On Algebraic Properties of Extreme Pathways in Metabolic Networks

Dimitrije Jevremovic[1], Cong T. Trinh[2], Friedrich Srienc[3], Daniel Boley[1]

Abstract

We give a concise development of some of the major algebraic properties of extreme pathways (pathways which cannot be the result of combining other pathways) of metabolic networks, contrasting them to those of elementary flux modes (pathways involving a minimal set of reactions). In particular, we show that an extreme pathway can be recognized by a rank test as simple as the existing rank test for elementary flux modes, without computing all the modes. We make the observation that, unlike elementary flux modes, the property of being an extreme pathway depends on the presence or absence of reactions beyond those involved in the pathway itself. Hence the property of being an extreme pathway is not a local property. As a consequence, we find that the set of all elementary flux modes for a network includes all the elementary flux modes for all its subnetworks, but that this property does not hold for the set of all extreme pathways.

keywords: algorithms, biochemical networks, linear algebra

[1]Department of Computer Science & Engineering, Univ. of Minnesota, Minneapolis, MN 55455, USA
[2]Energy Biosciences Institute, Univ. of California, Berkeley, CA 94720-5230, USA
[3]Biotechnology Institute, Univ. of Minnesota, St. Paul, MN 55108, USA

1 Introduction

The study of metabolic networks by means of their steady-state stoichiometry is by now a well established research activity, and the analysis of such networks by means of their elementary flux modes or pathways is recognized as a fundamental part of any study of such networks (Pfeiffer et al., 1999; Schuster & Hilgetag, 1994; Schuster et al., 1999; Schuster et al., 2002). The study of so-called extreme pathways has also been used by many researchers (Schilling et al., 2000; Papin et al., 2004). In this paper, we present mathematical algebraic properties of extreme pathways and elementary flux modes. We also introduce a new rank test to distinguish extreme pathways from elementary flux modes, in the form of an algebraic condition. Finally, we illustrate these properties within real metabolic networks. A formal definition of elementary modes and extreme pathways follows in the next section, but here we can remind the reader that an elementary mode is an admissible mode containing a minimal set of reactions (i.e., no reaction can be removed while maintaining admissibility). An extreme pathway is one which cannot be obtained by combining two other admissible pathways. Here admissible means that the pathway maintains internal mass-balances and satisfies the direction constraints on the reactions arising from thermodynamic conditions.

All algorithms that have been proposed and implemented for the computation of extreme pathways (Schilling et al., 2000) and elementary modes (Schuster et al., 2002; Schuster et al., 2000; Schuster et al., 1999; Schuster & Hilgetag, 1994), have been based on convex analysis and the Double Description Method (Fukuda & Prodon, 1996; Motzkin et al., 1953) for the computation of the extreme rays of a polyhedral cone. The extreme rays of the cone correspond to the extreme pathways of a metabolic network. Examples of such algorithms include the Canonical Basis Algorithm (Schuster et al., 2002) and the Nullspace Algorithm (von Kamp et al., 2006; Wagner, 2004; Urbanczik & Wagner, 2005b; Klamt et al., 2003; Gagneur & Klamt, 2004; Klamt et al., 2007; Terzer & Stelling, 2008) for elementary modes and variations for computing extreme pathways (Bell & Palsson, 2005). The Nullspace Algorithm has generally been found to be more efficient on most problems.

The Nullspace and Canonical Basis Algorithm were both implemented as software programs. Expa (Bell & Palsson, 2005) computes extreme pathways, Metatool (Pfeiffer et al., 1999; von Kamp et al., 2006) is capable of computing both extreme pathways and elementary flux modes, while a recent more complex software package CellNetAnalyzer (Klamt et al., 2007), a successor to Metatool, has in addition many other features designated for the analysis of metabolic and signaling networks. EFM Tool (Terzer & Stelling, 2008) is another recent implementation of Nullspace Algorithm in Java.

Estimation of the number of elementary modes and extreme pathways has also been examined (Klamt & Stelling, 2002; Yeung et al., 2007) in order to predict the complexity of the computational task to find all such metabolic pathways. Being a computationally demanding task, several approaches to parallel or distributed computation of elementary modes have been proposed through parallelization techniques (Klamt et al., 2005; Samatova et al., 2002; Lee et al., 2004), or algorithmic reformulations (Urbanczik & Wagner, 2005a; Urbanczik, 2007; Poolman et al., 2007; Terzer & Stelling, 2008).

Elementary modes and extreme pathways are used to analyze many aspects of metabolic

networks. We cite just a few examples. In (Stelling et al., 2002), elementary mode analysis was applied to examine the metabolic network robustness and cellular regulation of *Escherichia coli* central metabolism. The concept of extreme pathways was used to analyze the human red blood cell metabolism and determine the steady-state solution space under the given network properties (Wiback & Palsson, 2002). Another field of application of elementary modes, the use for identification of the most efficient pathway for the production of protein, was demonstrated again in the analysis of the metabolic network of *E. coli* (Vijayasankaran et al., 2005). In (Carlson & Srienc, 2004a; Carlson & Srienc, 2004b; Trinh et al., 2006) elementary modes were used to design a more efficient bacterium with a high yield of biomass. Elimination of five reactions in the metabolic network of metabolism of *E. coli* resulted in a collapsed network consisting of a single pathway producing biomass from glucose under aerobic growth conditions. *In vivo* implementation of this design in the lab on a glucose substrate resulted in biomass yields up to 30% higher than wild-type bacteria, close to theoretical predictions (Trinh et al., 2006). Elimination of two additional reactions resulted in a strain with minimal metabolic functionality that is optimized for efficient ethanol production anaerobically(Trinh et al., 2008).

The rest of this paper is organized as follows. In section 2 we review some basic theory regarding elementary modes and give some new simple derivations of a new simple rank test for the property of being an extreme pathway. In section 3 we give some simple consequences of the basic theory, including a simple self-contained derivation of a special case of the Nullspace Algorithm sufficient to show how it can be used to find both extreme pathways and elementary modes in one pass. In section 4 we give some simple examples to show how the set of extreme pathways for a network is qualitatively different from the set of elementary modes. In section 5 we give some further examples and some concluding remarks.

2 Basic Theory

We discuss the following standard stoichiometry problem. Let $N = (N_{\mathsf{irr}} \quad N_{\mathsf{rev}})$ be an $m \times q$ stoichiometry matrix for a given metabolic network, where the i, j-th entry is the amount of metabolite i produced by 1 unit of reaction j (with a negative value if the metabolite is consumed). The columns denoted 'irr' correspond to the irreversible reactions and those denoted 'rev' correspond to the reversible reactions. Let $\mathbf{x}$ be a q-vector of reaction fluxes, partitioned[1] as $\mathbf{x} = (\mathbf{x}_{\mathsf{irr}}\,;\,\mathbf{x}_{\mathsf{rev}})$ consistent with the partitioning of N.

Definition 1. We say $\mathbf{x}$ is an *admissible flux mode* or *pathway* if all of the following hold:

1. $N\mathbf{x} = \begin{pmatrix} N_{\mathsf{irr}} & N_{\mathsf{rev}} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x}_{\mathsf{irr}} \\ \mathbf{x}_{\mathsf{rev}} \end{pmatrix} = 0,$
2. $\mathbf{x}_{\mathsf{irr}} \geq 0$ (i.e. every entry in $\mathbf{x}_{\mathsf{irr}}$ is non-negative),
3. $\mathbf{x} \neq \mathbf{0}$.

We remark that the set of all admissible pathways is an open convex cone such that if $\mathbf{x}$ is admissible, so is $\alpha\mathbf{x}$ for any $\alpha > 0$.

Definition 2. Let $\bar{\mathrm{Z}}(\mathbf{x})$ denote the *indices* of the non-zero entries in the vector $\mathbf{x}$. We call $\mathbf{x}$ an *elementary [flux] mode* if there is no other admissible flux mode $\mathbf{y}$ whose indices of non-zero entries $\bar{\mathrm{Z}}(\mathbf{y})$ are a proper subset of $\bar{\mathrm{Z}}(\mathbf{x})$.

Definition 3. We call $\mathbf{x}$ an *extreme pathway* if $\mathbf{x}$ cannot be written as a *convex combination* of two other admissible flux modes, i.e., $\mathbf{x}$ cannot be written as $\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2$ where $\alpha_1, \alpha_2 > 0$ and $\mathbf{y}_1, \mathbf{y}_2$ are two different admissible flux modes (not scalar multiples of each other).

It is easy to see that the set of all admissible flux modes is a convex set, and we will refer to this set as the *flux cone* or *cone* associated with a given stoichiometry matrix N. If a given admissible flux mode $\mathbf{x}$ has a nonzero value for at least one irreversible reaction, then $\mathbf{x}$ is said to be an *irreversible flux mode* or *irreversible pathway*, since $-\mathbf{x}$ is not admissible. On the other hand, if $\mathbf{x}_{\mathsf{irr}} = \mathbf{0}$, then $-\mathbf{x}$ is still admissible and hence is a *reversible flux mode* or *reversible pathway*.

Definition 4. The flux cone is called a *pointed cone* if it has no admissible reversible pathway. Otherwise it is called a *non-pointed cone*.

Definition 5. Let A be an arbitrary $m \times q$ matrix. The right nullspace of A is the space of all vectors $\mathbf{v}$ such that $A\mathbf{v} = 0$. The *nullity* of A, denoted $\mathsf{nullity}(A)$ is the dimension of the right nullspace of A. Hence A has full column rank iff $\mathsf{nullity}(A) = 0$.

Proposition 1. Consider a standard stoichiometry problem as denoted in Def. 1. The associated cone is pointed if and only if $\mathsf{nullity}(N_{\mathsf{rev}}) = 0$.

[1] We use a ';' to denote vertical concatenation in column vectors and matrices, and a blank or a comma to denote horizontal concatenation in matrices, as inspired by Matlab®.

Proof. A pathway $\mathbf{x} = (\mathbf{x}_{\mathsf{irr}}\,;\,\mathbf{x}_{\mathsf{rev}})$ is both reversible and admissible iff $\mathbf{x}_{\mathsf{irr}} = 0$ and $N_{\mathsf{rev}}\mathbf{x}_{\mathsf{rev}} = 0$. Such an $\mathbf{x}$ exists iff N_{rev} has a nontrivial nullspace, i.e., $\mathsf{nullity}(N_{\mathsf{rev}}) > 0$. □

In the example of Fig. 1, N_{rev} consists of a single column, hence the cone is pointed.

Proposition 2. Consider a standard stoichiometry problem as denoted in Def. 1. Let $\mathbf{x}$ be an admissible flux mode. We assume without loss of generality that the entries in $\mathbf{x}$ and the columns of N are re-ordered so that $\mathbf{x}$ can be partitioned as

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_{\mathsf{irr}} \\ \hline \mathbf{x}_{\mathsf{rev}} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_a \\ \mathbf{x}_b \\ \hline \mathbf{x}_c \\ \mathbf{x}_d \end{pmatrix} = \begin{pmatrix} \mathbf{x}_a \\ \mathbf{0} \\ \hline \mathbf{x}_c \\ \mathbf{0} \end{pmatrix},$$

where all elements of $\mathbf{x}_a$ and $\mathbf{x}_c$ are non-zero and $\mathbf{x}_a > 0$. Partition the stoichiometry matrix consistent with the partitioning of $\mathbf{x}$:

$$N = \begin{pmatrix} N_{\mathsf{irr}} & N_{\mathsf{rev}} \end{pmatrix} = \begin{pmatrix} N_a & N_b & N_c & N_d \end{pmatrix}.$$

I. Then $\mathbf{x}$ is an extreme pathway if and only if

$$\mathsf{nullity}\begin{pmatrix} N_a & N_c & N_d \end{pmatrix} = 1.$$

II. The mode $\mathbf{x}$ is an elementary flux mode if and only if

$$\mathsf{nullity}\begin{pmatrix} N_a & N_c \end{pmatrix} = 1.$$

Proof. Suppose we are given an admissible flux mode $\mathbf{x} = (\mathbf{x}_a\,;\,\mathbf{0}\,;\,\mathbf{x}_c\,;\,\mathbf{0})$. First we note that if $\mathbf{x} = \mathbf{y} + \mathbf{z}$, the sum of two admissible flux modes, then the sign constraints imply that $\mathbf{y}_b = \mathbf{z}_b = 0$. To prove I, if $\mathsf{nullity}\begin{pmatrix} N_a & N_c & N_d \end{pmatrix} = 1$, then any admissible vector $\mathbf{z}$ such that $\mathbf{z}_b = 0$ must satisfy $(\mathbf{z}_a\,;\,\mathbf{z}_c\,;\,\mathbf{z}_d) \in \mathsf{nullspace}\begin{pmatrix} N_a & N_c & N_d \end{pmatrix}$, and hence must be a multiple of $\mathbf{x}$.

If $\mathsf{nullity}\begin{pmatrix} N_a & N_c & N_d \end{pmatrix} \geq 2$, then beside $\mathbf{x}$ there is a second admissible mode $\mathbf{y} = (\mathbf{y}_a\,;\,\mathbf{0}\,;\,\mathbf{y}_c\,;\,\mathbf{y}_d)$ not a scalar multiple of $\mathbf{x}$. Let $\alpha_* = \max\alpha$ subject to $\mathbf{z} = \mathbf{x} - \alpha\mathbf{y}$ is admissible. In this case, admissible means that $\mathbf{z}_a \geq 0$, due to non-negativity constraint for irreversible reactions corresponding to indices in a. Since $\mathbf{x}_a$ is entirely nonzero, we have that $\alpha_* > 0$. Hence we have $\mathbf{x} = \alpha_*\mathbf{y} + \mathbf{z}$, i.e., it is a convex combination of two different admissible vectors.

To prove II, assume that x is EFM. Consider a possible second admissible mode $\mathbf{y} = (\mathbf{y}_a\,;\,\mathbf{0}\,;\,\mathbf{y}_c\,;\,\mathbf{0})$, partitioned as $\mathbf{x}$, which is not a scalar multiple of $\mathbf{x}$. One could add a multiple of $\mathbf{y}$ to $\mathbf{x}$ to cancel out one nonzero element, while maintaining admissibility, yielding an admissible vector whose nonzero entries occupy a proper subset of positions 'a' and 'c'. Therefore, $\mathbf{x}$ is not elementary if and only if such a $\mathbf{y}$ exists. But such a $\mathbf{y}$ exists if and only if $\mathsf{nullity}\begin{pmatrix} N_a & N_c \end{pmatrix} \geq 2$. □

In the example of Fig. 1, consider the flux mode {r1,r2,r3}. To check that it is an elementary mode, we check the nullity of the first three columns of N, corresponding to the reactions with non-zero fluxes:

$$\mathsf{nullity}\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} = 1.$$

To check whether it is an extreme pathway for this network, we check the nullity of the same first three columns of N appended with all columns associated with reversible reactions not already included:

$$\mathsf{nullity}\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & -1 \end{pmatrix} = 2.$$

This indicated that {r1,r2,r3} is indeed an elementary mode, but not an extreme pathway. In fact, this pathway is the sum of the two extreme pathways {r1,r2,r5r} and {r3,–r5r}.

Remark 1. We remark that a, b, c, d could be considered index vectors consisting of the indices of the entries in $\mathbf{x}$ corresponding to nonzero fluxes for irreversible reactions, zero fluxes for irreversible reactions, nonzero fluxes for reversible reactions, zero fluxes for reversible reactions, respectively. To simplify the exposition, we assume without loss of generality that we have reordered the reactions to make these four groups contiguous.

Remark 2. We note that this Proposition implies that if a given flux mode is an extreme pathway, then it is also elementary, but not vice-versa.

3 Consequences of Basic Theory

The simple rank tests in the previous section give rise to two simple consequences for pointed cones. One is that a cone is pointed if and only if the set of extreme pathways is the unique minimal generating set for a network. A second consequence is a simple derivation for one special case of the Nullspace algorithm.

Proposition 3. Consider a standard stoichiometry problem as denoted in Def. 1. If the associated cone is pointed, then the set of all extreme pathways is a minimal generating set for all possible admissible modes.

Proof. Assume the cone is pointed. Let $\mathbf{x} = (\mathbf{x}_a\,;\,\mathbf{0}\,;\,\mathbf{x}_c\,;\,\mathbf{0})$ be an admissible flux mode. We wish to show it is a convex combination of extreme pathways which is an indicator that extreme pathways coincide with minimal generating set. Because the cone is pointed, some irreversible reaction must have a nonzero flux, i.e., $\mathbf{x}_a$ must be non-empty. If $\mathbf{x}$ were extreme we would be done, so assume it is not. By part I of Proposition 2, we can find a second admissible vector $\mathbf{y} = (\mathbf{y}_a\,;\,\mathbf{0}\,;\,\mathbf{y}_c\,;\,\mathbf{0})$.

We can follow the construction similar to that used to prove Proposition 2 to follow the line joining $\mathbf{x}$ and $\mathbf{y}$ to the boundary of the admissible region. We call those boundary points $\mathbf{z}$, $\mathbf{w}$. Being on the boundary, the parts $\mathbf{z}_a$, $\mathbf{w}_a$ must have some zero entries (corresponding to nonzero entries in $\mathbf{x}_a$).

Either the resulting vectors $\mathbf{w}$, $\mathbf{z}$ are extreme, or we can repeat this construction on each of them. But this construction can be repeated only a finite number of times because each construction introduces additional zeros among the "a" fluxes corresponding to irreversible reactions.

The specific construction of $\mathbf{z}$ and $\mathbf{w}$ proceeds as follows. Find $\lambda^+ = \max\lambda$ subject to the condition that $\mathbf{g}(\lambda) = (1-\lambda)\mathbf{x} + \lambda\mathbf{y}$ is admissible, and let $\lambda^- = \min\lambda$ subject to the same condition. Then $\mathbf{z} = \mathbf{g}(\lambda^-)$ and $\mathbf{w} = \mathbf{g}(\lambda^+)$. □

Since the example of Fig. 1 is a pointed cone, the network has a minimal generating set of the extreme pathways, namely: EP1={r1,r2,r5r}, EP2={r4,r2,r5r}, EP3={r3,–r5r}.

If the cone for a network is not pointed, then the extreme pathways do not form a minimal generating set for the system, but we can state the following.

Proposition 4. Consider a standard stoichiometry problem as denoted in Def. 1. The associated cone is pointed if and only if $\mathsf{nullity}(N_{\mathsf{rev}}) = 0$. If $\mathsf{nullity}(N_{\mathsf{rev}}) = 1$, then there is exactly one reversible admissible flux mode and it is the only extreme pathway (using Def. 3). If $\mathsf{nullity}(N_{\mathsf{rev}}) \geq 2$, then there are no extreme pathways, and there are at least 2 reversible elementary flux modes.

Proof. The case $\mathsf{nullity}(N_{\mathsf{rev}}) = 0$ was treated in Proposition 1.

If $\mathsf{nullity}(N_{\mathsf{rev}}) = 1$, then let $\mathbf{x}_c$ be a nonzero vector in $\mathsf{nullspace}(N_{\mathsf{rev}})$ (unique up to scalar multiple), and then $\mathbf{x} = (\mathbf{0}\,;\,\mathbf{x}_c)$ is the unique reversible pathway. It cannot be written as a convex combination of any other different modes, which are all irreversible. On the other

hand, the sum of an arbitrary multiple of $\mathbf{x}$ and any other admissible mode is still admissible, so no other mode can be extreme.

If $\mathsf{nullity}(N_{\mathsf{rev}}) \geq 2$, then at least two reversible flux modes of the form $\mathbf{x}$ above can be found, and arbitrary multiples of both of those could be added to any admissible flux mode with the result remaining admissible, so there cannot be any extreme pathways for this system. □

In order to arrive at the Nullspace Algorithm, we need two lemmas.

Lemma 1. Let S_k denote the standard stoichiometric model of Def. 1, where $N = (N_1 \quad N_2)$ is $m \times q$, the matrix N_1 of columns corresponding to the irreversible reactions is $m \times k$, and the part N_2 corresponding to the reversible reactions has full column rank $q - k$. Let S_{k+1} denote the same stoichiometric model, but where the $k + 1^{\text{st}}$ reaction is irreversible.

Suppose $\mathbf{x}$ is an extreme pathway with respect to S_{k+1}. Then either $\mathbf{x}$ is an extreme pathway with respect to S_k or it is a convex combination of exactly two extreme pathways of S_k.

Proof. Let $\mathbf{x}$ be an extreme pathway for S_{k+1}. If $x_{k+1} > 0$, then test I of Proposition 2 involves the same columns of N for both S_k and S_{k+1}, and thus $\mathbf{x}$ is an extreme pathway for S_k.

If $x_{k+1} = 0$, then let a be the vector of indices of the nonzero entries among $(x_1, \ldots, x_k)$, and let $c = (k+2, \ldots, q)$. Let b denote the indices of the zero entries among $(x_1, \ldots, x_k)$. By assumption $\mathsf{nullity}(N_a \quad N_c) = 1$. Appending one column can increase the nullity by at most one, hence $1 \leq \mathsf{nullity}(N_a \quad N_{k+1} \quad N_c) \leq 2$. If $\mathsf{nullity}(N_a \quad N_{k+1} \quad N_c) = 1$, then $\mathbf{x}$ was already extreme wrt S_k. So suppose $\mathsf{nullity}(N_a \quad N_{k+1} \quad N_c) = 2$. This means that the space $\mathcal{S}$ of all admissible vectors $\mathbf{v}$ with $\mathbf{v}_b = 0$ has dimension 2. We find a different admissible vector $\mathbf{y}$ such that $\mathbf{y}_b = 0$, and then define the boundary points $\mathbf{z} = \mathbf{g}(\lambda^-)$, $\mathbf{w} = \mathbf{g}(\lambda^+)$ where $\mathbf{g}$ is defined as in the proof of Proposition 3. The boundary points cannot be expressed as a convex sum of any other vector within $\mathcal{S}$ (else λ^+, λ^- could not be extreme points). Since $\mathbf{w}_b = \mathbf{z}_b = 0$, the two vectors $\mathbf{w}$, $\mathbf{z}$ cannot be expressed as a convex combination involving any admissible vector $\mathbf{v}$ not in $\mathcal{S}$ (i.e., with $\mathbf{v}_b \neq 0$). Hence $\mathbf{w}$, $\mathbf{z}$ must be extreme pathways wrt S_k. □

The example of Fig. 1 can be considered as S_4 in the notation of this Proposition, since the first four reactions are irreversible while the fifth reaction is reversible. The system S_5 would be the same system with all reactions irreversible (in this case the fifth reaction is irreversible in the direction Y→O, and hence denoted 'r5'). The flux mode {r1,r2,r3} is an extreme pathway within the network S_5, but as has been previously noted, it is not an extreme pathway for S_4. According to this Proposition, it is the convex combination of two extreme pathways of S_4, namely {r1,r2,r5r} and {r3,–r5r}.

Lemma 2. Let S_k be the standard stoichiometry problem as in Lemma 1, with m reversible reactions and $k = q - m$ irreversible reactions, and assume $\mathsf{rank}\, N = m$ and N_{rev} is square with full rank m. Then $R = \begin{pmatrix} I \\ -N_{\mathsf{rev}}^{-1} N_{\mathsf{irr}} \end{pmatrix}$ is the complete set of extreme pathways for S_k.

Proof. First we check admissibility:

1. $NR = (N_{\mathsf{irr}} \quad N_{\mathsf{rev}}) \begin{pmatrix} I \\ -N_{\mathsf{rev}}^{-1} N_{\mathsf{irr}} \end{pmatrix} = N_{\mathsf{irr}} - N_{\mathsf{irr}} = 0.$

2. The entries in each column of R corresponding to N_{irr} are part of the I part and hence non-negative, with at least one strictly positive entry.

Next we observe that R is $q \times n$, where $n = q - m = \mathsf{nullity}(N)$ and has full column rank. Hence any admissible flux mode $\mathbf{x}$ can be written in a unique way as $\mathbf{x} = R\boldsymbol{\lambda}$, and because of the presence of the I matrix, $\lambda \geq 0$. Hence each column of R is an extreme pathway according to Def. 3, and there are no others. □

The example of Fig. 1 corresponds to the model S_4 in the notation of Lemma 1, in which r5r is the only reversible reaction. The model S_3 is the same network in which the last two reactions are reversible, denoted r4r and r5r. The last two columns of N in Fig. 1 have full rank 2 and includes all the reversible reactions. In the notation of Lemma 2, we have

$$N_1 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad R = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline -1 & 1 & 0 \\ 0 & 1 & -1 \end{array}\right).$$

Each column of this R is an extreme pathway with respect to the system S_3 in which r4r and r5r are the two reversible reactions.

An elementary consequence of the preceding two lemmas is the Nullspace algorithm (von Kamp et al., 2006; Wagner, 2004; Urbanczik & Wagner, 2005b; Klamt et al., 2003; Gagneur & Klamt, 2004; Klamt et al., 2007; Terzer & Stelling, 2008) outlined in the following two Propositions.

Proposition 5. Let S_i denote a standard stoichiometry problem with i irreversible and r reversible reactions, so that $i + r = q$, and assume N_{rev} has full column rank $r = q - i$. The following procedure computes the complete set of extreme pathways.

1. Preprocess N by eliminating redundant rows, to reduce N to a matrix with full row rank, and assume the new N is $m \times q$ with full row rank m.

2. Select m linearly independent columns, including all of N_{rev} and order them last, calling this part N_2. Denote $N = (N_1 \quad N_2)$, where N_1 is $m \times (q - m)$. Let S_k denote the standard stoichiometry problem with this N, in which the first k reactions are irreversible, and the last $q - k$ are comprised of all reversible reactions and remaining $i - (q - m)$ irreversible reactions, where initially $k = q - m$.

3. For initial value $k = q - m$, form initial basis $R_k = \begin{pmatrix} I \\ -N_2^{-1} N_1 \end{pmatrix}$

4. For $k = (q-m), \ldots, i-1$,

   (a) Form convex combinations of every pair of columns of R_k with non-negative coefficients set to annihilate the $k+1^{\text{st}}$ entry of each resulting vector.

   (b) Collect all the columns of R_k which are admissible and extreme with respect to S_{k+1}, together with all the resulting convex combinations from the previous step which are admissible and extreme with respect to S_{k+1}. To check if the newly generated column is extreme we refer to the part I of Proposition 2.

Remark 3. Obviously, the detailed algorithm can be made much more efficient than sketched above. For example:

- If the sign of the $k+1^{\text{st}}$ entry in one column of a pair agrees with the sign of the corresponding entry in the other column, then no convex combination can annihilate the $k+1^{\text{st}}$ entry.
- Admissibility of a column of R_k with respect to S_{k+1} depends only on the sign of its $k+1^{\text{st}}$ entry.

Implementation details have been treated by many authors mentioned in the introduction and will be treated in a future paper.

We observe that the computation of elementary flux modes and extreme pathways are equivalent operations. We can compute the elementary flux modes using the same method used to compute extreme pathways. Thus, instead of computing the set of elementary flux modes directly, we will define the algorithm to compute the extreme pathways so that in each iteration the reaction processed corresponds to irreversible reaction. This is accomplished by introducing the idea of "temporary reversible reactions". Unlike in Proposition 5, here we will have to enforce the reversible reaction to become irreversible and iterate over all reactions, and not only over those corresponding to irreversibles.

Proposition 6. Let S_i denote a standard stoichiometry problem with i irreversible and r reversible reactions, so that $i + r = q$, and assume N_{rev} has full column rank $r = q - i$. The procedure sketched below computes all the elementary modes for a stoichiometric system with a pointed cone:

1. Preprocess N by eliminating redundant rows, to reduce N to a matrix with full row rank, and assume the new N is $m \times q$ with full row rank m.

2. Select m linearly independent columns, including all of N_{rev} and order them last, calling this part N_2. Denote $N = (N_1 \;\; N_2)$, where N_1 is $m \times (q-m)$. Let $\widetilde{S}_k$ denote the standard stoichiometry problem with this N, in which the last $q-k$ reactions are temporarily considered reversible. Initially $k = q - m$, and the first k columns are all irreversible.

3. For initial value $k = q - m$, form initial basis $R_k = \begin{pmatrix} I \\ -N_2^{-1} N_1 \end{pmatrix}$

4. For $k = (q-m), \ldots, q-1$,

   (a) Form convex combinations of every pair of columns of R_k with non-negative coefficients set to annihilate the $k+1^{\text{st}}$ entry of each resulting vector.

   (b) Collect all the columns of R_k which are admissible and elementary with respect to $\widetilde{S}_{k+1}$, together with all the resulting convex combinations from the previous step which are admissible and elementary with respect to $\widetilde{S}_{k+1}$. To check if the newly generated column is elementary we refer to the part II of Proposition 2.

Proof. (*informal sketch*) This algorithm has been well studied in the literature (Schuster & Hilgetag, 1994; Schuster et al., 1999; Schuster et al., 2002; Fukuda & Prodon, 1996; Motzkin et al., 1953) and hence we omit a formal proof. However we give some intuition with the following. The algorithm of this Proposition applied to the original stoichiometric model is equivalent to the algorithm of Proposition 5 applied to a modified stoichiometric model in which every reversible reaction has been split into two irreversible reactions, except for the futile cycles consisting of each reversible reaction and its inverse. One could replace any negative coefficient for one of the directions of a reversible reaction with the corresponding positive coefficient for its opposite reaction, though this will happen naturally during the course of this algorithm due to the presence of those futile cycles. Except for the futile cycles, the elementary modes for the two models coincide, and for the model with all irreversible reactions the condition of being elementary coincides with the condition of being extreme. □

The preceding Proposition encompasses the basic idea behind almost all variations of the Nullspace Algorithm that have been implemented in the literature, but without all the 'frills' necessary to make it run efficiently. But this is sufficient to notice that if all the irreversible reactions are ordered first, the procedures sketched in Propositions 5 and 6 coincide up to the point where the last irreversible reaction has been reached ($k = i - 1$, the number of irreversible reactions). Hence we can conclude that most variations of the Nullspace Algorithm can be used to compute both the extreme pathways and the elementary modes in one pass. It is only in passes $k \geq i$ that the reversible reactions are encountered and the procedure of Proposition 5 will differ from Proposition 6. We state this formally in the following.

Proposition 7. When applying the general procedure having the general form sketched in Proposition 6 to compute all elementary modes for a system S_i with a pointed cone and with i irreversible reactions which are ordered first, the collection of elementary modes R_i appearing when all the irreversible reactions have been processed (at the start of pass when $k = i$) is exactly the set of extreme pathways for the system S_i. □

Remark 4. The common practice is to treat all internal reversible reactions as two separate irreversible reactions. This usually has the effect of ensuring the resulting model has a pointed cone. But even if the original cone were pointed, this cone (and the associated extreme pathways) will differ from the original cone.

4 Subnetworks

We illustrate one consquence of the preceding theory. Consider the system of Fig. 1, and for the sake of exposition let the "O" represent oxygen, so that we may refer to this as the "aerobic" network. This network has three extreme pathways: EFM1={r1,r2,r5r}, EFM2={r4,r2,r5r}, EFM3={r3,–r5r}. We may consider the "anaerobic" network as the one where no external "O" is present, with the result that exchange reaction r5r is either absent or one-way outbound (hence denoted 'r5'). For this network, the extreme pathways are EFM4={r1,r2,r3} and EFM5={r4,r2,r3}. If r5 is present as an irreversible pathway outbound, then EFM1 and EFM2 are also present as extreme pathways. Observe that the first two extreme pathways for the anaerobic network, EFM4 & EFM5, do not appear among the extreme pathways for the aerobic network, but are still elementary flux modes for both networks. The property of being an elementary mode depends only on the interconnect among the reactions present in that mode, corresponding to those specific columns in N, regardless of the properties of the rest of the network in which this pathway is embedded. The property of being an extreme pathway depends on the the structure of the rest of the network, and specifically on the presence or absence of certain reversible reactions not present in the given mode. For our specific example, whether or not the pathway EFM4={r1,r2,r3} is extreme or not depends on the presence and reversibility of the exchange reaction r5r, which is not even part of this pathway.

If one were to compute all the *elementary modes* for the aerobic network, one would find these include all the elementary modes for any subnetwork, and in particular these would include the extreme pathways for the anaerobic network. But computing only the extreme pathways for a network does not necessarily yield all the extreme pathways for its subnetworks. On the other hand, it is not surprising to find the number of elementary modes to be much larger than the number of extreme pathways, since the former includes all elementary modes for all subnetworks.

We can illustrate a similar effect with the Calvin Cycle given in (Poolman et al., 2007) and listed in Fig. 2. Within this network, some reactions are inhibited by the presence of light, while some others are active only in the presence of light. This network has a pointed cone, as can be verified using Proposition 1, hence it makes mathematical sense to consider the extreme pathways for this network as is. One can use the Nullspace Algorithm or its variations to compute the pathways for the entire network, or for the sub-networks active under light conditions only, or dark conditions only. The entire network has 28 elementary modes, of which 8 are elementary modes for the "light-only" network and 2 are elementary modes for the "dark-only" network. The remaining 18 elementary modes for the entire network are inhibited under either light and dark conditions and hence may be biologically infeasible, but include all possible feasible modes in one computation. On the other hand, if one computes only the extreme pathways of the entire network, one will find that these will include both elementary modes for the "dark-only" network, but only one of the elementary modes for the "light-only" network. Hence, limiting one's attention to extreme pathways will necessitate the computation of pathways for each sub-network of interest separately from scratch.

5 Conclusions and Discussion

We conclude with some discussion illustrating what the rank tests show when applied to some other examples, together with some concluding remarks.

Example 1 (A model of *E. coli* central metabolism)**.** An example of a metabolic network usually studied is a model of the central metabolism for *E. coli* (Trinh et al., 2008), consisting of 70 reactions (19 reversible) and 68 metabolites (52 internal to the network). In this model we consider the anapleurotic pathway converting malate to pyruvate to be NADH dependent only. In addition, the reaction FEM9 catalyzed by pyruvate decarboxylase to convert PYR to ACA is not native in *E. coli* but cloned into *E. coli* through the plasmid pLOI297. For detailed discussion see (Trinh et al., 2008).

To give an application of elementary mode analysis, we use the Metatool software (von Kamp et al., 2006) on the *E. coli* network from (Trinh et al., 2008) to find a total of 38,001 elementary modes using glucose as the carbon source, of which 32,604 produce biomass and 5,010 are anaerobic. Using the theory developed in this paper, we can easily find that 2,739 of these are extreme pathways (1,191 are producing biomass and 978 are anaerobic which may or may not produce biomass). In (Trinh et al., 2008) the goal was to find pathways maximizing the production of ethanol as a biofuel for a given amount of glucose, while producing sufficient biomass to allow the cells to grow by deleting the inefficient pathways. Maximizing a single linear objective function such as ethanol production subject to the set of linear constraints in Def. 1 (plus the constraint that glucose consumption rate is 1 mole/L/hr) naturally leads to an extreme point in the polytope defining the feasible region, corresponding to an extreme pathway. But this pathway does not support cell growth, hence the need to trade off between the optimal solutions for two or more distinct objective functions. A resulting semi-optimal solution with a minimal number of reactions will be an elementary mode, which will generally be a convex combination of at least two extreme pathways representing the optimal solutions for each individual objective function. Having all elementary modes available allows one to explore many alternative knockouts to achieve similar performance objectives. Figure 3 shows the relative ethanol and biomass production of all the anaerobic modes, both extreme pathways and non-extreme elementary modes. Biomass yield is low because the result is shown for anaerobic growth conditions only. The investigation of engineering and biological applications of these modes is beyond the scope of this paper and will be the subject of separate papers.

Example 2 (Simple Pointed Cone)**.** We illustrate the rank test with the example from (Klamt & Stelling, 2003) and shown in Fig. 4. The elementary modes for this network are

- EFM1 = {2 R1, R3, 2 R5, –2 R7r, R8}
- EFM2 = {2 R1, R3, 2 R4, R8}
- EFM3 = {2 R1, R3, R5, R6, R9}
- EFM4 = {2 R1, R3, R4, R6, R7r, R9}
- EFM5 = {R1, R2r, R3, R6, R7, R9}

- EFM6 = {R1, –R2r, R4}
- EFM7 = {R1, –R2r, R5, –R7r}
- EFM8 = {2 R2r, R3, R8}

We remark that in the network of (Klamt & Stelling, 2003), the elementary modes EFM1 = EFM7+EMF8 and EFM2 = EFM6+EFM8 each yield an overall stoichiometry of 1A = 1P, while EFM3 = EFM5+EFM7 and EFM4 = EFM5+EFM6 each yield an overall stoichiometry of 2A = 1P, as previously noted in (Bernhard O. Palsson & Papin, 2003). To eliminate this discrepancy, we have modified reaction R8 from 1B = 1P to 2B = 1P. This change does not affect the set of reactions involved in each EFM, nor does it affect the observations we make here regarding the rank tests and extreme pathways.

We first observe that the two reversible reactions are independent, so Proposition 1 implies this is a pointed cone with extreme pathways. We can then apply Proposition 2 to each EFM to see if it is an extreme pathway with respect to this network. The result is that EFM5, EFM6, EFM7, EFM8 are extreme, and EFM1, EFM2, EFM3, EFM4 are not, with respect to this network. We observe that in (Klamt & Stelling, 2003) first the internal reversible reaction R7r is split into two irreversible reactions (R7f & R7b), obtaining the extreme pathways EFM3, EFM5, EFM6, EFM7, EFM8, and non-extreme elementary modes EFM1, EFM2, EFM4, plus the futile cycle {R7f, R7b}. By splitting this reaction, we have modified the network so that, while the set of elementary flux modes have not changed, Proposition 2 indicates that EFM3 is now an extreme pathway with respect to the modified network, consistent with (Klamt & Stelling, 2003). Indeed, one can observe that EFM3 is the sum of EFM5 and EFM7, by which the internal reversible reaction R7r is cancelled. Hence the property of being an extreme vs non-extreme pathway depends very much on the specific treatment of the reversible reactions within a network.

If the exchange flux R2r were removed, we would, of course, lose all the current extreme pathways EFM5–EFM8, since they involve R2r. But we would also notice that all the remaining modes EFM1–EFM4 would become extreme pathways regardless of whether R7r is split or not. That is, the property of EFM1–EFM4 being extreme pathways depends on the presence or absence of R2r, which is not present in any of these pathways.

We remark that condition II of Proposition 2 indicates that, unlike the extreme pathways, the set of elementary modes is not affected by splitting the reversible reactions, other than the futile cycles involving the split reactions themselves.

Example 3 (Human red blood cell metabolism)**.** It is useful to apply the nullity test in the analysis of the Human Red Blood Cell metabolic network that has been previously analyzed using extreme pathway analysis and elementary mode analysis and is well documented in the literature (Papin et al., 2004; Wiback & Palsson, 2002). The published results show that there exist 6,180 EFMs and 55 ExPas (extreme pathways) (Papin et al., 2004). It is important to note that the elementary mode analysis has been carried out with the network containing reversible reactions while the extreme pathways have been identified in a network where the internal reversible reactions have been separated into two distinct reactions operating in the opposite direction. The two approaches yield very different results and

the differences are revealed when the nullity test is applied. We analyzed the identically constructed metabolic network (Wiback & Palsson, 2002) that consists of 58 metabolites (39 of which are internal) and of 51 reactions (33 of which are reversible reactions). Among the 33 reversible reactions, 17 are reversible internal reactions; 16 are reversible exchange reactions. Using Metatool, elementary mode analysis on the network that includes all reversible reactions yields 6,180 EFMs as previously reported. The nullity test applied to these elementary modes identifies only a single extreme pathway that consists only of reversible reactions. This result is obtained because the flux cone is not pointed, and corresponds to the condition $\mathsf{nullity}(N_{\mathrm{rev}}) = 1$ using the notation of Proposition 4. To perform the analysis on the exactly identical network as previously published, we have split each of the 17 reversible internal reactions into two irreversible reactions. This guarantees also that the flux cone is pointed. Elementary mode analysis on the modified network identifies 6,198 EFMs. The nullity test applied to these EFMs identifies 55 ExPas, the same number as previously published.

Inspection of the 18 additional EFMs in the network case with separated internal reversible reactions reveals that 17 of these pathways are the futile cycles consisting of the two separate reactions derived from each reversible reaction. Out of these 17 futile cycles, 16 are also extreme pathways classified as Type III extreme pathways (Schilling et al., 2000). The 18^{th} pathway is a pathway consisting only of reversible reactions matching the single extreme pathway from the unmodified network, but opposite in direction. Thus, these 18 additional EFMs will not be calculated if the internal reversible reactions are not split into two separate reactions.

These examples demonstrate that the nullity test can accurately identify ExPas from calculated EFMs, and that the obtained results are consistent with previous reports. The differences in extreme pathways identified in the two types of networks emphasize the importance of the type of network that is subjected to the analysis. Therefore, the type of network (reversible reactions present or with reversible reactions split into two separate reactions) on which the analysis is performed should always be mentioned when the number of elementary modes or the number of extreme pathways is listed.

We have given a concise development of the basic theory regarding algebraic properties of elementary and extreme pathways and how to distinguish them. We have shown how this theory implies that the computation of only extreme pathways for a given network may not be applicable to subnetworks where some parts of the network are removed or inhibited.

Acknowldgements

This work was partially supported by NIH grant GM077529, NSF grant 0534286, and by the Univ. of Minn. Center for Bioinformatics and Computational Biology.

Author Disclosure Statement

No competing financial interests exist.

References

Bell, S., & Palsson, B. (2005). Expa: Program for calculating extreme pathways in biochemical reaction networks. *Bioinformatics*, *21*, 1739–1740.

Bernhard O. Palsson, N. D. P., & Papin, J. A. (2003). Development of network-based pathway definitions: the need to analyze real metabolic networks. *Trends in Biotechnology*, *21*, 195–198.

Carlson, R., & Srienc, F. (2004a). Fundamental escherichia coli biochemical pathways for biomass and energy production: Creation of overall flux states. *Biotechnology and bioengineering*, *86*, 149–162.

Carlson, R., & Srienc, F. (2004b). Fundamental escherichia coli biochemical pathways for biomass and energy production: Identification of reactions. *Biotechnology and bioengineering*, *85*, 1–19.

Fukuda, K., & Prodon, A. (1996). Double description method revisited. In M. Deza, R. Euler and I. Manoussakis (Eds.), *Combinatorics and computer science*, 91–111. Springer. Also tech. report, Mathematics, ETH, 1995.

Gagneur, J., & Klamt, S. (2004). Computation of elementary modes: a unifying framework and the new binary approach. *BMC Bioinformatics*, *5*.

Klamt, S., Gagneur, J., & von Kamp, A. (2005). Algorithmic approaches for computing elementary modes in large biochemical reaction networks. *Systems Biology, IEE Proceedings*, *152*, 249–255.

Klamt, S., Saez-Rodriguez, J., & Gilles, E. D. (2007). Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC Systems Biology*, *1*.

Klamt, S., & Stelling, J. (2002). Combinatorial complexity of pathway analysis in metabolic networks. *Molecular Biology Reports*, *29*, 233–236.

Klamt, S., & Stelling, J. (2003). Two approaches for metabolic pathway analysis. *Trends in Biotechnology*, *21*, 64–69.

Klamt, S., Stelling, J., Ginkel, M., & Gilles, E. (2003). FluxAnalyzer: Exploring structure, pathways, and flux distributions in metabolic networks on interactive flux maps. *Bioinformatics*, *19*, 261–269.

Lee, L., Varner, J., & Ko, K. (2004). Parallel extreme pathway computation for metabolic networks. *SAS'95, Static Analysis Symposium* (pp. 33–50). Springer.

Motzkin, T., Raiffa, H., Thompson, G., & Thrall, R. (1953). The double description method. In H. Kuhn and A. Tucker (Eds.), *Contributions to theory of games*, vol. II, 51–73. Princeton University Press.

Papin, J., Stelling, J., Price, N. D., Klamt, S., Schuster, S., & Palsson, B. (2004). Comparison of network-based pathway analysis methods. *Trends in Biotechnology*, *22*, 400–405.

Pfeiffer, T., Sanchez-Valdenebro, I., Nuno, J., Montero, F., & Schuster, S. (1999). Metatool: for studying metabolic networks. *Bioinformatics*, *15*, 251–257.

Poolman, M., Sebu, C., Pidcoc, M., & Fell, D. (2007). Modular decomposition of metabolic systems via null-space analysis. *Journal of Theoretical Biology*.

Samatova, N., Geist, A., Ostrouchov, G., & Melechko, A. (2002). Parallel out-of-core algorithm for genome-scale enumeration of metabolic systemic pathways. *Parallel and Distributed Processing Symposium., Proceedings International, IPDIPS 2002*, 185–192.

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*, 229–248.

Schuster, S., Dandekar, T., & Fell, D. (1999). Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends in Biotechnology*, *17*, 53–60.

Schuster, S., Fell, D., & Dandekar, T. (2000). A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nature Biotechnology*, *18*, 326–332.

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

Schuster, S., Hilgetag, C., Woods, J., & Fell, D. (2002). Reaction routes in biochemical reaction systems: Algebraic properties, validated calculation procedure and example from nucleotide metabolism. *Mathematical Biology*, *45*, 153–181.

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

Terzer, M., & Stelling, J. (2008). Large scale computation of elementary flux modes with bit pattern trees. *Bioinformatics*.

Trinh, C., Carlson, R., Wlaschin, A., & Srienc, F. (2006). Design, construction and performance of the most efficient biomass producing e. coli bacterium. *Metabolic Engineering*, *8*, 628–638.

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*, 3634–3643.

Urbanczik, R. (2007). Enumerating constrained elementary flux vectors of metabolic networks. *Systems Biology IET*, *1*, 274–279.

Urbanczik, R., & Wagner, C. (2005a). Functional stoichiometric analysis of metabolic networks. *Bioinformatics*, *21*, 4176–4180.

Urbanczik, R., & Wagner, C. (2005b). An improved algorithm for stoichiometric network analysis: theory and applications. *Bioinformatics*, *21*, 1203–1210.

Vijayasankaran, N., Carlson, R., & Srienc, F. (2005). Metablic pathway structures for recombinant protein synthesis in escherichia coli. *Applied Microbiology and Biotechnology*, *68*, 737–746.

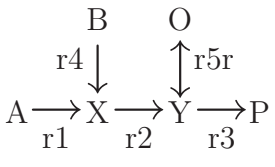
von Kamp, A., & Schuster, S. (2006). Metatool 5.0: fast and flexible elementary modes analysis. *Bioinformatics*, *22*, 1930–1931.

Wagner, C. (2004). Nullspace approach to determine the elementary modes of chemical reaction systems. *J. Phys. Chem.*, *108*, 2425–2431.

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

Yeung, M., Thiele, I., & Palsson, B. (2007). Estimation of the number of extreme pathways for metabolic networks. *BMC Bioinformatics*, *8*.

$$
\begin{array}{lrcl}
\text{r1} & \text{A} & \longrightarrow & \text{X} \\
\text{r2} & \text{X} & \longrightarrow & \text{Y} \\
\text{r3} & \text{Y} & \longrightarrow & \text{P} \\
\text{r4} & \text{B} & \longrightarrow & \text{X} \\
\text{r5r} & \text{Y} & \longleftrightarrow & \text{O}
\end{array}
$$



$$
N = \begin{array}{c}
\begin{pmatrix} \underbrace{\begin{matrix} 1 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{matrix}}_{N_{\mathsf{irr}}} & \underbrace{\begin{matrix} 0 \\ -1 \end{matrix}}_{N_{\mathsf{rev}}} \end{pmatrix} \\
\begin{matrix} \text{X} \\ \text{Y} \end{matrix}
\end{array}
$$

Figure 1: Small illustrative example, with its stoichiometry matrix N.

light	Rubisco	:	RuBP + co2	$\longrightarrow$	2 PGA
light	PGK	:	PGA + ATP	$\longrightarrow$	BPGA + ADP
light	G3Pdh	:	BPGA + nadph	$\longleftrightarrow$	GAP + Pi + nadp
light	FBPase	:	FBP	$\longrightarrow$	F6P + Pi
light	SBPase	:	SBP	$\longrightarrow$	S7P + Pi
light	Ru5PK	:	Ru5P + ATP	$\longrightarrow$	RuBP + ADP
both	StarchSynth	:	G1P + ATP	$\longrightarrow$	ADP + 2 Pi + starch
light	light	:	ADP + Pi	$\longrightarrow$	ATP
both	TPI	:	GAP	$\longleftrightarrow$	DHAP
both	Aldo	:	DHAP + GAP	$\longleftrightarrow$	FBP
both	TKL1	:	GAP + F6P	$\longleftrightarrow$	E4P + X5P
both	Aldo2	:	DHAP + E4P	$\longleftrightarrow$	SBP
both	TKL2	:	GAP + S7P	$\longleftrightarrow$	R5P + X5P
both	R5Piso	:	R5P	$\longleftrightarrow$	Ru5P
both	X5epi	:	X5P	$\longleftrightarrow$	Ru5P
both	PGI	:	F6P	$\longleftrightarrow$	G6P
both	PGM	:	G6P	$\longleftrightarrow$	G1P
both	StPase	:	Pi + starch	$\longrightarrow$	G1P
both	GAP:TPT	:	GAP + pi-cyt	$\longrightarrow$	Pi + gap-cyt
dark	Oxid	:	G6P + 2 nadp	$\longrightarrow$	R5P + 2 nadph + co2
both	PGA:TPT	:	PGA + pi-cyt	$\longrightarrow$	Pi + pga-cyt
both	DHAP:TPT	:	DHAP + pi-cyt	$\longrightarrow$	Pi + dhap-cyt
dark	TAL	:	F6P + E4P	$\longleftrightarrow$	GAP + S7P

Figure 2: Sample Calvin Cycle from (Poolman et al., 2007). External metabolites are those starting with a lower case letter.

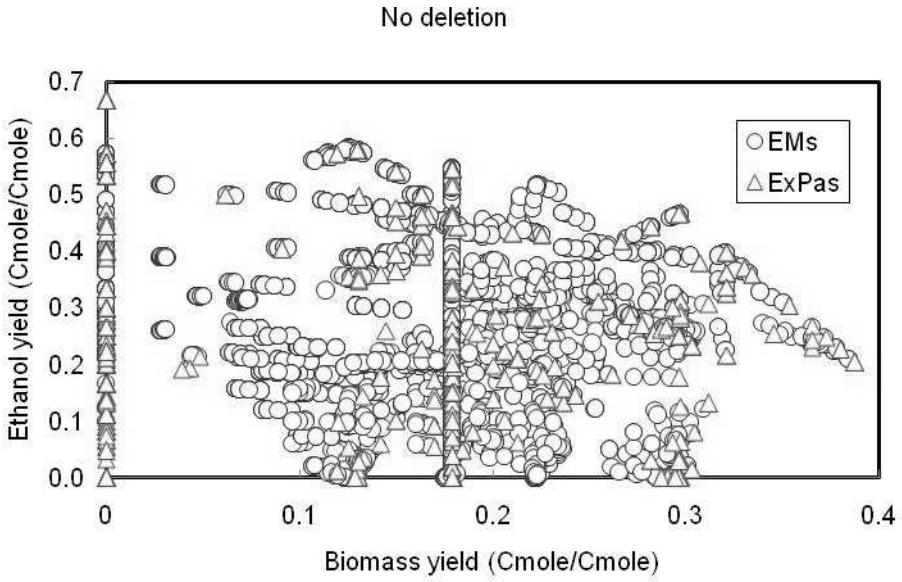


Figure 3: Relationship of ethanol and biomass yields corresponding to anaerobic elementary modes (EFMs, circles) and extreme pathways (ExPas, triangles).

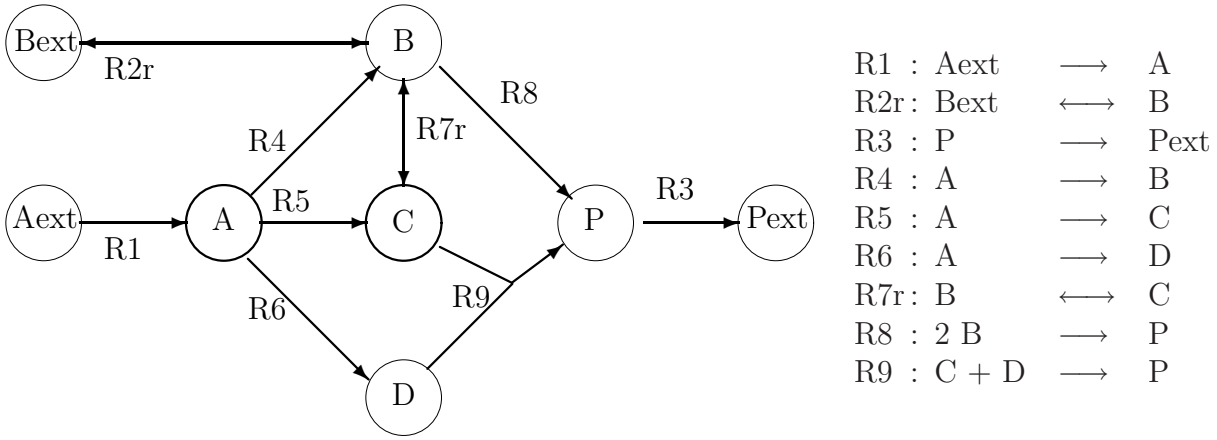


Figure 4: Example from (Klamt & Stelling, 2003) to illustrate the distinction between extreme pathways and elementary modes.