Uniform sampling of metabolic networks

Ben Cousins[1], Hulda S. Haraldsdóttir[2]* (hulda.haraldsdottir@uni.lu), Santosh Vempala[1]and **Ronan M. T. Fleming**[2]

[1]School of Computer Science, Algorithms and Randomness Center, Georgia Institute of Technology, Atlanta, Georgia; [2]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

http://stanford.edu/group/SOL/multiscale/

Project Goals: Development of a framework for mass conserved elementary kinetic modelling of metabolic networks [1, 2, 3, 4, 5]. This collaborative project was motivated by a need for data to constrain and validate metabolic models. Sampling algorithms have demonstrated applications in measurement and estimation of kinetic parameters, steady state fluxes and metabolite concentrations for biochemical systems [6, 7, 8].

Constraint-based metabolic modelling provides a framework to explore feasible steady state fluxes in metabolic networks. Physicochemical constraints imposed, e.g., by network topology, mass conservation and substrate availability are formulated as linear equalities and inequalities that define a high-dimensional convex set. Uniform sampling of this set provides an unbiased characterisation of the metabolic capabilities of a cell or organism [9]. However, uniform sampling of steady state metabolic flux sets has proven algorithmically challenging due to their high dimensionality and inherent anisotropy. Here, we evaluate the performance of a recently published sampling algorithm [10] on metabolic networks of increasing size. The algorithm is based upon the provably efficient hit-and-run random walk [11] and crucially uses a rounding preprocessing step to place the set of feasible metabolic fluxes in near-isotropic position. This algorithm converges to a uniform sampling distribution up to 25 times faster than a popular artificial centering hit-and-run algorithm [12]. We demonstrate the effects of improved convergence on predictions of the metabolic capabilities of *E. coli*.

References

[1] Sun Y, Fleming RM, Thiele I, Saunders MA (2013) Robust flux balance analysis of multiscale biochemical reaction networks. BMC Bioinformatics 14: 240.

[2] Noor E, Haraldsdóttir HS, Milo R, Fleming RMT (2013) Consistent Estimation of Gibbs Energy Using Component Contributions. PLoS Comput Biol 9: e1003098.

[3] Artacho FJA, Fleming RMT (2014) Globally convergent algorithms for finding zeros of duplomonotone mappings. Optimization Letters : 1–16.

[4] Artacho FJA, Fleming RMT, Vuong PT (2015) Accelerating the DC algorithm for smooth functions. arXiv:150707375 [math, q-bio] .

[5] Haraldsdóttir HS, Fleming RMT (2015) Identification of conserved moieties in metabolic networks by graph theoretical analysis of atom transition networks. Submitted .

[6] Yang L, Mahadevan R, Cluett W (2010) Designing experiments from noisy metabolomics data to refine constraint-based models. In: American Control Conference (ACC), 2010. pp. 5143–5148. doi:10.1109/ACC.2010.5530678.

[7] Schellenberger J, Zielinski DC, Choi W, Madireddi S, Portnoy V, et al. (2012) Predicting outcomes of steady-state 13c isotope tracing experiments using Monte Carlo sampling. BMC systems biology 6: 9.

[8] Khodayari A, Zomorrodi AR, Liao JC, Maranas CD (2014) A kinetic model of Escherichia coli core metabolism satisfying multiple sets of mutant flux data. Metabolic Engineering 25: 50–62.

[9] Schellenberger J, Palsson BØ (2009) Use of Randomized Sampling for Analysis of Metabolic Networks. Journal of Biological Chemistry 284: 5457–5461.

[10] Cousins B, Vempala S (2015) A practical volume algorithm. Mathematical Programming Computation : 1–28.

[11] Lovasz L, Vempala S (2006) Hit-and-run from a corner. In: SIAM journal on computing. Society for Industrial and Applied Mathematics, volume 35, pp. 985–1005.

[12] Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, et al. (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nature Protocols 6: 1290–1307.

This work was supported by the U.S. Department of Energy, Offices of Advanced Scientific Computing Research and the Biological and Environmental Research as part of the Scientific Discovery Through Advanced Computing program, grant #DE-SC0010429.