Simulation, Sensitivity Analysis, and Optimization of Bioprocesses using Dynamic Flux Balance Analysis

by

Jose Alberto Gomez

Abstract

Microbial communities are a critical component of natural ecosystems and industrial bioprocesses. In natural ecosystems, these communities can present abrupt and surprising responses to perturbations, which can have important consequences. For example, climate change can influence drastically the composition of microbial communities in the oceans, which in turn affects the entirety of the food chain, and changes in diet can affect drastically the composition of the human gut microbiome, making it stronger or more vulnerable to infection by pathogens. In industrial bioprocesses, engineers work with these communities to obtain desirable products such as biofuels, pharmaceuticals, and alcoholic beverages, or to achieve relevant environmental objectives such as wastewater treatment or carbon capture. Mathematical models of microbial communities are critical for the study of natural ecosystems and for the design and control of bioprocesses. Good mathematical models of microbial communities allow scientists to predict how robust an ecosystem is, how perturbed ecosystems can be remediated, how sensitive an ecosystem is with respect to specific perturbations, and in what ways and how fast it would react to environmental changes. Good mathematical models allow engineers to design better bioprocesses and control them to produce high-quality products that meet tight specifications.

Despite the importance of microbial communities, mathematical models describing their behavior remain simplistic and only applicable to very simple and controlled bioprocesses. Therefore, the study of natural ecosystems and the design of complex bioprocesses is very challenging. As a result, the design of bioprocesses remains experiment-based, which is slow, expensive, and labor-intensive. With high-throughput experiments large datasets are generated, but without reliable mathematical models critical links between the species in the community are often missed. The design of novel bioprocesses rely on informed guesses by scientists that can only be tested experimentally. The expenses incurred by these experiments can be difficult to justify. Predictive mathematical models of microbial communities can provide insights about the possible outcomes of novel bioprocesses and guide the experimental design, resulting in cheaper and faster bioprocess development.

Most mathematical models describing microbial communities do not take into account the internal structure of the microorganisms. In recent years, new knowledge of the internal structures of these microorganisms has been generated using high-throughput DNA sequencing. Flux balance analysis (FBA) is a modeling framework that incorporates this new information into mathematical models of microbial communities. With FBA, growth and exchange flux predictions are made by solving linear programs (LPs) that are constructed based on the metabolic networks of the microorganisms. FBA can be combined with the mathematical models of dynamical biosystems, resulting in dynamic FBA (DFBA) models. DFBA models are difficult to simulate, sensitivity information is challenging to obtain, and reliable strategies to solve optimization problems with DFBA models embedded are lacking. Therefore, the use of DFBA models in science and industry remains very limited.

This thesis makes DFBA simulation more accessible to scientists and engineers with DFBAlab, a fast, reliable, and efficient Matlab-based DFBA simulator. This simulator is used by more than a 100 academic users to simulate various processes such as chronic wound biofilms, gas fermentation in bubble column bioreactors, and beta-carotene production in microalgae. Also, novel combinations of microbial communities in raceway ponds have been studied. The performance of algal-yeast co-cultures and more complex communities for biolipids production has been evaluated, gaining relevant insights that will soon be tested experimentally. These combinations could enable the production of lipids-rich biomass in locations far away from power plants and other concentrated CO_2 sources by utilizing lignocellulosic waste instead.

Following reliable DFBA simulation, the mathematical theory required for sensitivity analysis of DFBA models, which happen to be nonsmooth, was developed. Methods to compute generalized derivative information for special compositions of functions, hierarchical LPs, and DFBA models were generated. Significant numerical challenges appeared during the sensitivity computation of DFBA models, some of which were resolved. Despite the challenges, sensitivity information for DFBA models was used to solve for the steady-state of a high-fidelity model of a bubble column bioreactor using nonsmooth equation-solving algorithms.

Finally, local optimization strategies for different classes of problems with DFBA models embedded were generated. The classes of problems considered include parameter estimation and optimal batch, continuous steady-state, and continuous cyclic steady-state process design. These strategies were illustrated using toy metabolic networks as well as genome-scale metabolic networks. These optimization problems demonstrate the superior performance of optimizers when reliable sensitivity information is used, as opposed to approximate information obtained from finite differences.

Future work includes the development of global optimization strategies, as well as increasing the robustness of the computation of sensitivities of DFBA models. Nevertheless, the application of DFBA models of microbial communities for the study of natural ecosystems and bioprocess design and control is closer to reality.

Thesis Supervisor: Paul I. Barton
Title: Lammot du Pont Professor of Chemical Engineering