



On the geometry of elementary flux modes

Frederik Wieder¹ · Martin Henk² · Alexander Bockmayr¹

Received: 26 October 2022 / Revised: 2 August 2023 / Accepted: 9 August 2023
© The Author(s) 2023

Abstract

Elementary flux modes (EFMs) play a prominent role in the constraint-based analysis of metabolic networks. They correspond to minimal functional units of the metabolic network at steady-state and as such have been studied for almost 30 years. The set of all EFMs in a metabolic network tends to be very large and may have exponential size in the number of reactions. Hence, there is a need to elucidate the structure of this set. Here we focus on geometric properties of EFMs. We analyze the distribution of EFMs in the face lattice of the steady-state flux cone of the metabolic network and show that EFMs in the relative interior of the cone occur only in very special cases. We introduce the concept of degree of an EFM as a measure how elementary it is and study the decomposition of flux vectors and EFMs depending on their degree. Geometric analysis can help to better understand the structure of the set of EFMs, which is important from both the mathematical and the biological viewpoint.

Keywords Metabolic networks · Elementary flux modes · Steady-state flux cone · Face lattice

Mathematics Subject Classification 92C42 Systems biology Networks · 52B05 Combinatorial properties of polytopes and polyhedra

Frederik Wieder, Martin Henk and Alexander Bockmayr contributed equally to this work.

✉ Alexander Bockmayr
alexander.bockmayr@fu-berlin.de

Frederik Wieder
f.wieder@fu-berlin.de

Martin Henk
henk@math.tu-berlin.de

¹ FB Mathematik und Informatik, Freie Universität Berlin, Arnimallee 6, 14195 Berlin, Germany

² Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany

1 Introduction

Constraint-based analysis of metabolic networks is an important area in computational biology (Bordbar et al. 2014; Fang et al. 2020). The stoichiometric and thermodynamic constraints that have to hold in a metabolic network at steady-state have led to the definition of the *steady-state flux cone*, which comprises all possible flux distributions over the network at steady-state.

An important concept to analyze the flux cone in a mathematically and biologically meaningful way are *elementary flux modes* (EFMs) (Schuster and Hilgetag 1994; Schuster et al. 2002), which provide an inner description of the flux cone consisting of a finite set of generating vectors (Gagneur and Klamt 2004; Wagner and Urbanczik 2005; Larhlimi and Bockmayr 2008; Jevremovic and Boley 2013). From a mathematical point of view, not all elementary flux modes are needed to describe the cone, i.e., the set of all EFMs does not form a minimal generating set, except for special cases (for example when all reactions are irreversible). Even for small networks, the number of elementary flux modes may be very large.

Larhlimi and Bockmayr (2009) introduced metabolic behaviors and studied outer descriptions of the flux cone based on *minimal metabolic behaviors* (MMBs), which are in a one-to-one correspondence with the minimal proper faces of the flux cone. Röhl and Bockmayr (2019) introduced the concept of a minimal set of elementary flux modes (MEMo) and gave an algorithm to compute such a set. A MEMo consists of an EFM from each minimal proper face of the flux cone together with a vector space basis of the lineality space. In general, a minimal proper face may contain many different EFMs.

The goal of this paper is to get a deeper understanding of the structure of the set of EFMs by further studying their geometric properties. We generalize the result of Larhlimi and Bockmayr (2009) and show that higher-dimensional faces of the flux cone can be characterized by metabolic behaviors. We study the distribution of EFMs over the faces of the flux cone and introduce as a measure of complexity the *degree* of a flux vector or EFM, which is the dimension of the inclusionwise minimal face containing it. The intuitive idea behind this is that the smaller the degree of a flux vector or EFM, the more elementary it is. More formally, we show that any flux vector of degree k can be decomposed into a sum of at most k EFMs of degree at most k . We prove upper bounds on the degree of EFMs and show that EFMs in the relative interior of the flux cone occur only in very special cases. We also study the effect of removing blocked reactions and redundant irreversibility constraints and deduce an upper bound on the cardinality of minimal metabolic behaviors.

2 Polyhedral cones

In this section we introduce some mathematical background on polyhedral cones. For further reading we refer to Lauritzen (2013), Schneider (1993) and Schrijver (1986).

A vector $x \in \mathbb{R}^n$ is a *linear combination* of the vectors $x^1, \dots, x^k \in \mathbb{R}^n$ if $x = \lambda_1 x^1 + \dots + \lambda_k x^k$, for some $\lambda_1, \dots, \lambda_k \in \mathbb{R}$. If, in addition

$$\left\{ \begin{array}{l} \lambda_1, \dots, \lambda_k \geq 0, \\ \lambda_1 + \dots + \lambda_k = 1, \end{array} \right\} \text{ we call } x \text{ a } \left\{ \begin{array}{l} \text{conic} \\ \text{affine} \\ \text{convex} \end{array} \right\} \text{ combination.}$$

For a nonempty subset $X \subseteq \mathbb{R}^n$, we denote by $\text{lin}(X)$ (resp. $\text{cone}(X)$, $\text{aff}(X)$, $\text{conv}(X)$) the *linear* (resp. *conic*, *affine*, *convex*) *hull of* X , i.e., the set of all linear (resp. conic, affine, convex) combinations of finitely many vectors of X .

A nonempty set $C \subseteq \mathbb{R}^n$ is a *convex cone*, if it is closed under conic combinations, i.e., $\lambda x + \mu y \in C$, for all $x, y \in C$ and $\lambda, \mu \geq 0$. A convex cone C is *polyhedral* if it is the solution set of a system of finitely many homogeneous linear inequalities, i.e., if

$$C = \{x \in \mathbb{R}^n \mid Ax \geq 0\},$$

for some matrix $A \in \mathbb{R}^{m \times n}$. If a cone C is the conic hull of a finite set $X = \{x^1, \dots, x^s\} \subset \mathbb{R}^n$, it is called *finitely generated* and the set X is called a *generating set of* C . By the well-known theorem of Farkas-Minkowski-Weyl (see e.g. (Schrijver 1986)), a convex cone is polyhedral if and only if it is finitely generated. For the rest of this paper we will only consider polyhedral cones and often simply write cone.

If $C = \{x \in \mathbb{R}^n \mid Ax \geq 0\}$, is a polyhedral cone, an inequality $ax \geq 0$, where a denotes a row of A and ax the inner product of a and x , is called an *implicit equality in* $Ax \geq 0$, if $ax = 0$, for all $x \in C$. Following (Schrijver 1986), we denote by $A^\#x \geq 0$ the system of implicit equalities in $Ax \geq 0$ and by $A^+x \geq 0$ the remaining inequalities.

If removing an inequality $ax \geq 0$ from $Ax \geq 0$ does not change the associated cone C , the inequality is called *redundant*. If there are no redundant inequalities, the description $Ax \geq 0$ is called *irredundant*.

The *dimension* $\text{dim}(C)$ of a cone C is the dimension of its affine hull $\text{aff}(C) = \{x \in \mathbb{R}^n \mid A^\#x = 0\}$ and is equal to $n - \text{rank}(A^\#)$. Since $0 \in C$, $\text{aff}(C)$ coincides with the linear hull $\text{lin}(C)$.

A vector $x \in C$ is in the *relative interior of* C , shortly $x \in \text{relint}(C)$, if there exists $\epsilon > 0$ such that $B_\epsilon(x) \cap \text{aff}(C) \subseteq C$, where $B_\epsilon(x)$ is the n -dimensional ball of radius ϵ centered at x . If $x \in C$ is not in the relative interior of C , it is in the *relative boundary of* C .

The *lineality space* of a cone $C = \{x \in \mathbb{R}^n \mid Ax \geq 0\}$ is given by $\text{lin. space}(C) = \{x \in \mathbb{R}^n \mid Ax = 0\}$, which is the inclusionwise maximal linear subspace contained in C . A cone C is called *pointed* if its lineality space is trivial, i.e., $\text{lin. space}(C) = \{0\}$. If a cone is pointed, it does not contain a line.

An inequality $ax \geq 0$ is called *valid for* C if $C \subseteq \{x \in \mathbb{R}^n \mid ax \geq 0\}$. A nonempty set $F \subseteq C$ is called a *face of* C if there exists an inequality $ax \geq 0$ valid for C such that $F = C \cap \{x \in C \mid ax = 0\}$. The hyperplane $\{x \in C \mid ax = 0\}$ is then called a *supporting hyperplane of* F . Alternatively, a face can be characterized as a nonempty

set $F \subseteq C$ with $F = \{x \in C \mid A_{I,\star}x = 0\}$, where $A_{I,\star}$ is the submatrix of A whose rows belong to the set $I \subseteq \{1, \dots, m\}$ (Schrijver 1986).

A polyhedral cone C has only finitely many faces, each face F of C is itself a polyhedral cone and $F' \subseteq F$ is a face of F if and only if F' is a face of C . A k -dimensional face will also be called a k -face. A cone C is pointed if and only if it has a 0-face, namely the origin.

A face $F \neq C$ of C is called a *facet* if it is inclusionwise maximal, i.e., there is no other face $F' \neq C$ such that $F \subset F'$. If the description $Ax \geq 0$ of C is irredundant, there is a 1–1 correspondence between the facets of C and the inequalities in $A^+x \geq 0$ (Schrijver 1986, Theor. 8.1). In particular, for every facet F there is an inequality $ax \geq 0$ from $A^+x \geq 0$ such that $F = \{x \in C \mid ax = 0\}$. We have $\dim(F) = \dim(C) - 1$ for every facet F of C , and every face of C (except C itself) is the intersection of facets of C .

The *face lattice* of a polyhedral cone C is the partially ordered set $L(C)$ of all faces of C , partially ordered by set inclusion (Henk et al. 2017; Ziegler 1994). Two polyhedral cones C, C' are *combinatorially equivalent* if there is a bijection from $L(C)$ to $L(C')$ that preserves the inclusion relation. The *combinatorial type* of a polyhedral cone is the equivalence class under combinatorial equivalence.

Proposition 2.1 *Let $C = \{x \in \mathbb{R}^n \mid Ax \geq 0\}$ be a polyhedral cone and $z \in C$. Let A_z^- be the submatrix of A whose rows correspond to the inequalities in $Ax \geq 0$ that are fulfilled with equality by z . Let F be the face of C defined by $F = \{x \in C \mid A_z^-x = 0\}$. Then*

- (i) F is the inclusionwise minimal face of C containing z ,
- (ii) $\dim(F) = n - \text{rank}(A_z^-)$,
- (iii) $z \in \text{relint}(F)$.

Proof For $x \in \mathbb{R}^n$, define $I(x) = \{i \in \{1, \dots, m\} \mid a_i x = 0\}$, where a_i is the i -th row in A . Let $F' = \{x \in C \mid A_{I,\star}x = 0\}$ be a face of C containing z . Then $A_{I,\star}z = 0$ and thus $I \subseteq I(z)$, which implies $F = \{x \in C \mid A_{I(z),\star}x = 0\} \subseteq F'$ and statement (i) follows.

For $x \in F$, we have $I(z) \subseteq I(x)$. Therefore, $I(z)$ has the minimal number of elements among $I(x)$, where $x \in F$. The statements (ii) and (iii) now follow from Proposition 4.3 in Lauritzen (2013) and its proof. □

If C is a cone with $\dim(\text{lin. space}(C)) = t \geq 0$, a face of dimension $t + 1$ is called a *minimal proper face* of C . For a pointed cone C , the minimal proper faces are the 1-faces, which are called the *extreme rays* of C . Equivalently, $\text{cone}(\{r\}) \subseteq C, r \neq 0$, is an extreme ray of C if and only if $r = x + y$ implies $x, y \in \text{cone}(\{r\})$, for all $x, y \in C$.

The *Minkowski sum* of two sets X and Y is defined as $X + Y = \{x + y \mid x \in X, y \in Y\}$. The next result states that any polyhedral cone can be decomposed into a Minkowski sum of its lineality space and a pointed cone.

Proposition 2.2 *Let $C \subseteq \mathbb{R}^n$ be a polyhedral cone, $L = \text{lin. space}(C)$. Let $G^1, \dots, G^s, s \geq 0$, be the distinct minimal proper faces of C and $g^i \in G^i \setminus L$, for $i = 1, \dots, s$. Let $P = \text{cone}(\{g^1, \dots, g^s\})$. Then*

- (i) P is a pointed cone and its extreme rays are $\text{cone}(\{g^1\}), \dots, \text{cone}(\{g^s\})$,
- (ii) $C = L + P = L + \text{cone}(\{g^1, \dots, g^s\})$, $L \cap P = \{0\}$ and if $L \cap \text{lin}(P) = \{0\}$ then $\dim(C) = \dim(L) + \dim(P)$.

Proof (i) By definition P is a finitely generated cone. Assume that P is not pointed. Then there exist $\lambda_i \geq 0, i = 1, \dots, s$, not all equal to zero, such that $0 = \sum_{i=1}^s \lambda_i g^i$. Hence, there exists $j \in \{1, \dots, s\}$ such that $-g^j \in P \subseteq C$ and so $g^j \in L$, contradicting our choice.

To see that g^1, \dots, g^s define extreme rays of P , assume without loss of generality that $\text{cone}(\{g^1\})$ is not an extreme ray of P . Then we can find $\mu_i \geq 0, 2 \leq i \leq s$, not all equal to zero, such that $g^1 = \sum_{i=2}^s \mu_i g^i$. As G^1 is a face, there exists a supporting hyperplane $H_a := \{x \in \mathbb{R}^n \mid ax = 0\}, a \in \mathbb{R}^n \setminus \{0\}$, such that $G^1 = H_a \cap C$ and $ax > 0$ for all $x \in C \setminus G^1$. Thus, from $0 = ag^1 = \sum_{i=2}^s \mu_i ag^i$ we conclude that $ag^k = 0$ for some $k \in \{2, \dots, s\}$. Therefore $g^k \in G^1$ and $G^k \subseteq G^1$, which leads to $G^k = G^1$, because G^1 is a minimal proper face. But then G^1, \dots, G^s are not distinct, which is a contradiction.

(ii) By Theorem 8.5 in (Schrijver 1986), we have $C = L + P$. Since L is a face of C , there exists $a \in \mathbb{R}^n \setminus \{0\}$ such that $ax = 0$ for all $x \in L$ and $ax > 0$ for all $x \in C \setminus L$. From $ag^i > 0, i = 1, \dots, s$, we get $ax > 0$, for all $x \in P \setminus \{0\}$, hence $L \cap P = \{0\}$. If $L \cap \text{lin}(P) = \{0\}$, we have $\text{lin}(C) = L \oplus \text{lin}(P)$, where $L \oplus \text{lin}(P)$ is the direct sum of the vector spaces L and $\text{lin}(P)$, and so $\dim(C) = \dim(L) + \dim(P)$. □

The combinatorial type of the pointed cone $P = \text{cone}(\{g^1, \dots, g^s\})$ in Proposition 2.2 is (in general) not uniquely determined. However, if we choose g^1, \dots, g^s such that $L \cap \text{lin}(P) = \{0\}$, i.e., all the g^i are contained in some linear subspace L' complementary to L , then the combinatorial type of P is independent of the choice of the g^i from the minimal proper faces. Observe that for any complementary space L' of $L, L' \cap G^i$ is a ray, i.e., $L' \cap G^i = \text{cone}(\{g^i\})$ for some $g^i \in L' \cap G^i$.

Proposition 2.3 *Let $C \subseteq \mathbb{R}^n$ be a polyhedral cone, $L = \text{lin. space}(C)$, and let P_1, P_2 be pointed cones with $L + P_1 = C = L + P_2$ and $L \cap \text{lin}(P_1) = \{0\} = L \cap \text{lin}(P_2)$. Then P_1 and P_2 are combinatorially equivalent.*

Proof Without loss of generality let $\dim(C) = n$ and $\dim(L) = t$. With $L'_j := \text{lin}(P_j)$, by Proposition 2.2 we have $\dim(L'_j) = n - t$ and $C \cap L'_j = P_j$, for $j = 1, 2$. Let u^1, \dots, u^{n-t} be a basis of L'_1 . As also L'_2 is a complement of L , there exist uniquely determined $v^1, \dots, v^{n-t} \in L'_2, w^1, \dots, w^{n-t} \in L$ such that $u^i = v^i + w^i, 1 \leq i \leq n - t$. Now let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the invertible linear map with

$$T(u) = u, \quad \forall u \in L, \quad T(u^i) = v^i, \quad 1 \leq i \leq n - t.$$

Then we get $T(C) = C$. To see this let $y = u + w \in C$ with $u \in L'_1$ and $w \in L$. Then we may write

$$y = \sum_{i=1}^{n-t} \lambda_i u^i + w, \quad \text{for some } \lambda_1, \dots, \lambda_{n-t} \in \mathbb{R}.$$

Thus

$$\begin{aligned}
 T(y) &= \sum_{i=1}^{n-t} \lambda_i v^i + w = \sum_{i=1}^{n-t} \lambda_i (u^i - w^i) + \left(y - \sum_{i=1}^{n-t} \lambda_i u^i \right) \\
 &= y - \sum_{i=1}^{n-t} \lambda_i w^i \in C + L = C,
 \end{aligned}$$

and vice versa. We conclude

$$T(P_1) = T(L'_1 \cap C) = T(L'_1) \cap T(C) = L'_2 \cap C = P_2.$$

Thus, P_1 and P_2 are affinely and thereby also combinatorially equivalent. □

We also point out that the relative interior of a polyhedral cone can easily be described by looking at the implicit equalities in $Ax \geq 0$:

Proposition 2.4 *Let $C = \{x \in \mathbb{R}^n \mid Ax \geq 0\} = \{x \in \mathbb{R}^n \mid A^-x = 0, A^+x \geq 0\}$ be a polyhedral cone. Then*

$$\text{relint}(C) = \{x \in \mathbb{R}^n \mid A^-x = 0, A^+x > 0\}.$$

Proof If $x \in C$ with $A^+x > 0$ then for any $y \in \text{lin}(C) = \{y \in \mathbb{R}^n \mid A^-y = 0\}$ there exists $\epsilon > 0$ such that $A^+(x + \epsilon y) > 0$. Hence $x \in \text{relint}(C)$. Conversely, let $x \in \text{relint}(C)$ and let a be an arbitrary row of A^+ . By definition of A^+ there exists $z \in C$ with $az > 0$. As $x \in \text{relint}(C)$, there exists $\epsilon > 0$ such $x - \epsilon z \in C$ and so $a(x - \epsilon z) \geq 0$. Thus $ax > 0$. □

3 Metabolic networks

A metabolic network $\mathcal{N} = (\mathcal{M}, \mathcal{R}, S, \text{Irr})$ is given by a set \mathcal{M} of (internal) metabolites, a set $\mathcal{R} = \text{Rev} \cup \text{Irr}$ of reversible and irreversible reactions, and a stoichiometric matrix $S \in \mathbb{R}^{m \times n}$, where $m = |\mathcal{M}|$ and $n = |\mathcal{R}|$. For $J \subseteq \mathcal{R}$, we denote by $S_{\star, J}$ the submatrix of S whose columns belong to J .

The network can be seen as a weighted hypergraph with the generalized incidence matrix S , where the metabolites are represented by nodes and the reactions by hyperarcs. A positive entry $S_{i,j} > 0$ in the stoichiometric matrix S indicates that reaction j produces metabolite i . If $S_{i,j} < 0$, metabolite i is consumed in reaction j .

Example 3.1 The metabolic network in Fig. 1 has the set of metabolites $\mathcal{M} = \{A, B, \dots, G, H\}$, the set of reversible reactions $\text{Rev} = \{1, 3, 4, 5, 9, 10, 11, 12\}$ and

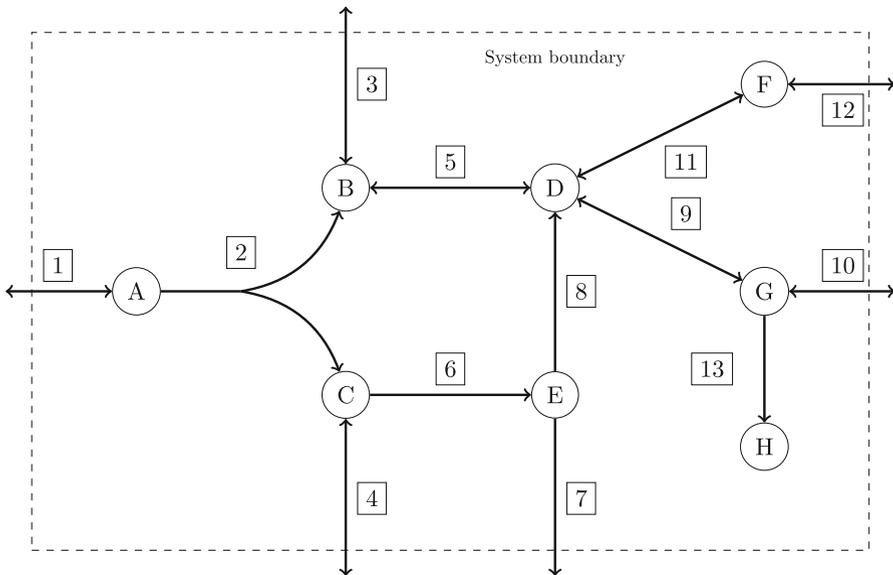


Fig. 1 Example of a metabolic network

the set of irreversible reactions $Irr = \{2, 6, 7, 8, 13\}$. The stoichiometric matrix is

$$S = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

where we assume that reversible reactions are oriented from left to right and from top to bottom. Furthermore, for simplicity, all stoichiometric coefficients are supposed to be 0, 1, or -1 .

3.1 Flux cones

In metabolic network analysis, we often assume that the network is in *steady-state*, i.e., for each internal metabolite the rate of production is equal to the rate of consumption. In matrix notation, the steady-state constraints can be written as $Sv = 0$, where $v \in \mathbb{R}^n$ denotes a *flux vector*. By adding the *thermodynamic irreversibility* constraints $v_j \geq 0$, for all $j \in Irr$, and setting

$$A = \begin{pmatrix} S \\ -S \\ I_{Irr, \star} \end{pmatrix} \tag{1}$$

we obtain the polyhedral cone

$$C = \{x \in \mathbb{R}^n \mid Ax \geq 0\} = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr}} \geq 0\}, \quad (2)$$

which is called the (steady-state) *flux cone* of \mathcal{N} . Here, v_{Irr} is the subvector of v whose components belong to Irr, and $I_{\text{Irr},\star}$ the submatrix of the $(n \times n)$ identity matrix I_n whose rows correspond to the irreversible reactions.

Implicit equalities The implicit equalities in the definition (2) of a flux cone C include all steady-state constraints $Sv = 0$. If any of the irreversibility constraints $v_j \geq 0$, $j \in \text{Irr}$, is an implicit equality, the corresponding reaction $j \in \text{Irr}$ is *blocked*, i.e., $v_j = 0$, for all $v \in C$. For some of our results, we will assume that there are no implicit equalities in $v_{\text{Irr}} \geq 0$, or equivalently that there are no blocked irreversible reactions. Blocked reactions can be determined by solving a linear optimization problem.

Redundant inequalities If in (2) one of the irreversibility constraints $v_j \geq 0$, $j \in \text{Irr}$, is redundant, the corresponding reaction j can be shifted from the set Irr of irreversible reactions to the set Rev of reversible reactions, without changing the flux cone C . The constraint $v_j \geq 0$ is then implied by the remaining constraints.

Example 3.2 In the metabolic network of Fig. 1, the irreversible reaction 13 is blocked, i.e., $v_{13} = 0$, for all $v \in C$, because there is no reaction to consume metabolite H.

The irreversibility constraint $v_6 \geq 0$ is redundant. There can be no flux from metabolite E to metabolite C because there is no reaction producing E.

The reversible reaction 1 cannot carry flux from right to left. Reaction 1 could be added to the set Irr of irreversible reactions, but then the inequality $v_1 \geq 0$ would be redundant.

If redundant inequalities are removed from the description of a flux cone, the resulting irredundant description is in general not unique, because it depends on the order in which the redundant constraints are removed.

Proposition 3.3 *Let $C = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr}} \geq 0\}$ be a flux cone such that none of the inequalities $v_j \geq 0$, $j \in \text{Irr}$, is redundant or an implicit equality. Then C has exactly $|\text{Irr}|$ facets and each facet F has the representation*

$$F = \{v \in C \mid v_j = 0\}, \text{ for some } j \in \text{Irr}. \quad (3)$$

Proof If there are no implicit equalities in $v_{\text{Irr}} \geq 0$ and A is given by (1), then $A^- = \begin{pmatrix} S \\ -S \end{pmatrix}$ and $A^+ = I_{\text{Irr},\star}$. Since there are no redundant inequalities in $v_{\text{Irr}} \geq 0$, the result follows from Theorem 8.1 in Schrijver (1986). \square

3.2 Elementary flux modes

In metabolic network analysis we are particularly interested in minimal functional units of the network. Let C be the flux cone of a metabolic network \mathcal{N} . A vector $e \in C \setminus \{0\}$ is called an *elementary flux mode* (EFM) (Schuster and Hilgetag 1994) if

it has inclusionwise minimal support, i.e., if

$$\forall v \in C \setminus \{0\} : \text{supp}(v) \subseteq \text{supp}(e) \implies \text{supp}(v) = \text{supp}(e), \tag{4}$$

where the *support* of $v \in \mathbb{R}^n$ is defined by $\text{supp}(v) = \{i \in \mathcal{R} \mid v_i \neq 0\}$. We say that a reaction $i \in \mathcal{R}$ is *active* in $v \in C$ if $i \in \text{supp}(v)$. By $\text{irr. supp}(v) := \text{supp}(v) \cap \text{Irr}$ we denote the *irreversible support* of v , i.e., the set of active irreversible reactions in v . Analogously, $\text{rev. supp}(v) := \text{supp}(v) \cap \text{Rev}$ denotes the *reversible support* of v .

To verify whether a given flux vector $v \in C$ is an EFM the *rank test* (Jevremovic et al. 2010; Urbanczik and Wagner 2005) can be applied, i.e.,

$$v \in C \text{ is an EFM if and only if } \text{rank}(S_{\star, \text{supp}(v)}) = |\text{supp}(v)| - 1. \tag{5}$$

Proposition 3.4 *Let $C = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr}} \geq 0\}$ be the flux cone of a metabolic network. Then*

$$|\text{irr. supp}(e)| \leq \text{rank}(S_{\star, \text{Irr}}) + 1,$$

for each EFM $e \in C$.

Proof Suppose the opposite. Then

$$\begin{aligned} |\text{supp}(e)| &= |\text{irr. supp}(e)| + |\text{rev. supp}(e)| \\ &\geq \text{rank}(S_{\star, \text{Irr}}) + 2 + |\text{rev. supp}(e)| \\ &\geq \text{rank}(S_{\star, \text{irr. supp}(e)}) + 2 + \text{rank}(S_{\star, \text{rev. supp}(e)}) \\ &\geq \text{rank}(S_{\star, \text{supp}(e)}) + 2, \end{aligned}$$

contradicting the rank test (5). □

3.3 Illustrative examples

To illustrate the theoretical results in the following sections through concrete examples, we will use the metabolic networks Pyruvate and Pentose Phosphate Pathway from the KEGG database (<https://www.genome.jp/kegg/pathway.html>, (Kanehisa and Goto 2000)) and Escherichia coli str. K-12 substr. MG1655 (E.coli core) from the BiGG database (King et al. 2016), where we removed the biomass reaction. The characteristics of these networks are summarized in Table 1. The EFMs were computed with `efmtool` (<https://csb.ethz.ch/tools/software/efmtool.html>, (Terzer 2009)). For polyhedral computations we used `polymake` (<https://polymake.org/>, (Assarf et al. 2017)).

4 Faces of the flux cone and metabolic behaviors

Given a metabolic network \mathcal{N} with flux cone C , a *metabolic behavior* (Larhlimi and Bockmayr 2009) of C is a nonempty set of irreversible reactions $D \subseteq \text{Irr}$ with

Table 1 Characteristics of the three example networks

	E. coli core	Pentose phosphate	Pyruvate
$(m, n) = (\mathcal{M} , \mathcal{R})$	(72, 94)	(34, 57)	(28, 81)
$ \text{Irr} $	48	19	40
$ \text{Rev} $	46	38	41
$\text{rank}(S)$	67	34	28
$n - \text{rank}(S)$	27	23	53
$\dim(C)$	23	23	53
$t = \dim(L)$	0	8	16
$\dim(P)$	23	15	37
$ \text{Facets} $	39	17	37
$ \text{blocked irr} $	8	0	0
$ \text{blocked rev} $	2	0	0
$ \text{EFMs} $	16673	5180	47854
$ \text{MMBs} $	1421	19	37

$|\mathcal{M}|$ and $|\mathcal{R}|$ denote the number of metabolites resp. reactions, which correspond to the number of rows resp. columns of the stoichiometric matrix. $|\text{Irr}|$ and $|\text{Rev}|$ denote the number of irreversible resp. reversible reactions of the network. $\text{rank}(S)$ is the rank of the stoichiometric matrix. The flux cone $C = L + P$ is the Minkowski sum of the lineality space L and a pointed cone P , with $\dim(C) = \dim(L) + \dim(P)$. $|\text{Facets}|$ is the number of facets of the flux cone, which is equal to the number of irreversibility constraints if none of these is redundant or an implicit equality. $|\text{blocked irr}|$ resp. $|\text{blocked rev}|$ describe the number of blocked irreversible resp. blocked reversible reactions. $|\text{EFMs}|$ is the number of EFMs and $|\text{MMBs}|$ the number of minimal metabolic behaviors (cf. Sect. 7)

$D = \text{irr. supp}(v)$, for some $v \in C$. A *minimal metabolic behavior* (MMB) is a metabolic behavior D for which there is no other metabolic behavior $D' \subsetneq D$. Larhlimi and Bockmayr (2009) have shown that minimal metabolic behaviors are in a 1–1 correspondence with the minimal proper faces of the flux cone C . In particular, if G is a minimal proper face and L the lineality space of C , then all flux vectors $v \in G \setminus L$ have the same irreversible support $D_G = \text{irr. supp}(v)$, which is a minimal metabolic behavior.

Proposition 4.1 *Let C be the flux cone of a metabolic network. Then each metabolic behavior is the union of MMBs.*

Proof Let $\emptyset \neq D \subseteq \text{Irr}$ be a metabolic behavior and let $v \in C$ with $D = \text{irr. supp}(v)$. Let L be the lineality space and $G^1, \dots, G^s, s \geq 0$, be the minimal proper faces of C . Since $D \neq \emptyset$ and $\text{irr. supp}(l) = \emptyset$, for all $l \in L$, we have $C \neq L$. Thus C has at least one minimal proper face and $s \geq 1$. By Proposition 2.2, $v = l + \sum_{i=1}^s \lambda_i g^i$, for some $l \in L$ and $\lambda_i \geq 0, g^i \in G^i \setminus L$, for $i = 1, \dots, s$. It follows $\text{irr. supp}(v) = \bigcup_{\lambda_i > 0} \text{irr. supp}(g^i) = \bigcup_{\lambda_i > 0} D^i$, where $D^i = \text{irr. supp}(g^i)$ is the MMB of the minimal proper face G^i , for $i = 1, \dots, s$. □

Next we generalize the characterization of minimal proper faces by MMBs (Larhlimi and Bockmayr 2009) to higher-dimensional faces.

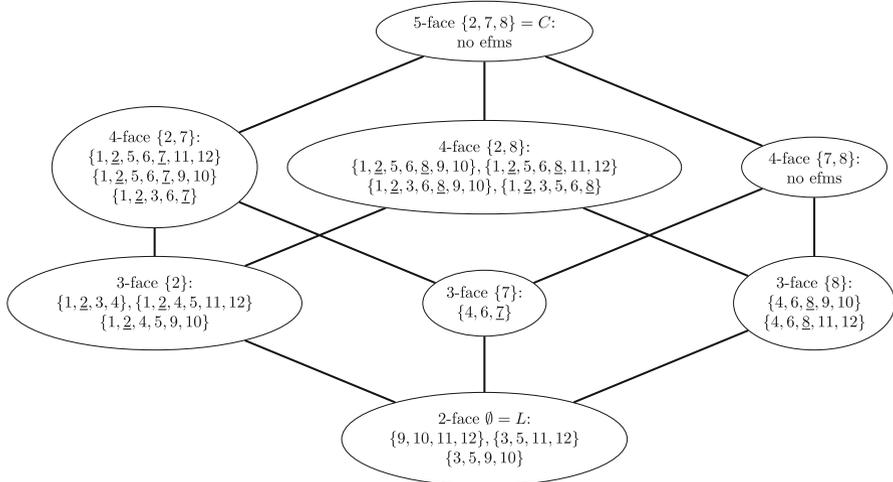


Fig. 2 Face lattice of the network in Fig. 1. Each node represents a face of the flux cone together with the corresponding metabolic behavior and the EFMs contained in that face. The active irreversible reactions are underlined. Edges connecting the nodes indicate the inclusion of lower-dimensional in higher-dimensional faces. The only 2-face is the lineality space L and the only 5-face is the entire flux cone C

Proposition 4.2 *Let C be the flux cone of a metabolic network and let F be a face of C . Then all $v \in \text{relint}(F)$ have the same irreversible support or equivalently share the same metabolic behavior.*

Proof Let $v, w \in \text{relint}(F)$. Assume w.l.o.g. that there exists $j \in \text{irr. supp}(w) \setminus \text{irr. supp}(v)$. Then $v_j = 0$, but $w_j > 0$, and hence for any $\lambda > 1$, we have $\lambda v + (1 - \lambda)w \notin C$. However, since $v, w \in \text{relint}(F)$, we know $\lambda v + (1 - \lambda)w \in F$, for some $\lambda > 1$. This shows $\text{irr. supp}(v) = \text{irr. supp}(w)$, which implies the statement. \square

Example 4.3 Consider the network in Fig. 1. If we remove the redundant irreversibility constraint $v_6 \geq 0$ and assume $6 \in \text{Rev}$, the MMBs of the network are $\{2\}, \{7\}, \{8\}$ (if $6 \in \text{Irr}$, the MMBs are $\{2\}, \{6, 7\}, \{6, 8\}$). The face lattice together with the EFMs contained in each face is shown in Fig. 2.

5 The degree of flux vectors

Let C be the flux cone of a metabolic network. We define the *degree* $\text{deg}(v)$ of a flux vector $v \in C$ as the dimension of the inclusionwise minimal face of C containing v , which for $v \neq 0$ is the unique face F of C with $v \in \text{relint}(F)$. By Proposition 2.1, we have

$$\text{deg}(v) = n - \text{rank}(A_v^-), \quad \text{where } A_v^- = \begin{pmatrix} S \\ -S \\ I_{\text{Irr} \setminus \text{irr. supp}(v), \star} \end{pmatrix}. \tag{6}$$

It follows that flux vectors in the lineality space L of C have degree $\dim(L)$ and flux vectors in minimal proper faces have degree $\dim(L) + 1$. A flux vector in the relative interior of C has degree $\dim(C)$.

A widely used technique in metabolic network analysis is to split reversible reactions into two irreversible reactions (Clarke 1980; Schilling et al. 2000; Papin et al. 2004; Gagneur and Klamt 2004; Röhl and Bockmayr 2019). In particular, this is applied in algorithms for EFM computation such as `efmtool` (Terzer 2009) or `EFMLRS` (Buchner and Zanghellini 2021). In these algorithms, splitting all reversible reactions leads to a pointed cone C' in a higher-dimensional space. The EFMs of the original flux cone C correspond to extreme rays in the reconfigured cone C' (Gagneur and Klamt 2004), which all have degree 1 in C' . Therefore, the degree of an EFM as defined by (6) has to be determined in the original flux cone C and not in C' .

Next we further characterize flux vectors in the relative interior of C .

Proposition 5.1 *Let $C = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr}} \geq 0\}$ be a flux cone with no implicit equalities in $v_{\text{Irr}} \geq 0$. For $v \in C$ we have $\deg(v) = \dim(C)$ if and only if $v_i > 0$ for all $i \in \text{Irr}$.*

Proof Direct consequence of Proposition 2.4. □

Although small examples with EFMs in the relative interior of the flux cone can be constructed, we note that real biological networks typically do not have EFMs where all unblocked irreversible reactions are active. In these cases, there are no EFMs in the relative interior of the cone, i.e., all EFMs lie on the relative boundary of C .

Next we prove an upper bound on the degree of flux vectors.

Proposition 5.2 *Let $C = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr}} \geq 0\}$ be the flux cone of a metabolic network with lineality space L . Then for each flux vector $v \in C$*

$$\deg(v) \leq \dim(L) + |\text{irr. supp}(v)|.$$

Proof By definition of the lineality space L , $t := \dim(L) = n - \text{rank}(A)$, with

$$A = \begin{pmatrix} S \\ -S \\ I_{\text{Irr}, \star} \end{pmatrix}.$$

By (6), we have $\deg(v) = n - \text{rank}(A_v^-)$, with $A_v^- = \begin{pmatrix} S \\ -S \\ I_{\text{Irr} \setminus \text{irr. supp}(v), \star} \end{pmatrix}$.

It follows $\text{rank}(A) - \text{rank}(A_v^-) = (n - t) - (n - \deg(v)) = \deg(v) - t$. Thus, at least $\deg(v) - t$ rows from A must be missing in A_v^- . This implies $|\text{irr. supp}(v)| \geq \deg(v) - t$ or $\deg(v) \leq t + |\text{irr. supp}(v)|$. □

By combining Proposition 3.4 and Proposition 5.2, we get an upper bound on the degree of EFMs.

Corollary 5.3 *Let $C = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr}} \geq 0\}$ be the flux cone of a metabolic network with lineality space L . Then for each EFM $e \in C$*

$$\deg(e) \leq \dim(L) + (\text{rank}(S_{\star, \text{Irr}}) + 1).$$

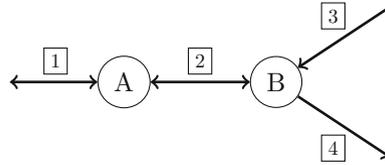


Fig. 3 Example network

Table 2 Maximum number of active irreversible reactions and maximum degree of EFMs together with the upper bounds from Proposition 3.4, 5.2 and Corollary 5.3

	E. coli core	Pentose phosphate	Pyruvate
$(\mathcal{M} , \mathcal{R})$	(72, 94)	(34, 57)	(28, 81)
$ \text{Irr} $	48	19	40
$ \text{Rev} $	46	38	41
$\text{rank}(S)$	67	34	28
$\text{rank}(S_{\star, \text{Irr}})$	41	16	24
$t = \dim(\text{lin. space}(C))$	0	8	16
$q = \max\{ \text{irr. supp}(e) : e \text{ EFM}\}$	23	9	10
$r = \text{rank}(S_{\star, \text{Irr}}) + 1$ (cf. Proposition 3.4)	42	17	25
$\max\{\text{deg}(e) : e \text{ EFM}\}$	6	14	24
$t + q$ (cf. Proposition 5.2)	23	17	26
$t + r$ (cf. Corollary 5.3)	42	25	41
$\dim(C)$	23	23	53
$ \text{EFMs} $	16673	5180	47854

The following example shows that the bound from Corollary 5.3 is sharp.

Example 5.4 The network in Fig. 3 contains 2 metabolites and 4 reactions, with $\text{Rev} = \{1, 2\}$ and $\text{Irr} = \{3, 4\}$. Given the stoichiometric matrix

$$S = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{pmatrix}$$

the network has the EFMs $e^1 = (1, 1, 0, 1)$, $e^2 = (-1, -1, 1, 0)$ and $e^3 = (0, 0, 1, 1)$, with $\text{deg}(e^1) = \text{deg}(e^2) = 1$ and $\text{deg}(e^3) = 2$. Note that $C = \text{cone}(\{e^1, e^2\})$ and $e^3 = e^1 + e^2 \in \text{relint}(C)$. Since there are no reversible flux vectors, we have $\dim(\text{lin. space}(C)) = 0$. Furthermore, $\text{rank}(S_{\star, \text{Irr}}) = 1$ and thus $\text{deg}(e^3) = \dim(\text{lin. space}(C)) + (\text{rank}(S_{\star, \text{Irr}}) + 1) = 2$.

For the example networks *E.coli core*, *Pentose Phosphate Pathway* and *Pyruvate* from Sect. 3.3, the maximum degree of EFMs and the upper bounds from Proposition 5.2 and Corollary 5.3 are given in Table 2. We can see that $l = \max\{|\text{irr. supp}(e)| : e \text{ EFM}\}$ is much smaller than the upper bound $\text{rank}(S_{\star, \text{Irr}}) + 1$ from Proposition 3.4. The actual degrees of the EFMs in these networks are summarized in Fig. 4.

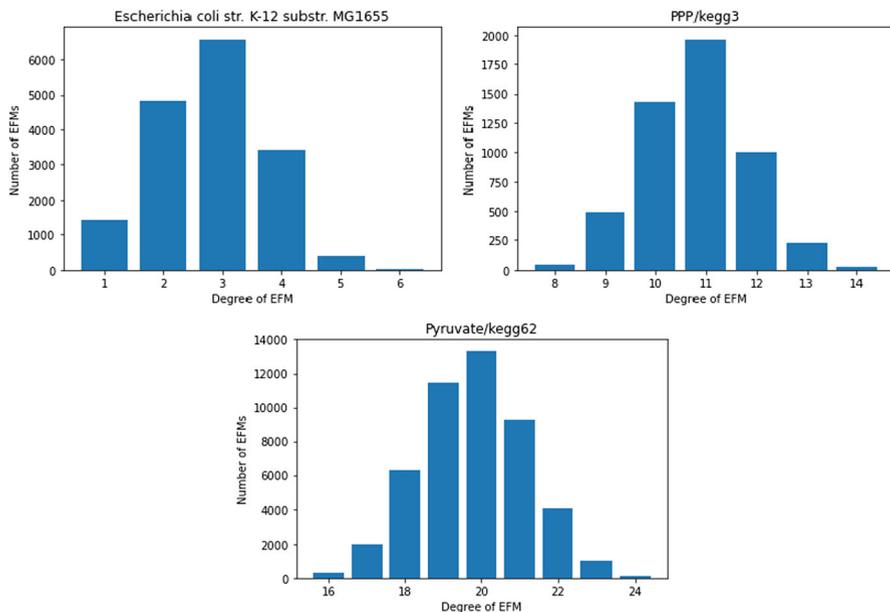


Fig. 4 Degree distribution of EFMs

The next proposition explains the scarcity of EFMs in the relative interior of a flux cone C , in the facets of C and in the faces of dimension $\dim(C) - 2$.

Proposition 5.5 *Let $C = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr}} \geq 0\}$ be a flux cone with no redundant inequalities or implicit equalities in $v_{\text{Irr}} \geq 0$. If $|\text{Irr}| > \text{rank}(S_{\star, \text{Irr}}) + q$, for some $q \in \{1, 2, 3\}$, then $\text{deg}(e) \leq \dim(C) - q$, for each EFM e of C .*

Proof By Prop. 3.4, $|\text{irr. supp}(e)| \leq \text{rank}(S_{\star, \text{Irr}}) + 1$, for each EFM e of C .

Assume $\text{deg}(e) = \dim(C) - (q - 1)$, for some $q \in \{1, 2, 3\}$. Then, by definition, the inclusionwise minimal face F containing e has dimension $\dim(C) - (q - 1) \geq \dim(C) - 2$. It follows that e resp. F is contained in exactly $q - 1$ facets of C . Here we use for $q = 3$ that a $(\dim(C) - 2)$ -face is contained in exactly two facets, cf. (Schrijver 1986, p.105).

By the hypothesis on the description of the cone, it follows $|\text{irr. supp}(e)| = |\text{Irr}| - (q - 1)$. So we get $|\text{Irr}| - (q - 1) \leq \text{rank}(S_{\star, \text{Irr}}) + 1$ or $|\text{Irr}| \leq \text{rank}(S_{\star, \text{Irr}}) + q$, in contradiction to the hypothesis $|\text{Irr}| > \text{rank}(S_{\star, \text{Irr}}) + q$. \square

In the proof we used that $(\dim(C) - q)$ -faces of a cone C are contained in exactly q facets of C , for $q = 0, 1, 2$. As the example of a 3-dimensional pointed cone with n facets shows, a $(\dim(C) - 3)$ -face (here the origin) can be contained in an arbitrary number of facets, and thus a similar argument does not hold for such faces. To limit the number of facets a face can be contained in, we introduce the concept of l -simple cones and use this for a generalization of Prop. 5.5.

A cone $C \subseteq \mathbb{R}^n$ is called l -simple for some $l \geq 1$, if every k -face of C is contained in at most $l \cdot (\dim(C) - k)$ facets of C , for all $k = \dim(\text{lin. space}(C)), \dots, \dim(C)$. Assuming that a flux cone is l -simple leads to another bound on the degree of EFMs.

Proposition 5.6 *Let $C = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr}} \geq 0\}$ be an l -simple cone with no redundant inequalities or implicit equalities in $v_{\text{Irr}} \geq 0$. Then for each EFM $e \in C$*

$$\text{deg}(e) \leq \dim(C) - \frac{|\text{Irr}| - (\text{rank}(S_{\star, \text{Irr}}) + 1)}{l}.$$

Proof By Prop. 3.4, $|\text{irr. supp}(e)| \leq \text{rank}(S_{\star, \text{Irr}}) + 1$ for each EFM $e \in C$ and thus at least $|\text{Irr}| - (\text{rank}(S_{\star, \text{Irr}}) + 1)$ entries of v_{Irr} are equal to zero. Hence e is contained in at least $|\text{Irr}| - (\text{rank}(S_{\star, \text{Irr}}) + 1)$ facets of C . Suppose $\text{deg}(e) = k$ and let $e \in F$, where F is a k -face of C . Since C is l -simple, F is contained in at most $l \cdot (\dim(C) - k)$ facets. It follows

$$|\text{Irr}| - (\text{rank}(S_{\star, \text{Irr}}) + 1) \leq l \cdot (\dim(C) - k)$$

or

$$\text{deg}(e) = k \leq \dim(C) - \frac{|\text{Irr}| - (\text{rank}(S_{\star, \text{Irr}}) + 1)}{l}.$$

□

Note that this bound is mainly theoretical because for the computation of l all faces of the flux cone have to be considered. Nevertheless $|\text{Irr}| \geq \text{rank}(S_{\star, \text{Irr}})$ and $|\text{Irr}|$ is typically significantly larger than $\text{rank}(S_{\star, \text{Irr}})$ (cf. Table 2).

6 Decomposing flux vectors

Elementary flux modes have been introduced as minimal functional units of a metabolic network (Schuster and Hilgetag 1994; Schuster et al. 2000). Any vector $v \in C$ in the flux cone C of a metabolic network can be decomposed into a conic combination

$$v = \sum_{i=1}^m \lambda_i e^i, \text{ with } \lambda_i > 0, \tag{7}$$

of elementary flux modes $e^1, \dots, e^m \in C$, see e.g. Lemma 1 in (Schuster et al. 2002). Note that in most cases, this decomposition is not unique.

Decomposing flux vectors into elementary flux modes has been widely used in metabolic pathway analysis, see e.g. (Poolman et al. 2004; Schwartz and Kanehisa 2005; Chan and Ji 2011; Jungers et al. 2011; Rügen et al. 2012; Kelk et al. 2012; Maarleveld et al. 2015; Oddsdóttir et al. 2015).

Here we show that the size and complexity of such a decomposition crucially depends on the degree of v . More precisely, we prove in Prop. 6.2 that any flux vector of degree k is a conic combination of at most k EFMs of degree at most k .

We first note that decomposing a flux vector belonging to some face $F \subseteq C$ is only possible by using vectors from the same face F .

Proposition 6.1 *Let $C = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr}} \geq 0\}$ be the flux cone of a metabolic network. Let $F = \{v \in C \mid v_I = 0\}$, for some $I \subseteq \text{Irr}$, be a face of C .*

If $v = \sum_{i=1}^m \lambda_i v^i, \lambda_i > 0$, is a conic decomposition of a flux vector $v \in F$ into flux vectors $v^1, \dots, v^m \in C$, then $v^1, \dots, v^m \in F$.

Proof Suppose $v^l \notin F$, for some $l \in \{1, \dots, m\}$. Then $v^l_j > 0$, for some $j \in I$. It follows $v_j = \sum_{i=1}^m \lambda_i v^i_j \geq \lambda_l v^l_j > 0$, and therefore $v \notin F$. \square

Next we show that any flux vector $v \in C$ of degree k can be decomposed into a conic combination of at most k EFMs of degree at most $t + 1$, where t is the dimension of the lineality space L of C . In other words, this means that any flux vector of degree k can be written as a non-negative sum of at most k EFMs, all belonging to the lineality space or to some minimal proper face of the flux cone. From each minimal proper face, at most one EFM is needed for the decomposition and all EFMs belonging to the same minimal proper face share the same minimal metabolic behavior (MMB). In the special case $k = t$, the EFMs in the decomposition all have degree t and belong to the lineality space (cf. Prop. 6.1).

Proposition 6.2 *Let $C \subseteq \mathbb{R}^n$ be the flux cone of a metabolic network and let $v \in C$ be a flux vector with $\text{deg}(v) = k$. Let $F \subseteq C, \dim(F) = k$, be the inclusionwise minimal face containing v . Then*

- i) *v can be decomposed into a conic combination $v = \sum_{i=1}^m \lambda_i e^i, \lambda_i > 0$, of at most $m \leq k$ EFMs $e^i \in F$, with $\text{deg}(e^i) \leq t + 1$, for $i = 1, \dots, m$, where t is the dimension of the lineality space L of C .
In the special case $k = t$ and $F = L$, we have $e^i \in L$ and $\text{deg}(e^i) = t$, for $i = 1, \dots, m$.*
- ii) *with probability 1, v does not allow for a conic decomposition into $m < k$ EFMs.*

Proof i): Since the EFMs generate the flux cone C (Schuster et al. (2002), Lemma 1), there exists a conic decomposition $v = \sum_{i=1}^m \lambda_i e^i, \lambda_i > 0$, for some EFMs $e^1, \dots, e^m \in C$. By Prop. 2.2 we can assume that e^1, \dots, e^m belong to the lineality space or to the minimal proper faces of C , i.e., $\text{deg}(e^i) \leq t + 1$, for $i = 1, \dots, m$. Using Carathéodory’s theorem, see e.g. (Lauritzen 2013, Prop. 3.14), we can also assume that the EFMs e^1, \dots, e^m in the conic decomposition (7) are linearly independent. By Prop. 6.1, it follows that $e^1, \dots, e^m \in F$. Since $\dim(F) = k$, there can be at most k linearly independent vectors in F , thus $m \leq k$ and the result follows.

ii): Using $k - 1$ EFMs one can generate at most a $(k - 1)$ -dimensional subset of the k -dimensional face F . Since there are only finitely many ways of choosing $k - 1$ EFMs in the set of all EFMs, the set of flux vectors $v \in F$ that can be decomposed into a combination of $k - 1$ EFMs is the finite union of sets of k -dimensional volume 0 and therefore itself a set of k -dimensional volume 0 in F . This implies the result. \square

Note that Prop. 6.2 applies also to EFMs in higher-dimensional faces, for which we get the following result:

Corollary 6.3 *Let $e \in C$ be an EFM with $\deg(e) = k \geq t + 2$, where t is the dimension of the lineality space L of C . Then e can be decomposed into a conic combination $e = \sum_{i=1}^m \lambda_i e^i$, $\lambda_i > 0$, $2 \leq m \leq k$, of at least 2 and at most k EFMs of degree strictly smaller than k .*

Since every k -face $F \subseteq C$ is the conic hull of the EFMs contained in F , we get the following lower bound on the number of EFMs in F .

Proposition 6.4 *Let $F \subseteq C$ be a face of the flux cone C with $\dim(F) = k$. Then F contains at least k EFMs. In particular, the cone C itself contains at least $\dim(C)$ EFMs.*

To conclude this section we illustrate the decomposition of flux vectors of different degree on a concrete example from flux balance analysis (Orth et al. 2010).

Example 6.5 We consider again the *E. coli* core model, now with the biomass reaction included. Using COBRApy (Ebrahim et al. 2013), we perform a flux balance analysis (FBA), where we determine the optimal growth rate while glucose uptake *Ex_glc__D_e* is limited by a lower bound of -10 (the number is negative because the exchange reactions in this model are oriented outwards). The remaining reactions have an upper bound of 1000 and lower bounds of -1000 resp. 0 for the reversible resp. irreversible reactions. The optimal growth rate of 0.917 is attained by a flux vector of degree 1, which is an extreme ray of the pointed flux cone and thus also an EFM.

Adding the lower bound 8.39 for the ATP-maintenance reaction *ATPM* leads to the standard version of the model when downloaded from the BiGG database (King et al. 2016). Now the optimal growth rate is 0.874 and the flux vector achieving this has degree 2. There is a unique decomposition into the 2 EFMs of degree 1 that span the 2-face containing this optimal solution.

If we perform a single knock-out of the gene *b1761 (gdhA)*, which blocks the reaction *GLUDy*, the new optimal growth rate is 0.851. The degree of the flux vector achieving this optimal growth rate changes to 3 and the 3-face containing this optimal solution is spanned by 3 EFMs of degree 1. Therefore, there is only one decomposition of this optimal solution into 3 EFMs of degree 1.

After knocking out a second gene *b3236 (mdh)*, which blocks the reaction *MDH*, the new optimal growth rate is 0.801. The degree of the flux vector achieving this optimal growth rate now is 4 and the 4-face containing this optimal solution is spanned by 6 EFMs of degree 1. Here there are 4 different decompositions of the optimal solution into 4 EFMs of degree 1.

It should be noted that the optimal solutions to FBA problems are generally not uniquely determined. We chose the solution that was returned by COBRApy. Different solutions with the same optimal value can lie in different faces of the flux cone. Therefore, they can have a different degree and thus also a different number of EFMs in their decomposition.

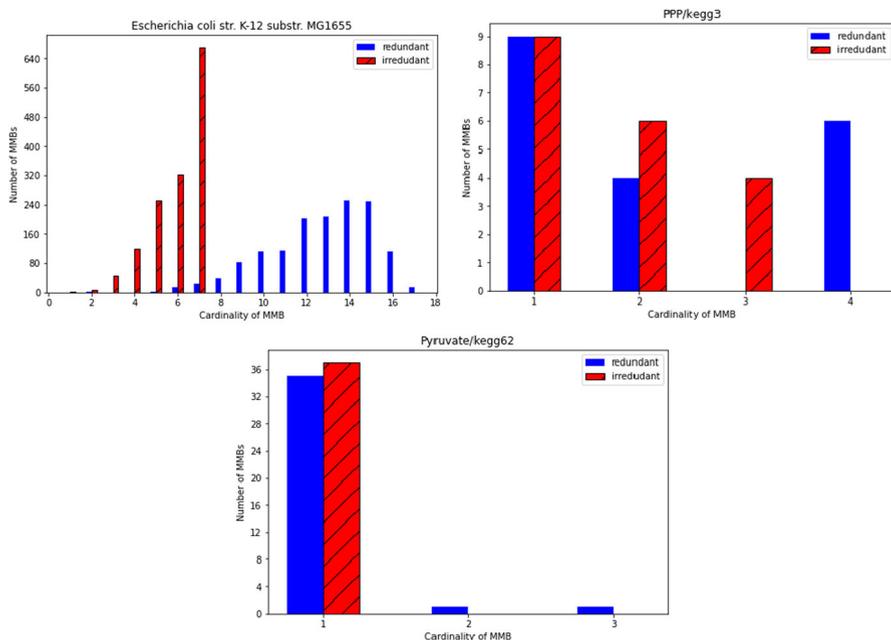


Fig. 5 Cardinalities of MMBs with and without redundant irreversibility constraints

7 The cardinality of minimal metabolic behaviors

We next prove an upper bound on the cardinality of MMBs.

Proposition 7.1 *Let C be the flux cone of a metabolic network \mathcal{N} with lineality space L . Then for each MMB D*

$$|D| \leq |\text{Irr}| - (\dim(C) - \dim(L)) + 1.$$

Proof By definition of an MMB, there exist a minimal proper face G of C , $\dim(G) = \dim(L) + 1$, and a flux vector $g \in G \setminus L$ such that $D = \text{irr. supp}(g)$. It follows that g is contained in at least $\dim(C) - (\dim(L) + 1)$ facets of C . Therefore at least $\dim(C) - (\dim(L) + 1)$ inequalities in $v_{\text{Irr}} \geq 0$ are satisfied by g with equality, which implies $|\text{irr. supp}(g)| = |D| \leq |\text{Irr}| - (\dim(C) - (\dim(L) + 1))$. \square

In general, MMBs often contain irreversible reactions for which the non-negativity constraint is redundant. If we remove redundant non-negativity constraints (i.e., shift the corresponding reactions from Irr to Rev) until the description contains no more redundant inequalities, this typically leads to much smaller cardinalities of MMBs.

For our example networks *E. coli* core, Pentose Phosphate Pathway and Pyruvate, the number of MMBs is 1421, 19 and 37 respectively. In Fig. 5 we compare the cardinalities of the MMBs in the original description of the flux cone and after removing redundant irreversibility constraints. If we start from the

original description, Prop. 7.1 provides the upper bounds 18, 5 and 4 respectively, while the actual maximal sizes of the MMBs are 17, 4 and 3. If we remove redundant non-negativity constraints in lexicographical order (i.e., the redundant non-negativity constraint corresponding to the irreversible reaction with the smallest index is removed first), the bounds become sharp, i.e., we get the bounds 9, 3 and 1 respectively, and these bounds coincide with the actual maximal sizes of the MMBs.

Proposition 7.2 *Let $C = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr}} \geq 0\}$ be the flux cone of a metabolic network with lineality space L and no redundant inequalities or implicit equalities in $v_{\text{Irr}} \geq 0$. Then the number of facets of C is equal to $\dim(C) - \dim(L)$ if and only if each MMB has cardinality one.*

Proof By Prop. 3.3, the number of facets is equal to $|\text{Irr}|$. Thus if $|\text{Irr}| = \dim(C) - \dim(L)$, then by Prop. 7.1, $1 \leq |D| \leq |\text{Irr}| - (\dim(C) - \dim(L)) + 1 = 1$, for each MMB D in C .

Conversely, if each MMB has cardinality 1, then each minimal proper face is the intersection of all facets of C but one. For all $i \in \text{Irr}$, $G^i = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr} \setminus \{i\}} = 0, v_i \geq 0\}$ is a face of C , with $\dim(G^i) \leq \dim(L) + 1$. Since $v_i \geq 0$ is not an implicit equality, there exists $g^i \in C$ with $g^i_i > 0$. Let $\bar{D}^i = \{j \in \text{Irr} \mid g^i_j > 0\}$ be the metabolic behavior defined by g^i . By Prop. 4.1, $\bar{D}^i, i \in \bar{D}^i$, is the union of MMBs, which by hypothesis all have cardinality 1. Thus, for all $i \in \text{Irr}$, $D^i = \{i\}$ is an MMB with the corresponding minimal proper face G^i , where $G^i \setminus L = \{v \in \mathbb{R}^n \mid Sv = 0, v_{\text{Irr} \setminus \{i\}} = 0, v_i > 0\}$ and $\dim(G^i) = \dim(L) + 1$. We conclude that the number of minimal proper faces of C is equal to the number of facets, which by Prop. 3.3 is equal to $|\text{Irr}|$.

It remains to prove that $|\text{Irr}| = \dim(C) - \dim(L)$. Let $U = \{u \in \mathbb{R}^n \mid Su = 0\}$ and $W = \{w \in \mathbb{R}^n \mid w_{\text{Irr}} = 0\}$. Then $U \cap W = L$ and since by hypothesis there are no implicit equalities, $\dim(C) = \dim(\text{aff}(C)) = \dim(U)$. From the dimension formula, we get $\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W) = \dim(C) + (n - |\text{Irr}|) - \dim(L)$ or $\dim(C) - \dim(L) = |\text{Irr}| - (n - \dim(U + W))$.

We claim $\dim(U + W) = n$, i.e., $U + W = \mathbb{R}^n$. For each minimal proper face $G^i, i \in \text{Irr}$, choose $e^i \in G^i \setminus L$ with $e^i_i = 1$. Then e^i_{Irr} is a unit vector, for all $i \in \text{Irr}$. Given $v \in \mathbb{R}^n$, let $u = \sum_{i \in \text{Irr}} v_i \cdot e^i$ and $w = \begin{pmatrix} v_{\text{Rev}} - u_{\text{Rev}} \\ 0 \end{pmatrix}$. Since $Se^i = 0, i \in \text{Irr}$, we get $Su = \sum_{i \in \text{Irr}} v_i \cdot Se^i = 0$ and thus $u \in U$. By definition, $w \in W$. For all $j \in \text{Irr}$, we have $u_j = \sum_{i \in \text{Irr}} v_i \cdot e^i_j = v_j e^i_j = v_j$, and thus $u_{\text{Irr}} = v_{\text{Irr}}$. Altogether, we get $u + w = \begin{pmatrix} u_{\text{Rev}} \\ u_{\text{Irr}} \end{pmatrix} + \begin{pmatrix} v_{\text{Rev}} - u_{\text{Rev}} \\ 0 \end{pmatrix} = \begin{pmatrix} v_{\text{Rev}} \\ v_{\text{Irr}} \end{pmatrix} = v$, which shows $U + W = \mathbb{R}^n$. □

8 Conclusion

In this paper, we studied geometric properties of elementary flux modes in metabolic networks. To structure the set of EFMs, we introduce the degree of an EFM and more generally of a flux vector. We show that EFMs of a smaller degree can be seen as

more elementary than those of a higher degree, since the EFMs of higher degree can be decomposed into ones of smaller degree. We use the degree of a given flux vector to predict the number of EFMs needed to decompose it and illustrate this with FBA solutions of different degree.

In previous work, see e.g. (Wagner and Urbanczik 2005), EFMs that are not extreme rays of a pointed flux cone were simply considered to be interior rays. We show that EFMs that belong to the (relative) interior of the cone (in the sense of Prop. 5.1) occur only rarely. Our bounds on the degree of EFMs as well as our computational results indicate that the EFMs rather lie in lower-dimensional faces of the flux cone.

Regarding future research, the new insights about the distribution of EFMs in the face lattice of the flux cone raise the question how these can be exploited algorithmically. For example, one could focus on enumerating EFMs of smaller degrees and omit the enumeration of EFMs of larger degrees in order to make EFM analysis for genome-scale metabolic network reconstructions more tractable.

Example 6.5 suggests that FBA solutions may be decomposed into a small number of EFMs. Deepening the understanding of the degree of FBA solutions, different EFM decompositions, and their biological relevance appears to be another promising topic for further research.

In Sect. 7, we analyze the cardinality of minimal metabolic behaviors and study the effect of removing redundant irreversibility constraints. Although irredundant descriptions of flux cones are not uniquely determined, further research is needed to determine the optimal handling of redundant constraints. Some problems benefit from the removal of redundant constraints, while others might benefit from adding them. For example, looking at the faces of the flux cone becomes easier when there are no redundant irreversibility constraints, because then every irreversible reaction corresponds to a facet of the flux cone. Conversely, the reconfiguration method to determine EFMs by splitting reversible reactions may benefit from adding redundant irreversibility constraints. Then fewer reversible reactions have to be split, which reduces the dimension of the reconfigured cone.

Acknowledgements FW was partially supported by project AA3-11* of MATH+, the Berlin Mathematics Research Center.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability Data was obtained from public domain resources as cited in the paper.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Assarf B, Gawrilow E, Herr K et al (2017) Computing convex hulls and counting integer points with polymake. *Math Program Comput* 9(1):1–38. <https://doi.org/10.1007/s12532-016-0104-z>
- Bordbar A, Monk JM, King ZA et al (2014) Constraint-based models predict metabolic and associated cellular functions. *Nat Rev Genet* 15(2):107–120. <https://doi.org/10.1038/nrg3643>
- Buchner BA, Zanghellini J (2021) EFMlrs: a Python package for elementary flux mode enumeration via lexicographic reverse search. *BMC Bioinform* 22(1):547. <https://doi.org/10.1186/s12859-021-04417-9>
- Chan SH, Ji P (2011) Decomposing flux distributions into elementary flux modes in genome-scale metabolic networks. *Bioinformatics* 27(16):2256–2262. <https://doi.org/10.1093/bioinformatics/btr367>
- Clarke B (1980) Stability of complex reaction networks. In: Prigogine I, Rice S (eds) *Advances in chemical physics*, vol 43. Wiley, pp 1–216. <https://doi.org/10.1002/9780470142622.ch1>
- Ebrahim A, Lerman J, Palsson B et al (2013) COBRApy: CONstraints-based Reconstruction and Analysis for Python. *BMC Syst Biol* 7:74. <https://doi.org/10.1186/1752-0509-7-74>
- Fang X, Lloyd CJ, Palsson BO (2020) Reconstructing organisms in silico: genome-scale models and their emerging applications. *Nat Rev Microbiol* 18(12):731–743. <https://doi.org/10.1038/s41579-020-00440-4>
- Gagneur J, Klamt S (2004) Computation of elementary modes: a unifying framework and the new binary approach. *BMC Bioinform* 5:175. <https://doi.org/10.1186/1471-2105-5-175>
- Henk M, Richter-Gebert J, Ziegler GM (2017) Basic properties of convex polytopes. In: Toth CD, O’Rourke J, Goodman JE (eds) *Handbook of discrete and computational geometry*, chap 15, 3rd edn. Chapman and Hall/CRC, Boca Raton. <https://doi.org/10.1201/9781315119601>
- Jevremovic D, Boley D (2013) Finding minimal generating set for metabolic network with reversible pathways. *BioSystems* 112(1):31–36. <https://doi.org/10.1016/j.biosystems.2013.02.003>
- Jevremovic D, Trinh C, Srien F et al (2010) On algebraic properties of extreme pathways in metabolic networks. *J Comp Biol*. <https://doi.org/10.1089/cmb.2009.0020>
- Jungers RM, Zamorano F, Blondel VD et al (2011) Fast computation of minimal elementary decompositions of metabolic flux vectors. *Automatica* 47(6):1255–1259. <https://doi.org/10.1016/j.automatica.2011.01.011>
- Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28(1):27–30. <https://doi.org/10.1093/nar/28.1.27>
- Kelk SM, Olivier BG, Stougie L et al (2012) Optimal flux spaces of genome-scale stoichiometric models are determined by a few subnetworks. *Sci Rep* 2:580. <https://doi.org/10.1038/srep00580>
- King ZA, Lu J, Dräger A et al (2016) BiGG models: a platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res* 44(D1):D515–522. <https://doi.org/10.1093/nar/gkv1049>
- Larhlimi A, Bockmayr A (2008) On inner and outer descriptions of the steady-state flux cone of a metabolic network. In: *Computational Methods in Systems Biology, CMSB 2008*, Rostock, Germany. Springer, LNBI 5307, pp 308–327. https://doi.org/10.1007/978-3-540-88562-7_22
- Larhlimi A, Bockmayr A (2009) A new constraint-based description of the steady-state flux cone of metabolic networks. *Discrete Appl Math* 157(10):2257–2266. <https://doi.org/10.1016/j.dam.2008.06.039>
- Lauritzen N (2013) *Undergraduate convexity: from Fourier and Motzkin to Kuhn and Tucker*. World Scientific Publishing Company, Singapore. <https://doi.org/10.1142/8527>
- Maarleveld TR, Wortel MT, Olivier BG et al (2015) Interplay between constraints, objectives, and optimality for genome-scale stoichiometric models. *PLoS Comput Biol* 11(4):e1004166. <https://doi.org/10.1371/journal.pcbi.1004166>
- Oddsóttir H, Hagrot E, Chotteau V et al (2015) On dynamically generating relevant elementary flux modes in a metabolic network using optimization. *J Math Biol* 71(4):903–920. <https://doi.org/10.1007/s00285-014-0844-1>
- Orth JD, Thiele I, Palsson B (2010) What is flux balance analysis? *Nat Biotechnol* 28(3):245–248. <https://doi.org/10.1038/nbt.1614>
- Papin J, Stelling J, Price N et al (2004) Comparison of network-based pathway analysis methods. *Trends Biotechnol* 22(8):400–405. <https://doi.org/10.1016/j.tibtech.2004.06.010>
- Poolman MG, Venkatesh KV, Pidcock MK et al (2004) A method for the determination of flux in elementary modes, and its application to *Lactobacillus rhamnosus*. *Biotechnol Bioeng* 88(5):601–612. <https://doi.org/10.1002/bit.20273>

- Röhl A, Bockmayr A (2019) Finding MEMo: minimum sets of elementary flux modes. *J Math Biol* 79(5):1749–1777. <https://doi.org/10.1007/s00285-019-01409-5>
- Rüger M, Bockmayr A, Legrand J et al (2012) Network reduction in metabolic pathway analysis: elucidation of the key pathways involved in the photoautotrophic growth of the green alga *Chlamydomonas reinhardtii*. *Metab Eng* 14(4):458–467. <https://doi.org/10.1016/j.ymben.2012.01.009>
- Schilling C, 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. *J Theor Biol* 203(3):229–248. <https://doi.org/10.1006/jtbi.2000.1073>
- Schneider R (1993) Convex bodies: the Brunn-Minkowski theory. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9781139003858>
- Schrijver A (1986) Theory of Linear and Integer Programming. Wiley. <https://doi.org/10.1002/net.3230200608>
- Schuster S, Hilgetag C (1994) On elementary flux modes in biochemical systems at steady state. *J Biol Syst* 2:165–182. <https://doi.org/10.1142/S0218339094000131>
- Schuster S, Fell DA, Dandekar T (2000) A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat Biotechnol* 18(3):326–332. <https://doi.org/10.1038/73786>
- Schuster S, Hilgetag C, Woods JH et al (2002) Reaction routes in biochemical reaction systems: algebraic properties, validated calculation procedure and example from nucleotide metabolism. *J Math Biol* 45:153–181. <https://doi.org/10.1007/s002850200143>
- Schwartz JM, Kanehisa M (2005) A quadratic programming approach for decomposing steady-state metabolic flux distributions onto elementary modes. *Bioinformatics* 21(Suppl 2):i204–205. <https://doi.org/10.1093/bioinformatics/bti1132>
- Terzer M (2009) Large scale methods to enumerate extreme rays and elementary modes. PhD thesis, ETH Zurich. <https://doi.org/10.3929/ethz-a-005945733>
- Urbanczik R, Wagner C (2005) An improved algorithm for stoichiometric network analysis: theory and applications. *Bioinformatics* 21(7):1203–1210. <https://doi.org/10.1093/bioinformatics/bti127>
- Wagner C, Urbanczik R (2005) The geometry of the flux cone of a metabolic network. *Biophys J* 89:3837–3845. <https://doi.org/10.1529/biophysj.104.055129>
- Ziegler GM (1994) Lectures on polytopes. Springer-Verlag, New York. <https://doi.org/10.1007/978-1-4613-8431-1>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.