SURYA GANGULI, Stanford
Statistical mechanics of deep learning

We will discuss several recent results in which ideas from statistical mechanics can yield both conceptual insights into the operation of deep networks as well as deliver better machine learning performance. These include: (1) Connecting the expressivity of deep networks to an order chaos transition in dynamic mean field theory, (2) Developing dynamically critical initial conditions for network weights that accelerate subsequent learning, (3) Creating generative models of complex data distributions by training neural networks to reverse the flow of time in entropy producing diffusion processes, and (4) Characterizing the non-convex geometry of deep network error landscapes using inspiration from the statistical mechanics of random landscapes.

References:

B. Poole, S. Lahiri, M. Raghu, J. Sohl-Dickstein, and S. Ganguli, Exponential expressivity in deep neural networks through transient chaos, Neural Information Processing Systems (NIPS) 2016.

J. Pennington, S. Schloenholz, and S. Ganguli, Resurrecting the sigmoid in deep learning through dynamical isometry: theory and practice, Neural Information Processing Systems (NIPS) 2017.

J. Sohl-Dickstein, N. Maheswaranathan, E. Weiss, S. Ganguli, Deep unsupervised learning using non-equilibrium thermodynamics, International Conference on Machine Learning (ICML) 2015.

Y. Dauphin, R. Pascanu, C. Gulcehre, K. Cho, S. Ganguli, Y. Bengio, Identifying and attacking the saddle point problem in high-dimensional non-convex optimization, Neural Information Processing Systems (NIPS), 2014.

M. Advani and S. Ganguli, Statistical mechanics of optimal convex inference in high dimensions, Physical Review X, 6, 031034, 2016.