OptForce Tutorial

Author: Sebastián N. Mendoza, Center for Mathematical Modeling, University of Chile. snmendoz@uc.cl

Reviewers(s): Chiam Yu Ng (Costas D. Maranas group), Lin Wang (Costas D. Maranas group), John Sauls

INTRODUCTION:

In this tutorial we will run optForce. For a detailed description of the procedure, please see [1]. Briefly, the problem is to find a set of interventions of size "K" such that when these interventions are applied to a wild-type strain, the mutant created will produce a particular target of interest in a higher rate than the wild-type strain. The interventions could be knockouts (lead to zero the flux for a particular reaction), upregulations (increase the flux for a particular reaction) and downregulations (decrease the flux for a particular reaction).

For example, imagine that we would like to increase the production of succinate in Escherichia coli. Which are the interventions needed to increase the production of succinate? We will approach this problem in this tutorial and we will see how each of the steps of OptForce are solved.

MATERIALS

EQUIPMENT

1. MATLAB
2. A solver for Mixed Integer Linear Programming (MILP) problems. For example, Gurobi.

EQUIPMENT SETUP

Use `changeCobraSolver` to choose the solver for MILP problems.

PROCEDURE

The proceduce consists on the following steps

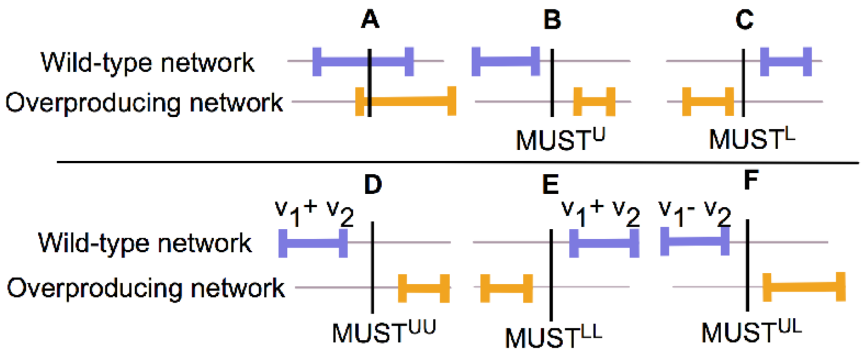
1) Maximize specific growth rate and product formation.

2) Define constraints for both wild-type and mutant strain:

3) Perform flux variability analysis for both wild-type and mutant strain.

4) Find must sets, i.e, reactions that MUST increase or decrease their flux in order to achieve the phenotype in the mutant strain.

Figure 1.



5) Find the interventions needed that will ensure a increased production of the target of interest

Now, we will approach each step in detail.

STEP 1: Maximize specific growth rate and product formation

First, we load the model. This model comprises only 90 reactions, which describe the central metabolism of E. coli [2].

Then, we change the objective function to maximize biomass ("R75"). We also change the lower bounds, so E. coli will be able to consume glucose, oxygen, sulfate, ammomium, citrate and glycerol.

```
changeCobraSolver('gurobi', 'ALL');
```

```
> Gurobi interface added to MATLAB path.
> Solver for LP problems has been set to gurobi.

> Gurobi interface added to MATLAB path.
> Solver for MILP problems has been set to gurobi.

> Gurobi interface added to MATLAB path.
> Solver for QP problems has been set to gurobi.

> Gurobi interface added to MATLAB path.
> Solver for MIQP problems has been set to gurobi.
> Solver gurobi not supported for problems of type NLP. Currently used: matlab
```

```
pathTutorial = which('tutorial_optForce.mlx');
pathstr = fileparts(pathTutorial);
cd(pathstr)

load('AntCore.mat');
model.c(strcmp(model.rxns, 'R75')) = 1;
model = changeRxnBounds(model, 'EX_gluc', -100, 'l');
model = changeRxnBounds(model, 'EX_o2', -100, 'l');
model = changeRxnBounds(model, 'EX_so4', -100, 'l');
model = changeRxnBounds(model, 'EX_nh3', -100, 'l');
model = changeRxnBounds(model, 'EX_cit', -100, 'l');
model = changeRxnBounds(model, 'EX_glyc', -100, 'l');
```

Then, we calculate the maximum specific growth rate and the maximum production rate for succinate.

```
growthRate = optimizeCbModel(model);
fprintf('The maximum growth rate is %1.2f', growthRate.f);
```

```
The maximum growth rate is 14.36
```

```
model = changeObjective(model, 'EX_suc');
maxSucc = optimizeCbModel(model);
fprintf('The maximum production rate of succinate is %1.2f', maxSucc.f);
```

```
The maximum production rate of succinate is 155.56
```

TIP: The biomass reaction is usually set to 1%-10% of maximum theoretical biomass yield when running the following steps, to prevent solutions without biomass formation.

1. Maximizing product formation
2. Finding MUST sets of second order
3. Finding FORCE sets

STEP 2: Define constraints for both wild-type and mutant strain

TIMING: This step should take a few days or weeks, depending on the information available for your species.

CRITICAL STEP: This is a manual task, so you should search for information in articles or even perform your own experiments. You can also make assumptions for describing the phenotypes of both strains which will make this task a little faster but make sure to have two strains different enough, because you should be able to find differences in reactions ranges.

We define constraints for each strain as follows:

1. The WT strain's biomass function ("R75") is constrained to near the maximum growth rate.
2. The mutant strain's biomass function is set to zero. Succinate export ('EX_suc') is forced to be the maximum as calculated previously.

```
constrWT = struct('rxnList', {{'R75'}}, 'rxnValues', 14, 'rxnBoundType', 'b')
```

```
constrWT = 
         rxnList: {'R75'}
       rxnValues: 14
    rxnBoundType: 'b'
```

```
constrMT = struct('rxnList', {{'R75', 'EX_suc'}}, 'rxnValues', [0, 155.55], ...
                  'rxnBoundType', 'b')
```

```
constrMT = 
         rxnList: {'R75'  'EX_suc'}
       rxnValues: [0 155.5500]
    rxnBoundType: 'b'
```

Step 3: Flux Variability Analysis

TIMING: This task should take from a few seconds to a few hours depending on the size of your reconstruction

We run the FVA analysis for both strains

```
[minFluxesW, maxFluxesW, minFluxesM, maxFluxesM, ~, ~] = FVAOptForce(model, ...
                                                                       constrWT, constrMT);
```

```
ans =
    'R75'    'EX_suc'

ans = 1
ans = 1
ans = 1
```

```
disp([minFluxesW, maxFluxesW, minFluxesM, maxFluxesM]);
```

```
  -90.1251   97.1300 -162.5000  100.0000
         0   86.0700         0  100.0000
         0   86.0700         0  100.0000
  -56.1567   86.0700 -200.0000   75.0000
   21.3033  163.5300 -100.0000  175.0000
   -3.0777  154.8640 -100.0000  175.0000
         0  151.5086         0  175.0000
         0  187.2551         0  262.5000
         0  169.5163         0  262.5000
  -10.0660  102.9449         0  175.0000
   10.0660   66.5714         0   87.5000
  -10.0660  102.9449         0  175.0000
  -48.9454    7.5600  -87.5000         0
  -53.9994    2.5060  -87.5000         0
  -53.9994    2.5060  -87.5000         0
   -2.5060   53.9994         0   87.5000
         0   86.0700         0  100.0000
         0   86.0700         0  100.0000
    9.7020  114.6466         0  175.0000
         0   56.5564         0   87.5000
   16.0264  145.2048         0  175.0000
   16.0264  145.2048         0  175.0000
    0.9344  130.1128         0  175.0000
   -5.6736  123.5048         0  175.0000
         0  118.0576         0  175.0000
    5.1940  123.2516         0  175.0000
  -98.1150  123.2516 -175.0000  175.0000
         0  151.5086         0  175.0000
         0  151.5086         0  175.0000
         0  254.5400         0  925.0000
         0  253.2493         0  925.0000
   -7.1960   94.6056         0  175.0000
         0   84.8467         0  282.6087
         0   84.8467         0  282.6087
         0  175.1064         0  336.3636
         0  175.1064         0  336.3636
   91.4130  107.1280         0         0
    9.4500    9.4500         0         0
    2.9400    2.9400         0         0
    3.9340    3.9340         0         0
   25.4520   56.8820         0         0
    3.2060    3.2060         0         0
    6.8320    6.8320         0         0
         0   15.7150         0         0
   -6.8880    8.8270         0         0
    0.6790   16.3940         0         0
         0   31.4300         0         0
```

```
    3.2620    3.2620         0         0
    4.5640    4.5640         0         0
    4.5640    4.5640         0         0
    7.2380   38.6680         0         0
    2.0440    2.0440         0         0
    5.6280    5.6280         0         0
    5.9920    5.9920         0         0
    3.8640    3.8640         0         0
    2.4640    2.4640         0         0
    1.8340    1.8340         0         0
    0.7560    0.7560         0         0
    1.2600    1.2600         0         0
    2.0440    2.0440         0         0
    1.2600    1.2600         0         0
   79.7324  200.0000         0  200.0000
         0  118.0576         0  175.0000
  -39.5563  353.9124 -342.8571  175.0000
         0  253.2493         0  925.0000
   40.6268  100.0000         0  100.0000
   15.0890  100.0000         0  100.0000
 -100.0000   84.8467 -100.0000  282.6087
         0  175.1064         0  336.3636
         0  101.8016         0  175.0000
  134.9718  407.3274         0  525.0000
   62.1267  100.0000         0  100.0000
   97.4820   97.4820         0         0
    3.2620    3.2620         0         0
   14.0000   14.0000         0         0
         0  175.1064         0  336.3636
  134.9718  407.3274         0  525.0000
         0  101.8016         0  175.0000
         0  253.2493         0  925.0000
 -100.0000  -40.6268 -100.0000         0
 -100.0000  -15.0890 -100.0000         0
 -100.0000   84.8467 -100.0000  282.6087
  -97.4820  -97.4820         0         0
 -100.0000  -62.1267 -100.0000         0
   -3.2620   -3.2620         0         0
         0  105.4230         0  155.5556
         0  105.4230         0  155.5556
   11.6200   11.6200         0         0
    5.0540    5.0540         0         0
    5.9920    5.9920         0         0
```

Now, the run the next step of OptForce.

Step 4: Find Must Sets

TIMING: This task should take from a few seconds to a few hours depending on the size of your reconstruction

First, we define an ID for this run. Each time you run the functions associated to the optForce procedure, some folders can be generated to store inputs used in that run. Outputs are stored as well. These folders will be located inside the folder defined by your run ID. Thus, if your runID is "TestOptForce", the structure of the folders will be the following:

```
├── CurrentFolder
|  ├── TestOptForce
|  |  ├── Inputs
```

```
| | └── Outputs
```

To avoid the generation of inputs and outputs folders, set `keepInputs = 0`, `printExcel = 0` and `printText = 0`.

Also, a report of the run is generated each time you run the functions associated to the optForce procedure. So, the idea is to give a different `runID` each time you run the functions, so you will be able to see the report (inputs used, outputs generated, errors in the run) for each run.

We define then our `runID`.

```
runID = 'TestOptForceM';
```

Fow now, only functions to find first and second order must sets are supported in this third step. As depicted in **Figure 1**, the first order must sets are MUSTU and MUSTL; and second order must sets are MUSTUU, MUSTLL and MUSTUL.

A) Finding first order must sets

We define constraints.

```
constrOpt = struct('rxnList', {{'EX_gluc', 'R75', 'EX_suc'}}, 'values', [-100, 0, 155.5]');
```

We then run the functions `findMustL` and `findMustU` that will allow us to find `mustU` and `mustL` sets, respectively.

i) MustL Set:

```
[mustLSet, pos_mustL] = findMustL(model, minFluxesW, maxFluxesW, 'constrOpt', constrOpt, ...
                                  'runID', runID, 'outputFolder', 'OutputsFindMustL', ...
                                  'outputFileName', 'MustL' , 'printExcel', 1, 'printText', 1,
                                  'printReport', 1, 'keepInputs', 1, 'verbose', 0);
```

```
Optimize a model with 710 rows, 798 columns and 2715 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 564 rows and 482 columns
Presolve time: 0.02s
Presolved: 146 rows, 316 columns, 957 nonzeros
Variable types: 273 continuous, 43 integer (43 binary)

Root relaxation: objective 9.748200e+01, 161 iterations, 0.01 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

*    0     0               0      97.4820000   97.48200  0.00%     -    0s

Explored 0 nodes (161 simplex iterations) in 0.05 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 97.482
Pool objective bound 97.482

Optimal solution found (tolerance 1.00e-12)
Best objective 9.748200000003e+01, best bound 9.748200000003e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2710 nonzeros
```

```
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 564 rows and 483 columns
Presolve time: 0.02s
Presolved: 146 rows, 315 columns, 954 nonzeros
Variable types: 273 continuous, 42 integer (42 binary)

Root relaxation: objective 9.141300e+01, 175 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

*    0     0               0      91.4130000   91.41300  0.00%     -    0s

Explored 0 nodes (175 simplex iterations) in 0.04 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 91.413
Pool objective bound 91.413

Optimal solution found (tolerance 1.00e-12)
Best objective 9.141300000002e+01, best bound 9.141300000002e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2705 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 564 rows and 484 columns
Presolve time: 0.02s
Presolved: 146 rows, 314 columns, 951 nonzeros
Variable types: 273 continuous, 41 integer (41 binary)

Root relaxation: objective 2.545200e+01, 174 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

*    0     0               0      25.4520000   25.45200  0.00%     -    0s

Explored 0 nodes (174 simplex iterations) in 0.04 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 25.452
Pool objective bound 25.452

Optimal solution found (tolerance 1.00e-12)
Best objective 2.545200000001e+01, best bound 2.545200000001e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2700 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 564 rows and 485 columns
Presolve time: 0.01s
Presolved: 146 rows, 313 columns, 948 nonzeros
Variable types: 273 continuous, 40 integer (40 binary)

Root relaxation: objective 1.162000e+01, 160 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
```

```
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

*    0     0               0      11.6200000   11.62000  0.00%     -    0s

Explored 0 nodes (160 simplex iterations) in 0.05 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 11.62
Pool objective bound 11.62

Optimal solution found (tolerance 1.00e-12)
Best objective 1.162000000003e+01, best bound 1.162000000003e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2695 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 564 rows and 486 columns
Presolve time: 0.02s
Presolved: 146 rows, 312 columns, 945 nonzeros
Variable types: 273 continuous, 39 integer (39 binary)

Root relaxation: objective 1.000350e+01, 186 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

*    0     0               0      10.0035000   10.00350  0.00%     -    0s

Explored 0 nodes (186 simplex iterations) in 0.04 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 10.0035
Pool objective bound 10.0035

Optimal solution found (tolerance 1.00e-12)
Best objective 1.000350000000e+01, best bound 1.000350000000e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2690 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 567 rows and 488 columns
Presolve time: 0.01s
Presolved: 143 rows, 310 columns, 933 nonzeros
Variable types: 272 continuous, 38 integer (38 binary)

Root relaxation: objective 9.450000e+00, 170 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

*    0     0               0       9.4500000    9.45000  0.00%     -    0s

Explored 0 nodes (170 simplex iterations) in 0.04 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 9.45
Pool objective bound 9.45

Optimal solution found (tolerance 1.00e-12)
Best objective 9.450000000029e+00, best bound 9.450000000029e+00, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2685 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
```

```
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 567 rows and 489 columns
Presolve time: 0.01s
Presolved: 143 rows, 309 columns, 930 nonzeros
Variable types: 272 continuous, 37 integer (37 binary)

Root relaxation: objective 7.238000e+00, 166 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

*    0     0               0       7.2380000    7.23800  0.00%     -    0s

Explored 0 nodes (166 simplex iterations) in 0.05 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 7.238
Pool objective bound 7.238

Optimal solution found (tolerance 1.00e-12)
Best objective 7.238000000012e+00, best bound 7.238000000012e+00, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2680 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 567 rows and 490 columns
Presolve time: 0.01s
Presolved: 143 rows, 308 columns, 927 nonzeros
Variable types: 272 continuous, 36 integer (36 binary)

Root relaxation: objective 6.832000e+00, 157 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

*    0     0               0       6.8320000    6.83200  0.00%     -    0s

Explored 0 nodes (157 simplex iterations) in 0.04 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 6.832
Pool objective bound 6.832

Optimal solution found (tolerance 1.00e-12)
Best objective 6.832000000029e+00, best bound 6.832000000029e+00, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2675 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 567 rows and 491 columns
Presolve time: 0.01s
Presolved: 143 rows, 307 columns, 924 nonzeros
Variable types: 272 continuous, 35 integer (35 binary)

Root relaxation: objective 5.992000e+00, 184 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time
```

```
*    0     0               0       5.9920000    5.99200  0.00%     -    0s

Explored 0 nodes (184 simplex iterations) in 0.04 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 5.992
Pool objective bound 5.992

Optimal solution found (tolerance 1.00e-12)
Best objective 5.992000000029e+00, best bound 5.992000000029e+00, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2670 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 567 rows and 492 columns
Presolve time: 0.01s
Presolved: 143 rows, 306 columns, 921 nonzeros
Variable types: 272 continuous, 34 integer (34 binary)

Root relaxation: objective 5.992000e+00, 157 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

*    0     0               0       5.9920000    5.99200  0.00%     -    0s
```

Note that the folder "TestOptForceM" was created. Inside this folder, two additional folders were created: "InputsMustL" and "OutputsMustL". In the inputs folder you will find all the inputs required to run the the function `findMustL`. Additionally, in the outputs folder you will find the `mustL` set found, which were saved in two files (.xls and .txt). Furthermore, a report which summarize all the inputs and outputs used during your running was generated. The name of the report will be in this format "report-Day-Month-Year-Hour-Minutes". So, you can mantain a chronological order of your experiments.

We display the reactions that belongs to the `mustL` set.

```
disp(mustLSet)
```

```
    'R11'
    'R26'
    'R37'
    'R38'
    'R39'
    'R40'
    'R41'
    'R42'
    'R43'
    'R46'
    'R48'
    'R49'
    'R50'
    'R51'
    'R52'
    'R53'
    'R54'
    'R55'
    'R56'
    'R57'
    'R58'
    'R59'
```

```
    'R60'
    'R61'
    'R73'
    'R74'
    'PSEUDOpyr_1'
    'PSEUDOpep_1'
    'PSEUDOco2_1'
```

ii) MustU set:

```
[mustUSet, pos_mustU] = findMustU(model, minFluxesW, maxFluxesW, 'constrOpt', constrOpt, ...
                                  'runID', runID, 'outputFolder', 'OutputsFindMustU', ...
                                  'outputFileName', 'MustU' , 'printExcel', 1, 'printText', 1,
                                  'printReport', 1, 'keepInputs', 1, 'verbose', 0);
```

```
Optimize a model with 710 rows, 798 columns and 2769 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 473 rows and 451 columns
Presolve time: 0.02s
Presolved: 237 rows, 347 columns, 1238 nonzeros
Variable types: 299 continuous, 48 integer (48 binary)

Root relaxation: objective 1.063553e+02, 182 iterations, 0.01 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  106.35533    0    2          -  106.35533      -     -    0s
*    0     0               0      97.4820000   97.48200  0.00%     -    0s

Cutting planes:
  Gomory: 1
  Implied bound: 1

Explored 0 nodes (204 simplex iterations) in 0.05 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 97.482
Pool objective bound 97.482

Optimal solution found (tolerance 1.00e-12)
Best objective 9.748199999997e+01, best bound 9.748199999997e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2764 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 473 rows and 452 columns
Presolve time: 0.03s
Presolved: 237 rows, 346 columns, 1235 nonzeros
Variable types: 299 continuous, 47 integer (47 binary)

Root relaxation: objective 1.063553e+02, 194 iterations, 0.01 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time
```

```
     0     0  106.35533    0    2          -  106.35533      -     -    0s
*    0     0               0      50.0770000   50.07700  0.00%     -    0s

Cutting planes:
  Gomory: 1
  Implied bound: 1

Explored 0 nodes (225 simplex iterations) in 0.07 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 50.077
Pool objective bound 50.077

Optimal solution found (tolerance 1.00e-12)
Best objective 5.007699999997e+01, best bound 5.007699999997e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2759 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 473 rows and 453 columns
Presolve time: 0.02s
Presolved: 237 rows, 345 columns, 1231 nonzeros
Variable types: 299 continuous, 46 integer (46 binary)

Root relaxation: objective 1.063553e+02, 191 iterations, 0.01 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  106.35533    0    2          -  106.35533      -     -    0s
*    0     0               0      31.9951818   31.99518  0.00%     -    0s

Cutting planes:
  Gomory: 1
  Implied bound: 1

Explored 0 nodes (235 simplex iterations) in 0.09 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 31.9952
Pool objective bound 31.9952

Optimal solution found (tolerance 1.00e-12)
Best objective 3.199518181815e+01, best bound 3.199518181815e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2754 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 477 rows and 455 columns
Presolve time: 0.02s
Presolved: 233 rows, 343 columns, 1216 nonzeros
Variable types: 298 continuous, 45 integer (45 binary)

Root relaxation: objective 1.063553e+02, 186 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  106.35533    0    2          -  106.35533      -     -    0s
*    0     0               0      25.3871818   25.38718  0.00%     -    0s

Cutting planes:
```

```
  Gomory: 1
  Implied bound: 1

Explored 0 nodes (207 simplex iterations) in 0.06 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 25.3872
Pool objective bound 25.3872

Optimal solution found (tolerance 1.00e-12)
Best objective 2.538718181814e+01, best bound 2.538718181814e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2749 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 481 rows and 457 columns
Presolve time: 0.02s
Presolved: 229 rows, 341 columns, 1201 nonzeros
Variable types: 297 continuous, 44 integer (44 binary)

Root relaxation: objective 1.063553e+02, 197 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  106.35533    0    2          -  106.35533      -     -    0s
     0     0   13.39362    0    2          -   13.39362      -     -    0s
*    0     0               0      13.3936250   13.39363  0.00%     -    0s

Cutting planes:
  Gomory: 1
  Implied bound: 1

Explored 0 nodes (417 simplex iterations) in 0.07 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 13.3936
Pool objective bound 13.3936

Optimal solution found (tolerance 1.00e-12)
Best objective 1.339362500000e+01, best bound 1.339362500000e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2744 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 484 rows and 459 columns
Presolve time: 0.03s
Presolved: 226 rows, 339 columns, 1189 nonzeros
Variable types: 296 continuous, 43 integer (43 binary)

Root relaxation: objective 1.063553e+02, 164 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  106.35533    0    2          -  106.35533      -     -    0s
*    0     0               0      13.3936250   13.39362  0.00%     -    0s

Cutting planes:
  Gomory: 1
  Implied bound: 1
```

```
Explored 0 nodes (201 simplex iterations) in 0.06 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 13.3936
Pool objective bound 13.3936

Optimal solution found (tolerance 1.00e-12)
Best objective 1.339362499996e+01, best bound 1.339362499996e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2739 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 487 rows and 461 columns
Presolve time: 0.02s
Presolved: 223 rows, 337 columns, 1177 nonzeros
Variable types: 295 continuous, 42 integer (42 binary)

Root relaxation: objective 1.063553e+02, 176 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  106.35533    0    2          -  106.35533      -     -    0s
*    0     0               0      13.3936250   13.39362  0.00%     -    0s

Cutting planes:
  Gomory: 1
  Implied bound: 1

Explored 0 nodes (207 simplex iterations) in 0.07 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 13.3936
Pool objective bound 13.3936

Optimal solution found (tolerance 1.00e-12)
Best objective 1.339362499997e+01, best bound 1.339362499997e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2734 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 497 rows and 468 columns
Presolve time: 0.02s
Presolved: 213 rows, 330 columns, 1138 nonzeros
Variable types: 291 continuous, 39 integer (39 binary)

Root relaxation: objective 1.063553e+02, 192 iterations, 0.01 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  106.35533    0    2          -  106.35533      -     -    0s
     0     0   29.81968    0    3          -   29.81968      -     -    0s
*    0     0               0      13.3936250   13.39363  0.00%     -    0s

Cutting planes:
  Gomory: 1
  Implied bound: 2
  Flow cover: 1

Explored 0 nodes (232 simplex iterations) in 0.08 seconds
Thread count was 4 (of 4 available processors)
```

```
Solution count 1: 13.3936
Pool objective bound 13.3936

Optimal solution found (tolerance 1.00e-12)
Best objective 1.339362500000e+01, best bound 1.339362500000e+01, gap 0.0000%
Optimize a model with 710 rows, 798 columns and 2729 nonzeros
Variable types: 708 continuous, 90 integer (90 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [5e-01, 1e+03]
Presolve removed 493 rows and 465 columns
Presolve time: 0.02s
Presolved: 217 rows, 333 columns, 1153 nonzeros
Variable types: 293 continuous, 40 integer (40 binary)

Root relaxation: objective 1.063553e+02, 177 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
```

Note that the folders "InputsMustU" and "OutputsFindMustU" were created. These folders contain the inputs and outputs of `findMustU`, respectively.

We display the reactions that belongs to the `mustU` set.

```
disp(mustUSet)
```

```
    'R21'
    'R22'
    'R23'
    'R24'
    'R33'
    'R34'
    'R35'
    'R36'
    'R69'
    'EX_pdo'
    'EX_nh3'
    'EX_so4'
    'SUCt'
```

B) Finding second order must sets

First, we define the reactions that will be excluded from the analysis. It is suggested to include in this list the reactions found in the previous step as well as exchange reactions.

```
constrOpt = struct('rxnList', {{'EX_gluc', 'R75', 'EX_suc'}}, 'values', [-100, 0, 155.5]');
exchangeRxns = model.rxns(cellfun(@isempty, strfind(model.rxns, 'EX_')) == 0);
excludedRxns = unique([mustUSet; mustLSet; exchangeRxns]);
```

Now, we run the functions for finding second order must sets.

i) MustUU:

```
[mustUU, pos_mustUU, mustUU_linear, pos_mustUU_linear] = ...
    findMustUU(model, minFluxesW, maxFluxesW, 'constrOpt', constrOpt, ...
               'excludedRxns', excludedRxns,'runID', runID, ...
               'outputFolder', 'OutputsFindMustUU', 'outputFileName', 'MustUU', ...
               'printExcel', 1, 'printText', 1, 'printReport', 1, 'keepInputs', 1, ...
```

```
                'verbose', 1);
```

```
Optimize a model with 1165 rows, 980 columns and 4128 nonzeros
Variable types: 800 continuous, 180 integer (180 binary)
Coefficient statistics:
  Matrix range     [5e-02, 2e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e-01, 2e+03]
Presolve removed 799 rows and 575 columns
Presolve time: 0.03s
Presolved: 366 rows, 405 columns, 1668 nonzeros
Variable types: 327 continuous, 78 integer (78 binary)

Root relaxation: objective 2.127107e+02, 268 iterations, 0.02 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

*    0     0               0     212.7106667  212.71067  0.00%     -    0s

Explored 0 nodes (397 simplex iterations) in 0.10 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 212.711
Pool objective bound 212.711

Optimal solution found (tolerance 1.00e-12)
Best objective 2.127106666667e+02, best bound 2.127106666667e+02, gap 0.0000%
Optimize a model with 1167 rows, 980 columns and 4132 nonzeros
Variable types: 800 continuous, 180 integer (180 binary)
Coefficient statistics:
  Matrix range     [5e-02, 2e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e-01, 2e+03]
Presolve removed 802 rows and 575 columns
Presolve time: 0.02s
Presolved: 365 rows, 405 columns, 1668 nonzeros
Variable types: 327 continuous, 78 integer (78 binary)

Root relaxation: objective 1.585013e+02, 246 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

*    0     0               0     158.5013333  158.50133  0.00%     -    0s

Explored 0 nodes (307 simplex iterations) in 0.05 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 158.501
Pool objective bound 158.501

Optimal solution found (tolerance 1.00e-12)
Best objective 1.585013333333e+02, best bound 1.585013333333e+02, gap 0.0000%
Optimize a model with 1169 rows, 980 columns and 4136 nonzeros
Variable types: 800 continuous, 180 integer (180 binary)
Coefficient statistics:
  Matrix range     [5e-02, 2e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e-01, 2e+03]
Presolve removed 803 rows and 575 columns
Presolve time: 0.02s
Presolved: 366 rows, 405 columns, 1674 nonzeros
Variable types: 327 continuous, 78 integer (78 binary)
```

```
Root relaxation: objective 1.237373e+02, 247 iterations, 0.00 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  123.73733    0    4          -  123.73733      -     -    0s
     0     0 infeasible    0               - infeasible      -     -    0s

Cutting planes:
  Gomory: 2
  MIR: 1
  Flow cover: 1

Explored 0 nodes (303 simplex iterations) in 0.07 seconds
Thread count was 4 (of 4 available processors)

Solution count 0

Model is infeasible
Best objective -, best bound -, gap -
a MustUU set was found
MustUU set was printed in MustUU.txt
MustUU set was also printed in MustUU_Info.txt
```

Note that the folders "InputsMustUU" and "OutputsFindMustUU" were created. These folders contain the inputs and outputs of `findMustUU`, respectively.

We display the reactions that belongs to the `mustUU` set

```
disp(mustUU);
```

```
    'R30'    'R65'
    'R31'    'R65'
```

ii) MustLL:

```
[mustLL, pos_mustLL, mustLL_linear, pos_mustLL_linear] = ...
    findMustLL(model, minFluxesW, maxFluxesW, 'constrOpt', constrOpt, ...
                'excludedRxns', excludedRxns,'runID', runID, ...
                'outputFolder', 'OutputsFindMustLL', 'outputFileName', 'MustLL', ...
                'printExcel', 1, 'printText', 1, 'printReport', 1, 'keepInputs', 1, ...
                'verbose', 1);
```

```
Optimize a model with 1165 rows, 980 columns and 4074 nonzeros
Variable types: 800 continuous, 180 integer (180 binary)
Coefficient statistics:
  Matrix range     [5e-02, 2e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e-01, 2e+03]
Presolve removed 799 rows and 578 columns
Presolve time: 0.02s
Presolved: 366 rows, 402 columns, 1633 nonzeros
Variable types: 324 continuous, 78 integer (78 binary)

Root relaxation: infeasible, 235 iterations, 0.01 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time
```

```
     0     0 infeasible    0               - infeasible      -     -    0s

Explored 0 nodes (235 simplex iterations) in 0.07 seconds
Thread count was 4 (of 4 available processors)

Solution count 0

Model is infeasible
Best objective -, best bound -, gap -
a MustLL set was not found
No mustLL set was not found. Therefore, no plain text file was generated
```

Note that the folders "InputsMustLL" and "OutputsFindMustLL" were created. These folders contain the inputs and outputs of `findMustLL`, respectively.

We display the reactions that belongs to the `mustLL` set. In this case, `mustLL` is an empty array because no reaction was found in the `mustLL` set.

```
disp(mustLL);
```

iii) MustUL:

```
[mustUL, pos_mustUL, mustUL_linear, pos_mustUL_linear] = ...
    findMustUL(model, minFluxesW, maxFluxesW, 'constrOpt', constrOpt, ...
                'excludedRxns', excludedRxns,'runID', runID, ...
                'outputFolder', 'OutputsFindMustUL', 'outputFileName', 'MustUL', ...
                'printExcel', 1, 'printText', 1, 'printReport', 1, 'keepInputs', 1, ...
                'verbose', 1);
```

```
Optimize a model with 1165 rows, 980 columns and 4101 nonzeros
Variable types: 800 continuous, 180 integer (180 binary)
Coefficient statistics:
  Matrix range     [5e-02, 2e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e-01, 2e+03]
Presolve removed 799 rows and 578 columns
Presolve time: 0.11s
Presolved: 366 rows, 402 columns, 1649 nonzeros
Variable types: 324 continuous, 78 integer (78 binary)

Root relaxation: objective 1.063553e+02, 243 iterations, 0.01 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  106.35533    0    2          -  106.35533      -     -    0s
     0     0 infeasible    0               - infeasible      -     -    0s

Cutting planes:
  Gomory: 1
  Flow cover: 1

Explored 0 nodes (249 simplex iterations) in 0.23 seconds
Thread count was 4 (of 4 available processors)

Solution count 0

Model is infeasible
Best objective -, best bound -, gap -
a MustUL set was not found
```

```
No mustUL set was not found. Therefore, no plain text file was generated
```

Note that the folders "InputsMustUL" and "OutputsFindMustUL" were created. These folders contain the inputs and outputs of `findMustUL`, respectively.

We display the reactions that belongs to the `mustUL` set. In this case, `mustUL` is an empty array because no reaction was found in the `mustUL` set.

```
disp(mustUL);
```

TROUBLESHOOTING 1: "I didn't find any reaction in my must sets"

TROUBLESHOOTING 2: "I got an error when running the `findMustX` functions (X = L or U or LL or UL or UU depending on the case)"

Step 5: OptForce

TIMING: This task should take from a few seconds to a few hours depending on the size of your reconstruction

We define constraints and we define K the number of interventions allowed, `nSets` the maximum number of sets to find, and `targetRxn` the reaction producing the metabolite of interest (in this case, succinate).

Additionally, we define the `mustU` set as the union of the reactions that must be upregulated in both first and second order must sets; and `mustL` set as the union of the reactions that must be downregulated in both first and second order must sets .

```
mustU = unique(union(mustUSet, mustUU));
mustL = unique(union(mustLSet, mustLL));
targetRxn = 'EX_suc';
biomassRxn = 'R75';
k = 1;
nSets = 1;
constrOpt = struct('rxnList', {{'EX_gluc','R75'}}, 'values', [-100, 0]);

[optForceSets, posOptForceSets, typeRegOptForceSets, flux_optForceSets] = ...
    optForce(model, targetRxn, biomassRxn, mustU, mustL, ...
             minFluxesW, maxFluxesW, minFluxesM, maxFluxesM, ...
             'k', k, 'nSets', nSets, 'constrOpt', constrOpt, ...
             'runID', runID, 'outputFolder', 'OutputsOptForce', ...
             'outputFileName', 'OptForce', 'printExcel', 1, 'printText', 1, ...
             'printReport', 1, 'keepInputs', 1, 'verbose', 1);
```

```
Optimize a model with 2062 rows, 1248 columns and 6270 nonzeros
Variable types: 978 continuous, 270 integer (270 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e+00, 1e+03]
Presolve removed 1228 rows and 437 columns
Presolve time: 0.05s
Presolved: 834 rows, 811 columns, 2973 nonzeros
Variable types: 678 continuous, 133 integer (133 binary)

Root relaxation: objective 1.555556e+02, 694 iterations, 0.03 seconds
```

```
    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  155.55556    0    4          -  155.55556      -     -    0s
H    0     0                       0.0000000  155.55556      -     -    0s
     0     0  155.55556    0    4    0.00000  155.55556      -     -    0s
     0     2  155.55556    0    3    0.00000  155.55556      -     -    0s

Explored 100 nodes (5400 simplex iterations) in 0.54 seconds
Thread count was 4 (of 4 available processors)

Solution count 1: 7.66478e-09
Pool objective bound 7.66478e-09

Optimal solution found (tolerance 1.00e-12)
Best objective 7.664780810046e-09, best bound 7.664780810046e-09, gap 0.0000%
set n 1 was found

optForce found 1 sets
Sets found by optForce were printed in OptForce.txt
```

Note that the folders "InputsOptForce" and "OutputsOptForce" were created. These folders contain the inputs and outputs of `optForce`, respectively.

We display the reactions found by `optForce`

```
disp(optForceSets)
```

```
    'R21'
```

The reaction found was "SUCt", i.e. a transporter for succinate (a very intuitive solution).

Next, we will increase k and we will exclude "SUCt" from upregulations to find non-intuitive solutions.

TIP: Sometimes the product is at the end of a long linear pathway. In that case, the recomendation is to also exclude most reactions on the linear pathway. Essential reactions and reactions not associated with any gene (i.e. spontaneous reacitons) should also be excluded.

We will only search for the 20 best solutions, but you can try with a higher number.

We will change the runID to save this second result (K = 2) in a diffetent folder than the previous result (K = 1)

```
k = 2;
nSets = 20;
runID = 'TestOptForceM2';
excludedRxns = struct('rxnList', {{'SUCt'}}, 'typeReg','U');
[optForceSets, posOptForceSets, typeRegOptForceSets, flux_optForceSets] = ...
    optForce(model, targetRxn, biomassRxn, mustU, mustL, ...
             minFluxesW, maxFluxesW, minFluxesM, maxFluxesM, ...
             'k', k, 'nSets', nSets, 'constrOpt', constrOpt, ...
             'excludedRxns', excludedRxns, ...
             'runID', runID, 'outputFolder', 'OutputsOptForce', ...
             'outputFileName', 'OptForce', 'printExcel', 1, 'printText', 1, ...
             'printReport', 1, 'keepInputs', 1, 'verbose', 1);
```

```
Optimize a model with 2062 rows, 1248 columns and 6270 nonzeros
Variable types: 978 continuous, 270 integer (270 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
```

```
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e+00, 1e+03]
Presolve removed 1187 rows and 440 columns
Presolve time: 0.05s
Presolved: 875 rows, 808 columns, 3052 nonzeros
Variable types: 676 continuous, 132 integer (132 binary)

Root relaxation: objective 1.555556e+02, 524 iterations, 0.03 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  155.55556    0    2          -  155.55556      -     -    0s
H    0     0                       0.0000000  155.55556      -     -    0s
     0     0  155.55556    0    4    0.00000  155.55556      -     -    0s
     0     2  155.55556    0    4    0.00000  155.55556      -     -    0s
*  180     8              27       0.0000001  155.55556      -  12.1    0s
*  218    29              25       0.0000001  155.55556      -  12.9    0s
*  551    57              35      47.8260870  155.55556   225%  13.8    0s

Cutting planes:
  Cover: 1
  Implied bound: 4
  Inf proof: 3

Explored 918 nodes (15060 simplex iterations) in 1.29 seconds
Thread count was 4 (of 4 available processors)

Solution count 4: 47.8261 7.67391e-08 7.09832e-08 7.66478e-09
Pool objective bound 47.8261

Optimal solution found (tolerance 1.00e-12)
Best objective 4.782608696419e+01, best bound 4.782608696419e+01, gap 0.0000%
set n 1 was found
Optimize a model with 2063 rows, 1248 columns and 6272 nonzeros
Variable types: 978 continuous, 270 integer (270 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e+00, 1e+03]
Presolve removed 1187 rows and 440 columns
Presolve time: 0.02s
Presolved: 876 rows, 808 columns, 3054 nonzeros
Variable types: 676 continuous, 132 integer (132 binary)

Root relaxation: objective 1.555556e+02, 524 iterations, 0.02 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  155.55556    0    3          -  155.55556      -     -    0s
H    0     0                       0.0000000  155.55556      -     -    0s
     0     0  155.55556    0   10    0.00000  155.55556      -     -    0s
     0     0  155.55556    0    3    0.00000  155.55556      -     -    0s
     0     2  155.55556    0    3    0.00000  155.55556      -     -    0s
*  107     3              26       0.0000000  155.55556      -  33.0    0s
*  415    15              20       0.0000000  155.55556      -  25.5    1s
*  527    34              16       0.0000001  155.55556      -  25.7    1s
* 1248   147              29       0.0000001  155.55556      -  26.9    2s
* 1450   142              42      14.2857143  155.55556   989%  27.0    2s
* 1925    45              72      47.8260870  133.33333   179%  24.5    3s

Cutting planes:
  Gomory: 1
  Cover: 2
  Implied bound: 5
  Inf proof: 4
```

```
Explored 2126 nodes (54745 simplex iterations) in 3.87 seconds
Thread count was 4 (of 4 available processors)

Solution count 7: 47.8261 14.2857 7.09741e-08 ... 7.66478e-09
Pool objective bound 47.8261

Optimal solution found (tolerance 1.00e-12)
Best objective 4.782608696417e+01, best bound 4.782608696417e+01, gap 0.0000%
set n 2 was found
Optimize a model with 2064 rows, 1248 columns and 6274 nonzeros
Variable types: 978 continuous, 270 integer (270 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e+00, 1e+03]
Presolve removed 1187 rows and 440 columns
Presolve time: 0.06s
Presolved: 877 rows, 808 columns, 3056 nonzeros
Variable types: 676 continuous, 132 integer (132 binary)

Root relaxation: objective 1.555556e+02, 524 iterations, 0.01 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  155.55556    0    4          -  155.55556      -     -    0s
H    0     0                       0.0000000  155.55556      -     -    0s
     0     0  155.55556    0    4    0.00000  155.55556      -     -    0s
     0     2  155.55556    0    4    0.00000  155.55556      -     -    0s
*   13     6               4       0.0000000  155.55556      -  72.2    0s
*  392    22              50       0.0000000  155.55556      -  34.5    1s
*  480    25              30       0.0000001  155.55556      -  34.3    1s
* 1054   145              49      14.2857143  155.55556   989%  31.0    2s
* 1452   142              44      47.8260870  155.55556   225%  30.2    2s

Cutting planes:
  Cover: 1
  Implied bound: 1
  Inf proof: 1

Explored 1864 nodes (55345 simplex iterations) in 3.09 seconds
Thread count was 4 (of 4 available processors)

Solution count 6: 47.8261 14.2857 7.29017e-08 ... 7.68341e-09
Pool objective bound 47.8261

Optimal solution found (tolerance 1.00e-12)
Best objective 4.782608696419e+01, best bound 4.782608696419e+01, gap 0.0000%
set n 3 was found
Optimize a model with 2065 rows, 1248 columns and 6276 nonzeros
Variable types: 978 continuous, 270 integer (270 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e+00, 1e+03]
Presolve removed 1187 rows and 440 columns
Presolve time: 0.03s
Presolved: 878 rows, 808 columns, 3058 nonzeros
Variable types: 676 continuous, 132 integer (132 binary)

Root relaxation: objective 1.555556e+02, 524 iterations, 0.01 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time
```

```
     0     0  155.55556    0    2          -  155.55556      -     -    0s
H    0     0                         0.0000000  155.55556      -     -    0s
     0     0  155.55556    0    3    0.00000  155.55556      -     -    0s
     0     2  155.55556    0    3    0.00000  155.55556      -     -    0s
*    8     4               3       0.0000000  155.55556      -  70.9    0s
*  150    35               8       0.0000001  155.55556      -  28.6    0s
*  188    34              29       0.0000001  155.55556      -  27.9    0s
*  192    34              32      14.2857143  155.55556   989%  27.5    0s
*  471    19              35      47.8260870  136.84211   186%  23.3    0s

Cutting planes:
  Cover: 3
  Implied bound: 1

Explored 611 nodes (17553 simplex iterations) in 1.13 seconds
Thread count was 4 (of 4 available processors)

Solution count 6: 47.8261 14.2857 7.09741e-08 ... 7.68341e-09
Pool objective bound 47.8261

Optimal solution found (tolerance 1.00e-12)
Best objective 4.782608696419e+01, best bound 4.782608696419e+01, gap 0.0000%
set n 4 was found
Optimize a model with 2066 rows, 1248 columns and 6278 nonzeros
Variable types: 978 continuous, 270 integer (270 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e+00, 1e+03]
Presolve removed 1187 rows and 440 columns
Presolve time: 0.03s
Presolved: 879 rows, 808 columns, 3060 nonzeros
Variable types: 676 continuous, 132 integer (132 binary)

Root relaxation: objective 1.555556e+02, 524 iterations, 0.02 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  155.55556    0    3          -  155.55556      -     -    0s
H    0     0                         0.0000000  155.55556      -     -    0s
     0     0  155.55556    0    3    0.00000  155.55556      -     -    0s
     0     2  155.55556    0    3    0.00000  155.55556      -     -    0s
*  102    15              29       0.0000001  155.55556      -  15.7    0s
H  113    17                         0.0000001  155.55556      -  14.8    0s
*  117    21              13      14.2857143  155.55556   989%  14.8    0s
*  219    19              21      47.8260870  144.44444   202%  13.8    0s

Cutting planes:
  Inf proof: 2

Explored 413 nodes (6625 simplex iterations) in 0.50 seconds
Thread count was 4 (of 4 available processors)

Solution count 5: 47.8261 14.2857 7.09742e-08 ... 7.66479e-09
Pool objective bound 47.8261

Optimal solution found (tolerance 1.00e-12)
Best objective 4.782608696419e+01, best bound 4.782608696419e+01, gap 0.0000%
set n 5 was found
Optimize a model with 2067 rows, 1248 columns and 6280 nonzeros
Variable types: 978 continuous, 270 integer (270 binary)
Coefficient statistics:
  Matrix range     [5e-02, 1e+03]
  Objective range  [1e+00, 1e+00]
  Bounds range     [1e+00, 1e+03]
  RHS range        [1e+00, 1e+03]
```

```
Presolve removed 1187 rows and 440 columns
Presolve time: 0.04s
Presolved: 880 rows, 808 columns, 3062 nonzeros
Variable types: 676 continuous, 132 integer (132 binary)

Root relaxation: objective 1.555556e+02, 524 iterations, 0.01 seconds

    Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0  155.55556    0    4          -  155.55556      -     -    0s
H    0     0                       0.0000000  155.55556      -     -    0s
     0     0  155.55556    0    3    0.00000  155.55556      -     -    0s
     0     2  155.55556    0    3    0.00000  155.55556      -     -    0s
*  127    26              21       0.0000000  155.55556      -  32.4    0s
*  133    27              42       0.0000001  155.55556      -  32.8    0s
*  308    29              52      14.2857143  155.55556   989%  31.2    0s
*  341    26              65      47.8260870  155.55556   225%  30.3    0s

Cutting planes:
  Cover: 1
  Implied bound: 1
  Inf proof: 3
```

Note that the folders "InputsOptForce" and "OutputsOptForce" were created inside TestOptForce2. These folders contain the inputs and outputs of `optForce`, respectively.

We display the reactions found by `optForce`

```
disp(optForceSets)
```

```
    'R32'       'EX_o2'
    'EX_ac'     'EX_o2'
    'R70'       'R72'
    'R32'       'R72'
    'R70'       'EX_o2'
    'R72'       'EX_ac'
    'R4'        'R70'
    'R4'        'R32'
    'R4'        'EX_ac'
    'R18'       'EX_o2'
    'R4'        'R62'
    'R62'       'EX_glyc'
    'R50'       'R72'
    'R25'       'EX_glyc'
    'R38'       'R17'
    'R4'        'R26'
    'R6'        'EX_o2'
    'R41'       'R37'
    'EX_nh3'    'R2'
    'R4'        'R68'
```

TIMING

1. STEP 1 ~ 1-2 seconds
2. STEP 2: ~ 2-5 seconds
3. STEP 3: ~ 10-20 seconds

4. STEP 4: ~ 10-20 seconds

TROUBLESHOOTING

1) Problem: "I didn't find any reaction in my must sets"

Possible reason: the wild-type or mutant strain is not constrained enough.

Solution: add more constraints to your strains until you find differences in your reaction ranges. If you don't find any differences, it is better to change the approach and use another algorithm.

2) Problem: "I got an error when running the `findMust` functions"

Possible reason: inputs are not defined well or solver is not defined.

Solution: verify your inputs, use `changeCobraSolver`, verify that the global variable `CBT_MILP_SOLVER` is not empty. It should containg the identifier for a MILP solver.

ANTICIPATED RESULTS

In this tutorial some folders will be created inside the folder called "runID" to store inputs and outputs of the optForce functions (findMustU.m, findMustL.m, findMustUU.m, findMustLL.m, findMustUL.m, optForce.m)

In this case runID = 'TestOptForce', so inside this folder the following folders will be created:

```
├── CurrentFolder
|  ├── TestOptForceM
|  |  ├── InputsFindMustL
|  |  ├── OutputsFindMustL
|  |  ├── InputsFindMustU
|  |  ├── OutputsFindMustU
|  |  ├── InputsFindMustLL
|  |  ├── OutputsFindMustLL
|  |  ├── InputsFindMustUU
|  |  ├── OutputsFindMustUU
|  |  ├── InputsFindMustUL
|  |  ├── OutputsFindMustUL
|  |  ├── InputsOptForce
|  |  └── OutputsOptForce
```

The input folders contain inputs (.mat files) for running the functions to solve each one of the bilevel problems. Output folders contain results of the algorithms (.xls and .txt files) as well as a report (.txt) summarizing the outcomes of the steps performed during the execution of the optForce functions.

The optForce algorithm will find sets of reactions that should increase the production of your target. The first sets found should be the best ones because the production rate will be the highest. The last ones should be the worse because the production rete will be lower. Be aware that some sets could not guarante a minimum production rate for your target, so you always have to check the minimum production rate. You can do this using the function testOptForceSol.m. Some sets could allow a higher growth rate than others, so keep in mind this too when deciding which set is better.

Acknowledgments

I would to thanks to the research group of Costas D. Maranas who provided the GAMS functions to solve this example. In particular I would like to thank to Chiam Yu Ng who kindly provides examples for using GAMS.

References

[1] Ranganathan S, Suthers PF, Maranas CD (2010) OptForce: An Optimization Procedure for Identifying All Genetic Manipulations Leading to Targeted Overproductions. PLOS Computational Biology 6(4): e1000744. https://doi.org/10.1371/journal.pcbi.1000744.

[2] Maciek R. Antoniewicz, David F. Kraynie, Lisa A. Laffend, Joanna González-Lergier, Joanne K. Kelleher, Gregory Stephanopoulos, Metabolic flux analysis in a nonstationary system: Fed-batch fermentation of a high yielding strain of E. coli producing 1,3-propanediol, Metabolic Engineering, Volume 9, Issue 3, May 2007, Pages 277-292, ISSN 1096-7176, https://doi.org/10.1016/j.ymben.2007.01.003.