20th

# SHANNON MEMORIAL LECTURE

Friday, December 5, 2025

11 AM: Lecture
12 PM: Reception

Center for Memory and Recording Research (CMRR)
Jack Keil Wolf Auditorium
University of California, San Diego

## Rüdiger Urbanke

Rüdiger L. Urbanke is a Professor at the School of Computer and Communication Sciences (I&C) at EPFL, where he is a member of the Information Processing Group. His research spans the design and analysis of error-correcting codes for both classical and quantum communication, as well as the theoretical foundations of modern machine learning. He received his Dipl. Ing. from the Vienna University of Technology in 1990, followed by an M.Sc. and Ph.D. in Electrical Engineering from Washington University in St. Louis in 1992 and 1995, respectively. After working in the Mathematics of Communications Department at Bell Labs (1995–1999), he joined EPFL. He served as Associate Editor for the *IEEE Transactions on Information Theory* (2000–2004), President of the IEEE Information Theory Society (2017), and Dean of I&C (2021–2025). Dr. Urbanke is co-author of *Modern Coding Theory* (Cambridge University Press) and a recipient of the 2002, 2013, and 2021 IEEE Information Theory Society Best Paper Awards, the 2016 STOC Best Paper Award, the 2011 IEEE Koji Kobayashi Computers and Communications Award, the 2014 IEEE Richard W. Hamming Medal, and the 2023 Claude E. Shannon Award.

CMRR Center for Memory and Recording Research

QUALCOMM INSTITUTE

Information Theory and Applications Center

*Information-Theoretic Framework for Understanding Learning in Modern Machine-Learning Architectures*


### ABSTRACT

I will present an information-theoretic framework for understanding learning in modern machine-learning architectures. The framework treats learning as universal prediction under log loss, with performance characterized by regret bounds. A central concept is architecture-dependent model complexity, which is defined through the probability volume of models near the data-generating process. This volume can be approximated by spectral properties of the Hessian or Fisher Information Matrix. I will discuss how successful learning architectures balance effective complexity, enabling them to adaptively fit a wide range of functions, even in over-parameterized settings. This perspective provides insights into the roles of inductive biases, stochastic gradient descent, and the nature of flat minima, while also offering a unified view of online and batch learning, as well as supervised and generative modeling.


*Joint work with Meir Feder and Yaniv Fogel and with the help of Ido Atlas, all Tel Aviv University.*

UC San Diego