Proton shuttle testing with sparse flux balance analysis

Author: Ronan Fleming, Ines Thiele, University of Luxembourg.

Reviewer:

INTRODUCTION

We consider a biochemical network of m molecular species and n biochemical reactions. The biochemical network is mathematically represented by a stoichiometric matrix $S \in \mathscr{Z}^{m \times n}$. In standard notation, flux balance analysis (FBA) is the linear optimisation problem

$$\begin{aligned} \min_{v} \quad & \rho(v) \equiv c^T v \\ \text{s.t.} \quad & Sv = b, \\ & l \leq v \leq u, \end{aligned}$$

where $c \in \Re^n$ is a parameter vector that linearly combines one or more reaction fluxes to form what is termed the objective function, and where a $b_i < 0$, or $b_i > 0$, represents some fixed output, or input, of the ith molecular species. A typical application of flux balance analysis is to predict an optimal non-equilibrium steady-state flux vector that optimises a linear objective function, such biomass production rate, subject to bounds on certain reaction rates.

In this tutorial, we demonstrate how to predict the minimal number of active reactions that are still consistent with an optimal objective derived from the result of a standard flux balance analysis problem. In each case, the corresponding problem is a cardinality minimisation problem that we term *sparse flux balance analysis*

$$\begin{aligned} \min_{v} \quad & \|v\|_0 \\ \text{s.t.} \quad & Sv = b \\ & l \leq v \leq u \\ & c^T v = \rho^\star \end{aligned}$$

where the last constraint is optional and represents the requirement to satisfy an optimal objective value $\rho^\star$ derived from any solution to a flux balance analysis (FBA) problem. This approach is used to check for minimal sets of reactions that either should be active, or should not be active in a flux balance model that is representative of a biochemical network.

In particular, this tutoriall illustrates the use of sparse flux balance analysis to compute the minimal set of reactions that must be active to produce ATP

TIMING

A minimal solution to sparse flux balance analysis problem can be obtained in < 10 seconds. The time consuming part is comparing the predictions with the biochemical literature to assess whether the predictions are consistent with biochemical network funcion or not, as such, the process of refining a model to increase its biochemical fidelity can take days or weeks.

PROCEDURE

Loading and examining the properties of Recon3.0model

We are going to focus here on testing the biochemical fidelity of Recon3.0model, so load it, unless it is already loaded into the workspace

```
clear %model
if ~exist('modelOrig','var')
    %filename='Recon1.0';
    %filename='Recon2.0';
    %filename='Recon2.0model';
    %filename='Recon2.04model';
    %filename='HMR2.0'
    %filename='Recon2.2model';
    %filename='Recon3.0';
    filename='Recon3.0model';
    directory='~/work/sbgCloud/programReconstruction/projects/recon2models/data/reconXC
    model = loadIdentifiedModel(filename,directory);
    model.csense(1:size(model.S,1),1)='E';
    modelOrig = model;
else
    model=modelOrig;
end
```

Setting the numerical tolerance

Implementation of sparse flux balance analysis with any numerical optimisation solver, requires a tolerance to be set that distinguished between zero and non-zero flux, based on the numerical tolerance of the currently installed optimisation solver. Typically 1e-6 will suffice, except for multiscale models.

```
feasTol = getCobraSolverParams('LP', 'feasTol');
```

Testing for activity of ATP synthase with all exchanges closed

Detect the ATP synthase reaction and if there is none already, add one.

```
atpsynthaseBool=strcmp(model.rxns,'ATPS4mi') | strcmp(model.rxns,'ATPS4m');% | strcmp(m
if ~any(atpsynthaseBool)
    fprintf('Could not find ATP synthase reaction, adding one.')
    if ~strcmp(filename,'HMR2.0')
        %model = addReaction(model, 'ATPMnew', 'h2o[c] + atp[c] -> h[c] + adp[c] + pi[c
        model = addReaction(model, 'ATPS4m', '4.0 h[c] + adp[m] + pi[m] -> h2o[m] + 3.0
    else
        %model = addReaction(model, 'ATPMnew', 'm02040c + m01371c -> m02039c + m01285c
        model = addReaction(model, 'ATPS4m', '4.0 m02039c + m01285m + m02751m -> m02040
    end
    atpsynthaseBool=strcmp(model.rxns,'ATPS4m');
    fprintf('%s %s\n',model.rxns{atpsynthaseBool},' is the ATP synthase reaction')
else
    fprintf('%s %s\n',model.rxns{atpsynthaseBool},' is the ATP synthase reaction')
end
```

```
ATPS4mi  is the ATP synthase reaction
```

Display the size of the model

```
[nMet,nRxn] = size(model.S);
fprintf('%6s\t%6s\n','#mets','#rxns'); fprintf('%6u\t%6u\t%s%s\n',nMet,nRxn,' totals in
```

```
 #mets     #rxns
  5835     10600     totals in Recon3model
```

Display the constraints

```
minInf = -1000;
maxInf =  1000;
printConstraints(model, minInf, maxInf);
```

```
MinConstraints:
maxConstraints:
```

Identify the exchange reactions(s) heuristically

```
if ~isfield(model,'SIntRxnBool')
    model = findSExRxnInd(model,size(model.S,1),1);
end
```

Maximise the atp synthase reaction

```
model.c(:)=0;
model.c(atpsynthaseBool)=1;
osenseStr='max';
```

Choose to minimize the zero norm of the optimal flux vector

```
minNorm='zero';
```

Allow thermodynamically infeasible fluxes

```
allowLoops=1;
```

Select the approximate step functions when minimising the zero norm of the flux vector

```
% zeroNormApprox='cappedL1';% : Capped-L1 norm
% zeroNormApprox='exp';%Exponential function
% zeroNormApprox='log';%Logarithmic function
% zeroNormApprox='SCAD';%Smoothly clipped absolute deviation function
% zeroNormApprox='lp-';%L_p norm with p<0
% zeroNormApprox='lp+';%L_p norm with 0<p<1
zeroNormApprox='all';% test all approximations avialable and return the best one
```

Close all external reactions

```
model.lb(~model.SIntRxnBool)=0;
model.ub(~model.SIntRxnBool)=0;
```

Run sparse flux balance analysis on the model with all exchanges closed

```
sparseFBAsolutionBounded = optimizeCbModel(model, osenseStr, minNorm, allowLoops, zeroN
```

Check to see if there is a non-zero flux through the ATP synthase reaction

```
fprintf('%g%s\n',sparseFBAsolutionBounded.v(atpsynthaseBool),' flux through the ATP syn
```

```
0 flux through the ATP synthase reaction
```

Display the sparse flux solution, but only the non-zero fluxes, above a specified cutoff.

```
cutoff=feasTol;
for n=1:nRxn
    if abs(sparseFBAsolutionBounded.v(n))>cutoff
        formula=printRxnFormula(model, model.rxns{n}, 0);
        fprintf('%10g%15s\t%-60s\n',sparseFBAsolutionBounded.v(n),model.rxns{n}, formul
    end
end
```

ANTICIPATED RESULTS

In a model for flux balance analysis, there should be zero flux through the ATP synthase reaction if all external reaction bounds are zero.

TROUBLESHOOTING

If there is non-zero flux through the ATP synthase reaction, with all external reaction bounds zero, then the bounds on one of the reactions in each of the minimal sets needs to be set to eliminate flux in one direction. Each of the minimal sets corresponds to net flux around a stoichiometrically balanced cycle, which is thermodynamically infeasible [fleming_variational_2012]. Steady-state mass balance constraints do not enforce thermodynamic constraints. In lieu of such constraints, the bounds on reactions can be set based on the biochemical literature to eliminate net flux around a stoichiometrically balanced cycle. In a model, with all external reactions blocked (bounds are set to zero), maximising the ATP synthase reaction while minimising the cardinality of all internal reactions, using sparse flux balance analysis can be used to find any such cycle of minimal cardinality (minimal number of active reactions). By further constraining the bounds to convert one reversible reaction in each such cycle to an irreversible reaction, thermodynamically infeasible flux around cycles, such as those involving the ATP synthase reaction, can be eliminated. The following sections of this tutorial illustrate how to test different parts of the model for thermodynamically infeasible flux through the ATP synthase reaction.

Testing for activity of ATP synthase with all exchanges closed and all internal reactions reversible

Fully open all internal reactions

```
model.lb(model.SIntRxnBool)=-1000;
model.ub(model.SIntRxnBool)=1000;
```

Run sparse flux balance analysis on the model with all exchanges closed and all internal reactions reversible

```
sparseFBAsolutionUnBounded = optimizeCbModel(model, osenseStr, minNorm, allowLoops, zer
```

Check to see if there is a non-zero flux through the ATP synthase reaction

```
fprintf('%g%s\n',sparseFBAsolutionUnBounded.v(atpsynthaseBool),' flux through the ATP s
```

```
1000 flux through the ATP synthase reaction
```

Display the sparse flux solution, but only the non-zero fluxes, above a specified cutoff.

```
cutoff=feasTol;
for n=1:nRxn
    if abs(sparseFBAsolutionUnBounded.v(n))>cutoff
        formula=printRxnFormula(model, model.rxns{n}, 0);
        fprintf('%10g%15s\t%-60s\n',sparseFBAsolutionUnBounded.v(n),model.rxns{n}, form
    end
end
```

```
       400          ADK1m    atp[m] + amp[m]  <=> 2 adp[m]
       200         G5SDym    h[m] + nadph[m] + glu5p[m]  <=> nadp[m] + pi[m] + glu5sa[m]
       200         GLU5Km    glu_L[m] + atp[m]  <=> adp[m] + glu5p[m]
      -200     P45027A15m    o2[m] + h[m] + nadph[m] + xol7ah3[m]  <=> 2 h2o[m] + nadp[m] + xol7ah2al[m]
       400           PPAm    h2o[m] + ppi[m]  <=> h[m] + 2 pi[m]
       200          r0074    h2o[m] + nad[m] + glu5sa[m]  <=> 2 h[m] + nadh[m] + glu_L[m]
       200        RE1804M    nad[m] + xol7ah3[m]  <=> h[m] + nadh[m] + xol7ah2al[m]
       400       HMR_3966    h2o[m] + atp[m]  <=> h[m] + amp[m] + ppi[m]
      1000         ATPS4mi    adp[m] + pi[m] + 4 h[i]  <=> h2o[m] + 3 h[m] + atp[m]
       400     CYOR_u10mi    2 h[m] + 2 ficytC[m] + q10h2[m]  <=> q10[m] + 2 focytC[m] + 4 h[i]
       400    NADH2_u10mi    5 h[m] + nadh[m] + q10[m]  <=> nad[m] + q10h2[m] + 4 h[i]
       200         CYOOm2i    o2[m] + 8 h[m] + 4 focytC[m]  <=> 2 h2o[m] + 4 ficytC[m] + 4 h[i]
```

ANTICIPATED RESULTS

In a model for flux balance analysis, there might be non-zero flux through the ATP synthase reaction if all external reaction bounds are zero and all internal reactions reversible. This indicates the imporance of appropriately constrained internal reaction bounds.

Testing for activity of ATP synthase with all exchanges closed and all transport reactions reversible

Identify all of the transport reactions

```
allTransportRxnBool=transportReactionBool(model);
fprintf('%u%s\n',nnz(allTransportRxnBool),' transport reactions in total.');
```

```
4230 transport reactions in total.
```

Revert to original Recon3.0model reaction bounds

```
model.lb=modelOrig.lb;
model.ub=modelOrig.ub;
```

Open all transport reactions (which might include an external reaction, e.g., a biomass reaction)

```
model.lb(allTransportRxnBool)=-1000;
model.ub(allTransportRxnBool)=1000;
```

Close all external reactions

```
model.lb(~model.SIntRxnBool)=0;
model.ub(~model.SIntRxnBool)=0;
```

Run sparse flux balance analysis on the model with all exchanges closed

```
sparseFBAsolutionBounded = optimizeCbModel(model, osenseStr, minNorm, allowLoops, zeroN
```

Check to see if there is a non-zero flux through the ATP synthase reaction

```
fprintf('%g%s\n',sparseFBAsolutionBounded.v(atpsynthaseBool),' flux through the ATP syn
```

```
1000 flux through the ATP synthase reaction
```

Display the sparse flux solution, but only the non-zero fluxes, above a specified cutoff.

```
cutoff=feasTol;
for n=1:nRxn
    if abs(sparseFBAsolutionBounded.v(n))>cutoff
        formula=printRxnFormula(model, model.rxns{n}, 0);
        fprintf('%10g%15s\t%-60s\n',sparseFBAsolutionBounded.v(n),model.rxns{n}, formul
    end
end
```

```
       500          ADK1m    atp[m] + amp[m]  <=> 2 adp[m]
       500         LDH_Lm    nad[m] + lac_L[m]  <=> h[m] + nadh[m] + pyr[m]
      -500       L_LACDcm    2 ficytC[m] + lac_L[c]  <=> 2 h[c] + pyr[c] + 2 focytC[m]
       500       L_LACtcm    lac_L[c]  <=> lac_L[m]
       500           PPAm    h2o[m] + ppi[m]  -> h[m] + 2 pi[m]
      -500         PYRt2m    h[c] + pyr[c]  <=> h[m] + pyr[m]
      -250          r2398    h[c] + lys_L[m] + citr_L[c]  <=> h[m] + lys_L[c] + citr_L[m]
      -250          r2410    h[c] + lys_L[c] + citr_L[m]  <=> h[m] + lys_L[m] + citr_L[c]
       500       HMR_3966    h2o[m] + atp[m]  -> h[m] + amp[m] + ppi[m]
      1000        ATPS4mi    adp[m] + pi[m] + 4 h[i]  <=> h2o[m] + 3 h[m] + atp[m]
       500     CYOR_u10mi    2 h[m] + 2 ficytC[m] + q10h2[m]  <=> q10[m] + 2 focytC[m] + 4 h[i]
       500    NADH2_u10mi    5 h[m] + nadh[m] + q10[m]  <=> nad[m] + q10h2[m] + 4 h[i]
```

Testing for activity of ATP synthase with all exchanges closed and all mitochondrial transport reactions reversible

Identify all of the transport reactions involving the cytoplasm and mitochondrial matrix

```
originCompartment='c';
destinationCompartment='m';
unidirectionalBool=0;
cmTransportRxnBool=transportReactionBool(model,originCompartment,destinationCompartment
fprintf('%u%s\n',nnz(cmTransportRxnBool),' transport reactions involving the cytoplasm
```

```
491 transport reactions involving the cytoplasm and mitochondrial matrix.
```

Revert to original Recon3.0model reaction bounds

```
model.lb=modelOrig.lb;
model.ub=modelOrig.ub;
```

Open all transport reactions (which might include an external reaction, e.g., a biomass reaction)

```
model.lb(cmTransportRxnBool)=-1000;
model.ub(cmTransportRxnBool)=1000;
```

Close all external reactions

```
model.lb(~model.SIntRxnBool)=0;
model.ub(~model.SIntRxnBool)=0;
```

Run sparse flux balance analysis on the model

```
sparseFBAsolutionBounded = optimizeCbModel(model, osenseStr, minNorm, allowLoops, zeroN
```

Check to see if there is a non-zero flux through the ATP synthase reaction

```
fprintf('%g%s\n',sparseFBAsolutionBounded.v(atpsynthaseBool),' flux through the ATP syn
```

```
1000 flux through the ATP synthase reaction
```

Display the sparse flux solution, but only the non-zero fluxes, above a specified cutoff.

```
cutoff=feasTol;
for n=1:nRxn
    if abs(sparseFBAsolutionBounded.v(n))>cutoff
        formula=printRxnFormula(model, model.rxns{n}, 0);
        fprintf('%10g%15s\t%-60s\n',sparseFBAsolutionBounded.v(n),model.rxns{n}, formul
    end
end
```

```
       500          ADK1m    atp[m] + amp[m]  <=> 2 adp[m]
       500         LDH_Lm    nad[m] + lac_L[m]  <=> h[m] + nadh[m] + pyr[m]
      -500       L_LACDcm    2 ficytC[m] + lac_L[c]  <=> 2 h[c] + pyr[c] + 2 focytC[m]
       500       L_LACtcm    lac_L[c]  <=> lac_L[m]
       500           PPAm    h2o[m] + ppi[m]  -> h[m] + 2 pi[m]
      -500         PYRt2m    h[c] + pyr[c]  <=> h[m] + pyr[m]
      -250          r2402    h[c] + arg_L[m] + citr_L[c]  <=> h[m] + arg_L[c] + citr_L[m]
      -250          r2411    h[c] + arg_L[c] + citr_L[m]  <=> h[m] + arg_L[m] + citr_L[c]
       500       HMR_3966    h2o[m] + atp[m]  -> h[m] + amp[m] + ppi[m]
      1000        ATPS4mi    adp[m] + pi[m] + 4 h[i]  -> h2o[m] + 3 h[m] + atp[m]
       500     CYOR_u10mi    2 h[m] + 2 ficytC[m] + q10h2[m]  -> q10[m] + 2 focytC[m] + 4 h[i]
       500    NADH2_u10mi    5 h[m] + nadh[m] + q10[m]  -> nad[m] + q10h2[m] + 4 h[i]
```

Testing for activity of ATP synthase with all exchanges closed and all plasma membrane transport reactions reversible

Identify all of the transport reactions across the plasma membrane

```
originCompartment='e';
destinationCompartment='c';
unidirectionalBool=0;
ecTransportRxnBool=transportReactionBool(model,originCompartment,destinationCompartment
fprintf('%u%s\n',nnz(ecTransportRxnBool),' transport reactions across the plasma membra
```

Revert to original Recon3.0model reaction bounds

```
model.lb=modelOrig.lb;
model.ub=modelOrig.ub;
```

Open all transport reactions (which might include an external reaction, e.g., a biomass reaction)

```
model.lb(ecTransportRxnBool)=-1000;
model.ub(ecTransportRxnBool)=1000;
```

Close all external reactions

```
model.lb(~model.SIntRxnBool)=0;
model.ub(~model.SIntRxnBool)=0;
```

Run sparse flux balance analysis on the model

```
sparseFBAsolutionBounded = optimizeCbModel(model, osenseStr, minNorm, allowLoops, zeroN
```

Check to see if there is a non-zero flux through the ATP synthase reaction

```
fprintf('%g%s\n',sparseFBAsolutionBounded.v(atpsynthaseBool),' flux through the ATP syn
```

Display the sparse flux solution, but only the non-zero fluxes, above a specified cutoff.

```
cutoff=feasTol;
for n=1:nRxn
    if abs(sparseFBAsolutionBounded.v(n))>cutoff
        formula=printRxnFormula(model, model.rxns{n}, 0);
        fprintf('%10g%15s\t%-60s\n',sparseFBAsolutionBounded.v(n),model.rxns{n}, formul
    end
end
```

Testing for activity of ATP synthase with all exchanges closed and peroxisomal transport reactions reversible

Identify all of the transport reactions across the plasma membrane

```
originCompartment='c';
destinationCompartment='x';
unidirectionalBool=0;
cxTransportRxnBool=transportReactionBool(model,originCompartment,destinationCompartment
fprintf('%u%s\n',nnz(cxTransportRxnBool),' transport reactions across the peroxisome me
```

Revert to original Recon3.0model reaction bounds

```
model.lb=modelOrig.lb;
model.ub=modelOrig.ub;
```

Open all transport reactions (which might include an external reaction, e.g., a biomass reaction)

```
model.lb(cxTransportRxnBool)=-1000;
model.ub(cxTransportRxnBool)=1000;
```

Close all external reactions

```
model.lb(~model.SIntRxnBool)=0;
model.ub(~model.SIntRxnBool)=0;
```

Run sparse flux balance analysis on the model

```
sparseFBAsolutionBounded = optimizeCbModel(model, osenseStr, minNorm, allowLoops, zeroN
```

Check to see if there is a non-zero flux through the ATP synthase reaction

```
fprintf('%g%s\n',sparseFBAsolutionBounded.v(atpsynthaseBool),' flux through the ATP syn
```

Display the sparse flux solution, but only the non-zero fluxes, above a specified cutoff.

```
cutoff=feasTol;
for n=1:nRxn
    if abs(sparseFBAsolutionBounded.v(n))>cutoff
        formula=printRxnFormula(model, model.rxns{n}, 0);
        fprintf('%10g%15s\t%-60s\n',sparseFBAsolutionBounded.v(n),model.rxns{n}, formul
    end
end
```

Testing for activity of ATP synthase with all exchanges closed and lysosomal transport reactions reversible

Identify all of the transport reactions across the plasma membrane

```
originCompartment='c';
destinationCompartment='l';
unidirectionalBool=0;
clTransportRxnBool=transportReactionBool(model,originCompartment,destinationCompartment
fprintf('%u%s\n',nnz(clTransportRxnBool),' transport reactions across the lysosomal mem
```

Revert to original Recon3.0model reaction bounds

```
model.lb=modelOrig.lb;
model.ub=modelOrig.ub;
```

Open all transport reactions (which might include an external reaction, e.g., a biomass reaction)

```
model.lb(clTransportRxnBool)=-1000;
model.ub(clTransportRxnBool)=1000;
```

Close all external reactions

```
model.lb(~model.SIntRxnBool)=0;
model.ub(~model.SIntRxnBool)=0;
```

Run sparse flux balance analysis on the model

```
sparseFBAsolutionBounded = optimizeCbModel(model, osenseStr, minNorm, allowLoops, zeroN
```

Check to see if there is a non-zero flux through the ATP synthase reaction

```
fprintf('%g%s\n',sparseFBAsolutionBounded.v(atpsynthaseBool),' flux through the ATP syn
```

Display the sparse flux solution, but only the non-zero fluxes, above a specified cutoff.

```
cutoff=feasTol;
for n=1:nRxn
    if abs(sparseFBAsolutionBounded.v(n))>cutoff
        formula=printRxnFormula(model, model.rxns{n}, 0);
        fprintf('%10g%15s\t%-60s\n',sparseFBAsolutionBounded.v(n),model.rxns{n}, formul
    end
end
```

REFERENCES

[fleming_cardinality_nodate] Fleming, R.M.T., et al., Cardinality optimisation in constraint-based modelling: illustration with Recon 3D (submitted), 2017.

[sparsePaper] Le Thi, H.A., Pham Dinh, T., Le, H.M., and Vo, X.T. (2015). DC approximation approaches for sparse optimization. European Journal of Operational Research 244, 26–46.