#### Stoichiometric vector and matrix

- The stoichiometric coefficients of a reaction are collected to a vector  $s_r$
- In  $s_r$  there is a one position for each metabolite in the metabolic system, and the stoichiometric co-efficient of the reaction are inserted to appropriate positions, e.g. for the reaction

$$r: A + B \mapsto 2C,$$

	•	0	
	•	0	
	A	-1	
	•	0	
$s_r =$	•	0	
	B	-1	
	В	-10	
	В	$-1\\0\\0$	
	В • С	$\begin{array}{c} -1 \\ 0 \\ 0 \\ 2 \end{array}$	

## Example: stoichiometric matrix

• Consider the set of reactions from								
the penthose-phospate pathway:	S =							
• The stoichiometric matrix is a	eta G6P	$\left\lceil -1 \right\rceil$	0	0	0	1	0	-1
10-by-7 matrix:	$\alpha G6P$	0	0	0	0	-1	-1	0
$R_1: \beta G6P + NADP^+ \stackrel{zwf}{\Rightarrow} 6PGL + NADPH$	$\beta F6P$	0	0	0	0	0	1	1
$R_2: 6PGL + H_2O \stackrel{pgl}{\Rightarrow} 6PG$	6PGL	1	-1	0	0	0	0	0
$R_3: 6PG + NADP^+ \stackrel{gnd}{\Rightarrow} R5P + NADPH$	6PG	0	1	-1	0	0	0	0
$R_4: \text{R5P} \stackrel{rpe}{\Rightarrow} \text{X5P}$	R5P	0	0	1	-1	0	0	0
$R_5: \alpha G6P \stackrel{gpi}{\Leftrightarrow} \beta G6P$	X5P	0	0	0	1	0	0	0
$R_6: \alpha G6P \stackrel{gpi}{\Leftrightarrow} \beta F6P$	$NADP^+$	-1	0	-1	0	0	0	0
$R_7: \beta G6P \Leftrightarrow \beta F6P$	NADPH	1	0	1	0	0	0	0
	$H_2O$	0	-1	0	0	0	0	0

## Systems equations

In a network of n metabolites and r reactions, the dynamics of the system are characterized by the systems equations

$$\frac{dX_i}{dt} = \sum_{j=1}^r s_{ij} v_j, \text{ for } i = 1, \dots, n$$

- $X_i$  is the concentration of the *i*th metabolite
- $v_j$  is the rate of the *j*th reaction and
- $s_{ij}$  is the stoichiometric coefficient of *i*th metabolite in the *j*th reaction.

Intuitively, each system equation states that the rate of change of concentration of a metabolite is the sum of metabolite flows to and from the metabolite.

### Systems equation example

- Assume our example metabolic network has the following rate vector v = (1, 1, 0, 0, 1, 0, 0)
- Let us compute the rate of change for metabolites

$$R_{1}: \beta G6P + NADP^{+} \stackrel{zwf}{\Rightarrow} 6PGL + NADPH$$

$$R_{2}: 6PGL + H_{2}O \stackrel{pgl}{\Rightarrow} 6PG$$

$$R_{3}: 6PG + NADP^{+} \stackrel{gnd}{\Rightarrow} R5P + NADPH$$

$$R_{4}: R5P \stackrel{rpe}{\Rightarrow} X5P$$

$$R_{5}: \alpha G6P \stackrel{gpi}{\Rightarrow} \beta G6P$$

$$R_{6}: \alpha G6P \stackrel{gpi}{\Rightarrow} \beta F6P$$

$$R_{7}: \beta G6P \stackrel{gpi}{\Rightarrow} \beta F6P$$

$$\frac{d\beta G6P}{dt} = -1v_{R_1} + 1v_{R_5} - 1v_{R_7} = 0$$

$$\frac{d\alpha G6P}{dt} = -1v_{R_5} - 1v_{R_6} = -1 \Rightarrow \text{ net consumption!}$$

$$\frac{d\beta F6P}{dt} = 1v_{R_6} + 1v_{R_7} = 0$$

$$\frac{d6GPL}{dt} = 1v_{R_1} - 1v_{R_2} = 0$$

$$\frac{d6PG}{dt} = 1v_{R_2} - 1v_{R_3} = 1 \Rightarrow \text{ net production!}$$

$$\frac{dR5P}{dt} = 1v_{R_3} - 1v_{R_4} = 0$$

$$\frac{dX5P}{dt} = 1v_{R_1} + 1v_{R_3} = 1 \Rightarrow \text{ net production!}$$

$$\frac{dNADPH}{dt} = 1v_{R_1} + 1v_{R_3} = 1 \Rightarrow \text{ net production!}$$

$$\frac{dNADPH}{dt} = -1v_{R_1} - 1v_{R_3} = -1 \Rightarrow \text{ net consumption!}$$

#### Steady state analysis

• The requirements a steady state, i.e. non-changing concentrations

$$\frac{dX_i}{dt} = \sum_{j=1}^r s_{ij} v_j = 0, \text{ for } i = 1, \dots, n$$

constitute a set of linear equations constraining to the reaction rates  $v_j$ .

• We can write this set of linear constraints in matrix form with the help of the stoichiometric matrix S and the reaction rate vector  $\mathbf{v}$ 

$$\frac{d\mathbf{X}}{dt} = S\mathbf{v} = \mathbf{0},$$

• A reaction rate vector **v** satisfying the above is called a *flux* vector.

## Null space of the stoichiometrix matrix (1/2)

• Any flux vector **v** that the cell can maintain in a steady-state is a solution to the system of equations

$$S\mathbf{v} = \mathbf{0}$$

• The null space of the stoichiometric matrix

$$\mathcal{N}(S) = \{\mathbf{u} | S\mathbf{u} = 0\}$$

contains all valid flux vectors

• Therefore, studying the null space of the stoichiometric matrix can give us important information about the cell's capabilities

## Null space of the stoichiometric matrix (2/2)

The null space  $\mathcal{N}(S)$  is a linear vector space, so all properties of linear vector spaces follow, e.g:

- $\mathcal{N}(S)$  contains the zero vector, and closed under linear combination:  $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{N}(S) \implies \alpha_1 \mathbf{v}_1 + \alpha \mathbf{v}_2 \in \mathcal{N}(S)$
- The null space has a basis {k<sub>1</sub>,..., k<sub>q</sub>}, a set of q ≤ min(n, r) linearly independent vectors, where r is the number of reactions and n is the number of metabolites.
- The choice of basis is not unique, but the number q of vector it contains is determined by the rank of S.

### Null space and feasible steady state rate vectors

- The kernel  $K = (\mathbf{k}_1, \dots, \mathbf{k}_q)$  of the stoichiometric matrix formed by the above basis vectors has a row corresponding to each reaction.
- K characterizes the feasible steady state reaction rate vectors: for each feasible flux vector  $\mathbf{v}$ , there is a vector  $\mathbf{b} \in \mathbb{R}^q$  such that  $K\mathbf{b} = \mathbf{v}$
- In other words, any steady state flux vector is a linear combination

 $b_1\mathbf{k}_1+\cdots+b_q\mathbf{k}_q$ 

of the basis vectors of  $\mathcal{N}(N)$ .

### Singular value decomposition of S (1/3)

- A basis for the null space can be obtained via the singular value decomposition (SVD)
- The SVD of S is the product  $S = U\Sigma V^T$ , where
  - U is a  $m \times m$  orthonormal matrix, where r first columns are the eigenvectors of the column space of S, and m r last columns span the left null space of S.
  - $-\Sigma = diag(\sigma_1, \sigma_2, \dots, \sigma_r) \text{ is } m \times n \text{ matrix containing the singular values } \sigma_i$ on its diagonal
  - -V is a  $n \times n$  orthonormal matrix, where r first columns span the row space of S, and n - r last columns span the null space of S

### Singular value decomposition of S (2/3)

The subspaces spanned by the columns of U are interpreted as follows:

- The set of *r m*-dimensional eigenvectors of the column space of *S* can be seen as prototypical or 'eigen-' reactions: all reaction stoichiometries in the metabolic system can be expressed as linear combinations of the eigen-reactions.
- The m r vectors  $u_{r+l}$  spanning the left null space of S represent conservation relations between metabolites or *pools* of metabolites whose concentration stays invariant.



### Singular value decomposition of S (3/3)

The subspaces spanned by the columns of V are interpreted as follows:

- The set of *r n*-dimensional eigenvectors of the row space of *S* can be seen as systems equations of prototypical 'eigen-' metabolites: all systems equations of the metabolism can be expressed as their linear combinations
- The set of n r *n*-dimensional vectors spanning the null space are flux vectors that can operate in steady state, i.e. statifying  $S\mathbf{v}_l = 0, l = r + 1, ..., n$ : these can be taken as the kernel K used to analyze steady state fluxes.



#### Basis steady state flux modes from SVD

• A basis for the null space is thus obtained by picking the n - r last columns of V from the SVD of S:

$$K = [v_{r+1}, \ldots, v_n]$$

- In MATLAB, the same operation is performed directly by the command null(S).
- Let us examine the following simple system



#### Basis steady state flux modes from SVD

- The two flux modes given by SVD for our example system
- All steady state flux vectors can be expressed as linear combinations of these two flux modes

$$K = \begin{bmatrix} 0.2980 & 0.4945 \\ 0.2980 & 0.4945 \\ 0.5772 & -0.0108 \\ -0.2793 & 0.5053 \\ 0.5772 & -0.0108 \\ -0.2793 & 0.5053 \end{bmatrix}$$



# Null space of PPP

• Consider again the set of reactions										
from the penthose-phospate	S =									
pathway	eta G6P	$\left\lceil -1 \right\rceil$	0	0	0	1	0	-1	0	0
• The stoichiometric matrix is a	$\alpha G6P$	0	0	0	0	-1	-1	0	1	0
10-by-9 matrix	$\beta F6P$	0	0	0	0	0	1	1	0	0
$R_1: \beta G6P + NADP^+ \stackrel{zwf}{\Rightarrow} 6PGL + NADPH$	6PGL	1	-1	0	0	0	0	0	0	0
$R_2: 6PGL + H_2O \stackrel{pgl}{\Rightarrow} 6PG$	6PG	0	1	-1	0	0	0	0	0	0
$R_3: 6PG + NADP^+ \stackrel{g_{nu}}{\Rightarrow} R5P + NADPH$	R5P	0	0	1	-1	0	0	0	0	0
$R_4: \text{R5P} \xrightarrow{\Rightarrow} \text{X5P}$ $R_5: \alpha \text{G6P} \xrightarrow{gpi} \beta \text{G6P}$	X5P	0	0	0	1	0	0	0	0	-1
$R_6: \alpha G6P \stackrel{gpi}{\Leftrightarrow} \beta F6P$	$NADP^+$	-1	0	-1	0	0	0	0	0	0
$R_7: \ \beta \text{G6P} \stackrel{gpi}{\Leftrightarrow} \beta \text{F6P}$	NADPH	1	0	1	0	0	0	0	0	0
$R_8 :\Rightarrow \alpha G6P$	$H_2O$	0	-1	0	0	0	0	0	0	0
$R_9: X5P \Rightarrow$										

## Null space of PPP

- Null space of this system has only one vector K = (0, 0, 0, 0, 0.5774, -0.5774, 0.5774, 0, 0, 0)<sup>T</sup>
- Thus, in a steady state only reactions  $R_5$ ,  $R_6$  and  $R_7$  can have non-zero fluxes.
- The reason for this is that there are no producers of NADP<sup>+</sup> or H<sub>2</sub>O and no consumers of NADPH.
- Thus our PPP is effectively now a dead end!

 $R_{1}: \beta G6P + NADP^{+} \stackrel{zwf}{\Rightarrow} 6PGL + NADPH$   $R_{2}: 6PGL + H_{2}O \stackrel{pgl}{\Rightarrow} 6PG$   $R_{3}: 6PG + NADP^{+} \stackrel{gnd}{\Rightarrow} R5P + NADPH$   $R_{4}: R5P \stackrel{rpe}{\Rightarrow} X5P$   $R_{5}: \alpha G6P \stackrel{gpi}{\Rightarrow} \beta G6P$   $R_{6}: \alpha G6P \stackrel{gpi}{\Rightarrow} \beta F6P$   $R_{7}: \beta G6P \stackrel{gpi}{\Rightarrow} \beta F6P$   $R_{8}:\Rightarrow \alpha G6P$   $R_{9}: X5P \Rightarrow$ 

## Null space of PPP

- To give our PPP non-trivial (fluxes different from zero) steady states, we need to modify our system
- We add reaction  $R_{10} :\Rightarrow H_2O$  as a water source
- We add reaction  $R_{11}$ : NADPH  $\Rightarrow$ NADP<sup>+</sup> to regenerate NADP<sup>+</sup> from NADPH.
- We could also have removed the metabolites in question to get the same effect

 $R_{1}: \beta G6P + NADP^{+} \stackrel{zwf}{\Rightarrow} 6PGL + NADPH$   $R_{2}: 6PGL + H_{2}O \stackrel{pgl}{\Rightarrow} 6PG$   $R_{3}: 6PG + NADP^{+} \stackrel{gnd}{\Rightarrow} R5P + NADPH$   $R_{4}: R5P \stackrel{rpe}{\Rightarrow} X5P$   $R_{5}: \alpha G6P \stackrel{gpi}{\Leftrightarrow} \beta G6P$   $R_{6}: \alpha G6P \stackrel{gpi}{\Leftrightarrow} \beta F6P$   $R_{7}: \beta G6P \stackrel{gpi}{\Leftrightarrow} \beta F6P$   $R_{8} :\Rightarrow \alpha G6P$   $R_{9}: X5P \Rightarrow$   $R_{10}: \Rightarrow H_{2}O$   $R_{11}: NADPH \Rightarrow NADP^{+}$ 

### Enzyme subsets of PPP

From the kernel, we can immediately identify enzyme subsets that operate with fixed flux ratios in any steady state:

- reactions  $\{R_1 R_4, R_8 R_{11}\}$  are one subset:  $R_{11}$  has double rate to all the others
- $\{R_6, R_7\}$  are another:  $R_6$  has the opposite sign of  $R_7$
- $R_5$  does not belong to non-trivial enzyme subsets, so it is not forced to operate in lock-step with other reactions

	0.2727	0.1066
	0.2727	0.1066
	0.2727	0.1066
	0.2727	0.1066
	0.3920	-0.4667
=	-0.1193	0.5733
	0.1193	-0.5733
	0.2727	0.1066
	0.2727	0.1066
	0.2727	0.1066
	0.5454	0.2132

K

#### Basis steady state flux modes from SVD

The kernel matrix obtained from SVD suffers from two shortcomings, illustrated by our small example system

- Reaction reversibility constraints are violated: in  $v_{svd1}$ , R5 operates in wrong direction, in  $v_{svd2}$ ,  $R_4$ operates in wrong direction
- All reactions are active in both flux modes, which makes visual interpretation impossible for all but very small systems
- The flux values are all non-integral



#### Choice of basis

- SVD is only one of the many ways that a basis for the null space can be defined.
- The root cause for hardness of interpretation is the orthonormality of matrix V in SVD  $S=U\Sigma V^T$ 
  - The basis vectors are orthogonal:  $v_{svd1}^T v_{svd2} = 0$
  - The basis vectors have unit length  $||v_{svd1}|| = ||v_{svd1}|| = 1$
- Neither criteria has direct biological relevance!

## Biologically meaningful pathways

- From our example system, it is easy  $V_1$  to find flux vectors that are more meaningful than those given by SVD
- Both pathways on the right statisfy the steady state requirement
- Both pathways obey the sign restrictions of the system
- One can easily verify (by solving b form the equation Kb = v) that they are linear combinations of the flux modes given by SVD, e.g.  $v_1 = 0.0373v_{svd1} + 1.997v_{svd2}$



### Elementary flux modes

The two pathways are examples of elementary flux modes

The study of elementary flux modes (EFM) and concerns decomposing the metabolic network into components that

- can operate independently from the rest of the metabolism, in a steady state,
- any steady state can be described as a combination of such components.



## Representing EFMs

• Elementary flux modes are given as  $V_1$  reaction rate vectors

 $\mathbf{e}=(e_1,\ldots,e_n),$ 

• EFMs typically consists of many zeroes, so they represent pathways in the network given by the non-zero components

$$P(\mathbf{e}) = \{j | e_j \neq 0\}$$



## Properties of elementary flux modes

The following properties are statisfied by EFMs:

- (Quasi-) Steady state
- Thermodynamical feasibility. Irreversible reactions need to proceed in the correct direction. Formally, one requires  $e_j \ge 0$  and that the stoichiometric coefficients  $s_{ij}$  are written with the sign that is consistent with the direction
- Non-decomposability. One cannot remove a reaction from an EFM and still obtain a reaction rate vector that is feasible in steady state. That is, if **e** is an EFM there is no vector **v** that satisfies the above and  $P(\mathbf{v}) \subset P(\mathbf{e})$

These properties define EFMs up to a scaling factor: if **e** is an EFM  $\alpha$ **e**,  $\alpha > 0$  is also an EFM.



#### EFMs and steady state fluxes

- Any steady state flux vector  $\mathbf{v}$  can be represented as a non-negative combination of the elementary flux modes:  $\mathbf{v} = \sum_{j} \alpha_j \mathbf{e}_j$ , where  $\alpha_j \ge 0$ .
- However, the representation is not unique: one can often find several coefficient sets α that satisfy the above.
- Thus, a direct composition of a flux vector into the underlying EFPs is typically not possible. However, the *spectrum* of potential contributions can be analysed

## EFMs of PPP

- One of the elementary flux modes of our PPP system is given below
- It consist of a linear pathway through the system, excluding reactions  $R_6$  and  $R_7$
- Reaction  $R_{11}$  needs to operate with twice the rate of the others



## EFMs of PPP

- Another elementary flux modes of our PPP system
- Similar linear pathway through the system, but excluding reactions  $R_5$  and using  $R_7$  in reverse direction
- Again, reaction  $R_{11}$  needs to operate with twice the rate of the others



## EFMs of PPP

- Third elementary flux mode contains only the small cycle composed of  $R_5$ ,  $R_7$ and  $R_6$ .  $R_6$  is used in reverse direction
- A yet another EFM would be obtained by reversing all the reactions in this cycle



Juho Rousu 28

## Building the kernel from EFMs

- In general there are more elementary flux modes than the dimension of the null space
- Thus a linearly independent subset of elementary flux modes suffices to span the null space
- In our PPP system, any two of the three EFMs together is linearly independent, and can thus be taken as the representative vectors

		Г			
	$R_1$	0	1	1	
	$R_2$	0	1	1	
	$R_3$	0	1	1	
	$R_4$	0	1	1	
	$R_5$	1	1	0	
EFM =	$R_6$	-1	0	1	
	$R_7$	1	0	-1	
	$R_8$	0	1	1	
	$R_9$	0	1	1	
	$R_{10}$	0	1	1	
	$R_{11}$	0	2	2	

## Software for finding EFMs

- From small systems it is relatively easy to find the EFMs by manual inspection
- For larger systems this becomes impossible, as the number of EFMs grows easily very large
- Computational methods have been devised for finding the EFMs by Heinrich & Schuster, 1994 and Urbanczik and Wagner, 2005
- Implemented in MetaTool package

#### Conservation relations

- As chemical reactions do not create or destroy matter, they obey conservation relations
- The counts of substrate and product molecules are balanced
- In the example reaction  $r: A + B \mapsto 2C$ , the sum A + B + 2C = const is constant.
- Other conserved quantitites:
  - Elemental balance: for each element species (C,N,O,P,...) the number of elements is conserved
  - Charge balance: total electrical charge, the total number of electrons in a reaction does not change.
  - Moiety balancing: it is possible to write balances for larger chemical moieties such as the co-factors (NAD,NADH, ATP, ADP,...)

#### Conservation relations from the stoichiometric matrix

- From the stoichiometric matrix conservation relations of metabolites can be found by examining the left null space of S, i.e. the set  $\{l|lS = 0\}$
- A basis spanning the left null space can be obtained from SVD  $S = U\Sigma V^T$ : the last m - r columns of the matrix U span the left null space, where r is the rank of S
- In MATLAB the basis can be computed by the command null(S').



## Conservation in PPP

The left null space of our PPP system only contains a single vector, stating that the sum of NADP<sup>+</sup> and NADPH is constant in all reactions.

