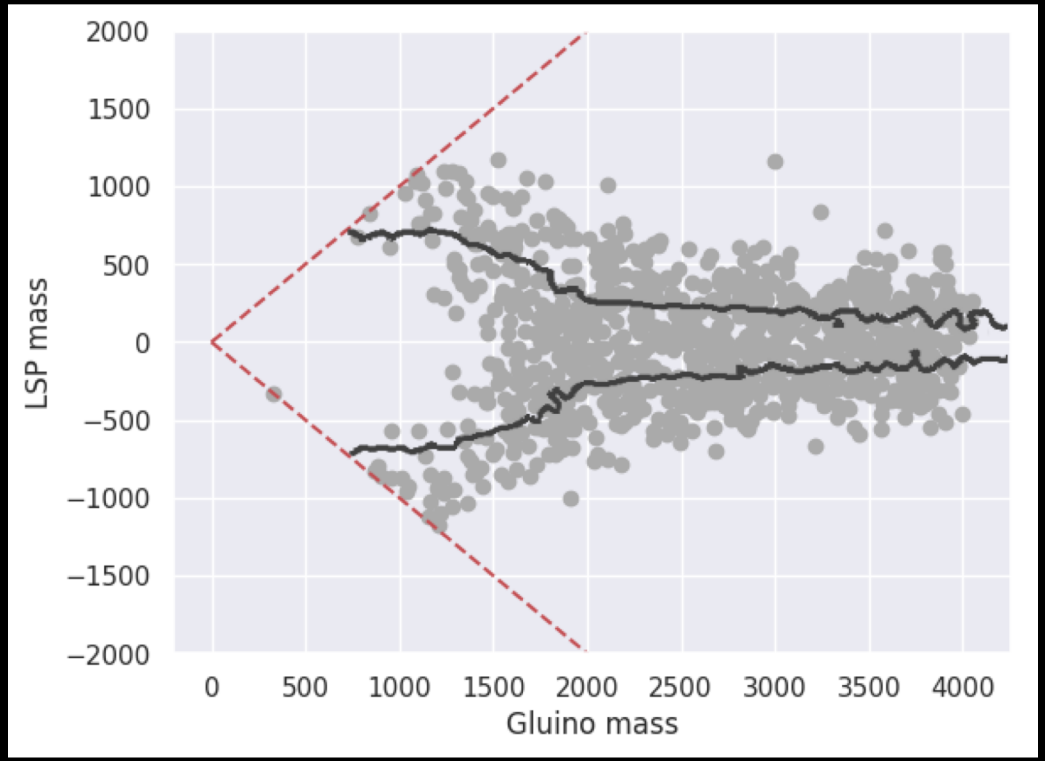
Meaning that with active learning we can explore **richer physics**





Backup slides


Random sampling
Active learning
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
−10
−5
0
5
10


2000
1500
1000
500
0
−500
−1000
−1500
−2000
LSP mass
0
500
1000
1500
2000
2500
3000
3500
4000
Gluino mass

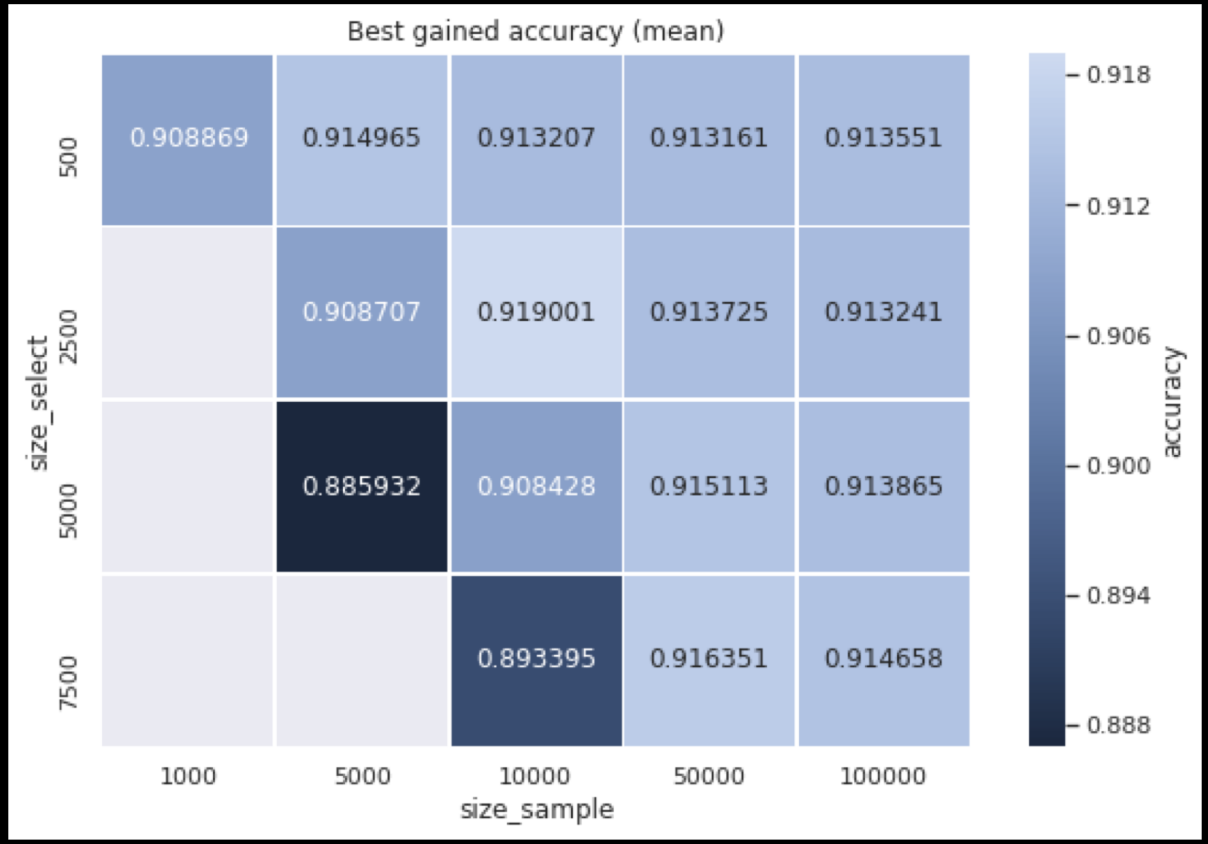
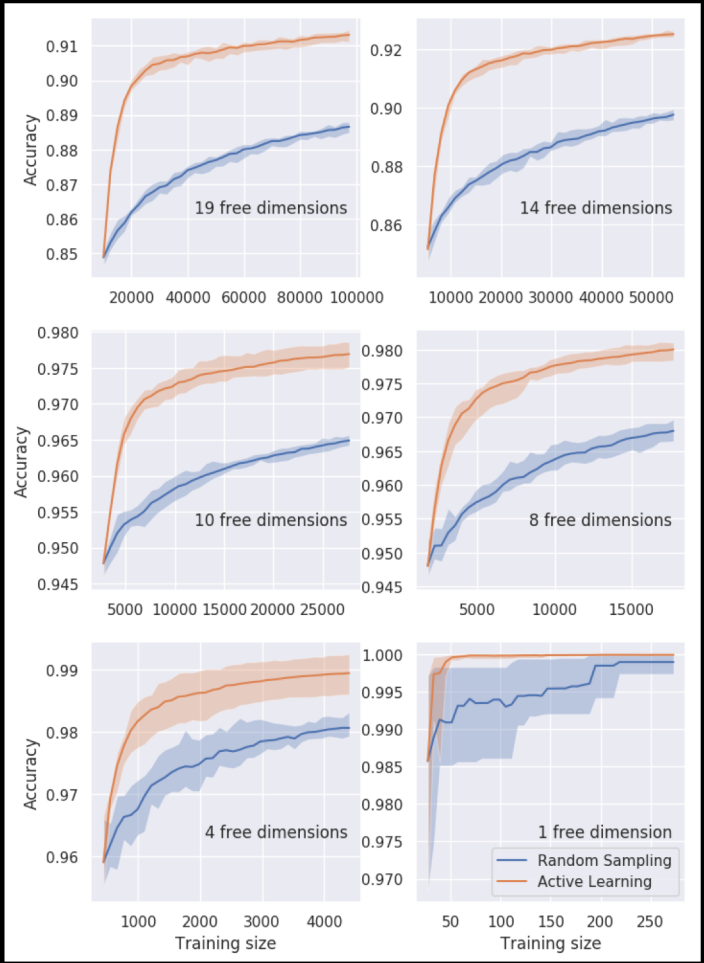
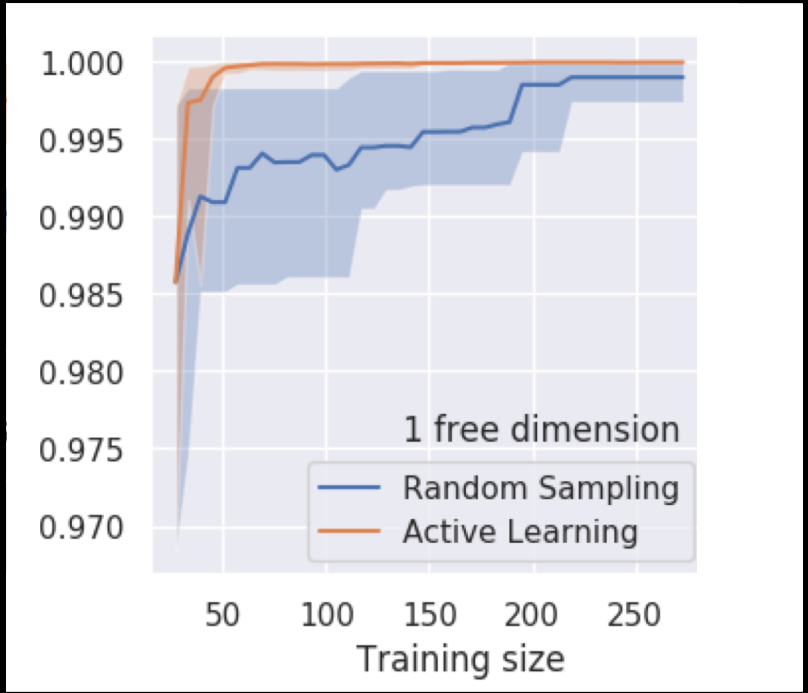

Best gained accuracy (mean)
size_select
500
2500
5000
7500
0.908869
0.914965
0.913207
0.913161
0.913551
0.908707
0.919001
0.913725
0.913241
0.885932
0.908428
0.915113
0.913865
0.893395
0.916351
0.914658
1000
5000
10000
50000
100000
size_sample
accuracy
0.918
0.912
0.906
0.900
0.894
0.888

TABLE III. Configuration for the active learning procedures in Section III A.

	III.A.1	III.A.2	III.A.3
Initial dataset	10,000		
Step size	2,500		
#candidates	remaining pool	100,000	
Maximum size	until pool empty	100,000	
Committee size	100		25
#iterations	7		
#test points	1,000,000		

# Fixed	5	9	11	15	18
M1 (GeV)	-	1750	1750	1750	1750
M2 (GeV)	-	1750	1750	1750	1750
M3 (GeV)	-	-	-	-	-
mL1 (GeV)	1750	1750	1750	1750	1750
mL3 (GeV)	1750	1750	1750	1750	1750
mE1 (GeV)	1750	1750	1750	1750	1750
mE3 (GeV)	1750	1750	1750	1750	1750
mQ1 (GeV)		-	-	-	1750
mQ3 (GeV)	-	-	-	1750	1750
mU1 (GeV)	-	-	-	-	1500
mU3 (GeV)	-	-	-	3000	3000
mD1 (GeV)	-	-	-	-	2000
mD3 (GeV)	-	-	-	2000	2000
At	-	-	3200	3200	3200
Ab	-	-	-	2000	2000
Atau	2000	2000	2000	2000	2000
mu (GeV)	-	200	200	200	200
mA2 (GeV^2)	-	-	10^7	10^7	10^7
tan(beta)	-	10	10	10	10


Accuracy
19 free dimensions
14 free dimensions
10 free dimensions
8 free dimensions
4 free dimensions
1 free dimension
Random Sampling
Active Learning
Training size
Training size


1 free dimension
Random Sampling
Active Learning
Training size

$$\text{time metric} = \frac{\text{time taken by random sampling}}{\text{time taken by active learning}}.$$

# free parameters	1μs	10μs	100μs	1ms	10ms	100ms	1s	10s	100s	1000s	$\lim_{t\to\infty}$
19	0.0112	0.1101	0.9674	4.3736	6.7504	7.1383	7.1796	7.1837	7.1842	7.1842	7.1842
14	0.0068	0.0672	0.6149	3.3239	5.9413	6.4491	6.5047	6.5103	6.5109	6.511	6.511
10	0.0046	0.0456	0.4282	2.6664	5.587	6.2742	6.3524	6.3603	6.3611	6.3612	6.3612
8	0.0021	0.0213	0.2058	1.5164	4.1751	5.0627	5.1727	5.1839	5.1851	5.1852	5.1852
4	0.0005	0.0053	0.0525	0.4784	2.5256	4.4148	4.7718	4.8107	4.8146	4.815	4.815
1	<0.0001	0.0001	0.0012	0.012	0.1152	0.8196	2.1106	2.5051	2.5528	2.5577	2.5582