

STATISTICS COLLOQUIUM

MONDAY, APRIL 16, 2007

TALK: 4:00 PM — SCIENCE CENTER RM. B-10

RECEPTION: 5:15 PM — SCIENCE CENTER, 7TH FLOOR

“High-Dimensional NMR Spectral Peak Analysis for Protein Structure Studies”

Nicoleta Seban

Georgia Institute of Technology

School of Industrial & Systems Engineering

ABSTRACT

The primary research objectives of this project is to develop a statistical technique for identification, characterization and estimation for high-dimensional NMR spectrum. The leading application is protein structure determination using nuclear magnetic resonance (NMR). Most of the current research relevant to our application is developed for two-dimensional NMR frequency data. But NMR spectroscopy data for protein structure determination are currently produced even in four dimensions. If there are still a few methods for NMR spectra identification in three-dimensional spectroscopy NMR data, higher dimensions are not yet addressed. Moreover, the problem of mixtures of spectra or overlapped spectra remains still a problem at all dimensions.

Our statistical method takes a novel overall perspective: (1) It applies to NMR spectroscopy data in three, four and even higher dimensions; (2) It incorporates a preliminary step for separating the signal from the background using a method that adapts for sharp changes in the data and non-homogeneous signal; (3) The location, width and the amplitude of the NMR spectra identified above a noise level-dependent threshold are estimated using an iterative algorithm; (4) It detects mixtures of spectra at some confidence level.