E C E/MATH/STAT 888 Syllabus

Topics in Mathematical Data Science
"High-Dimensional Probability and Statistics"

COURSE INFORMATION

Topics in Mathematical Data Science
"High-Dimensional Probability and Statistics"
E C E/MATH/STAT 888 (3 Credits)
Fall 2021-2022

Official Course Description: Advanced topics in the mathematical foundations of data science.

This topic: In *Fall 2021*, this course will provide a rigorous, self-contained introduction to the area of high-dimensional probability and statistics from a non-asymptotic perspective, aimed at graduate students in mathematics, statistics, computer science and engineering. It will include a focus on core methodology and theory (tail bounds, concentration of measure, uniform laws, random matrices) as well as in-depth exploration of particular model classes (sparse linear models, graphical models, community detection). No statistics or programming background will be assumed. Prior exposure to graduate-level probability (e.g., MATH 733 or ECE 730 or STAT 709) is highly recommended.

Requisites: Graduate/professional standing or member of the Pre-Masters Mathematics (Visiting International) Program.

Instruction Mode: Classroom Instruction
Attributes and Designations: Graduate Attribute

Departments:	MATHEMATICS / STATISTICS / ELECTRICAL & COMPUTER ENGINEERING
College:	Letters and Science / Engineering
Canvas URL:	https://canvas.wisc.edu
Course website:	https://people.math.wisc.edu/~roch/hdps/

Location and Schedule: MoWeFr 2:25PM - 3:15PM in VAN VLECK B239

How the Credit Hours are Met
This class meets for three 50-minute class periods each week over the semester and carries the expectation that students will work on course learning activities (reading, writing, problem sets, studying, etc) for about 2 hours out of classroom for every class period. The syllabus includes more information about meeting times and expectations for student work.

Instructor:	Prof. Sebastien Roch	Use Canvas for all communications, office hours held on Zoom
	Office hours	TBA

GRADING AND COURSE MATERIALS

Grading: Assignments (100%)

Homework assignments (33%): One or two homework assignments. To be submitted on Canvas. Graded for completion and one or two problems graded for accuracy. Homework is meant to be challenging and to push you to learn the material in depth. Do not expect to know what to do immediately. Try hard enough to solve the problems on your own using only the course notes and your own lecture notes. You can discuss the material with others and consult other sources but *you must specify your collaborators and sources on each homework*. Every student must submit their own written version of the homework solutions. *No late submissions.*

Scribing assignment (33%): Each student is expected to scribe one lecture. A sign-up sheet and template will be provided on Piazza. Submissions must be via BOX and include a compiled PDF, the LaTeX source, and whatever figures are needed. Please give real bibliographical citations for the papers that we mention in class. A first draft of the scribe notes is due 72 hours after the lecture (at the latest). They will be posted on the course website.

Short report assignment (33%): This assignment will consists in picking a talk of your choice from a list of virtual workshops related to the course that will be posted on Canvas. Write a 3-5 page summary of one of the main results in the talk. You do not need to provide a full proof, but you should define concepts carefully and state the main result formally. You should highlight at least one key technical lemma, ideally with a proof or proof sketch. Make sure to include appropriate bibliographic references. Submit your summary on or before the last day of classes

Exams: No exams.

Topics

The following topics will be covered in this course (time permitting):

1. BOOTCAMP ON STATISTICAL THEORY FOR MATHEMATICIANS. Probability models. Point estimation. Hypothesis testing. Introduction to R.
2. NON-ASYMPTOTIC METHODS IN HIGH-DIMENSIONAL PROBABILITY. Basic tail bounds. Concentration of measure. Metric entropy. Applications to random matrices and random graphs.
3. SELECTED APPLICATIONS IN HIGH-DIMENSIONAL STATISTICS. Covariance estimation. Principal component analysis in high dimensions. Community detection. Sparse linear models in high dimensions. Graphical models for high-dimensional data.
4. LOWER BOUNDS. Basic background in information theory. Minimax risk. Le Cam's and Fano's methods.

Required textbook, software and other course materials

Course material will be taken from the following required textbooks:

[Pa] Panaretos, V. M. (2016). Statistics for Mathematicians: A Rigorous First Course. Germany: Springer International Publishing.
(e-copy: https://search.library.wisc.edu/catalog/9912241369602121)

[Ve] Vershynin, R. (2018). High-Dimensional Probability: An Introduction with Applications in Data Science. United Kingdom: Cambridge University Press.
(e-copy: https://www.math.uci.edu/~rvershyn/papers/HDP-book/HDP-book.html)

[Wa] Wainwright, M. J. (2019). High-Dimensional Statistics: A Non-Asymptotic Viewpoint. United Kingdom: Cambridge University Press.

[Wo] Wood, S. N. (2015). Core Statistics. United Kingdom: Cambridge University Press.
(e-copy: https://www.maths.ed.ac.uk/~swood34/core-statistics.pdf)

Tentative schedule

A detailed plan is available on the course website: https://people.math.wisc.edu/~roch/hdps/

Course Learning Outcomes

Students will be able to:

1. Apply advanced mathematical concepts to solve a variety of data science problems.
2. Analyze rigorously the mathematical properties of methods used in data science.

ACADEMIC POLICIES

COVID

Students of the class are expected to comply with the University's current COVID rules and policies that are maintained here: https://covidresponse.wisc.edu (see in particular https://covidresponse.wisc.edu/faq/).
Students who do not comply with these rules can be asked to leave the classroom, and students who repeatedly fail to comply will be referred to the Office of Student Conduct and Community Standards. Any student who requires an exemption to current policies must contact the McBurney Office, as instructors do not have the authority to grant such exceptions.

ACADEMIC INTEGRITY

By enrolling in this course, each student assumes the responsibilities of an active participant in UW-Madison's community of scholars in which everyone's academic work and behavior are held to the highest academic integrity standards. Academic misconduct compromises the integrity of the university. Cheating, fabrication, plagiarism, unauthorized collaboration, and helping others commit these acts are examples of academic misconduct, which can result in disciplinary action. This includes but is not limited to failure on the assignment/course, disciplinary probation, or suspension. Substantial or repeated cases of misconduct will be forwarded to the Office of Student Conduct & Community Standards for additional review. For more information, refer to https://conduct.students.wisc.edu/academic-integrity/.

ACCOMMODATIONS FOR STUDENTS WITH DISABILITIES

McBurney Disability Resource Center syllabus statement: "The University of Wisconsin-Madison supports the right of all enrolled students to a full and equal educational opportunity. The Americans with Disabilities Act (ADA), Wisconsin State Statute (36.12), and UW-Madison policy (Faculty Document 1071) require that students with disabilities be reasonably accommodated in instruction and campus life. Reasonable accommodations for students with disabilities is a shared faculty and student responsibility. Students are expected to inform faculty [me] of their need for instructional accommodations by the end of the third week of the semester, or as soon as possible after a disability has been incurred or recognized. Faculty [I], will work either directly with the student [you] or in coordination with the McBurney Center to identify and provide reasonable instructional accommodations. Disability information, including instructional accommodations as part of a student's educational record, is confidential and protected under FERPA." http://mcburney.wisc.edu/facstaffother/faculty/syllabus.php.

DIVERSITY & INCLUSION

Institutional statement on diversity: "Diversity is a source of strength, creativity, and innovation for UW-Madison. We value the contributions of each person and respect the profound ways their identity, culture, background, experience, status, abilities, and opinion enrich the university community. We commit ourselves to the pursuit of excellence in teaching, research, outreach, and diversity as inextricably linked goals.

The University of Wisconsin-Madison fulfills its public mission by creating a welcoming and inclusive community for people from every background – people who as students, faculty, and staff serve Wisconsin and the world." https://diversity.wisc.edu/