

Physical Math Seminar

Statistical mechanics of learning and optimization in neural networks



SURYA GANGULI
(Stanford University)

ABSTRACT:

Statistical mechanics and neural network theory have long enjoyed fruitful interactions. We will review some of our recent work in this area and then focus on two vignettes. First, we will analyze the high dimensional geometry of neural network error landscapes that happen to arise as the classical limit of a dissipative many-body quantum optimizer. We will use the Kac-Rice formula and the replica method to calculate the number, location, energy levels, and Hessian eigenspectra of all critical points of any index and find optimal annealing schedules for optimization. Second, we will reveal a new implicit bias of stochastic gradient descent (SGD) that forces highly overparameterized neural networks to converge to saddle points, not local minima. Moreover, these saddle points: (1) correspond to very simple networks with far fewer effective degrees of freedom than their parameter count, and (2) possess superior generalization performance relative to local minima. Intriguingly, this implicit bias arises through position dependent diffusion terms in accurate Langevin dynamics models of SGD that cause freezing at these simple, high performing saddle points.

Bio:

Surya Ganguli triple majored in physics, mathematics, and EECS at MIT, completed a PhD in string theory at Berkeley, and a postdoc in theoretical neuroscience at UCSF. He is now an associate professor of Applied physics at Stanford where he leads the Neural Dynamics and Computation Lab. He has also been a visiting researcher at both Google and Meta AI and is currently a Venture Partner at a16z. His research spans the fields of neuroscience, machine learning and physics, focusing on understanding and improving how both biological and artificial neural networks learn striking emergent computations. He has been awarded a Swartz-Fellowship in computational neuroscience, a Burroughs-Wellcome Career Award, a Terman Award, two NeurIPS Outstanding Paper Awards, a Sloan fellowship, a James S. McDonnell Foundation scholar award in human cognition, a McKnight Scholar award in Neuroscience, a Simons Investigator Award in the mathematical modeling of living systems, an NSF career award, and a Schmidt Science Polymath Award.

TUESDAY, MARCH 19, 2024

2:30 PM – 3:30 PM

Building 4, Room 149

NOTE: New Location

<https://math.mit.edu/pms/>