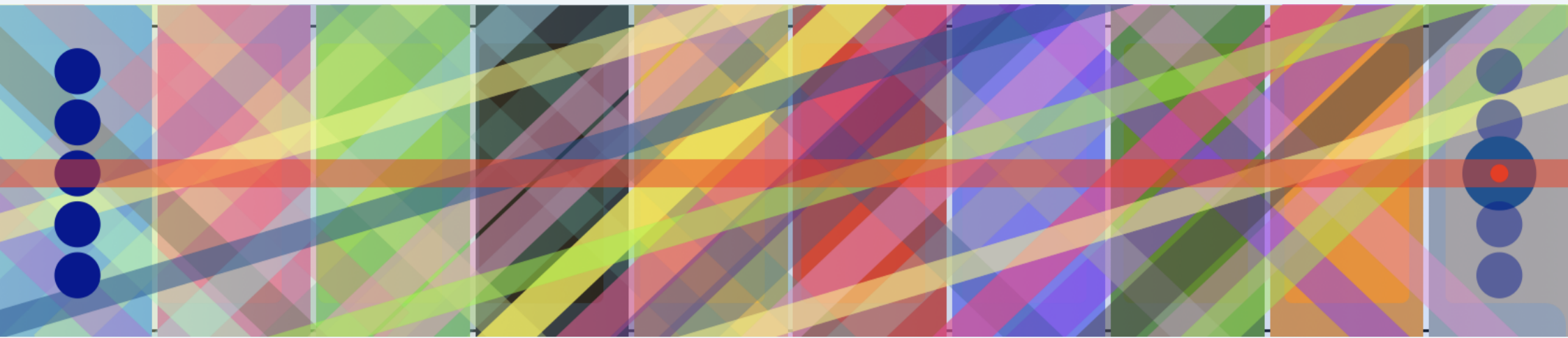


# Theoretical physics courses



## Mathematical problems in Machine Learning

*Andrea Montanari (Stanford University)*

*Mon 19th, Tue 20th, Wed 21st, Fri 23rd of June from 10:00 to 12:30. In person at IPhT and online.*

Despite their empirical success, the principles underlying modern deep learning models remain mysterious. These models are trained by optimizing highly non-convex objectives, using a variety of gradient-based algorithms that -at best- are only guaranteed to converge to local optima. The model complexity is often large or comparable to the sample size, and hence many choices of the model parameters exist that perform equally well on the training data, but not all of them generalize to unseen data. Over the last few years, an informal scenario has emerged that captures these phenomena. I will describe its elements and explain a few examples in which this scenario can be made more precise.

A rough outline of the lectures:

1. Why does overparametrization help optimization?
2. Why doesn't overparametrization hurt generalization?
3. The puzzle of architecture.
4. Generative modeling.

To receive the latest news on this course and the video-conference links, please subscribe to its newsletter, as explained at the website [courses.ipht.fr](https://courses.ipht.fr). An open, non-interactive livestream will also be available at [youtube.com/ipht-tv](https://youtube.com/ipht-tv).