## A Brief Course Description for ACM 270-1

# Special Topics on Data-Driven Multiscale Model Reduction

**Time:** 10:30AM to 11:50AM, Wed and Friday, Winter Quarter, 2018

**Place**: 314 Annenberg

**Instructor**: Thomas Y. Hou, email: hou@cms.caltech.edu

**Course Description**: In this seminar course, I will cover several recent research developments on multiscale modeling and computation for multscale partial different equations, stochastic differential equations with random and multiscale coefficients, nonlinear multiscale time series, and discrete networks and graphs. A common theme in these problems is to find data-driven problem-dependent multoscale basis functions to represent the solutions of these problems, which typically live in a very high dimensional solutions space. By constructing a data-driven multiscale basis, we can compress the solution space effectively. Once the multiscale basis is constructed in the offline stage, one can solve the same problem using the multiscale basis in a much lower dimensional solution space in the online stage for different forcing functions or boundary conditions, which can lead to considerable computational saving.

The course will emphasize the interplay between physical modeling and information science. Recently developed tools in information science such as optimization, compressed sensing, sparsity and low rank approximation will be used to construct low dimensional multiscale basis functions to solve complex high dimensional physical problems with multiscale features. Conversely, recent developments in multiscale analysis and modeling will be used to extract important intrinsic physical patterns and features from massive data that we collect from complex physical or biological systems. By integrating these two research areas, we can explore new exciting research opportunities that lie in the intersection of these two disciplines.

The first half of the class will focus on multiscale analysis and computation for multiscale partial differential equations (PDEs), which include linear and nonlinear elliptic PDEs, parabolic PDEs, and hyperbolic PDEs, multiphase flows in heterogeneous porous media with applications in geophysical and fluid flows. We begin by reviewing the classical homogenization theory and numerical homogenization, and then move to various multiscale methods that do not assume scale separation and periodic structures, and finish by reviewing some recent development of multiscale methods that are guaranteed to converge with almost optimal rate of convergence for multiscale problems with continuous spectrum of scales. Applications of multiscale analysis to develop efficient solvers for graph Laplacians and discrete network will be also discussed.

The second half of the class is devoted to studying some recent developments of adaptive data analysis for nonlinear multiscale time series. The aim is to extract some intrinsic physical patterns such as instantaneous frequency from a nonlinear time series. This can be considered as a data-driven nonlinear Fourier Transform with nonlinear phase functions extracted from the data. This gives rise to an adaptive time-frequency representation by exploiting the intrinsic sparsity structure of the time series based on a data-driven multiscale basis. Two complementary methods will be discussed. The first one is formulated as a nonlinear optimization problem by using nonlinear matching pursuit or nonlinear basis pursuit. This can be considered as a special case of nonlinear compressed sensing where one needs to identify the data-driven basis under which we have a sparse representation of the signal. The second approach is based on the Synchrosqueezed wavelets based on continuous Wavelet Transform . Applications to various physical and biomedical data will be presented. Of particular interest is the recent study of embedding a one-dimensional time series into a high-dimensional space and study their intrinsic geometric and topological structure. This idea has found some promising applications to biomedical data. Finally, we will discuss some recent attempts to use Machine Learning and stochastic approximation methods to extract the data-driven basis and to solve the nonlinear nonconvex optimization problems.

It is likely that we cannot finish all the materials in the Winter Quarter. Some of the course materials will be postponed to ACM 270-2 in the Spring Quarter of 2018. As a continuation of the above topics, I will spend more time in exploring research topics in deep learning and stochastic approximation. The aim is to understand the underlying mathematical structure and theory in deep learning and to identify research opportunities in using machine learning and stochastic algorithms to solve complex problems that arise from physical and information sciences. An interesting topic will be how to use data to construct reduced models without solving PDEs.

**Course Materials.** Except for the classical homogenization theory where we have some textbooks, the course mainly relies on research papers published in recent years. By going through these research papers together, the students are exposed to the frontiers of these research topics. We do not assume that the students have prior knowledge in the background of these topics. We will cover these materials from the introductory level and gradually move to the research topics by motivating the main ideas and highlighting the major technical issues without going into the full details of the technical analysis.

**Course Policy**. This class is meant to be an informal research seminar course where we can discuss and explore recent research topics in data-driven multiscale model reduction and adaptive data analysis. There will be a small number of course projects for this course and one paper presentation. The course projects are designed to give the students a chance to implement the methods that they learn in the class so that they understand the advantages and limitations of these methods and identify potential improvement of these methods. Some of these projects may have the potential to turn into a research paper. Each student is asked to present one research paper of his or her choice related to the course materials. This is an excellent way to learn the relevant research topic from the perspective of an original researcher.