



Mathematisches Institut, Sidlerstrasse 5, CH-3012 Bern

^b
**UNIVERSITÄT
BERN**

Philosophisch-
naturwissenschaftliche Fakultät
Departement Mathematik und Statistik
Mathematisches Institut

Mathematical Colloquia

Monday, 23 March 2020

17:15 h, Lecture Room 119

Prof. Dr. Caroline Uhler, ETH Zurich

Multi-Domain Data Integration: From Observations to Mechanistic Insights

Abstract:

Massive data collection holds the promise of a better understanding of complex phenomena and ultimately, of better decisions. An exciting opportunity in this regard stems from the growing availability of perturbation / intervention data (manufacturing, advertisement, education, genomics, etc.). In order to obtain mechanistic insights from such data, a major challenge is the integration of different data modalities (video, audio, interventional, observational, etc.). Using genomics as an example, I will first discuss our recent work on coupling autoencoders in the latent space to integrate and translate between data of very different modalities such as sequencing and imaging. I will then present a framework for integrating observational and interventional data for causal structure discovery and characterize the causal relationships that are identifiable from such data. We end by a theoretical analysis of autoencoders linking overparameterization to memorization. In particular, I will characterize the implicit bias of overparameterized autoencoders and show that such networks trained using standard optimization methods implement associative memory. Collectively, our results have major implications for planning and learning from interventions in various application domains.