Bayesian parameter estimation of processes that arise in chemical kinetics

Petar M. Djurić (Stony Brook University, Stony Brook, NY, USA)
Mónica Fernández-Bugallo, Joaquín Míguez

The traditional methodology of describing time evolutions of chemical reactions is by sets of coupled differential equations of chemical concentrations. This methodology is valid for spatially homogeneous systems and for large number of involved molecules. It is generally acknowledged, however, that a more general approach is based on a stochastic formulation of the chemical kinetics. Many phenomena, in fact, can only be described by stochastic processes, and they include very important molecular mechanisms that regulate critical cellular processes like signal transduction and gene expression. The stochastic processes involve many important parameters of biological interest, which often need to be estimated from observations. In this work, motivated by research on the kinetics of tissue proteolysis, which is a family of cell-mediated processes, we address the problem of estimating important parameters of stochastic processes that are used to build the full model of tissue proteolysis. These processes are Markov chains whose transition probabilities are defined by some common, and often, unknown parameters. Given a sequence of observations representing the current number of molecules of the chemical species in the system, the objective is to estimate the unknowns. It is easy to show that even in the simplest case of unimolecular unidirectional reactions, the problem is highly nonlinear. We propose to estimate the unknowns by using Bayesian theory implemented by particle filtering. With particle filtering, the a posteriori probability distributions of the unknowns are approximated by a discrete random measure of particles and their weights. The discrete measure is updated after every measurement, and it is used to compute various types of estimates. We investigate two sets of problems, which are characterized by the frequency of the made measurements. The data from each set imply different strategies for their processing.