ALPhA Summer Weeks 7 & 8 Presentation

BRADEN KRONHEIM

JULY 22, 2019

Summary

- Generated dataset of 500,000 softsusy points labeled valid/not valid
  - Trained on dataset using both normal neural net and the BNN
- Generated 538,759 mass data points from softsusy using BNN to pre-screen points
  - Trained a normal neural network on all of the mass data
  - Tried to train a BNN on all the mass data
  - Trained a BNN on just the higgs data
- Generated heatmaps of cross section when filtering out invalid softsusy points and Higgs mass outside of 124-126 GeV
  - Examined impact of BNN uncertainty
  - Examined impact of uniform scaling

Summary (cont.)

- Ran other general tests:
  - PCA on high percent error, bad spread cross section points
  - Compared leading order cross sections to neural net predictions
  - Created a sort of relevance detector for masses
- BNN performance improvements:
  - Added more @tf.function decorators
  - Determined GPU now offers speedup
  - Fixed PReLU implementation
- Ctrl + S freezes command prompt, Ctrl + Q unfreezes it

Softsusy valid/not valid dataset

- 500,000 points
  - 68.17% valid, 31.83% invalid
  - 20% of data used for test, 16% for validation, 64% for training
  - Binary classification, 1 is valid, 0 is invalid

Metric used to split	Precision	Recall	f1
Mean - 3sd	**0.969**	0.891	0.928
Mean	0.940	0.938	**0.939**
Mean + 3sd	0.900	**0.968**	0.933

Metric used to split	Valid flagged invalid/ total flagged invalid	Valid flagged valid/ total flagged valid	Valid flagged invalid/ total valid
Mean	2.81%	**94.8%**	6.23%
Mean + 3sd	**1.20%**	90.8%	**2.57%**

Softsusy mass data

- Generated by screening 1,250,000 parameter combinations through the valid/not valid classifier with mean + 3sd as the split metric

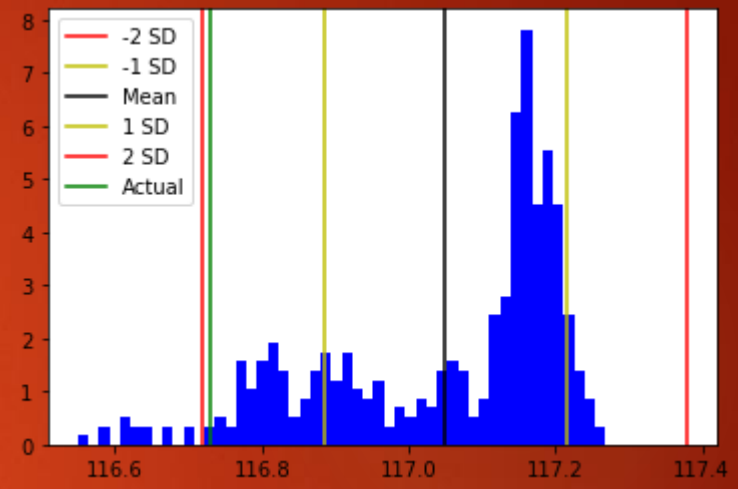
Particle	% error	Particle	% error	Particle	% error
MW	0.00281%	Chi_2+	0.833%	stau_1	5.42%
h0	0.516%	d_L	0.608%	nu_tau_L	2.52%
H0	0.827%	u_L	0.611%	d_R	0.900%
A0	1.96%	s_L	0.608%	u_R	0.734%
H+	0.827%	c_L	0.611%	s_R	0.900%
g	1.81%	b_1	2.73%	c_R	0.734%
Chi_1	118.1%	t_1	3.50%	b_2	1.48%
Chi_2	9.91%	e_1	1.22%	t_2	1.40%
Chi_1+	23.5%	nue_L	1.24%	e_R	2.47%
Chi_3	6.68%	mu_L	1.22%	mu_R	2.47%
Chi_4	3.41%	numu_L	1.24%	stau_2	1.63%

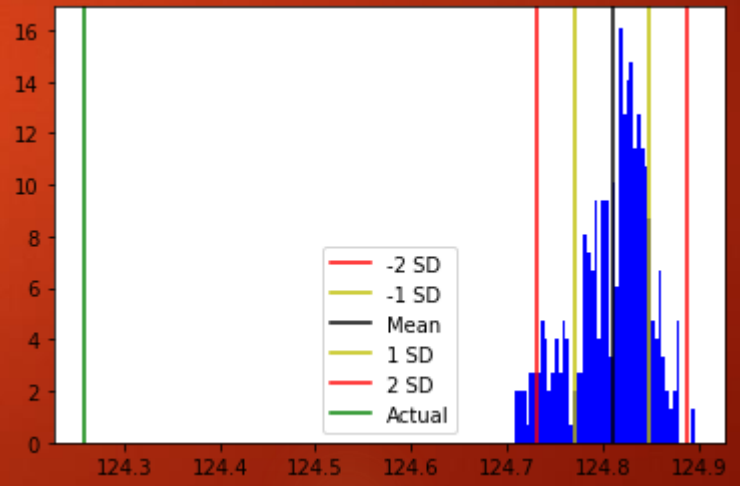
Softsusy mass data (cont.)

Degenerate masses	Percent difference	Degenerate masses	Percent difference
d_L, s_L	0.000933%	d_R, s_R	0.0019%
u_L, c_L	0.000923%	u_R, c_R	0.00028%
e_1, mu_L	0.015%	e_R, mu_R	0.039%
nue_L, numu_L	0.015%		

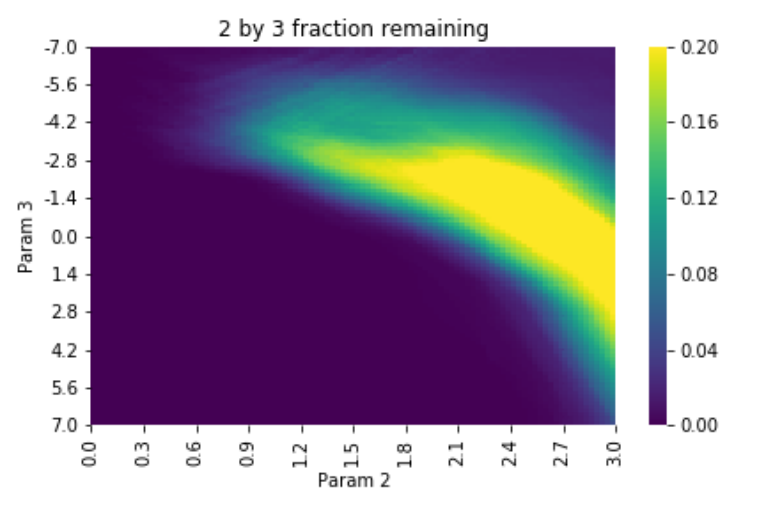
Softsusy mass data (cont.)

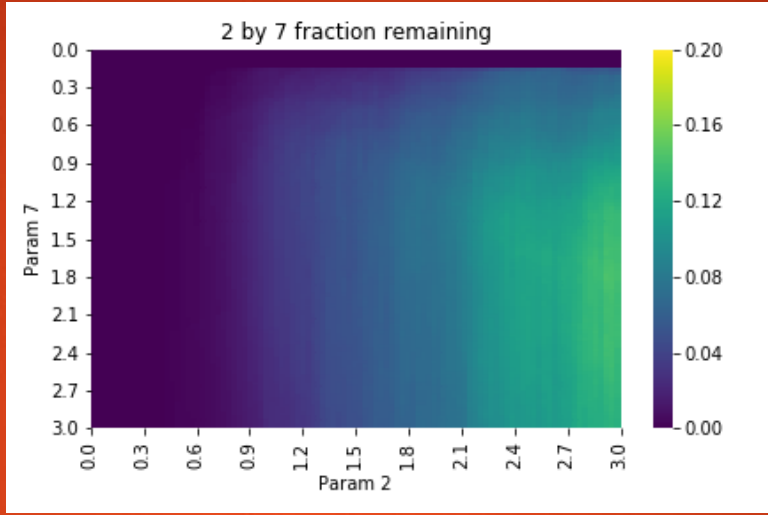
- BNN did not effectively train on the full mass data, the step size to obtain non zero acceptance rate was very small.
  - Trained on just the higgs data instead
  - After 2050 epochs:
    - Percent error test: 0.312%
    - % inside 3 sd test: 57.2%
    - Epoch 2050 squared error train: 0.00516
    - Epoch 2050 squared error validate: 0.02358

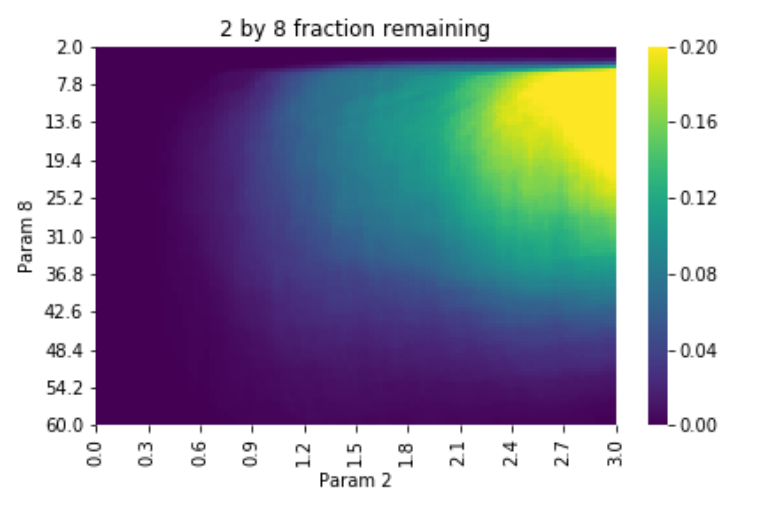


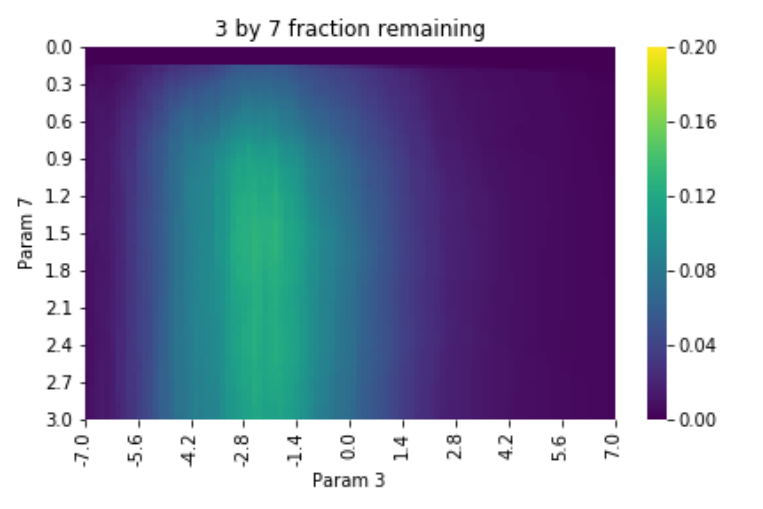


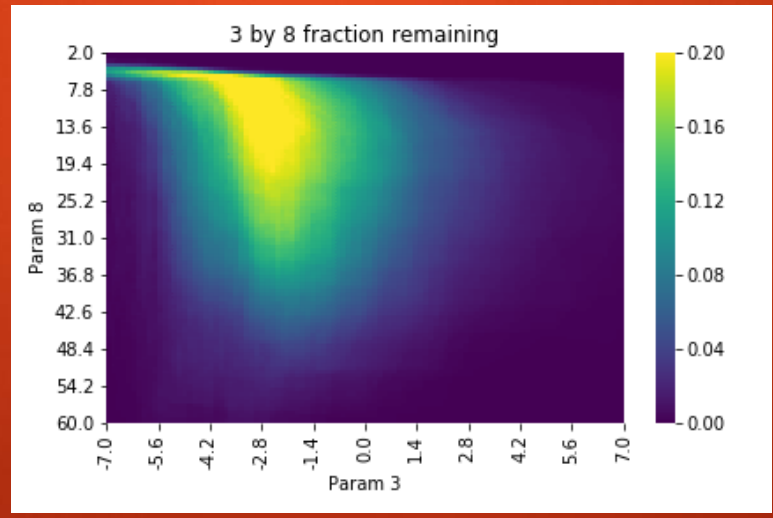
Heatmaps:

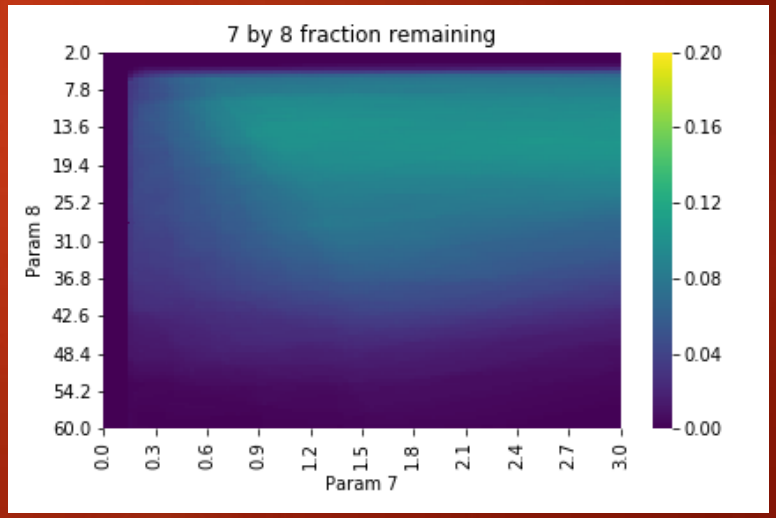

2 by 3 fraction remaining
Param 3
-7.0
-5.6
-4.2
-2.8
-1.4
0.0
1.4
2.8
4.2
5.6
7.0
Param 2
0.0
0.3
0.6
0.9
1.2
1.5
1.8
2.1
2.4
2.7
3.0
0.20
0.16
0.12
0.08
0.04
0.00


2 by 7 fraction remaining
Param 7
0.0
0.3
0.6
0.9
1.2
1.5
1.8
2.1
2.4
2.7
3.0
Param 2
0.0
0.3
0.6
0.9
1.2
1.5
1.8
2.1
2.4
2.7
3.0
0.20
0.16
0.12
0.08
0.04
0.00


2 by 8 fraction remaining
Param 8
2.0
7.8
13.6
19.4
25.2
31.0
36.8
42.6
48.4
54.2
60.0
Param 2
0.0
0.3
0.6
0.9
1.2
1.5
1.8
2.1
2.4
2.7
3.0
0.20
0.16
0.12
0.08
0.04
0.00

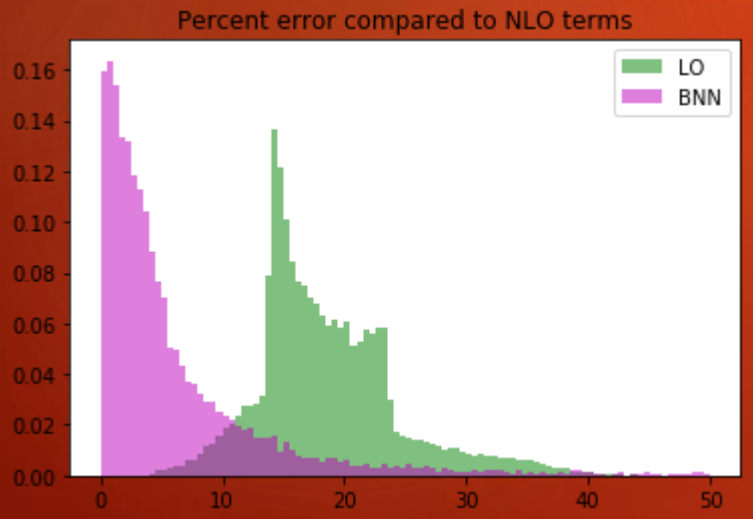

3 by 7 fraction remaining
Param 7
0.0
0.3
0.6
0.9
1.2
1.5
1.8
2.1
2.4
2.7
3.0
Param 3
-7.0
-5.6
-4.2
-2.8
-1.4
0.0
1.4
2.8
4.2
5.6
7.0
0.20
0.16
0.12
0.08
0.04
0.00

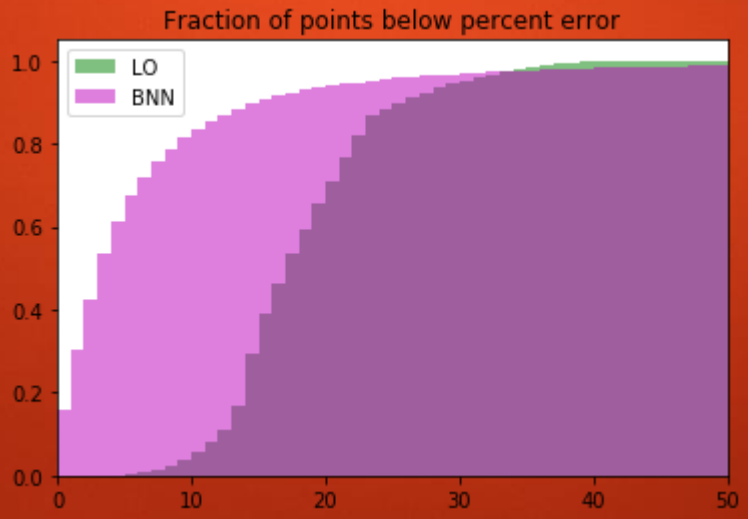

3 by 8 fraction remaining
Param 8
2.0
7.8
13.6
19.4
25.2
31.0
36.8
42.6
48.4
54.2
60.0
Param 3
-7.0
-5.6
-4.2
-2.8
-1.4
0.0
1.4
2.8
4.2
5.6
7.0
0.20
0.16
0.12
0.08
0.04
0.00

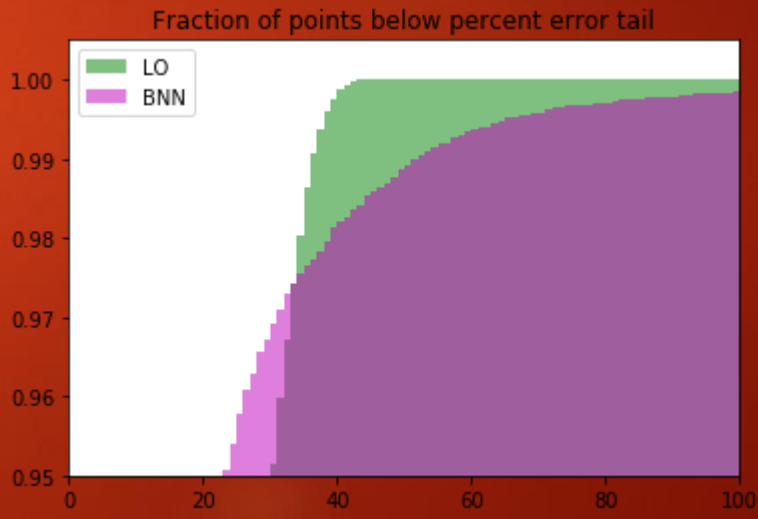

7 by 8 fraction remaining
Param 8
2.0
7.8
13.6
19.4
25.2
31.0
36.8
42.6
48.4
54.2
60.0
Param 7
0.0
0.3
0.6
0.9
1.2
1.5
1.8
2.1
2.4
2.7
3.0
0.20
0.16
0.12
0.08
0.04
0.00

General Tests

- PCA on points with a high percent error and bad distribution fit with cross section network revealed that these are all randomly scattered, as the principle eigenvalue was only ever responsible for about 8% of the variance
- LO vs. BNN prediction for cross sections:



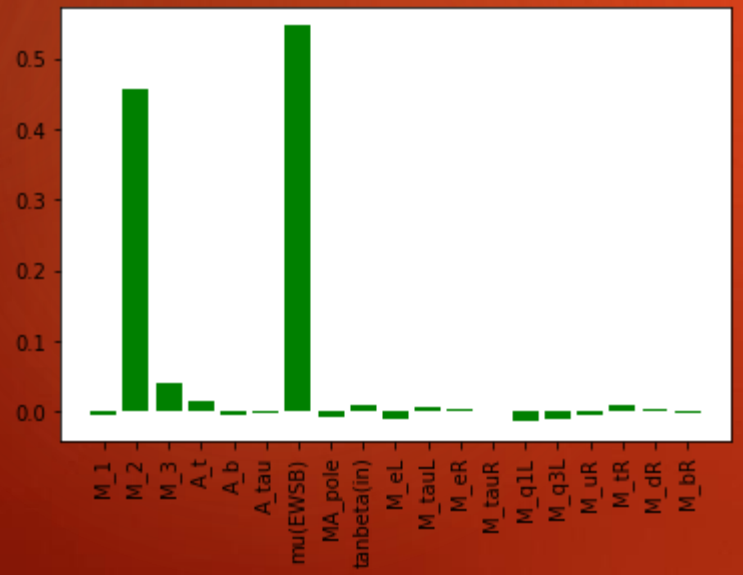




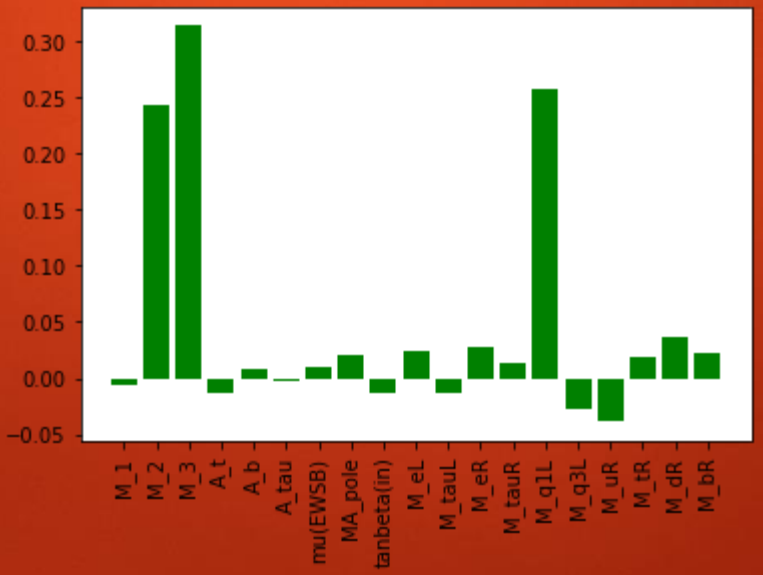
General Tests (cont.)

- Relevance detectors:
  - Trained a normal neural net on a single mass
  - Added a layer at the beginning connecting each input to a single neuron with a weight and a tanh activation function
  - Trained for a few epochs, ploted activated values of first layer

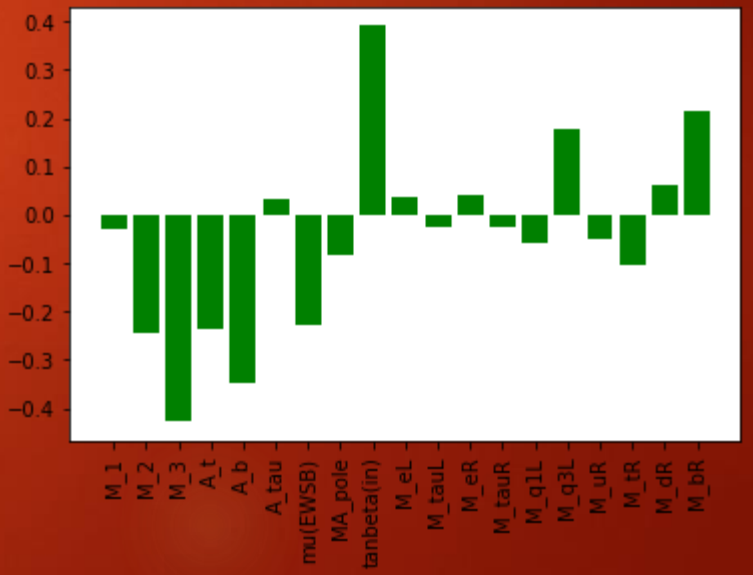
Chi_1 +



d_l



b_1



Goals for next week

- Modify PReLU to accept square root of slope
  - This allows it to be negative, prevents a discontinuity as 0
  - Normal tensorflow PReLU allows the parameter to be negative, giving a negative slope
- Pretrain a normal neural net with Relu, train a BNN using the modified PReLU
- Integrate relevance detector into BNN
- Regenerate heatmaps using BNN higgs predictor
- Storm T&I and make them turn the second node on