1 Chaos in random recurrent neural networks

At a high level, this project is about understanding the transition to chaos in the dynamical system that arises from a randomly-initialized recurrent neural network. Assuming for ease of analysis that all the weights of the recurrent neural network are sampled independently from a zero-mean Gaussian, we can take the thermodynamic limit as the number of neurons $N \to \infty$ and observe a critical phase transition into chaos. As it turns out, first studying the discrete dynamical system formed by the RNN gives results that resemble that of the continuous case quite well.

The interesting about this line of inquiry is that it resembles the "deep information propagation" work in arXiv:1611.01232 and arXiv:1711.00165, where layers of wide, randomly-initialized neurons are iteratively passed to the next layer, and the output is analyzed for correlation/criticality to the input. However, viewing this as a continuous-time dynamical system allows you to generalize beyond the discrete layers of a neural network to continuous layers, hence the connections to the recent work on Neural ODEs as well. In particular, studies of how criticality and chaos can occur in these systems can inform how best to train these systems (even outside of the random Gaussian weight regime).

Without the background in some of this previous work, the notion of connecting a neuron's output to itself may be somewhat confusing. It may be helpful to add a diagram comparing an unrolled RNN with discrete layers to the $dt = 1$ case, and an infinite-layer unrolled RNN to the continuous-time case. In addition, it would help to illustrate a fully-connected neural network, since my assumption is that every postsynaptic neuron h_i is connected to every presynaptic neuron h_j, and moreover each set is the same size. In the discrete case, the only difference is that the postsynaptic neuron values are evaluated at time $t+1$, hence $h_i^{(t+1)}$ whereas the presynaptic neuron values are evaluated at time t, hence $h_j^{(t)}$. Generalizing to the continuous case, we let t range continuously and define the dynamical system in the same way as in the discrete case.

I'm personally curious about the connections between this work and some of the "deep information propagation" work. This connection instantly brings up some possible extensions:

- Consider extending the analysis beyond the tanh function, if you can – try some other common activation functions like ReLU.
- Explicitly analyze correlations and see what kind of Gaussian process results in the $N \to \infty$ limit.
- What happens if we initialize a neural network at criticality (edge of chaos) and train it, vs. if we initialize it in the stable regime v. the chaotic regime?
- How can we add in the effect of randomly-sampled biases? Perhaps the deep information propagation work can illuminate that.

Otherwise, the draft looks good to me. The graphics (particularly Figures 3 and 4) are clear and well-made. I would add some additional background for the reader unfamiliar with RNNs on how to go from discrete fully-connected networks to RNNs to continuous-time dynamical systems defined by a RNN. Moreover, I would add some discussion about the thermodynamic limit and why we specifically choose the weights to be drawn from a Gaussian with variance scaling inversely as the number of neurons (law of large numbers?)