Kinetic parameter estimation in metabolic networks with GPGPU

Ali Khodayari¹, Pierre-Yves Taunay², Costas D. Maranas¹, and William Brouwer¹
¹Department of Chemical Engineering, The Pennsylvania State University-University Park
²Research Computing and Cyberinfrastructure, The Pennsylvania State University-University Park

1. Abstract

Estimation of kinetic parameters has been known to be the most challenging step in the construction of a kinetic model of microbial organisms. From a computational point of view, identification of the kinetic parameters can become impractical as the size of the model increases, due to the inherent non-linearity and non-convexity of the kinetic expressions and the stiffness of the ODEs describing the conservation of mass.

In this study, we develop an optimization-based parameter identification approach. This approach maximizes correlation between model predictions and measurements of experimental flux (also called steady-state flux) reaction rates) for a large-scale kinetic model, which relies on parallel implementation of the Newton-Raphson method. We call it a metabolic model composed of 64 reactions, 34 metabolites, and which includes 21 substrate-level regulatory interactions representing the core metabolism of E. Coli. The recently introduced Ensemble Modeling (EM) approach is used to construct a kinetic model of microbial organisms, bypassing the need for detailed kinetic expressions for all reactions by decomposing metabolic reactions into the elementary mechanisms.

The solver is designed and implemented using CUDA for Nvidia graphics processing units (GPU), in order to accelerate the overall process. The application initiated a large set of equations using the Boost::Spirit C++ framework, finds an analytic Jacobian J, where F corresponds to the numerical evaluation of the set of equations. The linear system is solved using the GMRES algorithm from the CUSP library, which is dedicated to computations and algorithms for sparse matrices on GPU. Successive updates of the parameter set, Jacobian matrix, and functions, as well as the system solver are all implemented on GPU. The result shows using CUDA for the numerical solver highly accelerates the convergence of kinetic parameters, satisfying the available flux measurements through wild-type and mutant strains.

2. Kinetic modeling of metabolic networks

- Restricted phenotype
- Determines metabolite concentrations
- Captures dynamics
- Captures regulatory interactions
- Challenges
- Deliberate experiments needed to estimate kinetic parameters
- (in vitro measurements do not always represent in vivo conditions)
- Unknown form of the kinetic rate laws
- Kinetic parameters and the functional form of kinetic rate laws may change with genetic and environmental perturbations
- Not yet scalable to genome-scale

3. Ensemble modeling of metabolic networks

- Reference steady state fluxes
- Obtain free energy of elementary steps
- Sample reversibility of elementary steps
- Calculate kinetic parameters
- Build an ensemble of models
- Predict the models in the ensemble

4. Screening step of the EM procedure

- Limitation
  - Screening may result in an empty ensemble
- Goal
  - Devise a procedure to identify optimal elementary kinetic parameters
- Challenges
  - Non-linearity of the parameters
  - Poorly-constrained
  - Ill-posedness

5. Schematic representation of the procedure

For a given metabolic network we can identify the optimal combination of the kinetic parameter values from the sampled models in the ensemble.

6. Solving the parameter identification problem in discrete space

- Generate an initial ensemble of models using the experimental measurements
- Predict the fluxes using model
- Select the best model
- From the best model, get sensitivities
- Add new set of flux data per iteration
- Minimize deviation from experimental data

7. A metabolic model of E. coli core metabolism

8. Objective function

Time dependent: integration of ordinary differential equations

\[ \mathbf{S} = \mathbf{E} \mathbf{P} \]

Decomposition of reaction \( j \)

\[ S = E \mathbf{K} S + \mathbf{E}_0 \]

The constructed ensemble for reaction \( j \):

\[ \text{model } 1: a_1 \mathbf{k}_1 + a_2 \mathbf{k}_2 + \ldots + a_m \mathbf{k}_m \]

\[ \text{model } n: b_1 \mathbf{k}_1 + b_2 \mathbf{k}_2 + \ldots + b_m \mathbf{k}_m \]

9. GPU implementation of Newton-Raphson

- Create Jacobian and function
- Initialization using the kinetic parameter values obtained in the previous iteration
- Minimize over kinetic parameters
- Subject to:
  - Kinetic parameters
  - Mixed acid metabolism
  - Thermodynamic feasibility and other constraints for reactions \( a \) + \( \mathbf{r} \)

10. Evaluation of the model performance

Node evaluation

Thread index \( x = 0, 1, 2, 3 \)

Reduction \( 0 \rightarrow 0 + 1 \)

Dedicated experiments needed to estimate kinetic parameters

Gibbs free energy of reaction \( j \)

Reaction rate

Steady-state: solving a system of non-linear equations

For a given metabolic network we can identify the optimal combination of the kinetic parameter values from the sampled models in the ensemble.

References

- Reference steady state fluxes
- Obtain free energy of elementary steps
- Sample reversibility of elementary steps
- Calculate kinetic parameters
- Build an ensemble of models
- Predict the models in the ensemble

Results

- Generate an initial ensemble of models using the experimental measurements
- Predict the fluxes using model
- Select the best model
- From the best model, get sensitivities
- Add new set of flux data per iteration
- Minimize deviation from experimental data

Conclusion

- Efficient parameter estimation and flux predictions in large-scale metabolic networks
- Potential for real-time flux predictions in microbial organisms

Acknowledgments

- Partial support from the National Science Foundation (NSF) under Grants PCM-0825745 and DBI-0613943
- Pennsylvania State University, Department of Chemical Engineering
- Research Computing and Cyberinfrastructure, The Pennsylvania State University