Mass Spectrometry Feature Extraction for Expression Proteomic Applications Using A Lévy Random Fields Model

Leanna L. House, Department of Mathematical Sciences, Durham University, Durham, DH1 3LE, UK

Merlise A. Clyde and Robert Wolpert Institute of Statistics and Decision Sciences, Duke University, Durham, NC 27705, USA

Expression proteomic analyses that use profiles generated by Matrix Assisted Laser Desorption Ionization, Time-of-Flight mass spectrometers (MALDI-TOF MS) require three components: 1) quality control, 2) feature extraction, and 3) inference. To develop a method that simultaneously addresses these components, we first focus on implementing a Lévy random fields model that extracts pertinent features from individual spectra. Specifically, features are identified by asserting that each multi-modal or peaked spectrum of protein ion abundances is influenced by one, m/z and resolution dependent, stochastic Gamma Process. Using a Bayesian hierarchical approach, this single spectrum model may naturally include data quality (e.g. background and scaling) parameters and extend to incorporate data from multiple spectra. Provided the extension, we may complete the final component for a proteomics analysis and make inference.