# Finding Minimal Generating Set for Metabolic Network with Reversible Pathways

Dimitrije Jevremović[a,*], Daniel Boley[a]

[a]*Computer Science & Engineering, University of Minnesota, Minneapolis 55455 MN, USA*

## Abstract

Elementary flux modes give a mathematical representation of metabolic pathways in metabolic networks satisfying the constraint of non-decomposability. The large cost of their computation shifts attention to computing a minimal generating set which is a conically independent subset of elementary flux modes. When a metabolic network has reversible reactions and also admits a reversible pathway, the minimal generating set is not unique. A theoretical development and computational framework is provided which outline how to compute the minimal generating set in this case. The method is based on combining existing software to compute the minimal generating set for a "pointed cone" together with standard software to compute the Reduced Row Echelon Form.

*Keywords:*
metabolic networks, elementary flux modes, minimal generating set, metabolic pathways

## 1. Introduction

One approach to analyze cellular metabolic networks of biochemical reactions is by means of the constraint-based stoichiometry network analysis. In this analysis, the quasi-steady state of a network is represented by a matrix of coefficients $\mathbf{S}$ in which the element $s_{ij}$ in the $i$-th row and $j$-th column is the molar amount of metabolite $i$ produced by a unit flux of reaction $j$ (a negative entry denotes consumption). The reactions can be reversible or irreversible; where the latter are constrained to a single direction only.

All the reactions in the network are characterized by their *flux rate* values which correspond to the speed of reaction execution. A set of reaction fluxes is collected into a *metabolic flux vector* $\mathbf{x}$, whose non-zero entries represent active reactions. In order to be feasible, the flux rates of irreversible reactions must be non-negative. The metabolic flux vector quantitatively corresponds to and describes a concept of a *metabolic pathway*. The quasi-steady state of the cellular metabolism imposes the constraint of mass-balance on the internal metabolites: net consumption must match net production. This leads to the equality constraints $\mathbf{S} \cdot \mathbf{x} = 0$, i.e., the flux vector must lie in the right nullspace of the stoichiometry matrix.

---

*jevrem@cs.umn.edu

The solutions of the stoichiometry equation $\mathbf{S} \cdot \mathbf{x} = 0$, which also satisfies the non-negativity constraints for the flux of its irreversible reactions, describe all possible metabolic states in which the metabolic network may be found. Geometrically this solution space corresponds to the polyhedral cone (Schrijver, 1988; Wagner and Urbanczik, 2005), and it may be fully generated by means of its extreme rays (Fukuda and Prodon, 1996). Extreme rays are conically independent set of vectors and in the convex analysis are also known as a *minimal generating set*. In the stoichiometry network analysis, alongside with the concept of the minimal generating set, stand the extreme pathways and elementary flux modes. It is important to say that the *minimal generating set*, *extreme pathways* and *elementary flux modes* are computed using the standard Double Description Method for the enumeration of extreme rays (i) when no reversible reactions are split, (ii) only internal reversible reactions are split, and (iii) all of the reversible reactions are split, into two irreversible components, respectively (Schuster and Hilgetag, 1994; Wagner and Urbanczik, 2005; Schilling et al., 2000; Llaneras and Pico, 2010; Jevremovic et al., 2010). Unlike the minimal generating set, in the original reaction space the extreme pathways and elementary flux modes are not necessarily conically independent which depends on the existence and number of the reversible reactions.

In the absence of reversible reactions, the *minimal generating set*, *extreme pathways* and *elementary flux modes* coincide, are uniquely defined, and correspond to the *extreme rays* of the polyhedral cone. Regarding the directionality of the metabolic pathways which the metabolic network accepts we distinguish two cases.

In the first case, if the metabolic network admits only irreversible pathways (i.e., every pathway contains at least one irreversible reaction), then the minimal generating set is unique, and the corresponding polyhedral cone is said to be *pointed*. On the other side, in the second case, if the metabolic network admits reversible pathways, the minimal generating set is no longer unique and the polyhedral cone is *not pointed*.

Regardless if the cone is pointed or not, the set of the elementary flux modes (or extreme pathways) is a superset of any minimal generating set, and some of the elementary flux modes (or extreme pathways) may lie in the interior of the cone. In addition, putatively exponential hardness and high computational cost of the algorithm used to compute elementary flux modes (Acuña et al., 2009, 2010) is another reason to shift the attention from extreme pathways and elementary flux modes to the minimal generating set.

Answering many questions requires the use of extreme pathways (Schilling and Palsson, 2000; Papin et al., 2002; Wiback et al., 2002; Wiback and Palsson, 2002; Price et al., 2002) and elementary flux modes (Stelling et al., 2002; Trinh et al., 2008; Trinh and Srienc, 2009; Trinh et al., 2009; Pérès et al., 2011; Flynn et al., 2012), however there are several applications one can answer with minimal generating sets. This situation especially arises in the case of genome-scale metabolic networks where the computation of elementary modes is prohibitively expensive (Larhlimi, 2008). Some simple structural properties may be observed, such as whether any reversible reaction appears only in one direction or only in irreversible pathways, or whether some reaction appears in no pathway at all. Flux coupling analysis, a procedure of determining dependencies between network reactions, can be accomplished using the minimal generating set vectors (Larhlimi and Bockmayr, 2006). Control-effective

analysis of individual reactions in the network was initially proposed on the basis of computed elementary flux modes (Stelling et al., 2002). However, minimal generating sets can be used to obtain an analogous control-effective metric, used in the regulatory network analysis and reaction importance assessment (Larhlimi, 2008). Minimal metabolic behaviors are exposed by the minimal generating set (Larhlimi and Bockmayr, 2009), but a simple method to compute it is still needed. In large genome-scale networks, where computation of entire minimal generating sets may be impractical, efforts have been made to compute the K-shortest minimal generating vectors (Rezola et al., 2011) (i.e., pathways involving as few reactions as possible). This was accomplished by means of solving several linear optimization problems and using existing methods for the computation of K-shortest elementary flux modes (de Figueiredo et al., 2009).

The problem of computing the minimal generating set for the metabolic network which admits reversible pathways is the topic of this manuscript. An earlier analysis of the metabolic networks with reversible pathways by means of two subnetworks, one with no reversible pathways and one with all reversible pathways, can be found in (Larhlimi and Bockmayr, 2009). The computation of the unique minimal generating set for a pointed cone can be accomplished using existing algorithms (von Kamp et al., 2006; Terzer and Stelling, 2008; Jevremovic et al., 2011) or using the general paradigm in (Jevremovic et al., 2010). But this is considerably more difficult when the cone is not pointed (i.e., there are reversible pathways). This situation can be recognized by computing the rank of the submatrix of $\mathbf{S}$ consisting of the reversible reactions (Jevremovic et al., 2010).

The major contribution of this paper is to provide a simple procedure to compute the minimal generating set for a stoichiometric network which has reversible pathways. The method is based on combining two existing algorithms: a method to compute the minimal generating set for a pointed cone, and a method to compute a nullspace of a matrix based on the Reduced Row Echelon Form, a classical method in linear algebra. All this is carried out without the necessity to compute all the elementary flux modes for any network. This paper is organized as follows. Section 2 gives a theoretical treatment of the representation of reversible and irreversible pathways and the decomposition of the original metabolic network into two subnetworks. Section 3 outlines the algorithm for the computation of the minimal generating set using two subnetworks. Section 4 uses a simple example to illustrate the method and show how the method exposes some of the structure of the network.

## 2. Theory

Let $\mathbf{S} = (\mathbf{A}, \mathbf{B}, \mathbf{C})$ be an $m \times n$ stoichiometry matrix with the $n$ columns (reactions) ordered so that $\mathbf{A}$ consists of the irreversible reactions (of which there are $n_i$) and $\mathbf{B}, \mathbf{C}$ consists of the reversible reactions (of which there are $n_r$). We assume the reversible reactions $(\mathbf{B}, \mathbf{C})$ form a matrix of rank $k_r$ and that $\mathbf{B}$ consists of $k_r$ columns which are independent, while $\mathbf{C}$ consists of $n_r - k_r$ columns. This implies that all the columns of $\mathbf{C}$ can be written as linear combinations of the columns $\mathbf{B}$: $\mathbf{C} = \mathbf{B}R$ for some $k_r \times (n_r - k_r)$ coefficient matrix $R$. We remark that the columns of $\mathbf{B}$ can be found by a variety of methods such as the Reduced Row Echelon Form (RREF) (Lay, 2012) where they appear as the "pivot" columns,

while the columns $\mathbf{C}$ appear as the "non-pivot" columns. Hence we will refer to $\mathbf{B}$ as the "pivot" columns. The standard RREF algorithm scans the matrix $\mathbf{S}$ left-to-right extracting independent columns $\mathbf{B}$, hence the choice of pivot columns varies depending on the order of columns (reactions) in the original $\mathbf{S}$, but once the latter is fixed, the former is also.

The matrix $(\mathbf{B}, \mathbf{C})$ has a nullspace of dimension $n_r - k_r$, and a suitable basis for this space is $N_R = \begin{pmatrix} -R \\ I \end{pmatrix}$. Any vector in this nullspace is a valid path for the subnetwork $(\mathbf{B}, \mathbf{C})$ and is a reversible path. By prepending zeros, we obtain

$$\widehat{N}_R = \begin{pmatrix} 0 \\ N_R \end{pmatrix} = \begin{pmatrix} 0 \\ -R \\ I \end{pmatrix},$$

which we will show is a minimal basis for the set of all reversible paths in the original network.

A column vector $\mathbf{x}$ is a valid path of the network represented by stoichiometry matrix $\mathbf{S}$ if and only if $\mathbf{S}\mathbf{x} = \mathbf{0}$ and the entries of $\mathbf{x}$ corresponding to irreversible reactions are non-negative. If we split $\mathbf{x} = (\mathbf{x}_a; \mathbf{x}_b; \mathbf{x}_c)$ to conform with $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ (where ";" denote vertical concatenation à la Matlab), then $\mathbf{x}$ is a valid path if and only if $\mathbf{A}\mathbf{x}_a + \mathbf{B}\mathbf{x}_b + \mathbf{C}\mathbf{x}_c = 0$ and $\mathbf{x}_a \geq \mathbf{0}$ (elementwise).

We have the following Lemmas:

**Lemma 1.** Any reversible pathway $\mathbf{x}$ for the stoichiometry matrix $\mathbf{S} = (\mathbf{A}, \mathbf{B}, \mathbf{C})$ split as above can be written in terms of the minimal generating set for the reversible subnetwork $(\mathbf{B}, \mathbf{C})$, as follows:

$$\mathbf{x} \equiv \begin{pmatrix} \mathbf{x}_a \\ \mathbf{x}_b \\ \mathbf{x}_c \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{x}_b \\ \mathbf{x}_c \end{pmatrix} = \widehat{N}_R \boldsymbol{\alpha} \equiv \begin{pmatrix} \mathbf{0} \\ -R \\ I \end{pmatrix} \boldsymbol{\alpha}$$

for some coefficient vector $\boldsymbol{\alpha}$. ∎

Since there are no sign constraints in the subnetwork represented by $(\mathbf{B}, \mathbf{C})$, the basis $N_R$ is the minimal generating set for all possible reversible paths. In fact any basis for the nullspace of $(\mathbf{B}, \mathbf{C})$ would be a minimal generating set, but we choose this specific one because each column in this basis has a minimal set of non-zeros, i.e., each is also an elementary flux mode. In this sense, we call this a "minimal basis" or "minimal generating set." There is still freedom to choose any set of $k_r$ independent columns of $(\mathbf{B}, \mathbf{C})$ (reversible reactions) to act as the basis $\mathbf{B}$.

**Lemma 2.** Any pathway $\mathbf{x}$ for the stoichiometry matrix $\mathbf{S} = (\mathbf{A}, \mathbf{B}, \mathbf{C})$ split as above can be written as the sum of a path involving just reactions indexing columns in $\mathbf{A}, \mathbf{B}$ and a reversible path involving just reactions indexing columns in $\mathbf{B}, \mathbf{C}$:

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_a \\ \mathbf{x}_b \\ \mathbf{x}_c \end{pmatrix} = \begin{pmatrix} \mathbf{x}_a \\ \tilde{\mathbf{x}}_b \\ \mathbf{0} \end{pmatrix} + \widehat{N}_R \mathbf{x}_c$$

where $\tilde{\mathbf{x}}_b = \mathbf{x}_b + R\mathbf{x}_c$.

**Proof.**

$$\mathbf{x} \equiv \begin{pmatrix} \mathbf{x}_a \\ \mathbf{x}_b \\ \mathbf{x}_c \end{pmatrix} = \begin{pmatrix} \mathbf{x}_a & + & \mathbf{0} \\ \mathbf{x}_b + R\mathbf{x}_c & - & R\mathbf{x}_c \\ \mathbf{0} & + & \mathbf{x}_c \end{pmatrix} = \begin{pmatrix} \mathbf{x}_a \\ \tilde{\mathbf{x}}_b \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} 0 \\ -R \\ I \end{pmatrix} \mathbf{x}_c.$$

As a valid path, $\mathbf{x}$ lies in the nullspace of $\mathbf{S}$, and so does the term $\widehat{N}_R\mathbf{x}_c$, hence the remaining term $\begin{pmatrix} \mathbf{x}_a \\ \tilde{\mathbf{x}}_b \\ \mathbf{0} \end{pmatrix}$ must also lie in the nullspace of $\mathbf{S}$ and hence is a valid path. Only the component $\mathbf{x}_a$ is subject to sign constraints. ∎

**Theorem 3.** Any pathway $\mathbf{x}$ for the stoichiometry matrix $\mathbf{S} = (\mathbf{A}, \mathbf{B}, \mathbf{C})$ split as above can be written as the sum of a linear combination of paths in the minimal generating set $\widehat{M}_I$ for the "pointed-cone" subnetwork represented by $(\mathbf{A}, \mathbf{B})$ together with a linear combination of the minimal generating set $\widehat{N}_R$ for the network of reversible reactions represented by $(\mathbf{B}, \mathbf{C})$.

**Proof.** Let $M_I$ be a matrix whose columns form the minimal generating set for the subnetwork $(\mathbf{A}, \mathbf{B})$. This network has no reversible pathways because the columns corresponding to the reversible reactions are linearly independent, i.e., the space of valid paths for this network is a pointed cone (Wagner and Urbanczik, 2005).

By Lemma 2, any path $\mathbf{x}$ through the entire network can be written as

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_a \\ \tilde{\mathbf{x}}_b \\ 0 \end{pmatrix} + \widehat{N}_R\mathbf{x}_c.$$

Now $\begin{pmatrix} \mathbf{x}_a \\ \tilde{\mathbf{x}}_b \end{pmatrix}$ is a path through the subnetwork $(\mathbf{A}, \mathbf{B})$ and hence can be written in terms of the minimal generating set for $(\mathbf{A}, \mathbf{B})$ as $M_I\boldsymbol{\beta}$ for some coefficient vector $\boldsymbol{\beta}$. Hence we have

$$\mathbf{x} = \widehat{M}_I\boldsymbol{\beta} + \widehat{N}_R\alpha, \quad \text{with} \quad \widehat{M}_I = \begin{pmatrix} M_I \\ 0 \end{pmatrix},$$

where we have extended $M_I$ with a block of zeros so that $(\mathbf{A}, \mathbf{B}) \cdot M_I = (\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \widehat{M}_I$. ∎

**Lemma 4.** The generating set for the network represented by the stoichiometry matrix $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ consisting of the union of the minimal generating set $\widehat{M}_I$ of matrices $(\mathbf{A}, \mathbf{B})$ and the minimal generating set $\widehat{N}_R$ of matrices $(\mathbf{B}, \mathbf{C})$, constructed as in Theorem 3, is minimal.

**Proof.** By Theorem 3, any given pathway $\mathbf{x} = (\mathbf{x}_a; \mathbf{x}_b; \mathbf{x}_c)$ can be written as $\mathbf{x} = \widehat{M}_I\boldsymbol{\beta} + \widehat{N}_R\alpha$. By Lemma 2 and its proof, the given pathway can be written as $\mathbf{x} = \begin{pmatrix} \mathbf{x}_a \\ \tilde{\mathbf{x}}_b \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} 0 \\ -R \\ I \end{pmatrix} \mathbf{x}_c$, which implies that the coefficients of the vectors in $\widehat{N}_R$ in the decomposition are equal to the elements in $\mathbf{x}_c$, hence $\alpha = \mathbf{x}_c$. Since $\widehat{N}_R$ consists of the identity matrix in its lower part, the residual vector $\mathbf{x} - \widehat{N}_R\mathbf{x}_c$ lies in the pointed polyhedral cone corresponding to $(\mathbf{A}, \mathbf{B})$ and therefore can be written in terms of $\widehat{M}_I$ in exactly one way: $\mathbf{x} - \widehat{N}_R\mathbf{x}_c = \widehat{M}_I\boldsymbol{\beta}$. Hence no member of the generating set $\widehat{M}_I \cup \widehat{N}_R$ can be written in terms of the other members of the generating set. So the generating set must be minimal. ∎

## 3. Discussion

Theorem 3 implies that we can compute a minimal generating set for a stoichiometry matrix $\mathbf{S}$ by the following procedure.

1. Collect all columns corresponding to irreversible reactions of $\mathbf{S}$ into matrix $\mathbf{A}$.
2. Use the Reduced Row Echelon Form (RREF) (or similar method) on the matrix of all reversible reactions to extract the matrices $\mathbf{B}$ and $\mathbf{C}$, where $\mathbf{B}$ has full column rank and $\mathbf{C}$ can be written as $\mathbf{B}R$ for some $R$, so that $(\mathbf{B}, \mathbf{C})$ consists of the columns corresponding to all the reversible reactions.
3. Compute a basis $N_R$ as the right nullspace of matrix $(\mathbf{B}, \mathbf{C})$.
4. Compute the minimal generating set $M_I$ for the subnetwork represented by $(\mathbf{A}, \mathbf{B})$ (a pointed cone).
5. Extending $M_I$ with a block of zeros to obtain $\widehat{M_I}$, we obtain a minimal generating set for all valid paths of $\mathbf{S}$, namely the columns of the combined matrix $(\widehat{M_I}, \widehat{N_R})$.

This procedure allows one to compute the minimal generating set for an arbitrary network by computing the minimal generating sets of two subnetworks, one of which is a pointed cone and the other one having reversible pathways is a non-pointed cone. Since it is well known that the minimal generating set of a network is almost always an order of magnitude smaller than the set of elementary flux modes (Larhlimi and Bockmayr, 2009), this procedure allows one to compute the minimal generating set for an arbitrary network at much less cost compared to an algorithm based on a full set of elementary flux modes (Llaneras and Pico, 2010).

We remark that, by construction, the minimal generating set $\widehat{M_I}$, of the subnetwork consisting only of irreversible pathways, is essentially unique once the basis $\mathbf{B}$ for the space of reversible reactions is chosen. They vary only in the combinations of reversible reactions. The patterns of irreversible reactions in the pathways of $\widehat{M_I}$ correspond to the minimal metabolic behaviors of (Larhlimi and Bockmayr, 2009).

However, there is quite a large freedom of choice for the minimal generating set $\widehat{N_R}$ corresponding to the subnetwork with reversible pathways. Any basis for the nullspace of $(\mathbf{B}, \mathbf{C})$ will do. By using the nullspace derived from the RREF, we can ensure that each reversible pathway in $\widehat{N_R}$ is minimal (i.e., has a minimal set of non-zero entries).

## 4. Example

### 4.1. Toy metabolic network

We illustrate the method with a small example derived from (Larhlimi and Bockmayr, 2009). We have made reactions $R6$, $R7$ reversible in order to illustrate some structure exposed by this method.

Stoichiometry Matrix **S**:

$$\mathbf{S} = \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \\ \end{array}$$

| | R1r | R2 | R3r | R4r | R5r | R6r | R7r | R8 | R9r | R10r | R11r | R12r | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | −1 | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | m1 |
| | _ | 1 | −1 | _ | −1 | _ | _ | _ | _ | _ | _ | _ | m2 |
| | _ | 1 | _ | −1 | _ | −1 | _ | _ | _ | _ | _ | _ | m3 |
| **S** = | _ | _ | _ | _ | 1 | _ | _ | 1 | −1 | _ | −1 | _ | m4 |
| | _ | _ | _ | _ | _ | 1 | −1 | −1 | _ | _ | _ | _ | m5 |
| | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | 1 | −1 | m6 |
| | _ | _ | _ | _ | _ | _ | _ | _ | 1 | −1 | _ | _ | m7 |
| | B | A | B | B | B | B | C | A | B | C | B | C | partition |

Here we show the partition into which each reaction ends up after extracting the reversible part and applying the RREF. The method consists of three steps as follows.

## 1. Get minimal generating set for reversible pathways subnetwork.

- Extract all columns from **S** corresponding to reversible reactions into matrix $(\mathbf{B}, \mathbf{C})$.

- Compute basis of right Nullspace of $(\mathbf{B}, \mathbf{C})$ using the Reduced Row Echelon Form (RREF). The rank of the matrix $(\mathbf{B}, \mathbf{C})$ is 7, so there will be 7 pivot columns. The dimension of the nullspace is 3, so the minimal generating set will have 3 entries.

- The pivot reactions correspond to what RREF identified as the pivot columns. The pivot columns form a basis for the entire column space of the reversible columns of **S**.

The columns have been ordered so that the pivot reversible reactions are: $\{R1r, R3r, R4r, R5r, R6r, R9r, R11r\}$. The resulting reversible minimal generating set has 3 entries labeled $R$ in formula (1) below. We remark that this minimal generating set is not unique, as discussed further below. In the following paragraphs, we show how the choice of pivot reversible reactions is not completely arbitrary, but depends on the particular partitioning among all the reversible reactions.

Notice in the $R$ columns of (1) how $R1r$ is isolated, and $R4r$, $R6r$, $R7r$ form combination disjoint from the remaining reactions. Specifically, the first $R$ column shows that the sum of $R6r$, $R7r$, minus $R4r$ is zero. No $R$ column involves $R1r$. Hence $R1r$ cannot be written as a linear combination of any other reaction, and any one of $R4r$, $R6r$, $R7r$ can be written as a linear combination of the other two, but none of these four reactions can be written as combinations of the remaining reversible reactions. This also shows that $R1r$ is not involved in any reversible pathway.

Regarding the irreversible reactions, there is one irreversible pathway containing just $R2$ and one containing just $R8$, hence those two correspond to the minimal metabolic behaviors in the sense of (Larhlimi and Bockmayr, 2009) [1]. Since $R1r$ only appears in the positive

---

[1] Recall this network has been modified from that of (Larhlimi and Bockmayr, 2009) for purposes of illustration

direction in a pathway also with $R2$, one might consider $R1$ to be irreversible with respect to this network and the two minimal metabolic behaviors to be $\{R1, R2\}$ and $\{R8\}$.

The $R$ columns of (1) is a generalized incidence matrix where two reactions are connected by an edge if they appear in a common path. This shows the set of reversible reactions can be partitioned into three disjoint connected subgraphs: $\{R1r\}$, $\{R4r, R6r, R7r\}$, $\{R3r, R5r, R9r, R10r, R11r, R12r\}$, with no connections between the subgraphs. Any pointed cone subset must have a certain number of members from each connected subgraph, where the number of members is equal to the rank of the corresponding columns of the stoichiometry matrix.

So any pointed cone subset in this example must include (in addition to all the irreversible reactions) $R1r$, and any 2 out of $\{R4r, R6r, R7r\}$, plus some 4 out of $\{R3r, R5r, R9r, R10r, R11r, R12r\}$ but not any combination. Varying the choice of reversible reactions not only yields a different pointed cone, but also a different minimal generating set for the reversible subnetwork. However, once the minimal generating set for the reversible part is chosen, the rest is all uniquely determined.

## 2. Get minimal generating set for irreversible pathways subnetwork.

- Combine irreversible reactions with the pivot reversible reactions to form a subnetwork with no reversible pathway (called a Pointed Cone Subset). In this example there are 2 irreversible reactions and 7 pivot reactions to form a subnetwork of 9 reactions.

- Compute minimal generating set for resulting reduced stoichiometry matrix $(\mathbf{A}, \mathbf{B})$ (using usual Nullspace algorithm).

The pointed cone subset has 9 reactions: $\{R2, R8\}$, $\{R1r\}$, $\{R4r, R6r\}$, $\{R3r, R5r, R9r, R11r\}$ (includes all irreversible reactions plus "pivot" reversible reactions, grouped by partitioning of the reactions induced from above.)

Applying the Nullspace algorithm yields 2 minimal generating vectors labeled $I$ in formula (1).

## 3. Combine the above two minimal generating sets.

All 5 paths in the two computed minimal generating set for this example are combined to represent the minimal generating set of the original metabolic network, a shown in (1).

Matrix (1):

| I | I | R | R | R | |
|---|---|---|---|---|---|
| . | 1 | . | . | . | $R1r$ |
| . | 1 | . | . | . | $R2$ |
| 1 | 1 | . | -1 | -1 | $R3r$ |
| -1 | 1 | -1 | . | . | $R4r$ |
| -1 | . | . | 1 | 1 | $R5r$ |
| 1 | . | 1 | . | . | $R6r$ |
| . | . | 1 | . | . | $R7r$ |
| 1 | . | . | . | . | $R8$ |
| . | . | . | 1 | . | $R9r$ |
| . | . | . | 1 | . | $R10r$ |
| . | . | . | . | 1 | $R11r$ |
| . | . | . | . | 1 | $R12r$ |

(1)

The label $I$ denotes an irreversible minimal generating vector derived from the pointed cone subnetwork $(\mathbf{A}, \mathbf{B})$, and $R$ denotes a reversible minimal generating vector derived from the reversible subnetwork $(\mathbf{B}, \mathbf{C})$.

This system has 21 EMs consisting of the 5 members of the minimal generating set (1) plus the following 16 EMs:

| R | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| . | . | 1 | 1 | 1 | 1 | . | . | 1 | 1 | 1 | 1 | . | . | 1 | 1 | $R1r$ |
| . | . | 1 | 1 | 1 | 1 | . | . | 1 | 1 | 1 | 1 | . | . | 1 | 1 | $R2$ |
| . | 1 | 2 | 1 | . | 1 | . | . | . | . | . | 1 | . | . | . | . | $R3r$ |
| . | . | . | . | 1 | . | -1 | . | . | . | 1 | . | -1 | . | . | . | $R4r$ |
| . | -1 | -1 | . | 1 | . | . | . | 1 | 1 | 1 | . | . | . | 1 | 1 | $R5r$ |
| . | . | 1 | 1 | . | 1 | 1 | . | 1 | 1 | . | 1 | 1 | . | 1 | 1 | $R6r$ |
| . | -1 | . | 1 | . | . | . | -1 | 1 | . | . | . | . | -1 | 1 | . | $R7r$ |
| . | 1 | 1 | . | . | 1 | 1 | 1 | . | 1 | . | 1 | 1 | 1 | . | 1 | $R8$ |
| -1 | . | . | . | 1 | 1 | 1 | 1 | 1 | 2 | . | . | . | . | . | . | $R9r$ |
| -1 | . | . | . | 1 | 1 | 1 | 1 | 1 | 2 | . | . | . | . | . | . | $R10r$ |
| 1 | . | . | . | . | . | . | . | . | . | 1 | 1 | 1 | 1 | 1 | 2 | $R11r$ |
| 1 | . | . | . | . | . | . | . | . | . | 1 | 1 | 1 | 1 | 1 | 2 | $R12r$ |

(2)

The third column of (2) is an elementary mode for the pointed cone subset $(\mathbf{A}, \mathbf{B})$.

To see that (1) is indeed the minimal generating set, we express all the EMs in (2) as linear combinations of the paths in (1):

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| . | 1 | 1 | . | . | 1 | 1 | 1 | . | 1 | . | 1 | 1 | 1 | . | 1 | $(\mathbf{A}, \mathbf{B})$ |
| . | . | 1 | 1 | 1 | 1 | . | . | 1 | 1 | 1 | 1 | . | . | 1 | 1 | $(\mathbf{A}, \mathbf{B})$ |
| . | -1 | . | 1 | . | . | . | -1 | 1 | . | . | . | . | -1 | 1 | . | $(\mathbf{B}, \mathbf{C})$ |
| -1 | . | . | . | 1 | 1 | 1 | 1 | 1 | 2 | . | . | . | . | . | . | $(\mathbf{B}, \mathbf{C})$ |
| 1 | . | . | . | . | . | . | . | . | . | 1 | 1 | 1 | 1 | 1 | 2 | $(\mathbf{B}, \mathbf{C})$ |

9

*4.2. Red Blood Cell metabolism*

We can apply the algorithm to find minimal generating set for the concrete example of the red blood cell metabolic network (Wiback and Palsson, 2002). This is also just a simple illustration of a simple biochemical question that can be answered with the minimal generating set: namely that certain reversible reactions operate only in one direction. This network has 32 internal reactions (17 reversible) and 19 external reactions (16 reversible), giving a total of 51 reactions (33 reversible). The network has 6,180 elementary flux modes, but the algorithm of this paper shows only 18 of these modes form the minimal generating set (with 1 reversible). In this example $\mathbf{A}$ consists of the 18 columns corresponding to the 18 irreversible reactions, $\mathbf{B}$ consists of 32 columns, and $\mathbf{C}$ has one column. The one reversible pathway is found to consist of reactions LD, TRA3r, TRA4r, TRA10r, TRA11r, TRA14r. A simple examination of the minimal generating set suffices to show that the following reversible reactions appear only in the positive direction: ALD, TPI, GAPDH, PGM, EN, LD, PGL, PRM, PNPase, TRA4r, TRA13r, TRA16r, TRA17r, and the following only in the negative direction: TRA7r, TRA12r, TRA14r.

Supplementary MATLAB script which facilitates the computation of minimal generating set using the software for the computation of elementary flux modes is given at http://elmocomp.sourceforge.net/mingen.zip. Software for the computation of elementary flux modes has a command line option to perform the processing of constraints only corresponding to irreversible reactions.

## 5. Conclusion

The minimal generating set represents a minimal subset of admissible elementary flux modes which fully characterizes the polyhedral cone corresponding to the given metabolic network. In this paper, a theoretical treatment of irreversible and reversible metabolic pathways is reviewed and a simple procedure to compute the minimal generating set for an arbitrary biochemical network is given. The proposed method reduces the problem of computing the minimal generating set for a network with reversible pathways to the simpler problem of computing the minimal generating set for the network without reversible pathways. The method is based on splitting a metabolic network into two subnetworks, one entirely reversible, and the other without reversible pathways.

## 6. Acknowledgement

## References

Acuña, V., Chierichetti, F., Lacroix, V., Marchetti-Spaccamela, A., Sagot, M., Stougie, L., 2009. Modes and cuts in metabolic networks: Complexity and algorithms. BioSystems 95 (1), 51–60.

Acuña, V., Marchetti-Spaccamela, A., Sagot, M., Stougie, L., 2010. A note on the complexity of finding and enumerating elementary modes. Biosystems 99 (3), 210–214.

de Figueiredo, L. F., Podhorski, A., Rubio, A., Kaleta, C., Beasley, J., Schuster, S., Planes, F. J., 2009. Computing the shortest elementary flux modes in genome-scale metabolic networks. Bioinformatics 25 (23), 3158–3165.

Flynn, C., Hunt, K., Gralnick, J., Srienc, F., 2012. Construction and elementary mode analysis of a metabolic model for Shewanella oneidensis MR-1. Biosystems 107 (2), 120–128.

Fukuda, K., Prodon, A., 1996. Double description method revisited. In: Deza, M., Euler, R., Manoussakis, I. (Eds.), Combinatorics and Computer Science. Springer, pp. 91–111, also tech. report, Mathematics, ETH, 1995.

Jevremovic, D., Trinh, C., Srienc, F., Boley, D., 2010. On algebraic properties of extreme pathways in metabolic networks. Journal of Computational Biology 17 (2), 107–119.

Jevremovic, D., Trinh, C. T., Srienc, F., Sosa, C., Boley, D., 2011. Parallelization of nullspace algorithm for the computation of elementary flux modes. Parallel Computing Journal 37 (6-7), 261–278.

Larhlimi, A., July 2008. New Concepts and Tools in Constraint-based Analysis of Metabolic Networks. Ph.D. thesis, Freie Universität, Berlin.

Larhlimi, A., Bockmayr, A., 2006. A new approach to flux coupling analysis of metabolic networks. In: Computational Life Sciences II. Vol. 4216 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg, pp. 205–215.

Larhlimi, A., Bockmayr, A., 2009. A new constraint-based description of the steady-state flux cone of metabolic networks. Discrete Applied Mathematics 157 (10), 2257–2266.

Lay, D., 2012. Linear Algebra and Its Applications, 4th Edition. Addison Wesley.

Llaneras, F., Pico, J., 2010. Which metabolic pathways generate and characterize the flux space? a comparison among elementary modes, extreme pathways and minimal generators. J of Biomedicine and Biotechnology 2010, article ID 753904.

Papin, J., Price, N., Edwards, J., Palsson, B., 2002. The genome-scale metabolic extreme pathway structure in haemophilus influenzae shows significant network redundancy. Journal of Theoretical Biology 215, 67–82.

Pérès, S., Vallée, F., Beurton-Aimar, M., Mazat, J., 2011. ACoM: A classication method for elementary ux modes based on motif nding. Biosystems 103, 410–419.

Price, N., Papin, J., Palsson, B., 2002. Determination of redundancy and systems properties of the metabolic network of helicobacter pylori using genome-scale extreme pathway analysis. Genome Research 12, 760–769.

Rezola, A., de Figueiredo, L., Brock, M., Pey, J., Podhorski, A., Wittmann, C., Schuster, S., Bockmayr, A., Planes, F. J., 2011. Exploring metabolic pathways in genome-scale networks via generating flux modes. Bioinformatics 27 (4), 534–540.

Schilling, C. H., Letscher, D., Palsson, B., 2000. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. Journal of Theoretical Biology 203 (3), 229–248.

Schilling, C. H., Palsson, B., 2000. Assessment of the metabolic capabilities of Haemophilus influenzae rd through a genome-scale pathway analysis. Journal of Theoretical Biology 203 (3), 249–283.

Schrijver, A., 1988. Theory of linear and integer programming. Wiley.

Schuster, S., Hilgetag, C., 1994. On elementary flux modes in biochemical reaction systems at steady state. Journal of Biological Systems 2 (2), 165–182.

Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., Gilles, E. D., 2002. Metabolic network structure determines key aspects of functionality and regulation. Nature 420 (6912), 190–193.

Terzer, M., Stelling, J., 2008. Large Scale computation of elementary flux modes with bit pattern trees. Bioinformatics.

Trinh, C., Srienc, F., 2009. Metabolic engineering of escherichia coli for efficient conversion of glycerol to ethanol. Applied and Environmental Microbiology 75 (21), 6696–6705.

Trinh, C., Unrean, P., Srienc, F., 2008. A minimal Escherichia coli cell for most efficient ethanol production from hexoses and pentoses. Applied and Environmental Microbiology 74 (12), 3634–3643.

Trinh, C., Wlaschin, A., Srienc, F., 2009. Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. Appl. Microbiol. Biotechnol 22 (5), 813–826.

von Kamp, A., Schuster, S., 2006. Metatool 5.0: fast and flexible elementary modes analysis. Bioinformatics 22 (15), 1930–1931.

Wagner, C., Urbanczik, R., 2005. The geometry of the flux cone of a metabolic network. Biophysics Journal 89 (6), 3837–3845.

Wiback, S., Mahadevan, R., Palsson, B., 2002. Reconstructing metabolic flux vectors form extreme pathways: defining the alpha-spectrum. Journal of Theoretical Biology (3), 313–324.

Wiback, S., Palsson, B., 2002. Extreme pathway analysis of human red blood cell metabolism. Biophysics Journal 83 (2), 808–818.