Exploring the Double Descent Curve in Deep Neural Networks

James Enouen The Ohio State University enouen.8@osu.edu

Abstract

This project intended to explore how the double descent curve from [1] is able to generalize to more complex models like deeper neural networks. Honestly, this work failed to get a perfectly comprehensive picture of exactly how the second descent interacts with multiple layers in a network; however, it manages to produce results manifesting the second descent and to give some future directions of further investigation.

I. INTRODUCTION

The goal of this project was to explore the double descent curve in its application to deep neural networks. The original goal was to thoroughly study how the double descent curve emerges throughout different datasets and different training regiments. This paper will walk through the journey of different approaches taken to attempt to find the double descent curve in this setting. Disappointingly, this paper spends more focus on 'how to train deep networks' than the originally desired 'investigating the second descent in deep networks.' In the end, there is no 'fairy tale ending' because none of the results are strongly conclusive; however, the final results indicate success in finding the second descent in the deep setting.

Let us begin with a review of what the double descent curve claims because it is the central focus of this paper. The paper [1] offers a new perspective on the classically believed bias-variance tradeoff. The new perspective suggests that we indeed get better results by adding more and more parameters/ hidden units to our model, despite the local maximum we approach as we reach the interpolation threshold of perfect fitting.



Fig. 1. Double descent curve from [1]

This goes against the traditionally believed idea that a model can overfit the training set and the only way to fix this problem is through regularization techniques. The paper suggests, moreover, that the minimum at the end of this curve generally tends to be lower than the local minimum attained in the bias-variance tradeoff curve. This advocates an overall different look on how to handle training models which is very interesting and consequently deserves study in applications like deep neural networks.

II. INSPIRTATION

This project began as a failure to easily replicate the results of the double descent curve in a deep neural network scenario. These original results were trained on a dataset similar to the one below. As can be seen in the tables below, the results are very far away from what one might expect for the double descent curve to react.

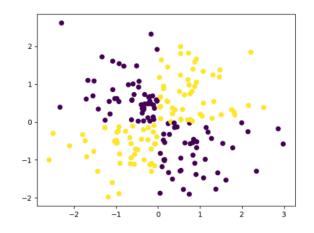


Fig. 2. Depiction of the dataset used

		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.433333	0.366667	0.366667	0.433333	0.366667	0.4	0.466667	0.6	0.366667	0.3	0.3	0.366667	0.333333	0.433333	0.256667	0.3	0.366667	0.3	0.3	0.3
2	0.433333	0.4	0.5	0.366667	0.5	0.333333	0.5	0.466667	0.3	0.333333	0.356667	0.466667	0.466667	0.366667	0.3	0.333333	0.333333	0.3	0.266667	0.333333
3	0.433333		0.366667	0.533333	0.5		0.366667	0.3	0.366667	0.433333	0.466667	0.466667	0.366667	0.366667	0.333333	0.333333	0.333333	0.466667	0.333333	0.333333
4	0.433333		0.466667	0.4	0.4		0.466667	0.5	0.366667	0.5	0.5	0.333333	0.266667	0.333333	0.333333	0.4	0.3	0.433333	0.433333	0.366667
5	0.433333		0.533333	0.633333	0.533333	0.5	0.5	0.3	0.533333	0.366667	0.433333	0.5	0.333333	0.3	0.266667	0.366667	0.366667	0.366667	0.333333	0.3
6	0.433333		0.466667	0.533333	0.566667	0.533333	0.466667	0.533333	0.5	0.366667	0.466667	0.366667	0.3	0.433333	0.466667	0.333333	0.466667	0.366667	0.3	0.333333
7	0.433333		0.533333	0.533333	0.466667	0.466667	0.466667	0.3	0.433333	0.333333	0.3333333	0.4	0.533333	0.366667	0.266667	0.333333	0.233333	0.4	0.5	0.366667
8	0.433333		0.533333	0.466667	0.466667	0.4	0.4	0.333333	0.466667	0.433333	0.466667	0.433333	0.366667	0.533333	0.466667	0.3	0.433333	0.3	0.2	0.433333
9	0.433333		0.466667	0.466667	0.466667	0.633333	0.466667	0.466667	0.366667	0.466667	0.233333	0.4	0.366667	0.366667	0.533333	0.333333	0.4	0.466667	0.366667	0.333333
10	0.433333	0.533333	0.533333	0.533333	0.466667	0.5	0.366667	0.533333	0.533333	0.533333	0.433333	0.266667	0.4	0.366667	0.3	0.433333	0.4	0.333333	0.5	0.4
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	97.02632	90.4348	86.63248	72.07665	42.35226	53.41749	44.37948	16.85936	29.74843	0.044949	0.034171	0.039822	0.036969	0.036587	0.03199	0.040733	0.035384	0.042595	0.040032	0.035306
2	97.02632	84.55031	66.92546	69.87594	66.13527	66.16225	31.06688	78.4379	43.06353	0.016425	26.71854	0.01475	0.014481	0.013263	0.01103	0.012402	0.012258	0.012836	0.014031	0.014288
3	97.02631	97.0406	94.18376	97.0405	74.9623	28.28021	75.22748	45.88446	32.18212	57.36574	70.80814	0.001043	20.66171	9.300972	0.012431	0.004979	0.006055	0.00598	0.012817	0.006177
4	97.02632	97.0406	97.0406	71,41319	84.91762	78.45159	77.04355	81.5393	9.534114	54.11592	48.44199	0.002894	32.91599	0.003194	0.002747	1.912341	0.003607	0.003559	0.003613	0.01038
5	97.02631	97.0406	97.0406	81.49667	97.0406	93.48235	89.71216	71.05329	95.64008	61.0632	51.87159	91.25308	0.001854	0.001982	0.002471	0.001704	0.002098	0.001493	0.00277	0.002816
6	97.02631	96.82388	97.0406	97.0406	85.74022	97.0405	90.81361	94.96538	85.15954	13.46222	73.0162	0.001634	0.00153	0.000121	0.001267	0.0012	0.001279	0.00186	0.001168	0.001176
7	97.02632	97.0406	97.0406	97.0406	94.9742	97.0408	97.0406	22.73616	88.9465	34.1413	20.99095	62.31834	85.74213	0.001063	0.00074	0.001247	0.001153	0.001424	0.003981	0.000922
8	97.02632	97.0406	97.0406	97.0406	97.0406	72.30798	74.09555	77.90444	97.0406	58.70549	85.5654	92.09752	20.87354	94.20922	0.001418	0.000819	0.000638	0.00224	0.000533	0.000117
9	97.02632	97.0406	97.0406	97.0406	97.0406	94.65105	97.0406	97.0406	78.8308	97.0406	31.57683	0.001222	90.49293	5.92E-05	64.05565	0.001126	3.73E-05	0.000665	0.000573	0.000916
10	97.02632	97.0406	97.0406	97.0406	97.0406	89.46602	76.16321	97.04061	97.0406	97.0406	78.4537	39.61018	0.000312	0.000568	0.000868	3.140363	0.000565	30.14213	0.001077	0.000286

Fig. 3. Training and testing results for original dataset

This clearly does not depict the pattern we were looking for, Because increasing the length of the network also increases the number of parameters, it should increase the complexity of the model. Hence, the results need to be 'flipped' as can be seen in the following two graphics. This is clearly an issue of the results which were gathered.



Fig. 4. Observed pattern of the validation errors



Fig. 5. Expected pattern of the validation errors from double descent perspective

A possibly even more fundamental issue is the fact that the more complex models had trouble scoring even better on the training loss function, which is fairly unreasonable. More advanced models have strictly greater functional complexity. This points to there being an issue in how these models are being trained. The rest of this paper are the steps taken to remedy this issue in light of the ultimate goal of finding the double descent curve in this setting.

III. EXPLORATIVE JOURNEY

A. Beginning Failures

The first steps taken to try to fix this issue was to try the same procedure on different datasets. Trying these on a sphere in two and four dimensions, the following results were generated:

87.24306	96.68315	96.68314	96.68315	96.68315	96.68316	96.68316	96.68316	96.68315	96.68316	96.68316	96.68316	96.68316	96.68314	96.68315	96.68316	0.466667	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333 (0.533333
48.06248	80.37458	96.68316	96.68316	96.68315	96.68315	96.68315	96.68315	96.68316	96.68315	96.68315	96.68315	96.68314	96.68316	96.68315	96.68315	0.233333	0.333333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333 (0.533333
3.13669	96.68315	96.68315	96.68315	96.68316	96.68315	96.68315	96.68316	96.68315	96.68315	96.68315	96.68315	96.68315	96.68315	96.68316	96.68315	0.066667	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333 (0.533333
0.017289	0.008277	96.68316	96.68315	96.68315	96.68315	96.68316	96.68315	96.68315	96.68315	96.68315	96.68316	96.68315	96.68315	96.68315	96.68315	0.033333	0.033333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333 (0.533333
0.014524	0.004567	0.001246	0.005367	96.68315	96.68314	96.68315	96.68315	96.68315	96.68315	96.68315	96.68315	96.68316	96.68315	96.68315	96.68315	0.033333	0.033333	0.066667	0.066667	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333 (0.533333
1.906065	0.004144	0.002203	0.005931	96.68315	96.68315	96.68315	96.68315	NaN	96.68316	96.68315	96.68316	96.68316	96.68315	96.68315	96.68316	0.1	0.033333	0.033333	0	0.533333	0.533333	0.533333	0.533333	0.466667	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333 (0.533333
0.014634	0.007854	0.002182	0.002086	0.005359	96.68314	12.29579	96.68316	96.68315	96.68315	96.68315	96.68316	NaN	96.68315	96.68316	96.68315	0.066667	0.033333	0.033333	0.033333	0.033333	0.533333	0.1	0.533333	0.533333	0.533333	0.533333	0.533333	0.466667	0.533333	0.533333 (0.533333
0.019759	0.003734	0.004459	0.000631	0.000247	0.000649	96.68316	0.000237	96.68315	96.68315	96.68315	96.68316	96.68315	96.68315	96.68315	96.68315	0.033333	0.066667	0.033333	0.033333	0.033333	0.033333	0.533333	0.033333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333 (0.533333
0.013969	0.003876	0.002795	0.001775	0.001822	0.000412	96.68315	96.68315	0.005343	0.000116	NaN	96.68315	96.68316	96.68316	96.68315	96.68316	0.033333	0.033333	0.1	0.066667	0.033333	0.033333	0.533333	0.533333	0.033333	0.033333	0.466667	0.533333	0.533333	0.533333	0.533333 (0.533333
0.02924	0.005688	0.002266	0.001158	0.001146	0.000359	0.005391	0.000359	0.000785	96.68315	96,68315	96.68314	96.68315	96.68315	96.68315	96.68316	0.033333	0.033333	0.033333	0.033333	0.066667	0.033333	0.066667	0.066667	0.033333	0.533333	0.533333	0.533333	0.533333	0.533333	0.533333 (0.533333
0.01366	0.003792	0.003064	0.005828	0.001086	0.003146	0.001276	0.003453	12.29569	96.68315	96.68315	96.68315	96.68316	96.68315	47.02259	96.68156	0.033333	0.033333	0.033333	0.066667	0.033333	0.033333	0.066667	0.033333	0.066667	0.533333	0.533333	0.533333	0.533333	0.533333	0.1	0.533333
0.017768	0.004443	0.002094	0.001027	96.68314	0.001366	0.005377	0.000295	0.000294	96.68315	0.00023	0.000208	96.68315	96.68316	NaN	96.68315	0.066667	0.033333	0.033333	0.033333	0.533333	0.033333	0.066667	0.033333	0.066667	0.533333	0.066667	0.033333	0.533333	0.533333	0.466667 (0.533333
				0.005682											96.68315	0.033333	0.033333	0.033333	0.033333	0.033333	0.066667	0.033333	0.033333	0.066667	0.533333	0	0.533333	0.033333	0.533333	0.533333 (0.533333
0.019078	0.005918	0.005132	0.002757	0.001349	0.000496	0.000709	0.000621	0.000819	96.68314	96.68316	96.68315	0.000907	96.68315	96.68315	96.68316	0.066667	0.033333	0.066667	0.033333	0.033333	0.033333	0.033333	0.066667	0.033333	0.533333	0.533333	0.533333	0.033333	0.533333	0.533333 (0.533333
				0.000702												0.033333	0.066667	0.066667	0.033333	0.033333	0.066667	0.033333	0.033333	0.033333	0.066667	0.033333	0.066667	0.533333	0.533333	0.033333	0.533333
				0.00076												0.033333	0.033333	0.033333	0.033333	0.033333	0.066667	0.066667	0.033333	0.066667	0.033333	0.033333	0.033333	0.533333	0.033333	0.033333	0.533333
				0.004214												0.033333	0.033333	0.033333	0.033333	0.033333	0.066667	0.066667	0.033333	0.066667	0.033333	0.033333	0.033333	0.533333	0.033333	0.066667	0.533333
				0.00089												0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.1	0.066667	0.066667	0.533333	0.533333 (0.533333
				0.001368												0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.066667	0.033333	0.033333	0.033333	0.033333	0	0.533333 (0.533333
				0.001412												0.066667	0.033333	0.033333	0.066667	0.033333	0.066667	0.033333	0.033333	0.033333	0.066667	0.033333	0.033333	0.033333	0.033333	0.033333	0.533333
				0.004700												0.022223	0.022222	0.022222	0.022222	0.022222	0.022222	0.033333	0.033333	0.033333	0.066667	0.066667	0.022222	0.022222	0.1	0.022222	0.066667

Fig. 6. 2D sphere training loss and testing error

96.9834	96.9834	96.98344	96.98344	96.98344	96.98346	96.98344	96.98344	96.98345	96.98345	96.98346	96.98344	96.98344	96.98344	96.98344	96.98345	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
96.9834	96.9834	96.98344	96.98344	96.98344	96.98344	96.98344	96.98344	96.98344	96.98344	96.98345	96.98344	96.98345	96.98345	96.98344	96.98344	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
52.8122	7 58.5175	96.98345	58.72007	96.98344	96.98345	96.98345	96.98344	96.98344	96.98344	96.98344	96.98344	96.98345	96.98344	96.98344	96.98345	0.166667	0.233333	0.466667	0.3	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
23.6474	34.8371	51.44621	NaN	73.14421	96.98344	96.98344	96.98345	96.98344	96.98344	96.98344	96.98344	96.98344	96.98344	96.98344	96.98344	0.166667	0.166667	0.266667	0.533333	0.333333	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
0.01518	37.4785	0.001244	96.98344	84.56888	96.98346	96.98345	12.59588	96.98344	96.98345	96.98346	96.98344	96.98344	96.98345	96.98346	96.98345	0.1	0.233333	0.033333	0.466667	0.533333	0.466667	0.466667	0.2	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
0.01650	0.00244	0.010131	11.24923	2.505554	52.80477	96.98344	5.086042	96.98346	96.98345	96.98344	96.98345	96.98344	96.98345	96.98344	96.98346	0.066667	0.133333	0.266667	0.133333	0.1	0.266667	0.466667	0.166667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
0.01911	6 0.00969	5 0.008307	0.001559	96.98344	96.98344	96.98344	0.006107	96.98345	45.72218	40.58935	96.98344	23.57624	96.98345	88.54199	96.98344	0.1	0.133333	0.133333	0.166667	0.466667	0.466667	0.466667	0.1	0.466667	0.066667	0.133333	0.466667	0.166667	0.466667	0.266667	0.466667
0.01327	0.00399	2 0.005975	0.001119	5.227005	96.98344	NaN	96.98344	68.13548	96.98345	96.98344	96.98345	65.36116	96.98344	96.98344	96.98344	0.133333	0.2	0.033333	0.033333	0.133333	0.466667	0.533333	0.466667	0.266667	0.466667	0.466667	0.466667	0.166667	0.466667	0.466667	0.466667
0.01195	6 0.00564	0.002682	0.006408	0.000839	0.005476	96.98344	96.98344	5.226889	96.98344	96.98346	96.98344	96.98344	96.98344	96.98344	96.98344	0.066667	0.1	0.2	0.233333	0.266667	0.166667	0.466667	0.466667	0.133333	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
0.01752	0.00453	2 0.006947	0.000599	0.005474	0.000916	NaN	96.98345	23.26373	5.283883	96.98346	96.98346	NaN	96.98345	5.226749	96.98345	0.066667	0.1	0.1	0.133333	0.133333	0.1	0.533333	0.466667	0.066667	0.2	0.466667	0.466667	0.533333	0.466667	0.066667	0.466667
0.01325	0.00300	0.001574	0.001692	0.002976	0.000659	0.000292	96.98344	96.98345	23.26375	96.98345	15.67062	96.98345	96.98344	96.98344	96.98344	0.066667	0.133333	0.133333	0.1	0.066667	0.166667	0.133333	0.466667	0.466667	0.1	0.466667	0.133333	0.466667	0.466667	0.466667	0.466667
0.01503	0.00418	L 0.001977	0.000793	96.98344	0.003064	0.000682	0.000395	96.98345	0.000297	96.98345	96.98345	89.97064	96.98341	96.98346	96.98344	0.133333	0.133333	0.066667	0.133333	0.466667	0	0.066667	0.133333	0.466667	0.1	0.466667	0.466667	0.366667	0.466667	0.466667	0.466667
0.01064	0.00578	0.006392	0.001435	0.010063	0.000745	0.000614	0.000458	0.000305	96.98345	0.000273	0.000281	5.232477	96.98344	96.98337	96.98344	0.1	0.133333	0.133333	0.066667	0.266667	0.1	0.166667	0.166667	0.066667	0.5	0.166667	0.1	0.1	0.466667	0.466667	0.466667
0.01188	0.00546	3 0.00211	0.002158	0.001707	0.000819	0.006244	0.00046	0.005341	96.98345	0.000294	96.98345	96.98345	0.000135	96.98344	96.98344	0.033333	0.066667	0.166667	0.1	0.166667	0.1	0.066667	0.066667	0.066667	0.466667	0.033333	0.466667	0.466667	0.066667	0.466667	0.466667
0.01621	0.00710	5 0.003204	0.002073	0.002464	0.000834	0.001524	5.226883	0.000243	33.81715	0.000476	NaN	96.98346	96.98345	96.98345	96.98344	0.033333	0.033333	0.1	0.133333	0.066667	0.133333	0.066667	0.166667	0.033333	0.2	0.1	0.533333	0.466667	0.466667	0.466667	0.466667
0.01563	0.00515	L 0.002774	0.003363	0.000706	0.005375	0.000385	0.000594	96.98344	NaN	96.98345	96.98344	0.000166	0.005802	96.98345	96.98344	0.066667	0.033333	0.1	0.1	0.1	0.066667	0.066667	0.133333	0.466667	0.533333	0.466667	0.466667	0.166667	0.1	0.466667	0.466667
0.02054	0.00574	0.002047	0.002088	0.001985	0.00069	0.000359	5.284004	0.001035	0.000306	0.000815	96.98344	96.98036	96.98345	0.002767	96.98344	0.1	0.066667	0.133333	0.033333	0.066667	0.1	0.1	0.133333	0.066667	0.066667	0.1	0.466667	0.466667	0.466667	0.1	0.466667
0.01476	0.00539	0.002416	0.002624	0.00174	0.001217	0.000448	0.000731	0.000372	0.00436	0.006907	5.226861	0.000314	NaN	96.98345	96.98345	0	0.133333	0.1	0.133333	0.1	0.166667	0.066667	0.133333	0.133333	0.133333	0	0.166667	0.066667	0.533333	0.466667	0.466667
0.01293	0.00536	0.002768	0.001802	0.001682	0.0014	0.001423	0.000403	0.000483	0.000206	96.98344	0.001405	0.000346	NaN	96.98345	NaN	0.033333	0.133333	0	0.1	0.2	0.066667	0.033333	0.133333	0.2	0.2	0.466667	0.133333	0.1	0.533333	0.466667	0.533333
0.01899	0.00694	5 0.002706	0.002818	0.001209	0.00234	0.003724	0.000813	0.00226	0.003522	0.000686	0.000842	0.000234	0.000352	0.000505	5.227009	0.033333	0.066667	0.133333	0.066667	0.066667	0.166667	0.1	0.1	0.066667	0.066667	0.166667	0.166667	0.066667	0.2	0.1	0.166667
0.06010	0.02080	5 0.009596	0.006534	0.004421	0.003131	0.002565	0.002094	0.001795	0.001437	0.003239	0.001252	0.001004	0.000741	0.001261	0.00088	0	0.066667	0.1	0.066667	0.033333	0.1	0.1	0.066667	0.066667	0.133333	0.166667	0.1	0.066667	0.2	0.1	0.1

Fig. 7. 4D sphere training loss and testing error

Both of these trials are not a huge step above the work started with. While the axis for these charts are now flipped from the previous images, they still have the issue of being in the wrong direction for the 'complexity' of the model. There are quite a few more failures of these first steps taken towards exemplifying the double descent curve. Notably, these charts have NaN as one of their values. This is due to numerical issues to do with logarithms and exponents. This happens to be because the loss function these training procedures are using is negative log likelihood. In retrospect, this was a little bit of an unreasonable loss function, considering the original paper was using mean squared error and zero-one loss as their functions. It is still interesting to see if this loss function has the potential to see the double descent curve itself, but this will be a future direction not included in this paper. Clearly, the training methods being used were insufficient to properly train a deep neural network which led to these results.

One of the first failures which unaligned the training procedure from the double descent paper was the size of the dataset. The goal of the second descent is to be able to perfectly fit the training set and the set being used was probably to large to feasibly do this with the number of parameters being used, hence the size of the training dataset was reduced and resulted in the following miniscule changes:

Starter Starter <t< th=""><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th></t<>																																
	94.97327	93.8775	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97328	94.97327	94.97327																
Sample Sample<	51.76485	71.39561	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327	94.97327																
Action Section Solution	94.97327	94.97327	69.92107	94.97327	94.97327	57.52935	94.97327	94.97327	94.97327	94.97327	94.97328	94.97327	94.97327	94.97327	94.97327	94.97327																
000336 000490 900336 900342<	76.085	9.451978	94,97327	0.005434	94,97327	94.97327	94,97327	94.97327	94,97327	94.97327	94,97327	94.97327	94,97327	94,97327	71.56166	94,97327	0.466667	0.166667	0.466667	0.066667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.3	0.466667
0.01560 0.00570 0.00570 7.53218 5.00516 5.00527 9.47327	0.019443	0.007032	0.005839	50.48221	94,97327	94,97327	54,26652	94,97327	94,97327	94,97327	94,97327	94,97327	94,97327	94,97327	94,97327	94,97327	0.066667	0.133333	0.066667	0.266667	0.466667	0.466667	0.1	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
0.01812 0.0076 0.0078 1.0077 0.00181 0.0078 0.007																	0.1	0.1	0.066667	0.033333	0.4	0.166667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.133333	0.466667
0.01332 0.0056 0.02134 9.4727 9.473																	0.066667	0.066667	0.133333	0	0.033333	0.033333	0.466667	0.466667	0.466667	0.166667	0.466667	0.166667	0.466667	0.466667	0.466667	0.466667
0.01343 0.0055 0.01143 0.0057 0.01567 0.06667																	0.133333	0.033333	0.033333	0.466667	0.2	0.066667	0.033333	0.1	0.133333	0.133333	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
0014167 (00357) 002156 002667 (004667 004667	0.010102																0.066667	0.166667	0.066667	0.166667	0.166667	0.133333	0.1	0.466667	0.133333	0.166667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
001215 0.00737 0.00186 0.0028 0.0016 0.00667 0.06667 0	0.013382																0.1	0.066667	0.066667	0.1	0.166667	0.066667	0.066667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.1	0.466667
0.02151 0.0075 0.0016 9.072 0.0001 9.072																	0.033333	0.066667	0.133333	0.1	0.066667	0.2	0.066667	0.466667	0.1	0.466667	0.466667	0.066667	0.1	0.466667	0.466667	0.466667
001197 000646 000114 000177 000646 000114 000177 000646 000114 000177 000646 000114 000177 000646 000114 000177 000646 000114 000177 000646 000114 000177 000646 000114 000177 000646 000114 000177 000646 000114 000177 000646 000114 000177 000646 000114 000177 000646 000117 000646 0001114 000177 000646 000111 000177 000646 0001111 000177 000646 0001111 000177 000646 0001111 000177 000646 0001111 000177 000646 0001111 000177 000646 0001131 0001313 001333 0001313 001333 0001111 000177 000164 000646 000646 0006667 01006667 0100667 0100667 0100667 0100667 0100667 0100667 0100667 0100667 0000113 0001333 00013333	0.014187																															
0000920 0000910 0000912 <t< td=""><td>0.012315</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></t<>	0.012315																															
0.003522 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00352 0.00353 0.00353 0.00353 0.00353 0.00353 0.00353 0.00353 0.00353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01352 0.00353 0.01352 0.00353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01353 0.01333 0.01353 0.01353 0.01333 <	0.011972																															
001322 000935 000273 000246 00085 00019 00076 00075 000075 000000 00000 00000 00000 00000 00000 0000																	0.1															
001296 00025 000319 000209 000259 000110 000695 000110 000695 000014 000014 000015 000152 00011 000697 00015 0000015 0000015 0000015 000000	0.013522	0.003915	0.002773	0.002609	0.000757	0.000989	0.00189	0.0054	0.000455	0.001701	0.009627	0.000366	12.97885	0.007191	0.000836	94.97327	0.1															
0033930 000429 0002735 0002735 0002735 0002735 0002735 0002745 000695 00072 000095 000025 000025 00015 000275 00075 000095 000095 00095 00095 00095 00095 00095 00095 00095 00095 00095 00	0.012145	0.005371	0.002703	0.001047	0.000706	0.005552	0.0011	0.000697	0.000724	0.000319	0.005192	0.000375	5.353239	24.5166	0.000207	94.97327																
001259 000139 000259 000259 000257 000255 000256 000256 000012 000526 000012 00052 000010 00051 00050 0010 00050 0010 00050 0010 00050 0010 00050 0010 00050 0010 00050 00000 000000 000000 000000 000000	0.013909	0.004269	0.002368	0.001746	0.001913	0.002521	0.005392	0.000441	0.000467	0.0017	0.000875	0.000432	0.000246	0.000261	NaN	94.97327	0.1															
002569 00085 00025 00025 00025 00025 00025 00025 00025 00025 00025 00025 00025 00005 00005 00005 00005 00005 00005 00005 00005 00005 00005 00005 00005 00005 000000	0.012425	0.006199	0.002748	0.004429	0.001398	0.000564	0.000623	0.000432	0.000452	0.000355	0.000321	0.000233	0.000226	0.00116	0.00592	NaN	0.033333															
0.016667 0.00379 0.002539 0.001816 0.0006667 0.01 0.1 0.1 0.1 0.1 0.1 0.0 0.066667 0.023333 0.066667 0.023333 0.066667 0.01 0.0 0.066667 0.1 0.0 0.06667 0.01 0.0 0.0 0.00012 0	0.015029	0.003875	0.002755	0.002037	0.002366	0.001139	0.000718	0.000538	0.005283	0.004291	0.000902	0.000858	0.000327	0.000335	0.0005	94.97328	0.1								0.1	0.1	0.2	0.166667	0.1			
	0.016867															94,97327	0.133333	0.1	0.066667	0.033333	0.166667	0.066667	0.2	0.066667	0.1	0.1	0.1	0.1	0.1	0	0.066667	0.466667
	0.041347																0.066667	0.1	0.133333	0.066667	0.133333	0.166667	0.1	0.133333	0.2	0.133333	0.133333	0.166667	0.133333	0.166667	0.233333	0.2

Fig. 8. Smaller dataset - 4D cylinder training loss and testing error

B. Minor Improvements (pretrain)

One of the characteristics that distinguishes the random Fourier features setting (where the double descent curve was discovered) from the neural network setting is the variance between the different layers of the model. The next step taken to fix this training procedure is an attempt to embrace this property that random Fourier features are fixing all previous layers and only training a final layer. For this, the models which got longer and longer by adding hidden layers would start off by using the original layers of the trained previous model. In this way, the goal was to reduce the variability by random initialization greatly and to smooth the results in the direction of longer and longer machines. This definitely made a notable improvement in the training losses of the models; however, it barely made a dent in the large scheme of our goal. This had nowhere near enough impact to 'flip' the results and barely added anything new to the boundary of our diagonal curve:

97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667
97.0406	78.82275	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	0.366667	0.3	0.633333	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667
NaN	41.86633	88.97706	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	0.366667	0.266667	0.3	0.633333	0.366667	0.633333	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667	0.366667
0.012198	72.71658	94.92845	63.44933	94.92841	97.0406	97.0406	97.0406	76.77616	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	0.033333	0.433333	0.366667	0.233333	0.366667	0.366667	0.633333	0.366667	0.233333	0.366667	0.366667	0.366667	0.366667	0.633333	0.366667
0.006706	42.94555	0.005746	64.59772	79.58968	52.525	97.0406	97.0406	97.0406	53.7965	97.0406	97.0406	97.0406	97.0406	97.0406	0.066667	0.333333	0.066667	0.2	0.3	0.266667	0.633333	0.366667	0.633333	0.233333	0.366667	0.633333	0.366667	0.633333	0.366667
0.01258	0.006216	NaN	0.002572	0.000462	97.0406	97.0406	64.59772	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	0.066667	0.066667	0.366667	0	0.066667	0.633333	0.366667	0.166667	0.366667	0.633333	0.633333	0.366667	0.366667	0.366667	0.633333
0.008363	0.007225	0.001215	0.002087	5.255765	0.000563	71.08364	NaN	0.000553	97.0406	23.45712	97.0406	97.0406	97.0406	97.0406	c	0	0.033333	0.066667	0.033333	0.033333	0.366667	0.366667	0.033333	0.633333	0.1	0.633333	0.366667	0.366667	0.366667
0.011245	0.002921	0.001269	0.000686	0.000711	0.000312	97.0406	0.000191	97.0406	97.0406	15.56102	NaN	97.0406	97.0406	97.0406	0.1	0.033333	0.1	0.033333	0.066667	0.1	0.633333	0.066667	0.633333	0.366667	0.133333	0.366667	0.366667	0.366667	0.366667
0.009282	0.002639	0.001106	0.000587	0.000329	0.001653	47.01648	97.0406	12.51313	97.0406	97.0406	97.0406	97.0406	97.0406	97.0406	0.033333	0.1	0.066667	0	0.066667	0.1	0.2	0.633333	0.066667	0.633333	0.366667	0.366667	0.366667	0.366667	0.366667
0.013227	0.003219	0.001814	0.001146	0.000375	0.005855	5.255657	0.000351	0.005496	97.0406	0.000276	97.0406	97.0406	97.0406	97.0406	0.033333	0.066667	0	0.1	0.033333	0.066667	0.033333	0	0.033333	0.366667	0.033333	0.4	0.366667	0.633333	0.366667
0.010092	0.003458	0.001123	0.001135	0.005437	0.001314	0.000601	0.000342	71.0834	97.0406	5.255729	93.4838	97.0406	97.0406	97.0406	0.066667	0	0.066667	0	0	0.033333	0.066667	0.033333	0.4	0.633333	0.166667	0.333333	0.366667	0.633333	0.366667
0.01062	0.003485	0.001471	0.001002	0.000844	0.000304	NaN	0.000319	0.003438	97.0406	0.00017	74.09519	0.000888	5.255618	0.000438	0.033333	0.066667	0.033333	0.066667	0.033333	0	0.366667	0.033333	0.166667	0.633333	0.033333	0.466667	0.033333	0.1	0
0.013093	0.004319	0.001832	0.000785	0.000481	0.000622	0.000668	0.00135	0.000262	5.255893	0.000415	0.000983	65.7258	97.0406	0.000474	0.033333	0.033333	0.1	0	0.033333	0	0.033333	0	0	0.033333	0	0.066667	0.266667	0.633333	0.066667
0.012253	0.003662	0.001486	0.001721	0.001569	0.000679	0.000314	0.000285	0.000523	0.000275	97.0406	0.000248	0.000167	NaN	97.0406	c	0.033333	0.033333	0.033333	0.066667	0	0.066667	0	0.1	0	0.633333	0.066667	0.066667	0.366667	0.633333
0.014678	0.003895	0.002443	0.000714	0.00075	0.000606	0.000808	0.001479	0.000423	0.000222	0.000423	NaN	0.000123	0.002861	0.000181	c	0.1	0	0.133333	0.066667	0.033333	0.066667	0.066667	0.066667	0.033333	0.033333	0.366667	0.033333	0.033333	0.033333
0.012599	0.004098	0.00206	0.001179	0.000676	0.000731	0.001268	0.000332	0.00022	0.003613	0.005526	0.000225	0.000296	9.64E-05	0.001096	0.033333	0	0.066667	0	0	0.1	0	0.066667	0	0.066667	0.033333	0.033333	0.033333	0.066667	0.066667
0.012266	0.004309	0.00272	0.001315	0.000676	0.00051	0.000471	0.000754	0.000324	0.000237	0.000257	0.000399	NaN	0.001021	0.000162	0.066667	0.066667	0	0.033333	0.033333	0.066667	0.133333	0.066667	0.066667	0.033333	0.066667	0.066667	0.366667	0	0.066667
0.013235	0.005061	0.002787	0.001424	0.000804	0.000705	0.000421	0.000427	0.00323	NaN	0.001604	0.000707	NaN	0.000364	NaN	0.033333	0	0.1	0.033333	0.066667	0	0.033333	0	0.033333	0.366667	0	0	0.366667	0.066667	0.366667
	0.006345									0.000348	0.000172	0.001258	0.000429			0.033333	0.066667	0.066667	0.066667			0.033333	0	0	0.033333	0.033333	0.066667	0	0
0.016389	0.005089	0.002071	0.001975	0.001302	0.001926	0.000595	0.000487	0.000348	0.000325	0.000416	0.001156	0.00047	0.000809	0.000146	0.033333	0.1	0.066667	0.066667	0.066667	0	0.066667	0.066667	0.066667	0.033333	0.1	0.033333	0.1	0	0.033333

Fig. 9. First pretrained training losses and testing errors

96.3384 96.3384 96.3384 96.3384 96.3384 96.3384 96.3384 96.3384 96.3384 96.3384 96.3384 96.3384 96.3383 0.3333 0.33333	33 0.533333 33 0.533333 33 0.533333 33 0.533333 33 0.533333 33 0.533333 33 0.166667
15 20569 0.01779 96.33842 95.33942 96.33942 96.33942 96.33942 96.33943	 33 0.533333 33 0.533333 33 0.533333 33 0.533333 33 0.533333 33 0.166667
0.00039 48.3595 0.00034 6.59920 98.33941 6.33941 96.33941	 33 0.533333 33 0.533333 33 0.533333 33 0.533333 33 0.166667
0.007143 0.002078 19.78029 0.00353 96.33944 96.33944 96.33944 96.33944 96.33944 96.33943	33 0.533333 33 0.533333 33 0.166667
0.0157 0.00157 0.00158 96.33843 95.33844 96.33843 0.33833 0.36667 0.53333 0.53833 0.53833 <t< td=""><td>33 0.533333 33 0.166667</td></t<>	33 0.533333 33 0.166667
0.001575 0.00223 0.00088 0.00118 6.72771 96.33944 96.33943 96.33944 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 <td< td=""><td>33 0.166667</td></td<>	33 0.166667
0.005131 0.00009 0.000756 0.006667 0.00538 0.000358 0.931435 5.15106 6.83943 96.33943 96.33942 95.33943 96.3394 96.3394 96.3394 96.3394 96.3394 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.33943 96.3394 96.3394 96.33943 96.33943 96.33943 96.3394 96.33943 96.339	
0.008918 0.002491 9.327754 0.00814 0.000999 0.000259 0.000329 0.000389 0.000302 5.151055 66.85142 0.000172 96.33943 96.33943 96.33944 96.33943 96.3	3 0.533333
0.000539 0.002304 NaN 0.001973 0.00211 0.000255 0.001024 96.33943 0.000573 0.00149 96.33943 NaN NaN 96.33944 0.033333 0.066667 0.033333 0.033333 0.033333 0.033333 0.033333 0.033333 0.033333 0.033333 0.033333 0.033333 0.00149 96.33944 0.0466667 0.046667 0.04667 0.046667 0.046667 0.04667 0.04667 0.04667 0.046667 0.04667 0.04667 0.046667 0.046667 0.046667 0.046667 0.0467 0.0467 0.04667 0.0467 0.0467 0.0467 0.0467 0.0467 0.0467 0.0467 0.0467 0.0467 0	
	33 0.533333
0.003833 0.003377 0.001474 0.0018 0.000522 0.000375 0.000255 0.000276 96.33943 0.000201 NaN 5.355803 96.33942 0.000543 96.33943 0 0.033333 0 0.066667 0.033333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.53333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.533333 0.066667 0.1 0.53333 0.066667 0.1 0.53333 0.066667 0.1 0.53333 0.066667 0.1 0.53333 0.06667 0.0 0.50700000000000000000000000000000	33 0.533333
	33 0.533333
0.009051 0.002532 0.001145 0.000789 0.000661 0.004076 0.0026 0.000451 0.00029 0.00143 96.33943 0.000169 96.33943 96.33943 0.66667 0.066667 0.066667 0.033333 0.2 0.1 0.066667 0.1 0.1 0.533333 0.33333 0.53333 0.53333 0.53333 0.53333 0.533333 0.53333 0.533333 0.53333 0.533333 0.533333 0.53330 0.53333 0.53333 0.53330 0.53330 0.53330000000000	33 0.533333
0.009338 0.002665 0.002041 0.000688 0.000708 0.000750 0.000734 0.000713 0.000459 0.003198 0.000128 0.005186 96.33945 96.33942 0.000159 0 0.033333 0.1 0.133333 0.033333 0 0.1 0.1 0.1066667 0.1 0.1 0.133333 0.53333 0.533333 0.533333 0.533333 0.533333 0.533333 0.533333 0.533333 0.533333 0.533333 0.533333 0.533333 0.533333 0.533333 0.533333 0.533333 0.53333 0.53333 0.5	
0.003133 0.003173 0.00124 0.006667 0.033333 0.00358 0.000364 0.000251 0.000128 0.000324 69.74277 0.000356 0.000302 0.000303 0.000143 0.066667 0.033333 0.066667 0.033333 0.033333 0.1 0.066667 0.033333 0.433333 0.1 0.066667 0.033333 0.433333 0.433333 0.433333 0.433333 0.433333 0.40144 0.4	33 0.1
0.01133 0.003363 0.001107 0.002282 0.000594 0.000529 0.00081 0.000254 0.000401 NaN 0.000554 0.00059 0.000638 95,73869 0.000126 0.033333 0.1 0.033333 0.166667 0.1 0.066667 0.033333 0.466667 0.166667 0.066667 0.1 0.5	83 (
0.010212 0.003701 0.00142 0.00158 0.00043 0.00043 0.00055 0.000248 0.00078 0.000432 0.00042 9633943 0.00017 9633943 0.00011 0.033333 0.1 0.06667 0.033333 0.1 0.033333 0.1 0.033333 0.1 0.133333 0.1 0.133333 0.06667 0.533333 0.00017 9633943	0.03333
0.010294 0.003521 0.00243 0.01077 0.000571 0.000427 0.000317 0.000428 0.000243 0.000512 0.000165 0.000904 9.57E-05 0.000103 9633948 0 0.1 0.133333 0.1 0.066667 0.033333 0.1 0.06667 0.00000 0.000000 0.000000000000	
0.011978 0.004097 0.001957 NaN 0.000892 0.00086 0.000346 0.000346 0.00029 5.151071 0.000238 0.04266 0.000192 0.000188 5.151066 0.033333 0.133333 0.133333 0.133333 0.133333 0.133333 0.133333 0.133333 0.1066667 0.033333 0.066667 0.03333 0.066667 0.03333 0.066667 0.03333 0.066667 0.03333 0.066667 0.03333 0.066667 0.03333 0.0000000000000000000000000000	.1 0.1
0.010325 0.005082 0.00244 0.001227 0.000991 0.000827 0.000995 0.000827 0.000995 0.00084 0.00126 9.48E-05 0.00086 0.000266 0.000226 0.000233 0.13333 0.166667 0.066667 0.033333 0.166667 0.046667 0.1 0.13333 0.13333 0.1 0.1 0.06667 0.0	
0.010995 0.003835 0.002467 0.000296 0.001408 0.000057 0.000225 0.000246 0.000377 0.005192 0.000479 0.000937 0 0.033333 0.1 0.1 0.1 0.1 0.1 0.003333 0.033333 0.046667 0.0	

Fig. 10. Second pretrained training losses and testing errors

C. Fair Improvements (MSE)

The next alteration to the training procedure was to use MSE instead of NLL. This had a very noticable impact on the results. This was the first dataset where the 'flip' was acheived. This is a vast improvement over the previous sets of models for this reason. The training loss curve made sense given the complexities of the models. The results are provided below:

69.77143	69.77	7143 6	59.77143	69.77143	69.77143	69.77143	69.77142	69.77143	69.77143	69.77143	69.77144	69.77142	69.77142	69.77143	69.77142	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
69.77143	42.8	8183 6	59.77142	56.34729	69.77142	69.77143	69.77144	69.77143	69.77143	69.77142	69.77143	69.77142	69.77143	69.77142	69.77142	0.466667	0.3	0.466667	0.333333	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
58.15588	10.00	0217 3	88.37239	9.746069	12.97496	69.77142	13.27538	69.77143	69.77143	69.77143	69.77142	69.77143	69.77143	69.77144	69.77143	0.333333	0.033333	0.233333	0.066667	0.033333	0.466667	0.1	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667	0.466667
0.009622	7.001	1967 3	3.943572	6.001435	3.783974	0.021507	54.21431	69.77143	69.77143	69.77143	49.81133	69.77143	69.77143	69.77142	69.77143	0.066667	0	0.1	0.033333	0	0	0.4	0.466667	0.466667	0.466667	0.266667	0.466667	0.466667	0.466667	0.466667
0.008135	0.001	1917 3	89.63919	5.972653	3.778234	3.555669	3.784409	69.77143	69.77143	69.77143	3.88237	3.778261	7.878795	69.77143	69.77142	0.1	0.033333	0.266667	0	0.033333	0.1	0	0.466667	0.466667	0.466667	0.033333	0.066667	0.066667	0.466667	0.466667
0.013157	4.001	437 0	0.000365	69.77143	0.000316	0.001096	2.002022	3.818824	0.000114	69.77142	69.77143	58.1982	5.637686	69.77143	69.77142	0.066667	0.033333	0.033333	0.466667	0.066667	0	0.2	0.033333	0	0.466667	0.466667	0.366667	0.1	0.466667	0.466667
0.005414	0.004	473 4	1.001297	0.000394	0.000366	3.88238	2.00028	6.75E-05	0.000167	36.71167	69.77143	0.000181	69,77143	69,77142	31,48942	0.1	0.033333	0	0.033333	0.133333	0.033333	0.1	0	0.133333	0.266667	0.466667	0	0.466667	0.466667	0.233333
										9.75E-05							0.133333		0.066667		0.166667		0		0.033333			0.066667	0.466667	0.066667
										3.881499								0.066667		0.133333	0.1		0.066667		0.066667	-	0.133333		0.466667	
										9.19E-05						0.066667				0.133333	0.066667							_	0.033333	
		_								0.000615						0.033333				0.066667			0.166667				0.133333		0.166667	_
										0.00013								0.033333						0.066667			0.066667		0.100007	
										0.000148				_		0.166667				0.066667						0.133333	0.033333		0.1	
0.013138										0.000166						0.166667				0.166667						0	0.1		0.133333	
0.007177										0.000207						0.066667		0.066667		0.166667					0.033333			0.1		0.033333
0.010536	0.003	8175 0	0.001158	0.000652	0.000575	0.000511	0.00024	0.000244	0.001123	0.000318	0.000125	0.000223	9.33E-05	0.000179	0.00018	0	0.066667	0.1	0.1	0.066667	0.133333	0.033333	0.066667	0.166667	0.066667	0.133333	0.033333	0.166667	0.033333	0.066667
0.009869	0.002	2395 0	0.002728	0.000712	0.000706	0.00038	0.000295	0.000327	0.000242	0.000112	8.08E-05	7.99E-05	0.000141	0.000486	5.31E-05	0.066667	0.1	0.133333	0.033333	0.133333	0.133333	0.166667	0.1	0.033333	0	0.033333	0.066667	0.166667	0.2	0.066667
0.009987	0.003	3521 0	0.001773	0.001024	0.000459	0.000586	0.000391	0.00016	0.000629	0.000192	0.000109	0.000195	0.000223	2.000069	0.000162	0.1	0.1	0.066667	0.1	0.1	0.133333	0.1	0.066667	0	0.1	0.033333	0.1	0.066667	0.066667	0.066667
0.011591	0.003	3559 0	0.001836	0.001192	0.000805	0.00091	0.0004	0.000495	0.000208	0.000188	0.00017	0.000351	0.000252	0.000167	0.000372	0.1	0.166667	0.1	0.1	0.1	0.1	0.166667	0.1	0.1	0	0.033333	0.133333	0.133333	0.066667	0.1
0.010274	0.003	3616 0	0.001761	0.001036	0.000738	0.00032	0.000433	0.000242	0.000276	0.000143	0.000143	0.000119	0.000106	0.000137	0.000388	0.1	0.133333	0.1	0.133333	0.066667	0.066667	0.1	0.166667	0.166667	0.1	0.166667	0.1	0.066667	0.1	0.033333

Fig. 11. First MSE training loss and testing error data

The shortcoming of this data is twofold. The first aspect is that neither are actually able to obtain the zero training loss which is required to hit the interpolation threshold. The second aspect which is in part a consequence of the first is how great the variance is in the testing accuracies of these models. There seems to be no clear cut pattern throughout these models. One could argue some small patterns like very good generalizability right after dropping off of the total inability to score well on the training set. This corresponds to the green ridge near the 'cliffs' of red and orange near the top and right of these

	1 04 04	04 04
69.77142 59.2559 69.77143 69.77143 69.77142 69.77142 69.77142 69.77142 69.77143 69.77143 69.77142 69.77143 69.77142 69.77143 69.77142 69.77143 69.77142 69.77143 69.77142 69.77143 69.77142 69.77143 69.77142 69.77143 69.77142 69.77143 69.77142 69.77143 69.77		
37.81438 59.15531 69.77142 52.8393 69.77143 69.77143 69.77143 69.77143 69.77143 69.77143 69.77143 69.77143 69.77144 0.3 0.466667 0.4 0.3 0.666667 0.4 0.4 0.4 0.4 0.4 0.4 0.4		
0.005307 5.880071 41.11911 69.77143 69.77142 69.77142 69.77142 69.77142 69.77143 69.7	0.4 0.4	0.4 0.4
0.004593 0.002273 0.000305 35.89147 2.000122 46.64647 69.77143 3.795375 69.77142 49.29451 63.01639 69.77143 1.970154 69.77143 0 0 0 0.233333 0 0.266667 0.4 0.0666667 0.4 0.33333 0.266667	0.4 0.4	0.4
0.005063 10.71914 0.000719 0.001481 0.00022 4.000214 69.77143 2.001372 22.1214 9.367101 5.73914 69.77142 3.10E-05 69.77143 69.77142 0.033333 0.2666667 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.4 0	0.4 0.4
0.004311 0.002755 1.796636 0.000242 2.001422 0.000373 35,74583 3.97315 5.732075 43,80491 69,77142 3.32E-05 33.75 0 0.066667 0.033333 0 0 0.0666667 0.133333 0.0666667 0.1 0.2 0.33333	0.4 0.4	0 0.233333
0.003927 0.002463 0.000664 5.34178 0.001318 9.90E-05 0.000179 2.001456 0.000177 0.000503 1.941434 69.77143 2.054751 69.77143 0.000184 0 0.066667 0.02 0.066667 0.033333 0 0.1 0.033333 0.033333 0.066667	0.4 0.1	0.4 0.033333
0.005035 0.001648 0.000764 0.000451 0.000385 0.000219 0.000104 0.00036 69.77142 0.000202 0.002275 148 69.77143 3.26E-05 69.77143 0.033333 0.000016 0.0000000000	0.6 0.4	0 0.4
0.004995 0.002136 0.000554 0.000396 0.000578 0.000715 0.000319 0.001603 0.000453 0.000119 1.951002 0.000124 0.000129 0.000176 0.066667 0.033333 0.033333 0.033333 0.133333 0.1 0.1 0.1 0.066667 0.033333 0.133333 0.1	0.033333 0.033333	0 0.066667
0.005666 0.001595 0.001801 0.000729 0.000424 0.000131 0.000131 0.000131 0.000134 0.000407 5.525-05 2.000066 0.001329 4.125-05 6.001086 0 0.066667 0.033333 0.133333 0.066667 0.066667 0.066667 0.066667 0.033333 0.066667	0.066667 0.033333	0.033333 0.166667
0.006781 0.001814 0.000728 0.000413 0.000621 0.000162 4.00082 0.00045 0.000194 9.955-05 0.000108 7.885-05 0.000105 132 0.000239 0.066667 0.033333 0.033333 0.066667 0.1 0.033333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.066667 0.1 0.03333 0 0.06667 0.1 0.03333 0 0.06667 0.1 0.03333 0 0.06667 0.1 0.03333 0 0.06667 0.1 0.03333 0 0.06667 0.1 0.03333 0 0.06667 0.1 0.03333 0 0.06667 0.1 0.03333 0 0.00000000000000000000000		
0.006488 0.00231 0.000847 0.00054 0.000313 0.000126 0.000316 0.000346 0.000316 0.00036 0.00036 0.00026 0.00026 0.00026 0.00026 0.00033 0.033333 0.033333 0.033333 0.033333 0.066667 0.1 0.2 0.1 0.233333 0.033333 0.133333 0.133333 0.000166 0.000316 0		0.133333 0.033333
	0.033333 0.133333	
	0.066667 0.033333	
0.007213 0.002059 0.001458 0.000814 0.000376 0.00041 0.000398 2.000265 0.000149 0.000388 0.000166 0.000116 0.000116 9.18E-05 4.01E-05 0.033333 0.1 0.166667 0.066667 0.1 0.166667 0.066667 0.133333 0.1 0.033333 0.1	0.033333 0.066667	0.033333 0.033333
0.007851 0.002358 0.001487 0.00138 0.000331 0.000338 0.000334 0.000314 0.000191 0.000146 0.000104 0.00118 0.000397 0.000145 0.000238 0.1 0.066667 0.1 0.033333 0.033333 0.033333 0.033333 0.033333 0.033333 0.0366667 0.1 0.1 0.066667	0.1 0.066667	0.133333 0.066667
0.007644 0.003189 0.001245 0.000792 0.000446 0.000364 0.000274 0.000192 0.000192 0.000159 7.73E-05 0.000239 0.000359 0.1 0.066667 0.033333 0.13333 0.13333	0.066667 0.033333	0.166667 0.033333
0.008562 0.003484 0.002333 0.000549 0.000518 0.000245 0.000166 0.000192 0.000371 0.000152 0.00012 0.001356 4.000195 8.07E-05 8.44E-05 0.1 0.1666667 0.0666667 0.066667 0.066667 0.033333 0.133333 0.133333	0.133333 0.166667	0.033333 0.033333
0.007887 0.002498 0.001309 0.000691 0.001045 0.000525 0.000418 0.000327 0.000479 0.000301 0.000584 0.00013 0.000637 6.59E-05 7.56E-05 0.066667 0.1 0.1 0.133333 0.033333 0.033333 0.056667 0.2 0.1 0.133333	0.066667 0.066667	0.1 0.133333
0.008846 0.002602 0.001889 0.000696 0.000688 0.000156 0.000156 0.000288 0.000156 0.000288 0.000184 0.000519 0.000123 0.000173 0.000112 0.033333 0.066667 0.133333 0.166667 0.13333 0.16667 0.13333 0.166667 0.13333 0.166667 0.13333 0.166667 0.13333 0.166667 0.13333 0.166667 0.13333 0.16667 0.1300000000000000000000000000000000000	0.133333 0.066667	0.133333 0.033333

Fig. 12. Second MSE training loss and testing error data

depictions. It is likely that these correspond to the first descent in the double descent curve, but the goal of this project is to see the second descent.

Additionally, this training procedure was unable to perform well on more complex datasets. The following results were for a six dimensional version of the checkerboard pattern that was used in the first dataset used. Clearly the models were not able to properly learn the training set:

127.9154	131.5488	69.64371	103.2037	125.3332	127.8245	116.4511	126.4197	108.718	69.74059	124.2454	72.90729	120.2849	138.3962	109.3653
109.2816	124.4461	71.48212	71.58243	125.808	124.7457	72.48469	126.9462	106.6367	92.63097	117.7299	76.16283	125.0287	135.1856	124.4788
150	119.4094	69.66331	122.1788	125.5485	83.36371	71.19849	98.445	125.7109	128.0051	128.0725	130.4486	108.5827	71.58723	126.967
98.071	73.61922	125.3795	97.55465	118.0288	71.45038	130	127.8525	126.1949	97.11473	87.97562	69.79673	126.213	118.5651	116.4469
61.39035	74.02924	109.1682	140.0621	124.0177	109.4018	137.18	130	75.25803	130	130	69.7552	118.3641	125.1752	149.9998
45.35609	85.57074	71.46879	127.1025	111.3608	125.901	69.82338	87.57318	87.91803	130	126.4905	124.1869	100.5957	129.9996	110.1592
42.40588	131.8298	124.058	133.9104	150	109.3342	150	101.1798	130	137.4014	122.3393	130	77.32324	124.3991	150
25.47083	53.50272	78.0564	100.5045	74.69188	128.0629	124.8819	130	138.1833	125.3344	82.03442	140.6159	130	119.0138	69.7047
32.4928	61.11149	74.67214	138.4706	114.3113	119.1248	150	74.14803	110.4635	124.7018	149.9997	125.0132	110.9086	130	138.8961
20.00051	28.35609	73.35744	130	130	107.9361	94.00268	109.9414	118.9189	104.3056	125.8788	127.2926	127.1147	124.9437	130
16.95728	35.30117	65.49026	55.21368	128.0965	70.61017	130	77.24823	130	130	118.6398	125.3575	127.989	130	117.3406
8.733603	33.77789	66.55247	70.94862	130	103.1691	72.25344	128.0854	77.08361	130	82.50465	130	150	130	130
12.00119	2.000202	32.58204	28.87501	75.9685	84.15002	130	150	74.74846	125.2897	126.7858	130	130	127.108	150
16.00064	16.00048	82.62085	78.82547	89.3929	124.5791	126.6642	150	130	140.6909	130	136.732	149.9991	140.2022	130
10.00088	26.59324	2.00051	53.23404	27.14286	130	127.0055	150	90.70845	140.2873	117.5416	130	130.481	130	130
6.000769	4.000229	13.0001	57.03561	85.34608	130	130	76.68822	130	76.80489	150	150	130	130	130
18.00067	6.000423	11.42863	14.00005	130	130.6442	71.30543	77.58648	150	150	130	124.1986	150	125.1498	130
4.000596	0.000269	26.25734	48.25274	52.15155	130	112.1349	120.0193	130	150	150	135.6105	114.9886	130	130
6.000683	14.00028	15.83338	130	108.3878	64.93118	117.7945	124.2897	130	130	130	130	130	126.7054	130
2.000683	0.000299	14.00007	150	38.00016	130	150	123.6849	115.8442	150	130	125.3657	150	130	150

Fig. 13. Failure of MSE on six-dimensional dataset (training loss)

D. Learning Rate Schemes

The work of many others including those in [2] was sufficiently inspiring to alter the training of these models with a new, adaptive learning rate. This paper discusses more advanced learning rate procedures, but the one implemented for the next results is simply a decreasing function over time. It was, however, able to drastically outperform the same procedure with a fixed learning rate. On the same dataset as the last figure from the previous section, the new procedure was able to get the following results:

																													_
87.0524	69.87143	69.87142	69.87143	69.87143	69.87143	69.87143	69.87142	69.87142	69.87142	69.87142	69.87143	69.87142	69.87142	69.87142	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667
56.4355	62.2568	69.87142	69.87143	61.10219	69.87142	69.87142	69.87141	69.87143	69.87142	69.87142	69.87143	69.87143	69.87143	69.87142	0.5	0.433333	0.566667	0.566667	0.4	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667
51.7508	63.3232	69.87142	58.32394	58.89801	66.60143	69.87142	69.87142	69.87142	69.87142	69.87141	52.97375	69.87143	69.86948	69.87143	0.6	0.466667	0.566667	0.433333	0.466667	0.566667	0.566667	0.566667	0.566667	0.566667	0.566667	0.533333	0.566667	0.566667	0.566667
54.0127	62.35619	46.89604	56.16024	56.31348	46.09124	36.81346	69.87143	69.87143	49.51162	69.87143	69.87142	69.87142	69.87142	69.17458	0.4	0.333333	0.566667	0.5	0.4	0.5	0.666667	0.566667	0.566667	0.366667	0.566667	0.566667	0.566667	0.566667	0.566667
43.3511	25.25117	43.69295	36.89626	57.17667	33.52914	66.80278	52.01388	69.87142	23.94778	69.79311	45.30954	69.87142	33.64698	43.80955	0.4	0.466667	0.366667	0.566667	0.366667	0.5	0.533333	0.4	0.566667	0.4	0.566667	0.366667	0.566667	0.466667	0.366667
36.2367	41.84794	33,27941	24,8644	38.06306	35.88536	10.68841	32,42402	27.90122	14,48759	14.51172	18.34719	20,72269	18,28099	69.87143	0.4	0.433333	0.5	0.566667	0.566667	0.6	0.566667	0.433333	0.466667	0.533333	0.433333	0.533333	0.433333	0.433333	0.566667
39,3520	7 36.66323	34.60865	22,93343	25,62839	25,71811	26,79816	16.85193	5.975017	24.43569	32,6615	69.87142	3.377932	29.37963	14,92132	0.433333	0.466667	0.466667	0.466667	0.433333	0.533333	0.433333	0.433333	0.433333	0.5	0.533333	0.566667	0.466667	0.4	0.466667
	24.36995													6,993646	0.466667		0.433333	0.6			0.533333		0.533333	0.4	0.5	0.433333	0.3	0.566667	
	26.91365									10.11592				4.069237	0.433333	0.466667	0.366667	0.333333				0.333333		0.533333				0.466667	0.433333
	5 34.37317														0.4		0.4		0.4		0.566667		0.566667		0.433333			0.533333	
	26.88684															0.366667						0.433333			0.366667				0.5
																						0.533333							0.5
	18.40326															0.366667									0.466667				0.5
38.039	5 24.10466	16.72168	16.57013	0.937861	7.48546	5.847515	2.718788	2.120227	4.125656	0.259509	6.269586	0.182622	0.101502	3.08365	0.3	0.433333	0.4	0.433333	0.366667	0.333333	0.4	0.333333	0.533333	0.366667	0.3	0.3		0.566667	0.6
28.8476	5 20.40363	8.976542	8.035038	6.905793	4.593979	7.649153	4.219534	8.058475	2.141757	2.926131	2.143963	0.349294	2.296376	0.430167	0.433333	0.633333	0.366667	0.333333	0.3	0.433333	0.466667	0.633333	0.533333	0.533333	0.466667	0.3	0.566667	0.366667	0.566667
35.0985	5 18.48015	11.37814	3.405656	7.298343	3.077727	0.902759	2.687955	0.503541	6.324107	0.244199	5.729555	2.116266	0.077417	2.08474	0.466667	0.433333	0.333333	0.366667	0.466667	0.333333	0.533333	0.366667	0.266667	0.5	0.433333	0.366667	0.566667	0.433333	0.4
31.7145	3 21.31707	10.67456	7.713533	1.763252	4.402596	3.34258	6.031193	0.216093	0.367503	0.389205	0.107189	0.32144	2.100587	2.70665	0.433333	0.466667	0.366667	0.333333	0.366667	0.4	0.5	0.4	0.433333	0.3	0.466667	0.333333	0.5	0.466667	0.5
19.5543	5 18.91216	14.80972	7.244877	4.280286	3.418557	4.446892	3.428035	1.142042	2.521719	5.867924	0.123093	0.101236	1.488616	0.084686	0.333333	0.433333	0.466667	0.4	0.433333	0.366667	0.4	0.633333	0.466667	0.4	0.466667	0.466667	0.6	0.566667	0.466667
25,1464	21,60944	12,70972	4,944423	3.571242	2.910841	4.155557	0.571102	4.126261	3.17159	0.180565	0.225728	5,716596	2.092778	0.143522	0.366667	0.4	0.4	0.333333	0.466667	0.3	0.466667	0.5	0.5	0.466667	0.666667	0.433333	0.533333	0.633333	0.433333
28.381	9,653046	6.636828	4.553305	8.33456	14,79787	2.485515	3.073153	0.19345	0.226798	6.293823	0.451114	0.175518	2.045691	0.082613	0.466667	0.366667	0.3	0.4	0.633333		0.433333		0.5	0.533333	0.5	0.5	0.566667	0.633333	0.266667
	18.26925														0.333333		0.433333					0.433333			0.5			0.566667	
20,3032	10.20323	11002400	2.023033	0.042070	2.020010	2.0.2755	21077425	0.1.1004	012-04207	0.203421	2120024	0.100110	0.075205	21203230	0.3333333	0.5	0.400000	0.00007	0.0000000	0.400007	0.000007	0.4555555	0.000007	0.0000000	0.5	0.400000	0.0000000	0.000007	0.000007

Fig. 14. Adapted learning rate training losses and testing errors

This data is clearly much more demonstrative of any type of second descent and achieved quite good training results. It seems almost that the valley of green possibly corresponds to the first descent. Hopefully extending the picture just a little would be sufficient to see the second descent. The next attempt was exactly this, running the same procedure for larger numbers of hidden layers and larger hidden unit sizes on a slightly simpler dataset. The results for this expanded chart are below:

E. Discussion

There is quite a lot of information in these two rectangles, so let us take this time to have a brief discussion about some of their features. A very prominent feature is the dark bands across the tops of these rectangles. This corresponds to the fact that,

39.9417	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286
31.47692	33.19711	34.74286	34.71488	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286
14.98706	16.58744	24.95724	34.74286	5.981618	34.74286	49.91169	32.78478	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286	34.74286
2.282926	5.94706	18.69151	7.781635	23.15639	18.75115	34.74286	24.64753	18.46542	34.74286	33.94915	24.28305	26.68446	34.74286	34.74286	34.74286	34.74286	34.74262	34.74286
2.247073	1.593039	2.292514	3.803061	7.291332	5.63022	4.222106	21.03856	10.62554	4.506479	32.81013	10.03653	34.74286	15.50355	2.372525	34.74286	34.74286	7.677993	34.74286
2.296677	6.267552	1.496338	1.315136	0.154206	5.100445	5.728635	4.937567	17.76291	4.864257	7.476812	1.758925	4.186625	2.528789	4.010234	14.46314	6.369112	28.7979	3.801031
5.373813	7.192982	6.869622	0.185548	1.287167	0.154183	0.555633	4.267349	0.846209	28.11708	9.443418	0.205194	5.932609	7.068625	2.802536	4.557863	1.9718	10.43444	11.42129
2.049911	3.266126	4.809151	0.242345	11.44112	2.21134	7.375144	0.207576	0.356377	1.099559	10.3798	2.137265	0.164463	0.196435	0.07413	1.938145	0.489154	5.23939	7.564335
31.03781	1.237054	1.241145	7.422771	0.17964	0.580466	1.112597	2.10807	1.899871	0.081484	1.75126	0.112124	0.026047	2.013572	2.088425	4.120001	15.19349	30.21604	2.06243
0.908437	2.802142	0.51364	0.370287	1.264288	2.078053	0.255236	0.162839	2.248136	0.153993	0.29939	0.42735	0.314579	3.522712	5.133423	0.099321	3.900885	0.06662	8.410292
2.098577	0.684358	0.315468	0.208121	0.788801	1.076674	0.066442	2.055835	0.039216	2.026525	0.189585	0.11605	3.451273	1.573526	0.061823	0.014995	5.329741	2.124883	0.06361
1.189087	1.265671	0.359632	0.308572	0.420412	0.153586	0.418844	0.71978	12.15858	0.654248	0.331832	5.89201	0.134371	2.184283	0.100736	5.355388	0.037253	0.121158	0.033781
1.844171	0.678073	1.053086	0.435192	0.23183	0.138739	0.195829	0.434994	0.589932	2.064511	0.108165	0.06464	0.92434	0.059804	0.193243	2.020034	0.446719	0.063969	0.015934
4.209053	2.342682	2.119465	0.510783	0.556164	0.350045	1.55069	0.161677	0.205248	0.052741	0.218655	1.77401	0.165593	3.956553	0.048979	2.050043	0.152669	0.040389	0.021217
1.928752	1.558422	2.506073	0.114623	1.456679	4.540766	0.35756	2.220635	0.062739	0.120547	0.104166	1.172507	1.958015	0.134788	0.079202	0.259371	0.046829	0.132919	2.031734
5.438932	1.513699	0.673597	0.442767	0.23942	0.153993	0.277395	2.129183	0.449645	0.112207	0.087891	0.357475	0.106922	0.146741	0.091046	0.073679	1.533473	0.158596	2.056636
5.143455	1.745564	0.689625	1.203583	0.203316	2.501852	0.386905	2.096021	0.059385	0.341083	0.085737	1.067446	0.060121	0.049433	0.068514	0.017595	0.064978	0.123348	0.036471
2.259663	0.86723	0.690636	1.312608	0.217436	1.793147	0.179106	1.826355	0.085773	2.426672	0.085666	0.114254	0.027877	0.025784	0.037866	0.056284	0.434645	0.072378	0.198752
1.873895	1.030586	0.550946	0.338037	0.213402	0.194999	0.241752	1.116793	0.151986	0.325657	0.069592	0.063259	0.179947	0.076672	0.11008	0.04366	0.013054	0.036346	1.879442
2.877115	1.010666	1.145105	1.980436	0.811049	0.659032	0.385791	0.06621	0.111129	0.080797	0.094188	0.054086	0.263589	0.056903	3.495325	1.950264	0.036949	0.027648	0.204878
2.10755	0.942446	0.553375	0.422572	0.226875	0.201947	0.460034	0.093659	0.287971	0.103183	0.075673	0.0394	0.272913	0.040174	0.165893	0.026701	0.043485	0.10349	0.063559
1.447685	3.067272											0.052183					0.069103	
2.824826	1.510198	0.454975	1.261757	0.158276	0.266442	0.17051	2.141584	0.199813	0.12988	0.170626	0.197427	0.19284	0.228809	0.050354	0.048216	0.084863	0.051222	0.166556
0.000297				1.215613								8.093704					_	0.050921
				0.728974											0.065717			
												0.133403				0.030363	0.052676	0.041732
				3.623433								0.082268					0.087382	
												0.243858						
												0.127862					0.247154	
4.262071												0.194716			0.072057			0.022615
0.000306				0.602972											0.201279			
	3.388699			0.901613		0.176854								0.098562		0.421858		0.047861
				0.417983								0.074063						0.042546
		0.752876		0.244927								0.095671						0.072217
	1.145579			0.352588		_								0.042119			0.076972	
				0.524667								2.101204					0.058927	
		1.458859			0.1835		0.207468				0.07798	0.06911			0.035982			
				0.593048								0.07141			0.121584			
3.348598	5.014014	1.633944	0.81119	0.180278	0.249446	0.240782	0.197854	0.138951	0.2030/4	0.2388/1	0.079958	0.082339	0.10024	0.070546	0.132377	0.150415	0.042049	0.030054

Fig. 15. Larger model space - adapted learning rate training losses

												0.392857						
0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857
0.5	_	0.285714			0.392857	_					0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857
0.214286	0.321429	0.357143	0.178571	0.464286	0.464286	0.392857	0.535714	0.5	0.392857	0.392857	0.25	0.428571	0.392857	0.392857	0.392857	0.392857	0.392857	0.392857
0.178571	0.214286	0.25	0.428571	0.25	0.357143	0.178571	0.535714	0.321429	0.178571			0.392857						0.392857
0.25	0.321429	0.392857	0.357143	0.178571	0.357143	0.357143	0.285714	0.535714	0.357143	0.5	0.357143	0.321429	0.178571	0.392857	0.285714	0.464286	0.642857	0.392857
0.25	0.25	0.428571	0.25	0.214286	0.214286	0.321429	0.25	0.5	0.321429	0.392857	0.25	0.321429	0.357143	0.285714	0.392857	0.357143	0.25	0.464286
0.178571	0.214286	0.392857	0.321429	0.464286	0.25	0.25	0.321429	0.321429	0.357143	0.392857	0.214286	0.25	0.25	0.142857	0.285714	0.392857	0.535714	0.321429
0.5	0.178571	0.25	0.285714	0.25	0.285714	0.392857	0.214286	0.285714	0.392857	0.464286	0.5	0.25	0.392857	0.321429	0.285714	0.428571	0.464286	0.321429
0.214286	0.285714	0.214286	0.178571	0.357143	0.178571	0.285714	0.428571	0.25	0.535714	0.285714	0.571429	0.25	0.392857	0.25	0.428571	0.5	0.285714	0.392857
0.142857	0.142857	0.321429	0.142857	0.428571	0.285714	0.214286	0.214286	0.178571	0.5	0.392857	0.392857	0.357143	0.214286	0.5	0.214286	0.321429	0.642857	0.25
0.107143	0.178571	0.285714	0.25	0.357143	0.357143	0.357143	0.25	0.285714	0.285714	0.357143	0.5	0.285714	0.25	0.357143	0.321429	0.357143	0.178571	0.357143
0.178571	0.321429	0.285714	0.214286	0.321429	0.285714	0.285714	0.25	0.285714	0.357143	0.25	0.464286	0.464286	0.178571	0.25	0.321429	0.321429	0.357143	0.178571
0.285714	0.214286	0.285714	0.178571	0.214286	0.285714	0.357143	0.178571	0.392857	0.178571	0.321429	0.392857	0.321429	0.428571	0.357143	0.392857	0.357143	0.5	0.464286
0.142857	0.25	0.25	0.285714	0.214286	0.392857	0.5	0.25	0.285714	0.464286	0.321429	0.214286	0.357143	0.25	0.25	0.5	0.464286	0.428571	0.392857
0.214286	0.214286	0.178571	0.214286	0.214286	0.25	0.214286	0.321429	0.285714	0.464286	0.428571	0.464286	0.464286	0.321429	0.321429	0.321429	0.285714	0.25	0.107143
0.214286	0.107143	0.25	0.25	0.214286	0.285714	0.285714	0.142857	0.25	0.357143	0.285714	0.357143	0.321429	0.285714	0.25	0.428571	0.357143	0.357143	0.392857
0.178571	0.178571	0.178571	0.25	0.285714	0.178571	0.214286	0.428571	0.321429	0.25	0.285714	0.321429	0.321429	0.392857	0.321429	0.285714	0.25	0.357143	0.25
0.142857	0.178571	0.178571	0.214286	0.285714	0.214286	0.25	0.214286	0.214286	0.214286	0.178571	0.392857	0.25	0.214286	0.428571	0.25	0.142857	0.392857	0.285714
0.142857	0.285714	0.25	0.357143	0.25	0.321429	0.321429	0.321429	0.214286	0.357143	0.25	0.357143	0.535/14	0.357143	0.142857	0.25	0.25	0.357143	0.607143
								0.214286					0.357143		0.25			0.607143
0.107143		0.142857	0.321429	0.214286	0.357143		0.285714	0.357143	0.428571	0.178571	0.214286		0.25	0.25	0.321429	0.285714	0.5	0.285714
0.107143	0.214286	0.142857 0.142857	0.321429	0.214286 0.285714	0.357143	0.321429	0.285714 0.214286	0.357143	0.428571 0.214286	0.178571 0.392857	0.214286 0.285714	0.392857	0.25 0.428571	0.25 0.321429	0.321429 0.392857	0.285714 0.428571	0.5 0.321429	0.285714 0.285714
0.107143 0.178571 0.25	0.214286 0.142857 0.25	0.142857 0.142857 0.25	0.321429 0.285714 0.321429	0.214286 0.285714 0.285714	0.357143 0.25 0.25	0.321429 0.357143 0.25	0.285714 0.214286 0.25	0.357143 0.5	0.428571 0.214286 0.428571	0.178571 0.392857 0.178571	0.214286 0.285714 0.25	0.392857 0.357143	0.25 0.428571 0.392857	0.25 0.321429 0.214286	0.321429 0.392857 0.178571	0.285714 0.428571 0.428571	0.5 0.321429 0.214286	0.285714 0.285714 0.285714
0.107143 0.178571 0.25 0.071429	0.214286 0.142857 0.25 0.142857	0.142857 0.142857 0.25 0.214286	0.321429 0.285714 0.321429 0.178571	0.214286 0.285714 0.285714 0.142857	0.357143 0.25 0.25 0.392857	0.321429 0.357143 0.25 0.214286	0.285714 0.214286 0.25 0.107143	0.357143 0.5 0.142857	0.428571 0.214286 0.428571 0.178571	0.178571 0.392857 0.178571 0.178571	0.214286 0.285714 0.25 0.25	0.392857 0.357143 0.535714	0.25 0.428571 0.392857 0.285714	0.25 0.321429 0.214286 0.285714	0.321429 0.392857 0.178571 0.357143	0.285714 0.428571 0.428571 0.535714	0.5 0.321429 0.214286 0.357143	0.285714 0.285714 0.285714 0.357143
0.107143 0.178571 0.25 0.071429 0.071429	0.214286 0.142857 0.25 0.142857 0.178571	0.142857 0.142857 0.25 0.214286 0.178571	0.321429 0.285714 0.321429 0.178571 0.142857	0.214286 0.285714 0.285714 0.142857 0.178571	0.357143 0.25 0.25 0.392857 0.178571	0.321429 0.357143 0.25 0.214286	0.285714 0.214286 0.25 0.107143 0.321429	0.357143 0.5 0.142857 0.214286 0.142857	0.428571 0.214286 0.428571 0.178571 0.285714	0.178571 0.392857 0.178571 0.178571 0.25	0.214286 0.285714 0.25 0.25 0.321429	0.392857 0.357143 0.535714 0.392857	0.25 0.428571 0.392857 0.285714 0.285714	0.25 0.321429 0.214286 0.285714 0.178571	0.321429 0.392857 0.178571 0.357143 0.392857	0.285714 0.428571 0.428571 0.535714	0.5 0.321429 0.214286 0.357143 0.321429	0.285714 0.285714 0.285714 0.357143 0.321429
0.107143 0.178571 0.25 0.071429 0.071429 0.107143	0.214286 0.142857 0.25 0.142857 0.178571 0.107143	0.142857 0.142857 0.25 0.214286 0.178571 0.178571	0.321429 0.285714 0.321429 0.178571 0.142857 0.107143	0.214286 0.285714 0.285714 0.142857 0.178571 0.178571	0.357143 0.25 0.25 0.392857 0.178571 0.142857	0.321429 0.357143 0.25 0.214286 0.214286 0.071429	0.285714 0.214286 0.25 0.107143 0.321429 0.285714	0.357143 0.5 0.142857 0.214286 0.142857	0.428571 0.214286 0.428571 0.178571 0.285714 0.25	0.178571 0.392857 0.178571 0.178571 0.25 0.321429	0.214286 0.285714 0.25 0.25 0.321429 0.071429	0.392857 0.357143 0.535714 0.392857 0.214286	0.25 0.428571 0.392857 0.285714 0.285714 0.285714	0.25 0.321429 0.214286 0.285714 0.178571 0.285714	0.321429 0.392857 0.178571 0.357143 0.392857 0.25	0.285714 0.428571 0.428571 0.535714 0.428571 0.214286	0.5 0.321429 0.214286 0.357143 0.321429 0.107143	0.285714 0.285714 0.285714 0.357143 0.321429
0.107143 0.178571 0.25 0.071429 0.071429 0.107143	0.214286 0.142857 0.25 0.142857 0.178571 0.107143 0.071429	0.142857 0.142857 0.25 0.214286 0.178571 0.178571 0.107143	0.321429 0.285714 0.321429 0.178571 0.142857 0.107143 0.285714	0.214286 0.285714 0.285714 0.142857 0.178571 0.178571 0.214286	0.357143 0.25 0.25 0.392857 0.178571 0.142857 0.178571	0.321429 0.357143 0.25 0.214286 0.214286 0.214286 0.071429 0.357143	0.285714 0.214286 0.25 0.107143 0.321429 0.285714 0.214286	0.357143 0.5 0.142857 0.214286 0.142857 0.107143 0.107143	0.428571 0.214286 0.428571 0.178571 0.285714 0.25 0.214286	0.178571 0.392857 0.178571 0.178571 0.25 0.321429 0.25	0.214286 0.285714 0.25 0.25 0.321429 0.071429 0.214286	0.392857 0.357143 0.535714 0.392857 0.214286 0.321429	0.25 0.428571 0.392857 0.285714 0.285714 0.285714 0.178571	0.25 0.321429 0.214286 0.285714 0.178571 0.285714 0.285714	0.321429 0.392857 0.178571 0.357143 0.392857 0.25 0.214286	0.285714 0.428571 0.535714 0.428571 0.428571 0.214286 0.178571	0.5 0.321429 0.214286 0.357143 0.321429 0.107143 0.535714	0.285714 0.285714 0.357143 0.321429 0.321429 0.25
0.107143 0.178571 0.25 0.071429 0.071429 0.107143 0.142857 0.107143	0.214286 0.142857 0.25 0.142857 0.178571 0.107143 0.071429 0.25	0.142857 0.142857 0.25 0.214286 0.178571 0.178571 0.107143 0.214286	0.321429 0.285714 0.321429 0.178571 0.142857 0.107143 0.285714 0.142857	0.214286 0.285714 0.285714 0.142857 0.178571 0.178571 0.214286 0.107143	0.357143 0.25 0.25 0.392857 0.178571 0.142857 0.178571 0.142857	0.321429 0.357143 0.25 0.214286 0.214286 0.071429 0.357143 0.071429	0.285714 0.214286 0.25 0.107143 0.321429 0.285714 0.214286 0.357143	0.357143 0.5 0.142857 0.214286 0.142857 0.107143 0.178571 0.142857	0.428571 0.214286 0.428571 0.178571 0.285714 0.25 0.214286 0.285714	0.178571 0.392857 0.178571 0.178571 0.25 0.321429 0.25 0.285714	0.214286 0.285714 0.25 0.25 0.321429 0.071429 0.214286 0.428571	0.392857 0.357143 0.535714 0.392857 0.214286 0.321429 0.142857	0.25 0.428571 0.392857 0.285714 0.285714 0.285714 0.178571 0.214286	0.25 0.321429 0.214286 0.285714 0.178571 0.285714 0.285714 0.428571	0.321429 0.392857 0.178571 0.357143 0.392857 0.25 0.214286 0.357143	0.285714 0.428571 0.428571 0.535714 0.428571 0.214286 0.178571 0.321429	0.5 0.321429 0.214286 0.357143 0.321429 0.107143 0.535714 0.107143	0.285714 0.285714 0.285714 0.357143 0.321429 0.321429 0.25 0.285714
0.107143 0.178571 0.25 0.071429 0.071429 0.107143 0.142857 0.107143 0.214286	0.214286 0.142857 0.25 0.142857 0.178571 0.107143 0.071429 0.25	0.142857 0.142857 0.25 0.214286 0.178571 0.178571 0.107143 0.214286 0.178571	0.321429 0.285714 0.321429 0.178571 0.142857 0.107143 0.285714 0.142857 0.071429	0.214286 0.285714 0.285714 0.142857 0.178571 0.214286 0.107143 0.178571	0.357143 0.25 0.25 0.392857 0.178571 0.142857 0.142857 0.142857 0.214286	0.321429 0.357143 0.25 0.214286 0.214286 0.071429 0.357143 0.071429 0.285714	0.285714 0.214286 0.25 0.107143 0.321429 0.285714 0.214286 0.357143 0.178571	0.357143 0.5 0.142857 0.214286 0.142857 0.107143 0.178571 0.142857	0.428571 0.214286 0.428571 0.178571 0.285714 0.25 0.214286 0.285714 0.285714	0.178571 0.392857 0.178571 0.178571 0.25 0.321429 0.25 0.285714 0.107143	0.214286 0.285714 0.25 0.321429 0.071429 0.214286 0.428571 0.178571	0.392857 0.357143 0.535714 0.392857 0.214286 0.321429 0.142857 0.357143	0.25 0.428571 0.392857 0.285714 0.285714 0.285714 0.214286 0.285714	0.25 0.321429 0.214286 0.285714 0.178571 0.285714 0.285714 0.428571 0.178571	0.321429 0.392857 0.178571 0.357143 0.392857 0.25 0.214286 0.357143	0.285714 0.428571 0.535714 0.535714 0.214286 0.178571 0.321429 0.392857	0.5 0.321429 0.214286 0.357143 0.321429 0.107143 0.535714 0.107143	0.285714 0.285714 0.357143 0.321429 0.321429 0.25 0.285714 0.321429
0.107143 0.178571 0.25 0.071429 0.071429 0.107143 0.142857 0.107143 0.214286 0.142857	0.214286 0.142857 0.142857 0.142857 0.178571 0.107143 0.071429 0.25 0.107143 0.142857	0.142857 0.142857 0.25 0.214286 0.178571 0.178571 0.107143 0.214286 0.178571 0.214286	0.321429 0.285714 0.321429 0.178571 0.142857 0.107143 0.285714 0.142857 0.071429 0.25	0.214286 0.285714 0.285714 0.142857 0.178571 0.178571 0.214286 0.107143 0.178571 0.142857	0.357143 0.25 0.392857 0.178571 0.142857 0.142857 0.214286 0.214286	0.321429 0.357143 0.25 0.214286 0.214286 0.071429 0.357143 0.071429 0.285714	0.285714 0.214286 0.25 0.107143 0.321429 0.285714 0.214286 0.357143 0.178571 0.142857	0.357143 0.5 0.142857 0.214286 0.142857 0.142857 0.178571 0.142857 0.142857 0.142857	0.428571 0.214286 0.428571 0.178571 0.285714 0.25 0.214286 0.285714 0.107143 0.107143	0.178571 0.392857 0.178571 0.178571 0.25 0.321429 0.25 0.285714 0.107143 0.25	0.214286 0.285714 0.25 0.25 0.321429 0.071429 0.214286 0.428571 0.178571 0.142857	0.392857 0.357143 0.535714 0.392857 0.214286 0.321429 0.142857 0.357143 0.214286	0.25 0.428571 0.392857 0.285714 0.285714 0.285714 0.178571 0.214286 0.285714 0.392857	0.25 0.321429 0.214286 0.285714 0.178571 0.285714 0.285714 0.428571 0.178571 0.25	0.321429 0.392857 0.178571 0.357143 0.392857 0.25 0.214286 0.357143 0.428571 0.392857	0.285714 0.428571 0.535714 0.535714 0.214286 0.178571 0.321429 0.392857	0.5 0.321429 0.214286 0.357143 0.321429 0.107143 0.535714 0.107143 0.285714 0.357143	0.285714 0.285714 0.357143 0.321429 0.321429 0.225 0.285714 0.321429 0.214286
0.107143 0.178571 0.25 0.071429 0.071429 0.107143 0.142857 0.107143 0.214286 0.142857	0.214286 0.142857 0.142857 0.142857 0.178571 0.107143 0.071429 0.25 0.107143 0.142857 0.142857	0.142857 0.142857 0.25 0.214286 0.178571 0.178571 0.107143 0.214286 0.178571 0.214286 0.214286	0.321429 0.285714 0.321429 0.178571 0.142857 0.107143 0.285714 0.142857 0.071429 0.25 0.142857	0.214286 0.285714 0.285714 0.142857 0.178571 0.178571 0.214286 0.107143 0.178571 0.142857 0.178571	0.357143 0.25 0.392857 0.178571 0.142857 0.142857 0.214286 0.214286 0.285714	0.321429 0.357143 0.25 0.214286 0.214286 0.071429 0.357143 0.071429 0.285714 0.214286 0.142857	0.285714 0.214286 0.25 0.107143 0.321429 0.285714 0.214286 0.357143 0.178571 0.142857 0.392857	0.357143 0.5 0.142857 0.214286 0.142857 0.142857 0.178571 0.142857 0.142857 0.142857	0.428571 0.214286 0.428571 0.178571 0.285714 0.25 0.214286 0.285714 0.107143 0.107143 0.285714	0.178571 0.392857 0.178571 0.178571 0.25 0.321429 0.25 0.285714 0.107143 0.25 0.285714	0.214286 0.285714 0.25 0.25 0.321429 0.071429 0.214286 0.428571 0.178571 0.142857 0.357143	0.392857 0.357143 0.535714 0.392857 0.214286 0.321429 0.142857 0.357143 0.214286 0.214286	0.25 0.428571 0.392857 0.285714 0.285714 0.285714 0.178571 0.214286 0.285714 0.392857 0.321429	0.25 0.321429 0.214286 0.285714 0.285714 0.285714 0.285714 0.428571 0.178571 0.25 0.428571	0.321429 0.392857 0.178571 0.357143 0.392857 0.25 0.214286 0.357143 0.428571 0.392857 0.25	0.285714 0.428571 0.535714 0.428571 0.214286 0.178571 0.321429 0.392857 0.25 0.357143	0.5 0.321429 0.214286 0.357143 0.321429 0.107143 0.535714 0.107143 0.285714 0.357143	0.285714 0.285714 0.357143 0.321429 0.321429 0.225 0.285714 0.321429 0.214286 0.392857
0.107143 0.178571 0.25 0.071429 0.071429 0.107143 0.142857 0.107143 0.214286 0.142857 0.142857 0.142857	0.214286 0.142857 0.142857 0.142857 0.178571 0.107143 0.071429 0.25 0.107143 0.142857 0.142857	0.142857 0.142857 0.214286 0.178571 0.178571 0.178571 0.107143 0.214286 0.178571 0.214286 0.214286 0.214286	0.321429 0.285714 0.321429 0.178571 0.142857 0.107143 0.285714 0.242857 0.071429 0.25 0.142857 0.178571	0.214286 0.285714 0.285714 0.142857 0.178571 0.178571 0.214286 0.107143 0.178571 0.142857 0.178571	0.357143 0.25 0.392857 0.178571 0.142857 0.178571 0.142857 0.214286 0.214286 0.285714 0.214286	0.321429 0.357143 0.25 0.214286 0.214286 0.071429 0.357143 0.071429 0.285714 0.214286 0.142857 0.071429	0.285714 0.214286 0.25 0.107143 0.321429 0.285714 0.214286 0.357143 0.178571 0.142857 0.392857 0.321429	0.357143 0.5 0.142857 0.214286 0.142857 0.142857 0.178571 0.142857 0.142857 0.142857 0.25	0.428571 0.214286 0.428571 0.285714 0.25 0.214286 0.285714 0.107143 0.107143 0.285714 0.428571	0.178571 0.392857 0.178571 0.178571 0.25 0.321429 0.25 0.285714 0.107143 0.25 0.285714 0.25	0.214286 0.285714 0.25 0.321429 0.071429 0.214286 0.428571 0.178571 0.142857 0.357143 0.214286	0.392857 0.357143 0.535714 0.392857 0.214286 0.321429 0.142857 0.357143 0.214286 0.214286 0.214286 0.214286 0.357143	0.25 0.428571 0.392857 0.285714 0.285714 0.285714 0.178571 0.214286 0.285714 0.392857 0.321429	0.25 0.321429 0.214286 0.285714 0.285714 0.285714 0.285714 0.428571 0.178571 0.25 0.4285714	0.321429 0.392857 0.178571 0.357143 0.392857 0.25 0.214286 0.357143 0.428571 0.392857 0.25	0.285714 0.428571 0.535714 0.21428571 0.214286 0.178571 0.321429 0.392857 0.25 0.357143 0.25	0.5 0.321429 0.214286 0.357143 0.321429 0.107143 0.535714 0.285714 0.357143 0.214286 0.285714	0.285714 0.285714 0.285714 0.357143 0.321429 0.21429 0.25 0.285714 0.321429 0.214286 0.392857 0.285714
0.107143 0.178571 0.25 0.071429 0.071429 0.107143 0.142857 0.107143 0.214286 0.142857 0.142857 0.107143 0.178571	0.214286 0.142857 0.25 0.178571 0.107143 0.071429 0.25 0.107143 0.142857 0.142857 0.25	0.142857 0.142857 0.25 0.214286 0.178571 0.178571 0.107143 0.214286 0.214286 0.214286 0.107143 0.035714	0.321429 0.285714 0.321429 0.178571 0.142857 0.107143 0.285714 0.142857 0.071429 0.25 0.142857 0.178571 0.142857	0.214286 0.285714 0.285714 0.142857 0.178571 0.214286 0.107143 0.178571 0.178571 0.178571 0.178571 0.25	0.357143 0.25 0.25 0.392857 0.178571 0.142857 0.142857 0.214286 0.214286 0.285714 0.214286 0.25	0.321429 0.357143 0.25 0.214286 0.214286 0.071429 0.357143 0.071429 0.285714 0.214286 0.142857 0.071429 0.321429	0.285714 0.214286 0.25 0.107143 0.321429 0.285714 0.214286 0.357143 0.178571 0.142857 0.392857 0.321429 0.142857	0.357143 0.5 0.142857 0.214286 0.142857 0.107143 0.178571 0.142857 0.142857 0.142857 0.25 0.357143	0.428571 0.214286 0.428571 0.178571 0.285714 0.255 0.214286 0.285714 0.107143 0.107143 0.285714 0.428571	0.178571 0.392857 0.178571 0.255 0.321429 0.25 0.285714 0.107143 0.25 0.285714 0.25 0.255 0.214286	0.214286 0.285714 0.25 0.321429 0.071429 0.214286 0.428571 0.178571 0.142857 0.357143 0.214286	0.392857 0.357143 0.535714 0.392857 0.214286 0.321429 0.142857 0.357143 0.214286 0.214286 0.214286 0.214286 0.357143	0.25 0.428571 0.392857 0.285714 0.285714 0.285714 0.214286 0.285714 0.392857 0.321429 0.285714	0.25 0.321429 0.214286 0.285714 0.285714 0.285714 0.285714 0.428571 0.178571 0.25 0.428571 0.285714 0.255	0.321429 0.392857 0.178571 0.357143 0.392857 0.25 0.214286 0.357143 0.428571 0.392857 0.25 0.285714 0.285714	0.285714 0.428571 0.535714 0.214286 0.178571 0.321429 0.392857 0.25 0.357143 0.25 0.255714	0.5 0.321429 0.214286 0.357143 0.321429 0.107143 0.535714 0.357143 0.285714 0.357143 0.214286 0.285714 0.392857	0.285714 0.285714 0.285714 0.357143 0.321429 0.225 0.285714 0.321429 0.214286 0.392857 0.285714 0.285714
0.107143 0.178571 0.25 0.071429 0.071429 0.107143 0.142857 0.107143 0.214286 0.142857 0.142857 0.107143 0.178571 0.142857	0.214286 0.142857 0.25 0.142857 0.178571 0.107143 0.071429 0.25 0.107143 0.142857 0.142857 0.25 0.178571	0.142857 0.142857 0.25 0.214286 0.178571 0.178571 0.214286 0.214286 0.214286 0.214286 0.107143 0.035714	0.321429 0.285714 0.321429 0.178571 0.142857 0.107143 0.285714 0.142857 0.071429 0.25 0.142857 0.142857 0.25	0.214286 0.285714 0.285714 0.142857 0.178571 0.178571 0.107143 0.142857 0.142857 0.178571 0.128571 0.255 0.214286	0.357143 0.25 0.25 0.392857 0.178571 0.142857 0.142857 0.214286 0.214286 0.285714 0.214286 0.25	0.321429 0.357143 0.25 0.214286 0.214286 0.071429 0.357143 0.071429 0.285714 0.214286 0.142857 0.071429 0.321429	0.285714 0.214286 0.25 0.107143 0.321429 0.285714 0.214286 0.142857 0.392857 0.321429 0.321429 0.32143	0.357143 0.142857 0.214286 0.142857 0.107143 0.178571 0.142857 0.142857 0.142857 0.142857 0.142857 0.357143 0.321429 0.178571	0.428571 0.214286 0.428571 0.285714 0.285714 0.285714 0.285714 0.107143 0.107143 0.285714 0.428571 0.142857	0.178571 0.392857 0.178571 0.25 0.321429 0.25 0.285714 0.107143 0.25 0.285714 0.25 0.214286 0.25	0.214286 0.285714 0.25 0.321429 0.214286 0.428571 0.142857 0.142857 0.357143 0.214286 0.214286	0.392857 0.357143 0.535714 0.392857 0.214286 0.321429 0.142857 0.357143 0.214286 0.214286 0.214286 0.357143 0.25	0.25 0.428571 0.392857 0.285714 0.285714 0.285714 0.214286 0.285714 0.392857 0.321429 0.3221429 0.285714 0.285714 0.107143	0.25 0.321429 0.214286 0.285714 0.285714 0.285714 0.428571 0.178571 0.25 0.428571 0.255 0.285714	0.321429 0.392857 0.178571 0.357143 0.392857 0.25 0.214286 0.357143 0.392857 0.25 0.285714 0.285714 0.357143	0.285714 0.428571 0.335714 0.335714 0.214286 0.178571 0.321429 0.322429 0.322429 0.325143 0.25 0.255714 0.428571	0.5 0.321429 0.214286 0.357143 0.357143 0.35714 0.107143 0.285714 0.357143 0.357143 0.285714 0.392857 0.357143	0.285714 0.285714 0.357143 0.321429 0.321429 0.285714 0.321429 0.321429 0.242426 0.392857 0.285714 0.357143 0.392857
0.107143 0.178571 0.25 0.071429 0.107143 0.142857 0.107143 0.214286 0.142857 0.142857 0.107143 0.178571 0.142857	0.214286 0.142857 0.25 0.142857 0.178571 0.07143 0.07143 0.07143 0.107143 0.107143 0.142857 0.142857 0.128571 0.178571	0.142857 0.142857 0.25 0.214286 0.178571 0.178571 0.178571 0.214286 0.214286 0.214286 0.214286 0.214286 0.214285 0.214285 0.214285	0.321429 0.285714 0.321429 0.178571 0.142857 0.107143 0.285714 0.142857 0.071429 0.25 0.142857 0.178571 0.142857 0.25 0.107143	0.214286 0.285714 0.285714 0.178571 0.178571 0.178571 0.178571 0.178571 0.178571 0.178571 0.255 0.214286 0.178571	0.357143 0.25 0.392857 0.178571 0.142857 0.142857 0.214286 0.214286 0.214286 0.214286 0.214286 0.215714	0.321429 0.357143 0.25 0.214286 0.071429 0.357143 0.071429 0.214286 0.142857 0.071429 0.214286 0.321429 0.321429 0.321429	0.285714 0.214286 0.25 0.321429 0.285714 0.24286 0.357143 0.178571 0.322857 0.322429 0.322429 0.321429 0.3257143 0.357143	0.357143 0.142857 0.214286 0.142857 0.107143 0.178571 0.142857 0.142857 0.142857 0.142857 0.142857 0.357143 0.321429 0.178571	0.428571 0.214286 0.428571 0.285714 0.285714 0.285714 0.285714 0.107143 0.107143 0.285714 0.428571 0.142857 0.285714 0.178571	0.178571 0.392857 0.178571 0.178571 0.25 0.321429 0.25 0.285714 0.107143 0.25 0.285714 0.25 0.214286 0.25 0.285714	0.214286 0.285714 0.25 0.321429 0.214286 0.428571 0.178571 0.12857 0.357143 0.214286 0.214286 0.321429	0.392857 0.357143 0.535714 0.392857 0.214286 0.321429 0.42857 0.357143 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286	0.25 0.428571 0.392857 0.285714 0.285714 0.285714 0.285714 0.285714 0.392857 0.321429 0.3221429 0.325714 0.107143 0.357143	0.25 0.321429 0.214286 0.285714 0.285714 0.285714 0.428571 0.178571 0.25 0.428571 0.255 0.285714	0.321429 0.392857 0.178571 0.357143 0.392857 0.25 0.214286 0.3257143 0.428571 0.392857 0.25 0.285714 0.357143 0.357143	0.285714 0.428571 0.335714 0.335714 0.214286 0.178571 0.321429 0.322857 0.357143 0.25 0.255 0.357143 0.285714	0.5 0.321429 0.214286 0.357143 0.357143 0.357143 0.107143 0.285714 0.357143 0.214286 0.285714 0.392857 0.357143 0.357143	0.285714 0.285714 0.325714 0.321429 0.321429 0.225 0.285714 0.321429 0.214286 0.322428 0.325714 0.357143 0.357143
0.107143 0.178571 0.25 0.071429 0.071429 0.107143 0.142857 0.142857 0.142857 0.142857 0.107143 0.142857 0.107143 0.142857	0.214286 0.142857 0.25 0.142857 0.178571 0.107143 0.07143 0.107143 0.107143 0.142857 0.142857 0.128571 0.178571 0.178571	0.142857 0.142857 0.25 0.214286 0.178571 0.178571 0.178571 0.214286 0.214286 0.214286 0.214286 0.214286 0.214285 0.214285 0.214285	0.321429 0.285714 0.321429 0.178571 0.142857 0.285714 0.285714 0.242857 0.071429 0.25 0.142857 0.142857 0.142857 0.142857 0.25 0.107143 0.035714	0.214286 0.285714 0.285714 0.142857 0.178571 0.214286 0.107143 0.178571 0.178571 0.178571 0.178571 0.214286 0.214286	0.357143 0.25 0.392857 0.178571 0.142857 0.178571 0.142857 0.214286 0.214286 0.245714 0.214286 0.255 0.178571	0.321429 0.357143 0.25 0.214286 0.071429 0.357143 0.071429 0.357143 0.071429 0.321429 0.321429 0.321429 0.321429 0.321429 0.321429	0.285714 0.214286 0.25 0.25 0.285714 0.242429 0.242426 0.214286 0.357143 0.32657 0.321429 0.327143 0.321429	0.357143 0.142857 0.214286 0.142857 0.107143 0.178571 0.142857 0.142857 0.142857 0.357143 0.321429 0.278571 0.321429	0.428571 0.214286 0.428571 0.285714 0.214286 0.214286 0.214286 0.285714 0.407143 0.285714 0.428571 0.142857 0.285714	0.178571 0.392857 0.178571 0.178571 0.321429 0.25 0.285714 0.25 0.285714 0.214286 0.25 0.214286 0.25 0.285714 0.357143	0.214286 0.285714 0.25 0.22 0.021429 0.071429 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286	0.392857 0.357143 0.535714 0.392857 0.214286 0.321429 0.42857 0.357143 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286	0.25 0.428571 0.392857 0.285714 0.285714 0.285714 0.325714 0.321428 0.285714 0.321429 0.285714 0.285714 0.3251143 0.357143	0.25 0.321429 0.214286 0.285714 0.285714 0.285714 0.428571 0.178571 0.178571 0.25 0.428571 0.285714 0.25 0.285714 0.214286	0.321429 0.392857 0.178571 0.357143 0.357143 0.32857 0.25 0.214286 0.257143 0.392857 0.25 0.285714 0.3257143 0.285714 0.2857143	0.285714 0.428571 0.428571 0.335714 0.242860 0.321429 0.392857 0.25 0.357143 0.25 0.255714 0.285714 0.285714	0.5 0.321429 0.214286 0.357143 0.3271429 0.357144 0.357144 0.2857143 0.285714 0.357143 0.214286 0.357143 0.352714 0.352714 0.352714 0.352714	0.285714 0.285714 0.357143 0.321429 0.321429 0.285714 0.321429 0.214286 0.3224286 0.325714 0.357143 0.357143 0.357143
0.107143 0.178571 0.25 0.071429 0.071429 0.107143 0.142857 0.142857 0.142857 0.142857 0.142857 0.142857 0.142857	0.214286 0.142857 0.25 0.142857 0.178571 0.07142 0.07142 0.07142 0.07143 0.142857 0.142857 0.178571 0.178571 0.178571 0.178571	0.142857 0.142857 0.25 0.214286 0.178571 0.178571 0.214286 0.214286 0.214286 0.214286 0.035714 0.178571 0.178571 0.178571	0.321429 0.285714 0.321429 0.178571 0.142857 0.285714 0.285714 0.28574 0.142857 0.142857 0.142857 0.142857 0.142857 0.142857 0.142857 0.25	0.214286 0.285714 0.285714 0.142857 0.178571 0.214286 0.214286 0.178571 0.178571 0.178571 0.178571 0.214286 0.214286 0.214286 0.214284 0.255	0.357143 0.25 0.392857 0.178571 0.142857 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286 0.21571 0.178571 0.178571	0.321429 0.357143 0.25 0.214286 0.071429 0.357143 0.357143 0.357142 0.214286 0.142857 0.321429 0.321429 0.321429 0.321429 0.321429 0.321429	0.285714 0.214286 0.25 0.25 0.285714 0.242429 0.242426 0.214286 0.357143 0.32657 0.321429 0.327143 0.321429	0.357143 0.5 0.142857 0.142857 0.142857 0.142857 0.142857 0.142857 0.142857 0.142857 0.357143 0.321429 0.337143 0.321429 0.357143	0.428571 0.214286 0.428571 0.285714 0.24285 0.24286 0.24286 0.285714 0.107143 0.285714 0.428571 0.142857 0.285714 0.178571 0.178571	0.178571 0.392857 0.178571 0.278571 0.321429 0.25 0.285714 0.107143 0.25 0.285714 0.214286 0.25 0.24286 0.25 0.2485714 0.357143 0.357143	0.214286 0.285714 0.25 0.25 0.321429 0.071429 0.214286 0.214286 0.214286 0.214286 0.214286 0.214286 0.214284 0.214295 0.21429	0.392857 0.357143 0.335714 0.32857 0.214286 0.321429 0.3214286 0.214286 0.214286 0.214286 0.214286 0.214286 0.314286 0.314286 0.314286 0.314286 0.325 0.392857	0.25 0.428571 0.392857 0.285714 0.285714 0.285714 0.3285714 0.321429 0.285714 0.325714 0.325714 0.357143 0.321429 0.5	0.25 0.321429 0.214286 0.2485714 0.285714 0.285714 0.285714 0.285714 0.285714 0.285714 0.285714 0.285714 0.214286 0.214286	0.321429 0.392857 0.378571 0.357143 0.392857 0.25 0.214286 0.357143 0.3285714 0.3285714 0.3285714 0.3242486 0.214286 0.214286 0.214285	0.285714 0.428571 0.428571 0.335714 0.242857 0.321429 0.392857 0.25 0.357143 0.25 0.285714 0.428571 0.4285714	0.5 0.321429 0.214286 0.357143 0.327143 0.327144 0.357144 0.357143 0.285714 0.357143 0.214286 0.357143 0.357143 0.357143 0.357143 0.357143 0.357143 0.357143	0.285714 0.285714 0.357143 0.321429 0.321429 0.285714 0.321429 0.214286 0.3224286 0.325714 0.357143 0.357143 0.357143

Fig. 16. Larger model space - adapted learning rate testing errors

especially in the one hidden unit case, the model is not becoming more complex despite having more parameters being thrown at it. A chain of multiple single nodes is unable to do anything better than a linear separation boundary if its activation is monotone. Consequently, this should not be able to do any better in classifying a training set and the same sort of phenomena holds while the hidden layers are quite small. Generally, as we go down and right we are getting better training losses which is what we expect, but the function is not as smooth as it should be. Optimally, the loss would strictly decrease when any step was made to the right or down. This is clearly not the case in our empirical setting and indicates that our validation dataset will be similarly bumpy. In the testing errors, we see the darkest red spots are at the top left corner (with the worst model) and scattered slightly after the boundary of the 'learning cliff' we can see in the training losses.

Beyond this, there is a mixture of greens and oranges with no immediately obvious patterns. If we recall that this is a continuous function with some sort of noise added upon it, we can start to make out some patterns in this data. There is a green sort of streak coming from the middle left of the data, corresponding to some models which generalized particularly well. We can also see a sort of valley in the top right quadrant near the middle right. It seems this valley could correspond to the first descent of the double descent curve. To see this one has to keep in mind that complexity is not a linear function of layer number and hidden unit size. Consequently, the second descent will not map out a linear curve in this plane and it will be hard to determine exactly what the 'peaks' which correspond to the interpolation maximum will manifest itself as in this plane. Nonetheless, we do start to see the double descent curve on models which are only a few layers deep (from the green section in the bottom left)

IV. CONCLUSION

Overall, this project turned out to be more about how to tune a deep neural network than it did about discovering the double descent curve. This is slightly disappointing, however, the last results produced are clearly the beginning of uncovering the double descent curve in this deep neural network setting. It is likely that other methodologies to help train deep neural networks could be useful in furthering this study so long as they do not inhibit the network's learning by regularization techniques. An example of an unhelpful technique is penalizing the network's weights. The double descent paradigm intends to achieve zero training loss which isn't feasible with a regularizing term on the size of the network's weights. Other schemes like an adapted learning rate which is able to get even lower scores on the training loss function is an example of a technique which would be useful.

As this project was limited by computational resources, it may also be true that for real applications as well, it becomes too difficult to actually go beyond the point of interpolation. Hence, these models would want to fit themselves snugly within the first descent. Ultimately, the double descent curve is a very promising outlook which seems like it will generalize to deeper network models. Unfortunately, the details are not worked out today by this project, but this should only inspire further work on the fascinating topic.

REFERENCES

- [1] Mikhail Belkin et al. "Reconciling modern machine learning and the bias-variance trade-off". In: *arXiv preprint arXiv:1812.11118* (2018).
- [2] Noam Shazeer and Mitchell Stern. "Adafactor: Adaptive Learning Rates with Sublinear Memory Cost". In: arXiv preprint arXiv:1804.04235 (2018).