Flux Variability analysis (FVA)

Authors : Vanja Vlasov, Systems Biochemistry Group, LCSB, University of Luxembourg.

Reviewers : Anne Richelle, Lewis Lab at University of California, San Diego.

Flux variability analysis (FVA) is a widely used computational tool for evaluating the minimum and maximum range of each reaction flux that can still satisfy the constraints using a double LP problem (i.e. a maximization and a subsequent minimization) for each reaction of interest [1].

$$
\begin{aligned}
v_{j,upper} / \; v_{j,lower} = & \underset{v}{max} / \min_{v} \quad v_j \\
\text{s.t.} \quad & Sv = 0, \\
& l \leq v \leq u
\end{aligned}
$$

where $v \in R^n$ is the vector of specific reaction rates (metabolic fluxes) and $v_{j,upper}$ and $v_{j,lower}$ are respectively the upper and lower values of each flux v_j satisfying the system of linear equations.

Depending on the size of the model you are using for the analysis, use:

- `fluxVariability()` function - for the low dimensional FVA;
- `fastFVA()` function - for the models with more than 1,000 reactions;
- distributedFBA.jl - for high dimensional FVA, models larger than 10,000 reactions [2].

EQUIPMENT SETUP

If necessary, initialize the cobra toolbox

```
initCobraToolbox(false) % false, as we don't want to update
```

```
     _____   _____   _____   _____     _____                |
    /  ___| /  _  \ |  _  \ |  _  \   / ___ \               |   COnstraint-Based Reconstruction and Analysis
    | |     | | | | | |_| | | |_| |   | |___| |              |   The COBRA Toolbox - 2021
    | |     | | | | |  _  { |  _  /   |  ___  |            |
    | |___  | |_| | | |_| | | | \ \   | |   | |            |   Documentation:
    \_____| \_____/ |_____/ |_|  \_\ |_|   |_|            |   http://opencobra.github.io/cobratoolbox
                                                           |

> Checking if git is installed ...  Done (version: 2.20.1).
> Checking if the repository is tracked using git ...  Done.
> Checking if curl is installed ...  Done.
> Checking if remote can be reached ...  Done.
> Initializing and updating submodules (this may take a while)... Done.
> Adding all the files of The COBRA Toolbox ...  Done.
> Define CB map output... set to svg.
> TranslateSBML is installed and working properly.
> Configuring solver environment variables ...
  - [-*--] ILOG_CPLEX_PATH: /usr/local/bin/cplex/ibm/ILOG/CPLEX_Studio1210/cplex/matlab/x86-64_linux
  - [-*--] GUROBI_PATH: /usr/local/bin/gurobi811/linux64/matlab
  - [-*--] TOMLAB_PATH: /usr/local/bin/tomlab/tomsym
  - [----] MOSEK_PATH: --> set this path manually after installing the solver ( see instructions )
  Done.
```

```
 > Checking available solvers and solver interfaces ...No valid license for the SNOPT solver


tomRun: ERROR! snopt solver returned empty Result

Run startup to set paths


 Done.
 > Setting default solvers ... Done.
 > Saving the MATLAB path ... Done.
   - The MATLAB path was saved as ~/pathdef.m.

 > Summary of available solvers and solver interfaces

          Support        LP      MILP        QP      MIQP       NLP
    ------------------------------------------------------------------
    gurobi           active               1         1         1         1         -
    ibm_cplex        active               1         1         1         1         -
    tomlab_cplex     active               1         1         1         1         -
    glpk             active               1         1         -         -         -
    mosek            active               0         -         0         -         -
    matlab           active               1         -         -         -         1
    pdco             active               1         -         1         -         -
    quadMinos        active               1         -         -         -         -
    dqqMinos         active               1         -         1         -         -
    cplex_direct     active               1         1         1         -         -
    cplexlp          active               1         -         -         -         -
    qpng             passive              -         -         1         -         -
    tomlab_snopt     passive              -         -         -         -         1
    lp_solve         legacy               1         -         -         -         -
    ------------------------------------------------------------------
    Total            -                   11         5         7         3         2

 + Legend: - = not applicable, 0 = solver not compatible or not installed, 1 = solver installed.


 > You can solve LP problems using: 'gurobi' - 'ibm_cplex' - 'glpk' - 'matlab' - 'pdco' - 'quadMinos' - 'c
 > You can solve MILP problems using: 'gurobi' - 'ibm_cplex' - 'glpk'
 > You can solve QP problems using: 'gurobi' - 'ibm_cplex' - 'pdco'
 > You can solve MIQP problems using: 'gurobi' - 'ibm_cplex'
 > You can solve NLP problems using: 'matlab'

> Checking for available updates ... skipped
```

For solving linear programming problems in FBA and FVA analysis, certain solvers are required. The present tutorial can run with glpk package, which does not require additional installation and configuration. Although, for the analysis of large models is recommended to use the GUROBI package.

```
changeCobraSolver ('ibm_cplex', 'LP');
```

```
 > changeCobraSolver: IBM ILOG CPLEX interface added to MATLAB path.
 > ibm_cplex (version 1210) is compatible and fully tested with MATLAB R2019a on your operating system.
```

```
changeCobraSolver ('ibm_cplex', 'QP');
```

```
 > changeCobraSolver: IBM ILOG CPLEX interface added to MATLAB path.
 > ibm_cplex (version 1210) is compatible and fully tested with MATLAB R2019a on your operating system.
```

PROCEDURE

In this tutorial, we will use the generic model of the human cellular metabolism, Recon2.0 [3]. Load the model

```
global CBTDIR
modelFileName = 'Recon2.0model.mat';
modelDirectory = getDistributedModelFolder(modelFileName); %Look up the folder for the
modelFileName= [modelDirectory filesep modelFileName]; % Get the full path. Necessary t
model = readCbModel(modelFileName);
```

```
Each model.subSystems{x} is a character array, and this format is retained.
```

The metabolites structures and reactions in Recon2.0 can be found in the Virtual Metabolic Human database (VMH, http://vmh.life).

Constrain the model to limit the availability of carbon and oxygen energy sources. Find the uptake exchange reactions using *findExcRxns*

```
[selExc, selUpt] = findExcRxns(model);
uptakes = model.rxns(selUpt);
```

Select from the set of reactions defined in *uptakes* those which contain a least one carbon in the metabolites involved in the reaction and set the flux values of these reactions to zero:

```
subuptakeModel = extractSubNetwork(model, uptakes);
hiCarbonRxns = findCarbonRxns(subuptakeModel,1);
modelalter = changeRxnBounds(model, hiCarbonRxns, 0, 'b');
```

Also close the other reaction related to the exchange of oxygen and energy sources:

```
energySources = {'EX_adp'; 'EX_amp(e)'; 'EX_atp(e)'; 'EX_co2(e)';...
    'EX_coa(e)'; 'EX_fad(e)'; 'EX_fe2(e)'; 'EX_fe3(e)'; 'EX_gdp(e)';...
    'EX_gmp(e)'; 'EX_gtp(e)'; 'EX_h(e)'; 'EX_h2o(e)'; 'EX_h2o2(e)';...
    'EX_nad(e)'; 'EX_nadp(e)'; 'EX_no(e)'; 'EX_no2(e)'; 'EX_o2s(e)'};
modelalter = changeRxnBounds (modelalter, energySources, 0, 'l');
```

For this tutorial, we will analyse the variability of several reactions of the human cellular metabolism in aerobic and anaerobic condition. For each simulation, the original model will be copied to a new model structure (e.g., *modelfva1* for aerobic condition and *modelfva2* for anaerobic condition). This preserves the constraints of the original model and allows to perform simulations with new constraints. Additionally, this method of renaming the model avoids confusion while performing multiple simulations at the same time.

```
% modelfva1 represents aerobic condition
modelfva1 = modelalter;
modelfva1 = changeRxnBounds(modelfva1, 'EX_glc(e)', -20, 'l');
modelfva1 = changeRxnBounds(modelfva1, 'EX_o2(e)', -1000, 'l');
% modelfva2 represents anaerobic condition
modelfva2 = modelalter;
modelfva2 = changeRxnBounds(modelfva2, 'EX_glc(e)', -20, 'l');
```

```
modelfva2 = changeRxnBounds(modelfva2, 'EX_o2(e)',  0, 'l');
```

1) Standard FVA

The full spectrum of flux variability analysis options can be accessed using the command:

```
% [minFlux, maxFlux, Vmin, Vmax] = fluxVariability(model,...
% optPercentage,osenseStr, rxnNameList, verbFlag, allowLoops, method);
```

The `optPercentage` parameter allows one to choose whether to consider solutions that give at least a certain percentage of the optimal solution (default - 100). Setting the parameters `osenseStr = 'min'` or `osenseStr = 'max'` determines whether the flux balance analysis problem is first solved with minimization or maximisation (default - 'max'). The `rxnNameList` accepts a cell array list of reactions to selectively perform flux variability upon (default - all reactions of the model). This is useful for high-dimensional models as computation of flux variability for all reactions can be time consuming:

```
% Selecting several reactions of the model that we want to analyse with FVA
rxnsList = {'DM_atp_c_'; 'ACOAHi'; 'ALCD21_D'; 'LALDO'; 'ME2m';...
    'AKGDm'; 'PGI'; 'PGM'; 'r0062'};
```

The `verbFlag` input determines the verbose output (default - false). `allowLoops` input determines whether loops are allowed in the solution (default - true). Note that `allowLoops==false` invokes a mixed integer linear programming implementation of thermodynamically constrained flux variability analysis for each minimization or maximisation of a reaction rate. The `method` parameter input determines whether are the output flux vectors also minimise the `0-norm`, `1-norm` or `2-norm` whilst maximising or minimising the flux through one reaction (default - 2-norm).

Run `fluxVariability()` on both models (`modelfva1`, `modelfva2`) to generate the minimum and maximum flux values of selected reactions (*rxnsList*) in the model.

```
% Run FVA analysis for the model with the constraints that simulates aerobic conditions
[minFlux1, maxFlux1, Vmin1, Vmax1] = fluxVariability(modelfva1, 100, 'max', rxnsList)
```

```
OPTIMAL
Error using solveCobraQP (line 1049)
[gurobi] Primal optimality condition in solveCobraQP not satisfied, residual = 0.00038861, while
feasTol = 1e-06

Error in fluxVariability>getMinNorm (line 707)
        solution = solveCobraQP(LPproblem, solverVarargin.QP{:});

Error in fluxVariability>calcSolForEntry (line 685)
        V = getMinNorm(LPproblem, LPsolution, numel(model.rxns), Flux, model, method,
solverVarargin);

Error in fluxVariability (line 557)
        [maxFlux(i), V] = calcSolForEntry(model, rxnID ,LPproblem, method, allowLoopsI, minNorm,
solverVarargin, heuristicSolutions{preCompMaxSols(i)}, -1);
```

```
% Run FVA analysis for the model with the constraints that
```

```
% simulates anaerobic conditions:
[minFlux2, maxFlux2, Vmin2, Vmax2] = fluxVariability(modelfva2, [], [], rxnsList)
```

The additional n × k output matrices `Vmin` and `Vmax` return the flux vector for each of the k ≤ n fluxes selected for flux variability.

You can further plot and compare the FVA results for the selected reaction from both models:

```
ymax1 = maxFlux1;
ymin1 = minFlux1;
ymax2 = maxFlux2;
ymin2 = minFlux2;

maxf = table(ymax1, ymax2)
minf = table(ymin1, ymin2)
maxfxs = table2cell(maxf);
minfxs = table2cell(minf);

figure
plot1 = bar(cell2mat(maxfxs(1:end, :)));
hold on
plot2 = bar(cell2mat(minfxs(1:end, :)));
hold off
xticklabels({'DM_atp_c_', 'ACOAHi', 'ALCD21__D', 'LALDO',...
             'ME2m', 'AKGDm', 'PGI', 'PGM', 'r0062'})
set(gca, 'XTickLabelRotation', -80);
yticks([-1000 -800 -600 -400 -200 0 200 400 600 800 1000])
xlabel('Reactions from the models')
ylabel('Fluxes')
legend({'Aerobic', 'Anaerobic'}, 'Location', 'southwest')
title('Variations in fluxes in the aerobic and anaerobic conditions')
```

2) Fast FVA

When the number of reaction on which you want to perform a flux variability is higher to 1000, we recommend using `fastFVA()` function instead of `fluxVariability()`. Note that for large models the memory requirements may become prohibitive.

The `fastFVA()` function only supports the CPLEX solver.

```
changeCobraSolver ('ibm_cplex', 'all', 1);
```

Run fastFVA analysis for the whole model (i.e. flux varaibility analysis performed on all reactions included in the model) with the constraints that simulates aerobic conditions:

```
[minFluxF1, maxFluxF1, optsol, ret, fbasol, fvamin, fvamax, statussolmin, statussolmax]
```

Run fast FVA analysis for the whole model with the constraints that simulates anaerobic conditions:

```
[minFluxF2, maxFluxF2, optsol2, ret2, fbasol2, fvamin2, fvamax2,...
    statussolmin2, statussolmax2] = fastFVA(modelfva2);
```

Plot the results of the fast FVA and compare them between the aerobic and anaerobic models:

```
ymaxf1 = maxFluxF1;
yminf1 = minFluxF1;
ymaxf2 = maxFluxF2;
yminf2 = minFluxF2;

maxf =table(ymaxf1, ymaxf2);
minf =table(yminf1, yminf2);

maxf = table2cell(maxf);
minf = table2cell(minf);

figure
plot3 = bar(cell2mat(maxf(1:end, :)));
hold on
plot4 = bar(cell2mat(minf(1:end, :)));
hold off
xticks([0 2000 4000 6000 8000 10600])
yticks([-1000 -800 -600 -400 -200 0 200 400 600 800 1000])
xlabel('All reactions in the model')
ylabel('Fluxes')
legend({'Aerobic', 'Anaerobic'})
title('Variations in fluxes in the aerobic and anaerobic conditions')
```

REFERENCES

[1] Gudmundsson, S., Thiele, I. Computationally efficient flux variability analysis. *BMC Bioinformatics.* 11, 489 (2010).

[2] Heirendt, L., Thiele, I., Fleming, R.M. DistributedFBA.jl: high-level, high-performance flux balance analysis in Julia. *Bioinformatics.* 33 (9), 1421-1423 (2017).

[3] Thiele, I., et al. A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.,* 31(5), 419–425 (2013).