COPE-FBA 2.0: Better and Faster Enumeration of the Optimal Solution Space of Genome-scale Stoichiometric Models

Timo Maarleveld

Centrum Wiskunde en Informatica, Life Sciences

VU University Amsterdam, IBIVU Systems Bioinformatics

CWI Scientific Meeting, 12-12-2014

Part of a story submitted to PLoS Computational Biology

Genome-scale stoichiometric models in a nutshell

1. **metabolism**, i.e. *the set of life-sustaining chemical transformations within the cells of living organisms*

Genome-scale stoichiometric models in a nutshell

1. **metabolism**, i.e. *the set of life-sustaining chemical transformations within the cells of living organisms*
2. **genome-scale**: many reactions and metabolites ($\geq$ 1000)

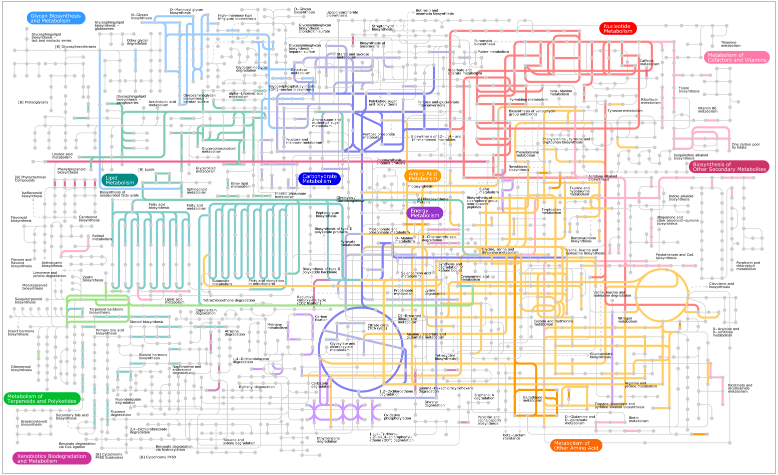
Genome-scale stoichiometric models in a nutshell

1. **metabolism**, i.e. *the set of life-sustaining chemical transformations within the cells of living organisms*
2. **genome-scale**: many reactions and metabolites ($\geq$ 1000)
3. **stoichiometric**: *only relative quantities*
   - A + B $\rightarrow$ C
   - no kinetics: ~~$k_1 \times A[t] \times B[t]$~~
   - assume steady-state: $\frac{dA}{dt} = \frac{dB}{dt} = \frac{dC}{dt} = 0$

Genome-scale stoichiometric models in a nutshell

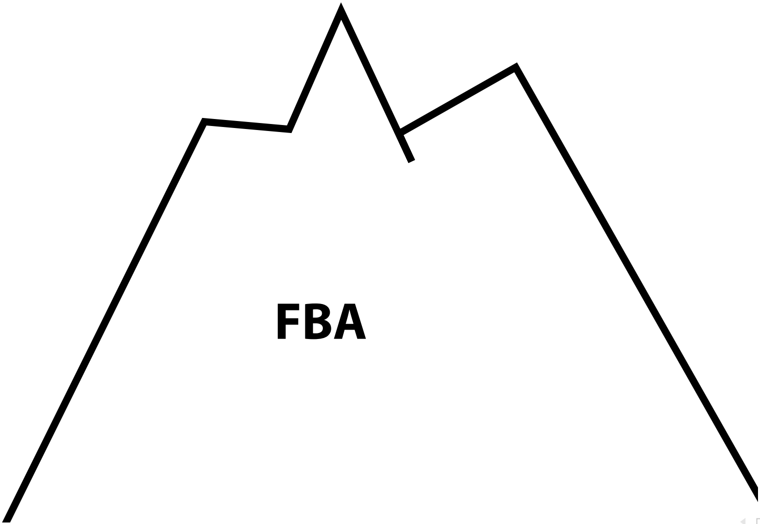
1. **metabolism**, i.e. *the set of life-sustaining chemical transformations within the cells of living organisms*
2. **genome-scale**: many reactions and metabolites (≥ 1000)
3. **stoichiometric**: *only relative quantities*
   - $A + B \rightarrow C$
   - no kinetics: ~~$k_1 \times A[t] \times B[t]$~~
   - assume steady-state: $\frac{dA}{dt} = \frac{dB}{dt} = \frac{dC}{dt} = 0$
4. predict steady-state flux distributions, i.e. *the rate of turnover of molecules through a metabolic pathway*

Genome-scale stoichiometric model

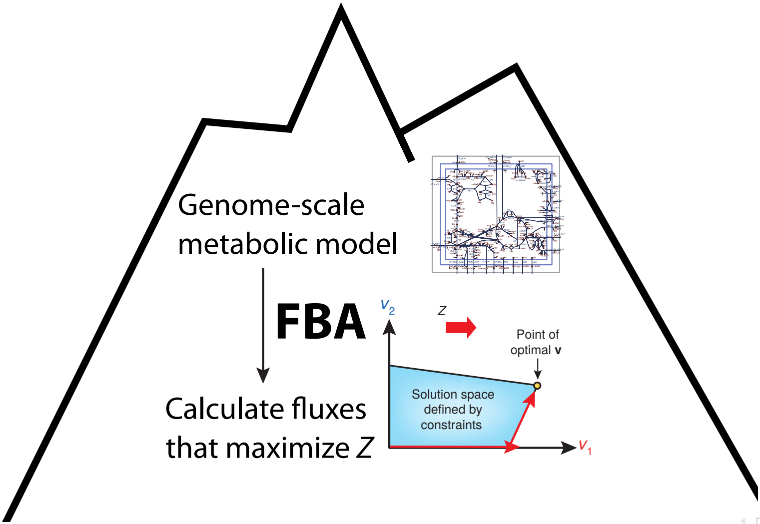


How can we simulate these models?

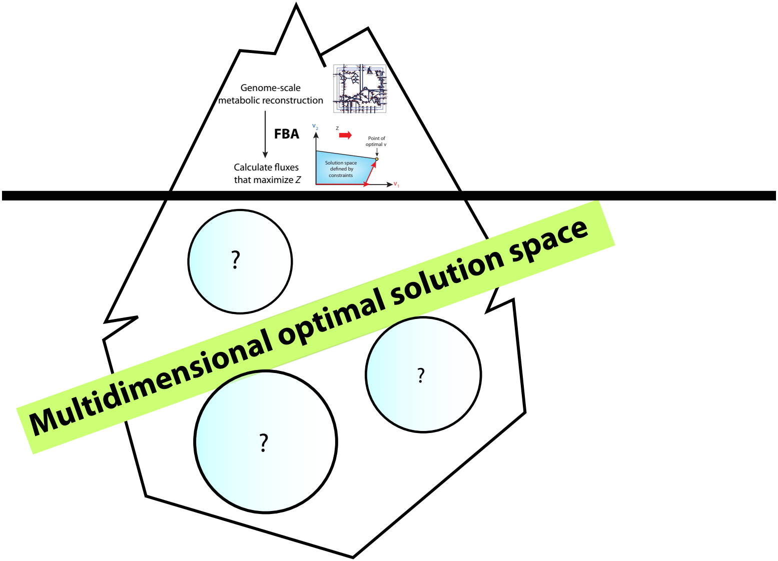
Flux **B**alance **A**nalysis

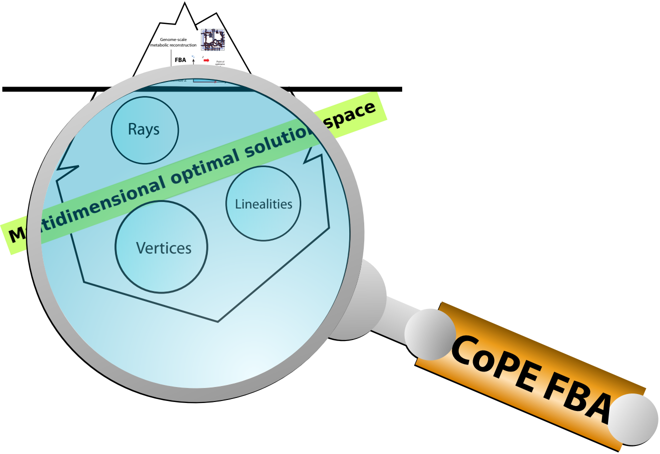


Flux Balance Analysis



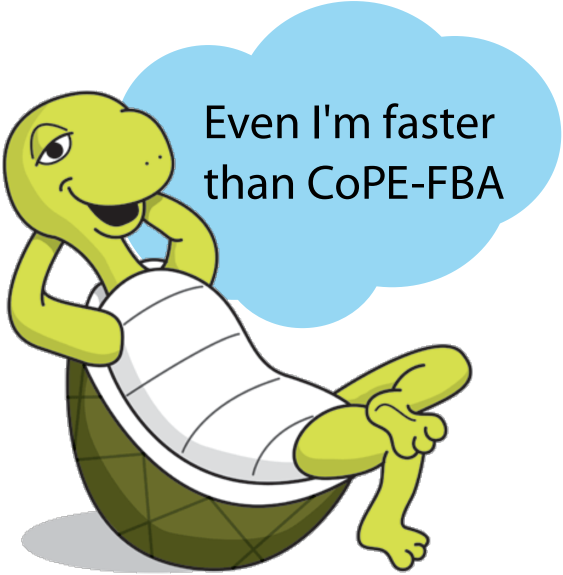
Tip of the iceberg



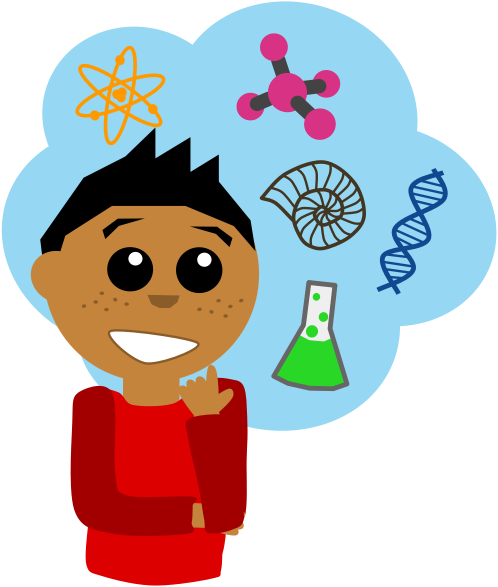


Topological understanding of the metabolic capacity in terms of metabolic flux routes

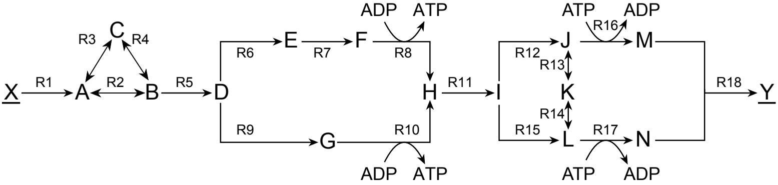
CoPE-FBA is slow!



Can we speed-up CoPE-FBA?



Toy model for stoichiometric modeling



- both X (source) and Y (sink) are fixed
- co-factors ATP and ADP

Flux **B**alance **A**nalysis formulation

Linear Program:

Maximize $Z_{obj} = \mathbf{c}^T \mathbf{J}$

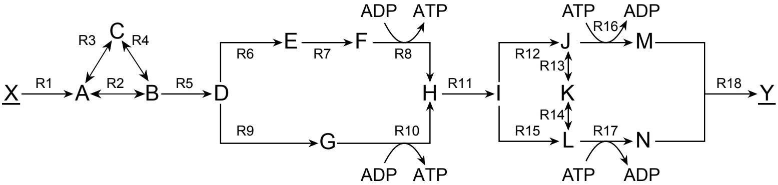
subject to,

$$\boldsymbol{NJ} = \mathbf{0}$$

$$\boldsymbol{J}^{min} \leq \boldsymbol{J} \leq \boldsymbol{J}^{max}$$

N: stoichiometric matrix
J: steady-state flux vector
c: vector of coefficients that represent the contribution of each flux in vector $\boldsymbol{J}$ to the objective function Z_{obj}.

Flux **B**alance **A**nalysis formulation, example



Maximize $Z_{obj} = \mathbf{c}^T \mathbf{J} = J_{18}$

subject to,

$\boldsymbol{NJ} = \mathbf{0}$

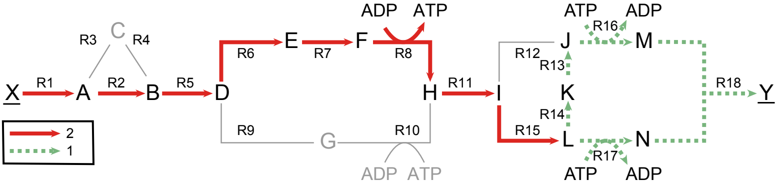
$-\infty \leq J_r \leq \infty \quad J_r \in$ reversible reactions

$0 \leq J_i \leq \infty \quad J_i \in$ irreversible reactions

$0 \leq J_1 \leq 2$

FBA example

Maximize $Z_{obj} \rightarrow J_{18} = 1$



with $\boldsymbol{J} = [2, 2, 0, 0, 2, 2, 2, 2, 0, 0, 2, 0, -1, -1, 2, 1, 1, 1]$

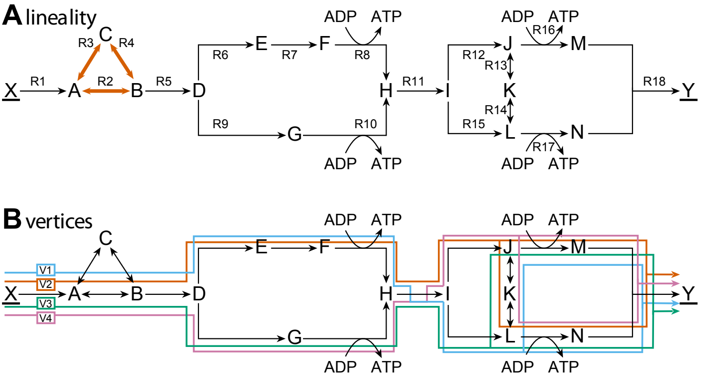
Optimal solution space characterization[1]

$$F_{opt} = \{\mathbf{J} : \mathbf{NJ} = \mathbf{0}, \boldsymbol{J}^{min} \leq \boldsymbol{J} \leq \boldsymbol{J}^{max}, \boldsymbol{c}^T \boldsymbol{J} = opt\}$$

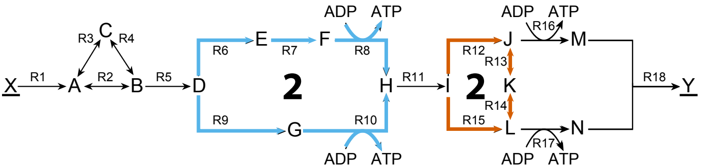
1. **vertices** --- optimal flux vectors
   - corner points of the optimal solution space
   - non-decomposable
   - no convex combination of other optimal flux vectors
2. **rays** --- irreversible cycles
3. **linealities** --- reversible cycles

[1] Kelk et al. 2012

Optimal solution space characterization *example*

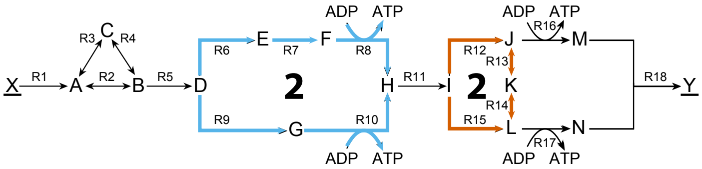


Subnetworks explain # vertices



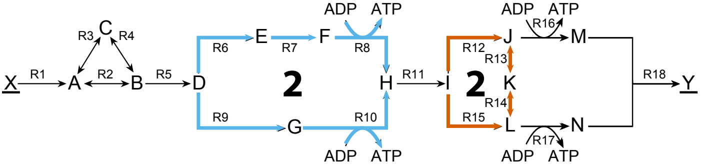
- $\mathbf{N_A J_A} = \mathbf{d} \neq \mathbf{0}$ i.e. input-output relationship
  1. $D + ADP \rightarrow H + ATP$
  2. $I \rightarrow 0.5(J + L)$

Subnetworks explain # vertices



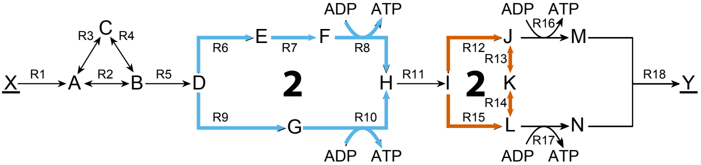
- $\mathbf{N_A J_A} = \mathbf{d} \neq \mathbf{0}$ i.e. input-output relationship
  1. $D + ADP \rightarrow H + ATP$
  2. $I \rightarrow 0.5(J + L)$
- Both subnetworks have 2 vertices
  1. {R6, R7, R8} and {R9, R10}
  2. {R12, R13, R14} and {R15, -R13, -R14}

Subnetworks explain # vertices



- $\mathbf{N_A J_A} = \mathbf{d} \neq \mathbf{0}$ i.e. input-output relationship
  1. $D + ADP \rightarrow H + ATP$
  2. $I \rightarrow 0.5(J + L)$
- Both subnetworks have 2 vertices
  1. {R6, R7, R8} and {R9, R10}
  2. {R12, R13, R14} and {R15, -R13, -R14} NOT: {R12, R15}

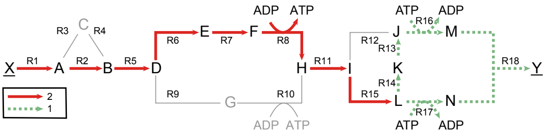
Subnetworks explain # vertices



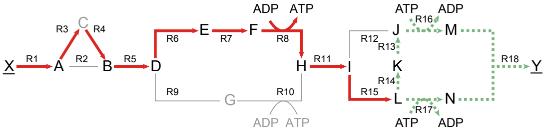
- $\mathbf{N_A J_A} = \mathbf{d} \neq \mathbf{0}$ i.e. input-output relationship
  1. $D + ADP \rightarrow H + ATP$
  2. $I \rightarrow 0.5(J + L)$
- Both subnetworks have 2 vertices
  1. {R6, R7, R8} and {R9, R10}
  2. {R12, R13, R14} and {R15, -R13, -R14} NOT: {R12, R15}
- Each subnetwork is an independent module with a fixed **d**: $2 \times 2 = 4$ network vertices

Not all non-decomposable flux vectors are vertices?

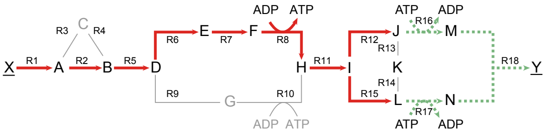
Are these vertices?



- Sometimes



- Sometimes



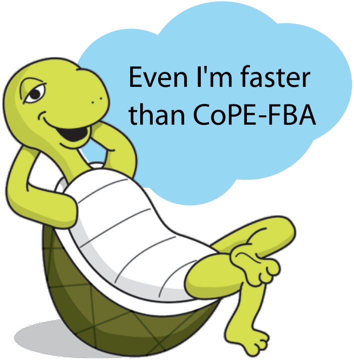
- Never

The extended list of problems

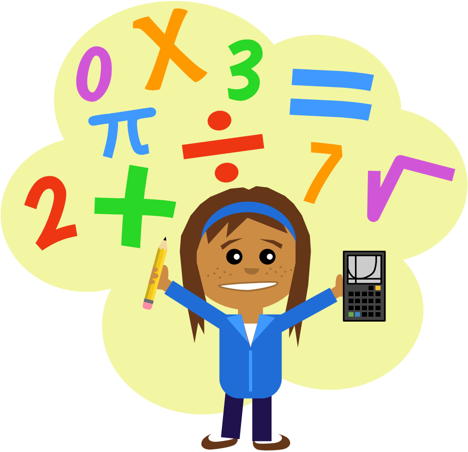
- Not all non-decomposable flux routes are vertices
- The decomposition is not unique

The extended list of problems

- Not all non-decomposable flux routes are vertices
- The decomposition is not unique
- CoPE-FBA is slow!

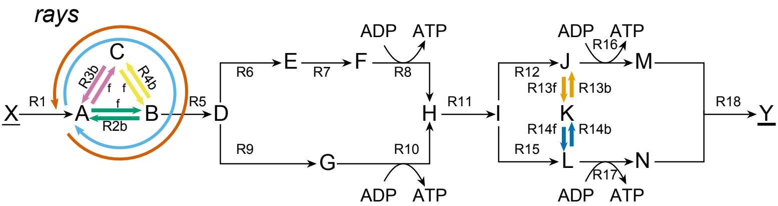


What if we split reversible reactions [2] ?



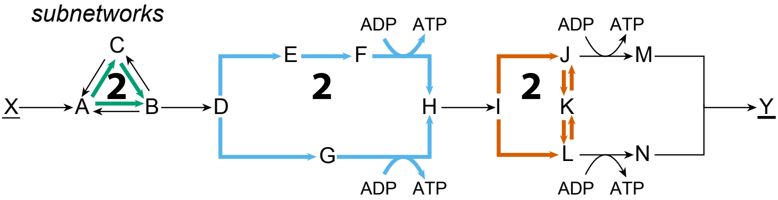
[2] Klamt and Stelling 2003, Wagner and Urbanczik 2005

Reversible-reaction splitting: the effect



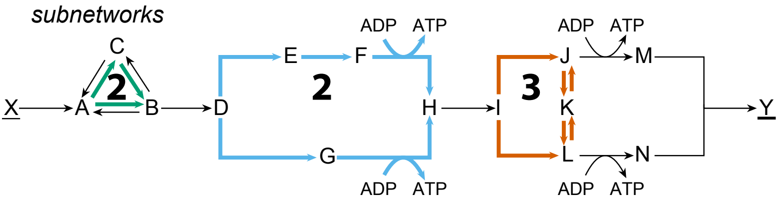
1. additional rays → ~~reversible reaction~~ and ~~linealities~~
2. # vertices ↑
   - ...
   - ...

Reversible-reaction splitting: the effect



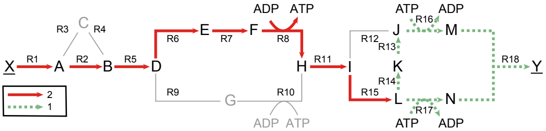
1. additional rays $\rightarrow$ ~~reversible reaction~~ and ~~linealities~~
2. # vertices ↑
   - # subnetworks ↑
   - ...

Reversible-reaction splitting: the effect

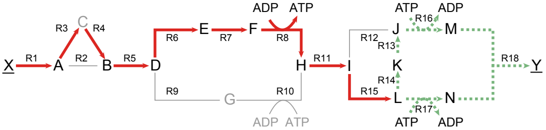


1. additional rays → ~~reversible reaction~~ and ~~linealities~~
2. # vertices ↑
   - # subnetworks ↑
   - convex combinations, i.e. reactions can't cancel out each other ($J_i \geq 0$)

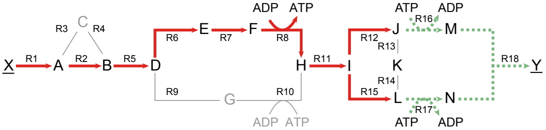
Are these vertices?



- Yes

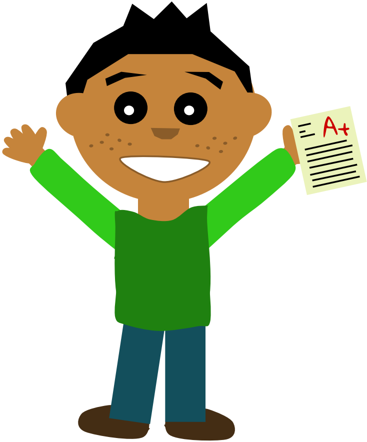


- Yes



- Yes

Reversible-reaction splitting yields all optimal non-decomposable paths in the optimum

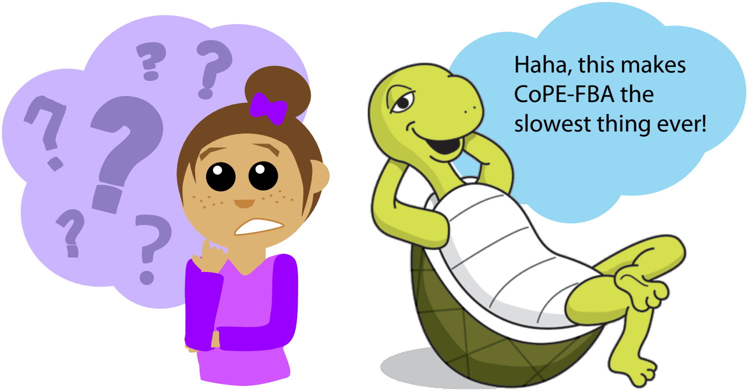


A Real Life Example: *E.coli* Growing on Glucose

Model	Toy		*E. coli* iAF1260	
Split	No	Yes	No	Yes
Growth condition			Aerobic	
Total reactions	12	23	2374	3226
Rays	0	7	26	604
Linealities	1	0	1	0
Vertices	4	12	839.808	**120.932.352**
Subnetworks	2	3	6	9

Model	*E. coli* iAF1260		*E. coli* iAF1260	
Split	No	Yes	No	Yes
Growth condition	Aerobic restricted		Anaerobic	
Total reactions	2374	3226	2374	3226
Rays	25	602	25	602
Linealities	1	0	1	0
Vertices	1.679.616	**40.310.784**	31104	**1.492.992**
Subnetworks	4	4	6	8

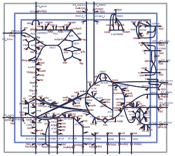
These simulations take forever right?



Can we speed-up CoPE-FBA?

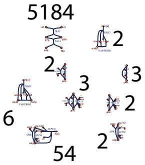
CoPE-FBA approach (in a nutshell)

iAF1260 (glc, aerobic)



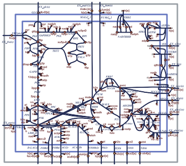
CoPE-FBA (weeks)

Vertices → 120.932.352 Subnetworks →

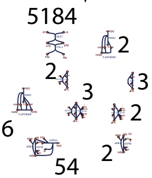


Can we do the reverse?

iAF1260 (glc, aerobic)



Subnetworks



Vertices

$6 + 3 + 5184 + 3 + 2 + 54 + 2 + 2 = 5256$

Alternative (minutes)

Enumerate 0.004% of the vertices to get all

Sounds simple, right?

Our solution: CoPE-FBA 2.0

Example: Successfully enumerated the 120.932.352 vertices in minutes rather than weeks!

"Faster than lightning"

"A new standard"

Summary

- **Faster:** CoPE-FBA 2.0, *"faster than lightning"*
- **Better:** We enumerate all non-decomposable routes in the optimum

- Topological understanding of the metabolic capacity
- Investigate flexibility

Acknowledgments

- Meike Wortel (VU)
- Brett Olivier (VU)
- Bas Teusink (VU)
- Frank Bruggeman (VU)
- "Scout" (`openclipart.org`)
- Arne Reimers (CWI)

Contact information: t.r.maarleveld@cwi.nl