Null space and Linear Programming

Mark Poolman

May 5, 2022

Recap - Structural Modelling

Null space: Encapsulates all possible steady-state solutions.

Enzyme subsets: Sets of reactions carrying flux in fixed ratio.

Elementary modes: Minimal, independent pathways in a system

Conserved cycles: Sets of metabolites whose total concentration is fixed.

Recap - Structural Modelling

Null space: Encapsulates all possible steady-state solutions.

Enzyme subsets: Sets of reactions carrying flux in fixed ratio.

Elementary modes: Minimal, independent pathways in a system

Conserved cycles: Sets of metabolites whose total concentration is fixed.

Recap - Structural Modelling

Null space: Encapsulates all possible steady-state solutions.

Enzyme subsets: Sets of reactions carrying flux in fixed ratio.

Elementary modes: Minimal, independent pathways in a system

Conserved cycles: Sets of metabolites whose total concentration is fixed.

Recap - Structural Modelling

Null space: Encapsulates all possible steady-state solutions.

Enzyme subsets: Sets of reactions carrying flux in fixed ratio.

Elementary modes: Minimal, independent pathways in a system

Conserved cycles: Sets of metabolites whose total concentration is fixed.

Recap - Structural Modelling

Null space: Encapsulates all possible steady-state solutions.

Enzyme subsets: Sets of reactions carrying flux in fixed ratio.

Elementary modes: Minimal, independent pathways in a system

Conserved cycles: Sets of metabolites whose total concentration is fixed.

Technical challenges with large models

Null space: Readily calculated, but can't analyse by inspection.

Enzyme subsets: Not as generally useful compared to small models.

Elementary modes: Impractical.

Conserved cycles: Not as generally useful compared to small models.

Technical challenges with large models

Null space: Readily calculated, but can't analyse by inspection.

Enzyme subsets: Not as generally useful compared to small models.

Elementary modes: Impractical.

Conserved cycles: Not as generally useful compared to small models.

Technical challenges with large models

Null space: Readily calculated, but can't analyse by inspection.

Enzyme subsets: Not as generally useful compared to small models.

Elementary modes: Impractical.

Conserved cycles: Not as generally useful compared to small models.

Technical challenges with large models

Null space: Readily calculated, but can't analyse by inspection.

Enzyme subsets: Not as generally useful compared to small models.

Elementary modes: Impractical.

Conserved cycles: Not as generally useful compared to small models.

Technical challenges with large models

Null space: Readily calculated, but can't analyse by inspection.

Enzyme subsets: Not as generally useful compared to small models.

Elementary modes: Impractical.

Conserved cycles: Not as generally useful compared to small models.

Other disadvantages of null-space analysis

- Provides a rather ‘unfocussed’ view of the system.
- Does not (implicitly) take into account thermodynamics.
- Hard to integrate experimental flux observations.
- Less interpretable for large (genome-scale) models.
- (Still very useful for validation).

Application of LP to metabolic networks - FBA

Linear programming calculates a specific solution to the equation:

$$\mathbf{Nv} = \mathbf{0}$$

Subject to some additional information supplied by the user - at least one flux value specified.

Application of LP to metabolic networks - FBA

$$\begin{aligned} \min/\max \quad &: \quad \mathbf{v}_{\text{targs}} && \longleftarrow \text{objective} \\ \text{subject to} \quad &\begin{cases} \mathbf{Nv} = \mathbf{0} & \longleftarrow \text{steady state} \\ \max_i \geq \mathbf{v}_i \geq \min_i & \longleftarrow \text{flux constraints} \end{cases} \end{aligned}$$

Typical Objectives:

- Maximise output(s) (need to fix input(s))
- FBA maximise growth rate for fixed input.
- Minimise input(s) (need to fix output(s))
- Minimise *all* reactions (need to fix input(s) and/or output(s))

Application of LP to metabolic networks - FBA

$$\begin{aligned} \min/\max \quad & : \quad \mathbf{v}_{\text{targs}} && \longleftarrow \text{objective} \\ \text{subject to} \quad & \begin{cases} \mathbf{Nv} = \mathbf{0} & \longleftarrow \text{steady state} \\ \max_i \geq \mathbf{v}_i \geq \min_i & \longleftarrow \text{flux constraints} \end{cases} \end{aligned}$$

Typical Objectives:

- Maximise output(s) (need to fix input(s))
- FBA maximise growth rate for fixed input.
- Minimise input(s) (need to fix output(s))
- Minimise *all* reactions (need to fix input(s) and/or output(s))

Application of LP to metabolic networks - FBA

$$\begin{aligned} \min/\max \quad & : \quad \mathbf{v}_{\text{targs}} && \longleftarrow \text{objective} \\ \text{subject to} \quad & \begin{cases} \mathbf{Nv} = \mathbf{0} & \longleftarrow \text{steady state} \\ \max_i \geq \mathbf{v}_i \geq \min_i & \longleftarrow \text{flux constraints} \end{cases} \end{aligned}$$

Typical Objectives:

- Maximise output(s) (need to fix input(s))
- FBA maximise growth rate for fixed input.
- Minimise input(s) (need to fix output(s))
- Minimise *all* reactions (need to fix input(s) and/or output(s))

Application of LP to metabolic networks - FBA

$$\begin{aligned} \min/\max \quad & : \quad \mathbf{v}_{\text{targs}} && \longleftarrow \text{objective} \\ \text{subject to} \quad & \begin{cases} \mathbf{Nv} = \mathbf{0} & \longleftarrow \text{steady state} \\ \max_i \geq \mathbf{v}_i \geq \min_i & \longleftarrow \text{flux constraints} \end{cases} \end{aligned}$$

Typical Objectives:

- Maximise output(s) (need to fix input(s))
- FBA maximise growth rate for fixed input.
- Minimise input(s) (need to fix output(s))
- Minimise *all* reactions (need to fix input(s) and/or output(s))

Application of LP to metabolic networks - FBA

$$\begin{array}{lll} \min/\max & : & \mathbf{v}_{\text{targs}} \qquad \longleftarrow \text{objective} \\ \text{subject to} & \left\{ \begin{array}{ll} \mathbf{Nv} = \mathbf{0} & \longleftarrow \text{steady state} \\ \max_i \geq \mathbf{v}_i \geq \min_i & \longleftarrow \text{flux constraints} \end{array} \right. \end{array}$$

Typical Objectives:

- Maximise output(s) (need to fix input(s))
- FBA maximise growth rate for fixed input.
- Minimise input(s) (need to fix output(s))
- Minimise *all* reactions (need to fix input(s) and/or output(s))

Application of LP to metabolic networks - FBA

$$\begin{array}{lll} \min/\max & : & \mathbf{v}_{\text{targs}} \quad \longleftarrow \text{objective} \\ \text{subject to} & \begin{cases} \mathbf{Nv} = \mathbf{0} & \longleftarrow \text{steady state} \\ \max_i \geq \mathbf{v}_i \geq \min_i & \longleftarrow \text{flux constraints} \end{cases} \end{array}$$

Typical flux constraints:

- $\min_i = \max_i \neq 0$: flux is fixed
- $\min_i = \max_i = 0$: reaction is knocked out.
- $\min_i = 0, \max_i \neq 0$: force irreversible L->R
- $\min_i \neq 0, \max_i = 0$: force irreversible R->L

Advantages of FBA

- Very fast.
- Integrates flux data.
- Easy to reformulate the problem and solve again.
- The reactions in a solution can be extracted from the main model for more detailed analysis.

Disadvantages of FBA

- Only provides a single solution.
- Potential for numerical instability (esp. if maximising).
- Potential for multiple optima.
- Choice of the objective is subjective (!)

Exploring the optimal space - constraint scanning

$$\begin{array}{lll} \text{minimise} & : & \mathbf{v}_{\text{targs}} \longleftarrow \text{objective} \\ \text{subject to} & \begin{cases} \mathbf{Nv} = \mathbf{0} & \longleftarrow \text{steady state} \\ \max_i \geq \mathbf{v}_i \geq \min_i & \longleftarrow \text{flux constraints} \end{cases} \end{array}$$

- Find a solution.
- Increment one (or more) of the constraints $\mathbf{v}_i$
- Solve again.
- Repeat to build up a set of solutions.
- Identify correlated responses in the solution set.

Example - identifying a catabolic core

A study of Salmonella spp.

- Antibiotic challenges generate a stress response.
- This increases the demand for ATP.
- How to identify which reactions will respond to this demand?

Example - identifying a catabolic core

Scan over a range of ATP demand fluxes (while synthesising biomass) and identify responding reactions.

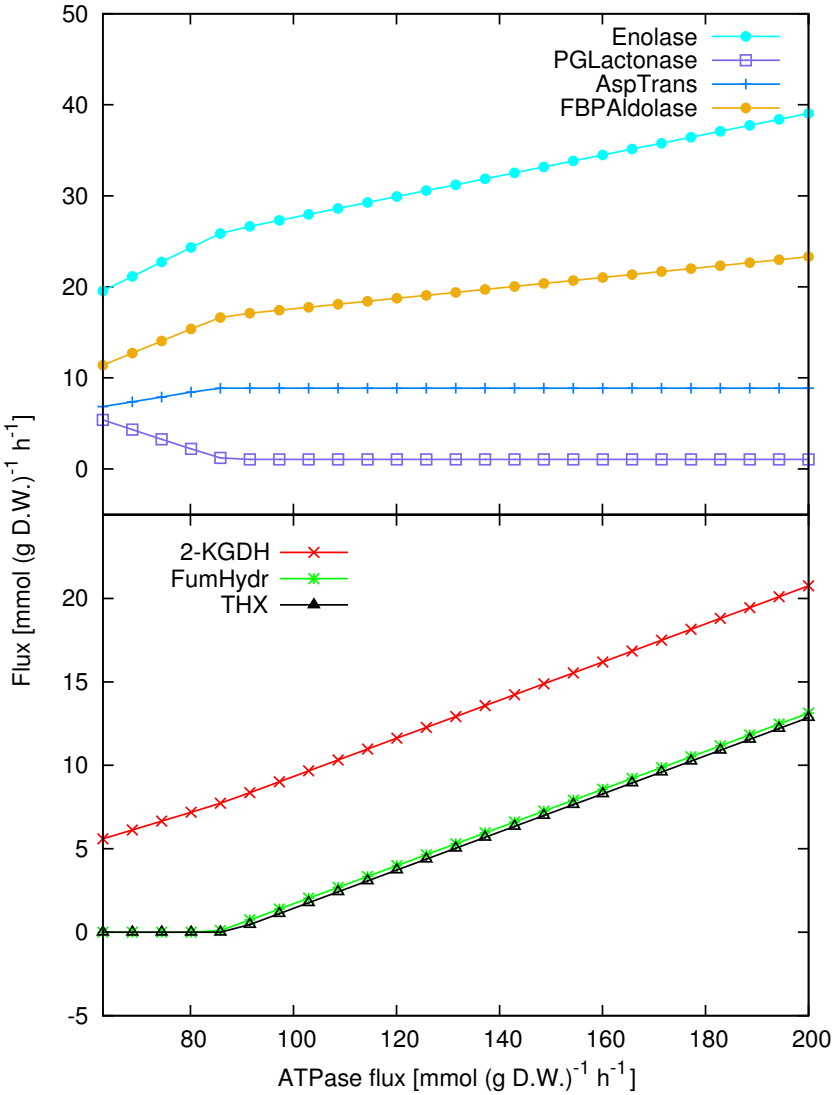
$$
\begin{array}{lll}
\text{minimise} & : & |\mathbf{v}| \quad \longleftarrow \text{objective} - \text{min. sum of fluxes} \\
\text{subject to} & \left\{ \begin{array}{ll} \mathbf{Nv} = \mathbf{0} & \longleftarrow \text{steady state constraint} \\ v_j = t_j & \longleftarrow \text{output transporters, constant} \\ v_{\text{ATPase}} = J_{\text{ATPase}} & \longleftarrow \text{ATP hydrolysis, varied} \end{array} \right.
\end{array}
$$

Example - identifying a catabolic core

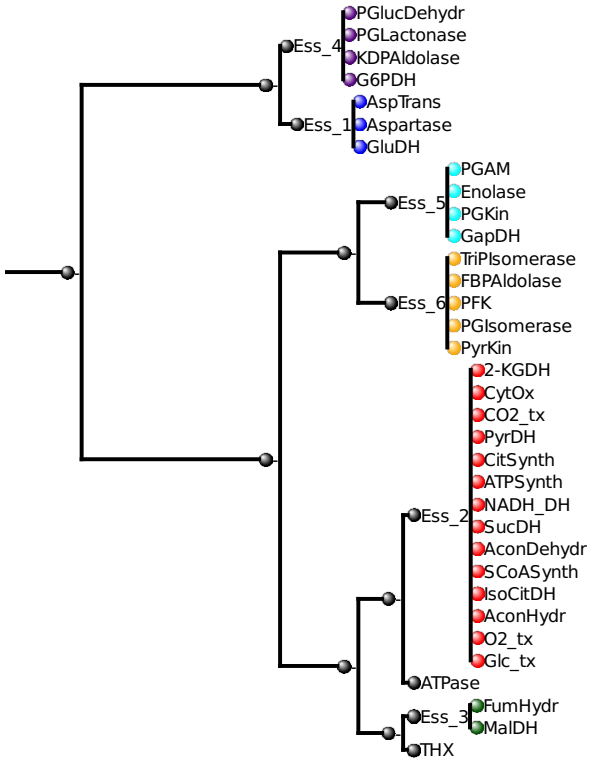
Scan over a range of ATP demand fluxes (while synthesising biomass) and identify responding reactions.

$$
\begin{array}{lll}
\text{minimise} & : & |\mathbf{v}| \qquad \longleftarrow \text{objective} - \text{min. sum of fluxes} \\
\text{subject to} & \left\{ \begin{array}{ll} \mathbf{Nv} = \mathbf{0} & \longleftarrow \text{steady state constraint} \\ v_j = t_j & \longleftarrow \text{output transporters, constant} \\ v_{\text{ATPase}} = J_{\text{ATPase}} & \longleftarrow \text{ATP hydrolysis, varied} \end{array} \right.
\end{array}
$$

Results - flux response

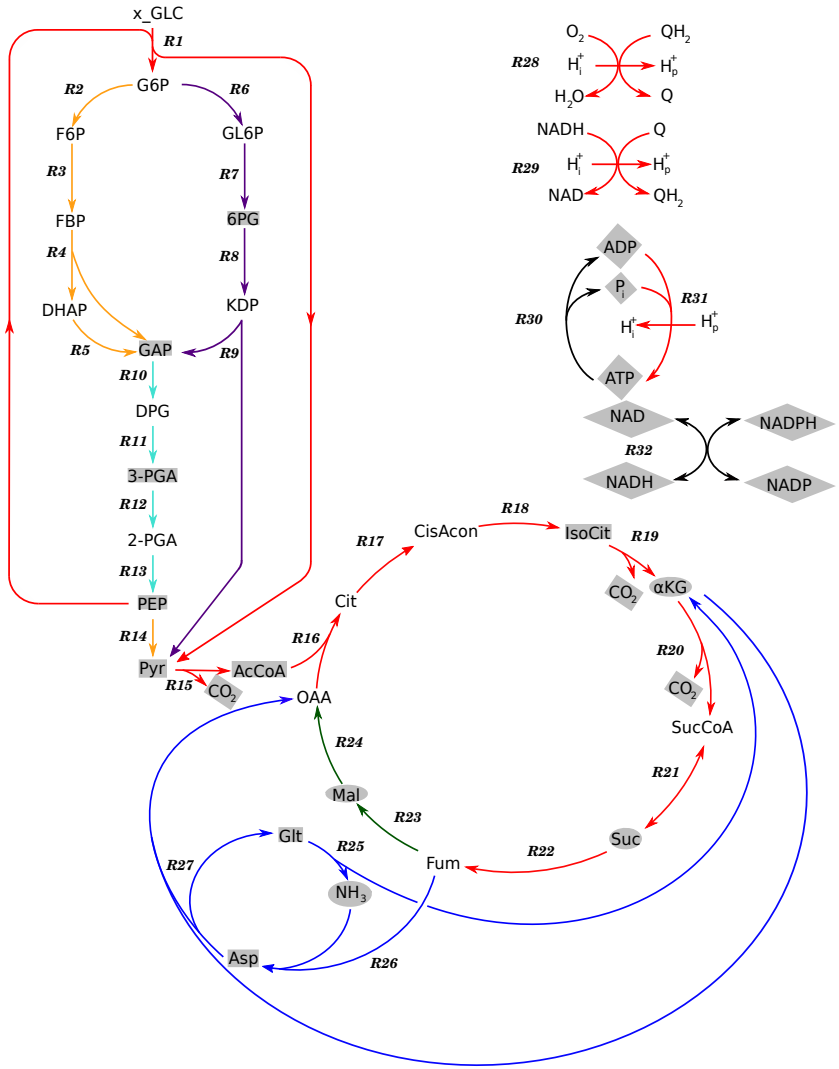
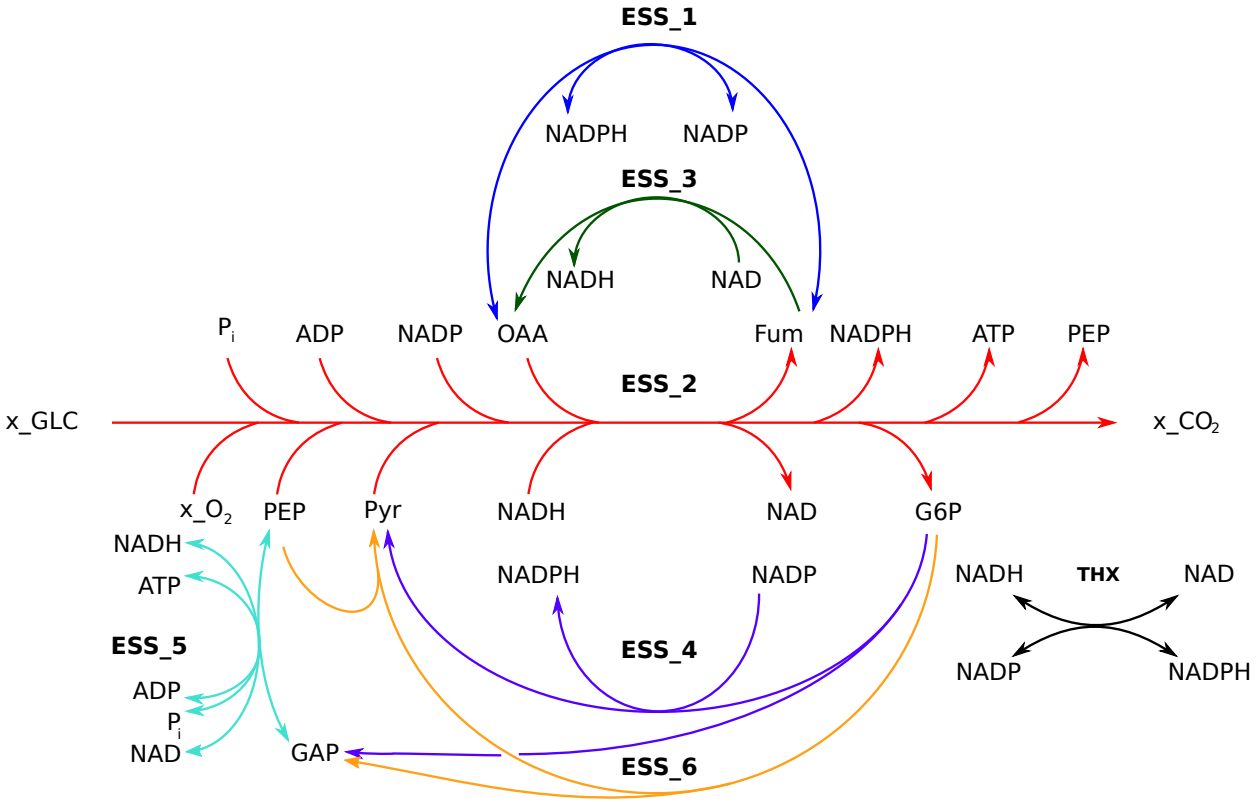


Results - flux correlations



- 33 reactions correlated with imposed ATPase.

Results - catabolic core



Results - condensed network