Supplement to: Elimination of Thermodynamically Infeasible Loops in Steady State Metabolic Models

Jan Schellenberger,[†] Nathan E. Lewis,[‡] and Bernhard Ø. Palsson[‡]*

[†]Bioinformatics and Systems Biology Program and [‡]Bioengineering Department, University of California, San Diego, California

Performance enhancements for ll-FVA

Several performance considerations can be made in order to speed up this computation:

- Null basis computation: The algorithm requires a set of basis vectors of the internal part of the stoichiometric matrix, S. N_i = null(S_{int}). By default, MATLAB will compute a dense orthonormal basis based on Singular Value Decomposition. It was found to be many fold faster to use a sparse representation of N_{int} based on LU decomposition. As an added benefit, calculating N_{int} in this fashion can be faster than computing the orthonormal basis.
- Elimination of unnecessary reactions: If there is no flux through a reaction, it can be eliminated from consideration for the loop law formulation. This elimination saves an indicator variable a_i and its corresponding G_i. Reactions can have no flux if lb = ub = 0, which is easy to check. A more complete check requires performing a Flux Variability Analysis (FVA) computation, which is expensive; however, FVA may be worthwhile if many loop-removing computations will be done. Depending on the MILP/MIQP solver used, this optimization may not bring much benefit as a pre-processor may catch these conditions automatically.
- Combining reactions with the same direction: Often in a network, reactions are coupled, such that their fluxes are either both active or both inactive. For example, this happens if two reactions are part of the same linear pathway. If two reactions i and j are coupled, then it is known that $a_i = a_j$ and one variable can be used instead. G_i and G_j must remain two separate variables, however. In rare instances, two reactions are coupled inversely where $v_i = -v_j$. In this case it is possible to combine a_i and a_j as $a_i = 1-a_j$. Coupled reactions can be easily computed because their rows in N_{int} will be similar (i.e., $N_{int}(i,:) = k\ N_{int}(j,:)$, where $k > 0$ implies positive coupling and $k < 0$ implies negative coupling).
- Flux Variability Analysis only: It is comparatively inexpensive to compute the classic FVA procedure before going to the loopless method. As a solution pool of loopless distributions is formed, they can be checked for optimality for any of the subsequent computations. If two reactions are in a linear pathway, then a flux solution that optimally uses one reaction will also optimally use the second. This saves time by eliminating expensive MILP computations. For the *E. coli* model, the number of computations is cut by a factor of 2.

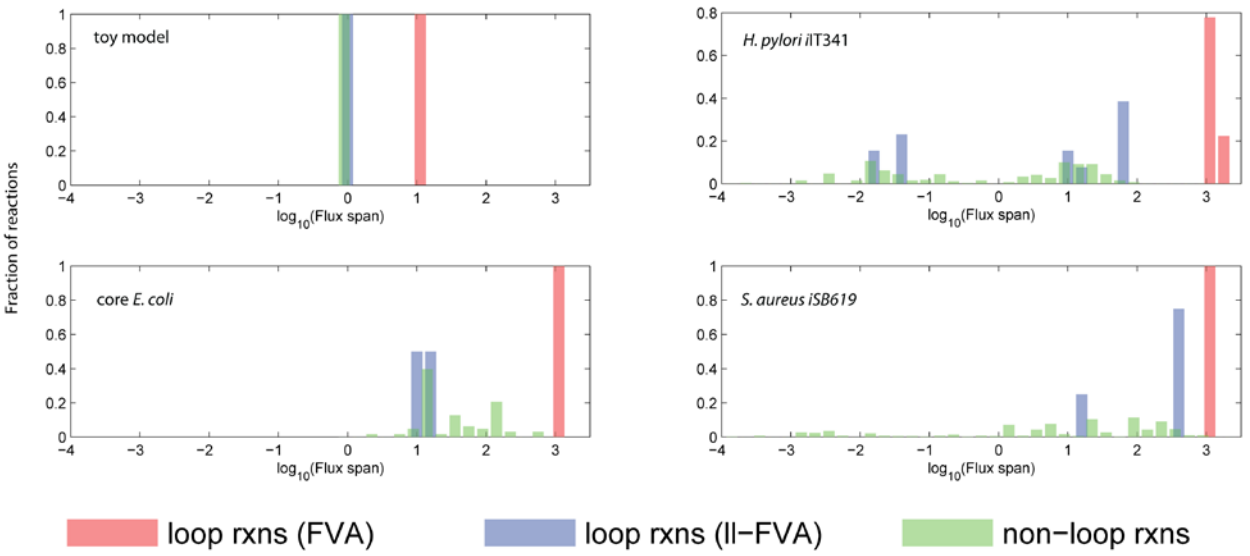
Effects of media choice and objective function on ll-FVA simulation time

Given the increase in computational time required for the usage of loop law constraints with flux variability analysis (ll-FVA), it is of interest to see if variations in model setup induce significant changes in computation time. To test this, four different minimal media sources were selected that lead to significant changes in flux distributions (i.e., glucose, lactate, acetate, and α-ketoglutarate). In addition, these were all tested using two different common objective functions (i.e., biomass optimization and

ATP synthesis optimization). These were tested using the *E. coli* core model and the genome-scale model of *E. coli* metabolism (iAF1260). For both of these models, the standard deviation for computation time was 4% and 35% of the mean running time for the various growth/objective conditions for the core model and iAF1260, respectively. On the other hand, when the simulations for glucose minimal media and biomass optimization were repeated multiple times using the iAF1260 model, the standard deviation for running time was only 1%. Thus, while there is some variation among different growth conditions and objective functions, this variation should not affect the feasibility of the calculations.

A comparison of non-loop and loop reactions before and after imposition of loop law constraints

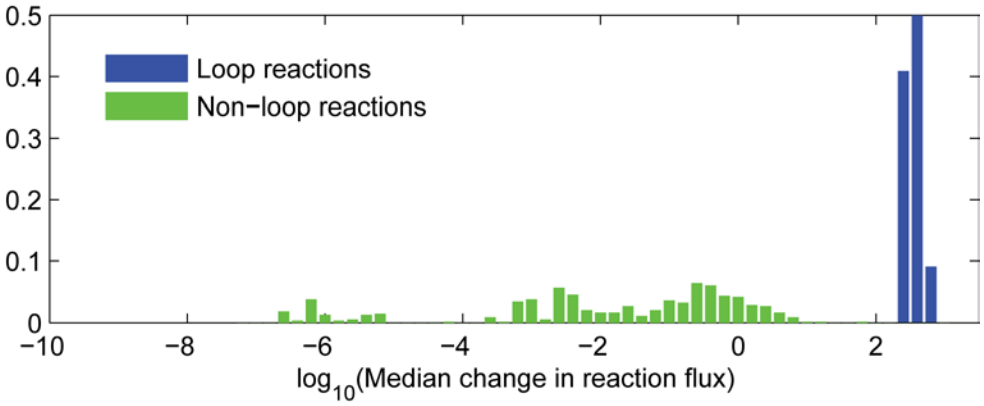
In constraint-based models, loop reactions represent a subset of the model that has unbounded fluxes. Even when the model is provided no substrate, these reactions can maintain a flux that is only limited by the upper bounds provided by the user. However, in the past, these loop fluxes have been ignored since they do not affect the biomass function. As detailed in the main text, the loop law constraints lead to improved flux predictions by disallowing the usage of these loops. Therefore, if these constraints are successful in this endeavor, it is expected that following the addition of loop law constraints, the feasible ranges of fluxes through loop reactions should be comparable to the feasible ranges of non-loop reactions. This was demonstrated on the *i*AF1260 model in the main text (Figure 3.b). As seen in Supplementary Figure 1, this is true for other models in Table 1.



Supplementary Figure 1. Ranges for loop fluxes are reduced to non-loop reaction levels following the imposition of loop law constraints. The ranges of feasible flux for each reaction were determined using FVA, and in all models, loop reactions (red) had higher flux ranges than non-loop reactions (green). When loop law constraints were added using ll-FVA, the ranges of feasible fluxes for loop reactions (blue) were within the range of non-loop reactions.

Non-loop reactions that change following use of MIQP ll-sampling

In all methods detailed in this study, only loop reactions are affected by loop law constraints. The only exception is the MIQP formulation of ll-sampling. While the MILP formulism dismantles loops by shutting off loop reactions that are not necessary for a given flux distribution, the MIQP formulism also allows non-loop reactions to change their flux to find the nearest loopless flux distribution. However, as demonstrated in Supplementary Figure 2, median adjustments in non-loop reactions are orders of magnitude smaller than the adjustments to loop reactions.



Supplementary Figure 2. In finding the nearest loopless flux distribution using ll-sampling with MIQP, non-loop reactions (green) are often moved. However, this movement is always much smaller than the changes seen by loop reactions (blue).