University of Cologne
Institute of Mathematics
Prof. Dr. Alexander Drewitz

e-mail: drewitz@math.uni-koeln.de
office: Mathematics 205

WS 2021:
(High-dimensional) probability and applications to data science

We will investigate certain topics in mostly high-dimensional probability theory. While our focus will be on the mathematical point of view, such techniques also play a fundamental role in applications to topics such as data science or statistical mechanics.

An emblematic example is the Johnson-Lindenstrauss lemma (see e.g. Section 2.9 in [BLM13] or Section 5.3 in [Ver18]): It states that for N points in a high-dimensional vector space one can find an orthogonal projection into a lower-dimensional space of dimension approximately $\log N$ which only slightly distorts distances between points. Thus, such a projection preserves the geometry and at the same time reduces the complexity of the problem.

Talks will be based on [BSS20] which can be downloaded at

`https://people.math.ethz.ch/~abandeira/BandeiraSingerStrohmer-MDS-draft.pdf`.

The seminar is aimed at BSc and MSc students. Participants are expected to have mastered the lectures 'Einführung in die Stochastik' as well as either 'Wahrscheinlichkeitstheorie I' or 'Wahrscheinlichkeitstheorie II'. In order to obtain the corresponding credit points, participants have to give a presentation on one of the available topics and actively contribute to the discussions of the remaining presentations.

Presentations can be given in English or German. At
`http://www.alt.mathematik.uni-mainz.de/Members/lehn/le/seminarvortrag`
you can find some advice on how to prepare a valuable seminar talk which you should take serious.

A preliminary meeting will take place on Wednesday, July 14, at 3:00 p.m. on zoom via the link
   `https://uni-koeln.zoom.us/j/95459748705?pwd=YnFKWisvenhHWVhEdFluNkd4NU80Zz09`
Students who intend to participate in the seminar are asked to notify the lecturer via email (see above) by the July 21, 2021, including

1. matriculation number,

2. relevant lectures attended and grades obtained.

The coordinates of the regular meetings are:

Room:	Seminarraum 1
Day & time:	Thursdays, 10:00 to 11:30 a.m.

References

[BLM13] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities.* Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.

[BSS20] A Bandeira, A Singer, and T Strohmer. Mathematics of data science. *Book draft available at https://people. math. ethz. ch/abandeira/BandeiraSingerStrohmer-MDS-draft. pdf,* 2020.

[Ver18] Roman Vershynin. *High-dimensional probability,* volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, Cambridge, 2018. Draft available at `https://www.math.uci.edu/~rvershyn/papers/HDP-book/HDP-book.html`.