Minimal cut sets in a metabolic network are elementary modes in a dual network

Kathrin Ballerstein 1,†, Axel von Kamp 2,†, Steffen Klamt 2,¶,†* and Utz-Uwe Haus 1,¶,‡

1Institute for Operations Research, Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland and 2Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, 39106 Magdeburg, Germany

Received on July 20, 2011; revised on November 22, 2011; accepted on November 30, 2011

1 INTRODUCTION

Modeling and analysis of complex metabolic networks is a central area of research in systems biology. Stoichiometric or constraint-based modeling relying on the assumption of steady state in the metabolites has become a key methodology to analyze functional properties of large- or even genome-scale metabolic networks (Price et al., 2004). One of these approaches is flux balance analysis (FBA), which uses Linear Programming in conjunction with a linear objective function to compute particular stationary flux vectors, corresponding e.g. to optimal growth (Edwards et al., 2000). Another approach for metabolic network and pathway analysis based on Elementary Modes (EMs), which are steady-state flux vectors involving a minimal set of reactions (Schuster et al., 2000). EMs analysis allows identification of balanced metabolic pathways and cycles and to study a multitude of functional network properties [for a review see Trinh et al. (2009)]. One important application of EMs is the computation of intervention strategies, e.g. for metabolic engineering (Trinh et al., 2008). A particular approach for rational (re)design of metabolic networks built upon EMs is the computation of Minimal Cut Sets (MCSs). In the original work (Klamt and Gilles, 2004), MCSs were introduced as minimal sets of reactions whose deletion will block the operation of a given objective or target reaction; i.e. removal of an MCS implies a zero flux for the target reaction in steady state. Generalizations of this definition were presented afterwards, e.g. for blocking the operation of arbitrary sets of EMs (Klamt, 2006) or/and for accounting for side constraints (Hädicke and Klamt, 2011). The problem of computing EMs results in the computation of extreme rays of convex polyhedral cones and is well studied in systems biology. Although a full enumeration is still not possible for most genome-scale models, considerable algorithmic improvements could be achieved over the last years (Gagneur and Klamt, 2004; Terzer and Stelling, 2008; Urbanczik and Wagner, 2005), and the computation of tens of millions of EMs is now, in principle, possible.

Regarding the enumeration of MCSs, three different approaches have been proposed so far (Fig. 1; upper half). The first is based on an a priori calculation of EMs with subsequent calculation of the minimal hitting sets of the (target) EMs. This approach is also known as computing a hypergraph’s transversals: the support patterns of the target EMs give rise to an undirected hypergraph and the minimal hitting sets (which are then the MCSs) correspond to those transversals whose entire set spans another hypergraph, the transversal hypergraph. The algorithm introduced in Klamt and Gilles (2004) represents one variant for computing all transversals of the hypergraph of a given set of EMs. However, this algorithm requires the enumeration of many partial solutions, which leads to an extensive running time and memory usage. As shown in Haus et al. (2008), the computation of MCSs as minimal hitting sets from a set of EMs can significantly be accelerated by using Berge’s Algorithm (Berge, 1989), which is a fast method to compute hypergraph transversals. Haus et al. (2008) presented also a completely different approach to compute MCSs, which is based on the Joint-Generation Algorithm (Friedman and Khachian, 1996) and generates both the EMs and the MCSs simultaneously. However, both techniques—hypergraph transversal (minimal hitting sets) as well as Joint-Generation Algorithm—either need the EMs of the metabolic network or generate them as a byproduct although they might not be of interest.
Fig. 1. Relation of primal and dual network with respect to EMs and MCSs.

Here we propose a method to compute MCSs directly without knowing the EMs. A brute-force procedure to achieve this could be to use FBA to test systematically all 1-, 2-, 3-, ..., combinations of knock-outs for their ability to imply the desired blocking of targeted flux vectors (Fig. 1). However, such an approach again becomes quickly prohibitive for MCSs with larger cardinalities. A slightly different concept are Minimal Direction Cuts introduced by Larhlimi and Bockmayr (2007). These cuts allow to cancel only one direction of reversible reactions, which can also be embedded in the model of MCSs by splitting reversible directions and target one of them. However, the concept of minimal direction cuts allows for a direct computation by means of the Farkas Lemma (Farkas, 1902). The enumeration of all minimal cuts is accomplished by solving iteratively (general) Mixed Integer Linear Programming problems.

In contrast, our approach relies on the notion of a dual network whose stoichiometric matrix is basically given by the transposed stoichiometric matrix of the original system. Based on the Farkas Lemma, we will show that the computation of MCSs is equivalent to finding the EMs in the dual network (Fig. 1). In particular, the duality implies also the other way around: EMs in the original metabolic network correspond to MCSs in the dual.

Additionally, our duality framework enables us to expand the concept of EMs and MCSs (originally defined for homogeneous constraints only) for systems with inhomogeneous constraints that can model e.g. boundaries for the reaction rates or for linear combinations of fluxes. The computation of EMs and MCSs remain the same in this generalization while the expressive power highly increases.

For reasons of representation, we will initially restrict ourselves to the MCS problem for a given set of target reactions which will be generalized in a later step. The duality theory, which enables direct computation of MCSs as EMs of a dual network, is presented in Section 2.2. Implementation details and computational results for benchmark problems are provided in Sections 2.3 and 3.1. Generalizations together with an application example are described in the Section 3.2.

2 METHODS

2.1 Basic concepts

A metabolic reaction network consists of \( m \) metabolites and \( n \) biochemical reactions for which we denote the index sets by \( M \) and \( N \), respectively. We model such a network via the \( m \times n \) stoichiometric matrix \( N \) where each column encodes one reaction by storing the respective stoichiometric coefficients of the metabolites consumed and produced. Each reaction may operate with a certain rate and the full set of \( n \) reaction rates is collected in the (flux) vector \( r \). The reactions usually divide into a set of irreversible and a set of irreversible reactions, denoted by \( \text{Rev} \) and \( \text{Irrev} = N^\prime \setminus \text{Rev} \), respectively. Irreversible reactions can only proceed in the forward direction, hence, \( r_i \geq 0 \) for all reactions \( i \in \text{Irrev} \).

Given such a metabolic network we say that it is in steady state if each metabolite is equally consumed and produced by the reactions. The set of flux vectors keeping the network in steady state is described as the following convex polyhedral cone, also called flux cone:

\[
F = \{ r \in \mathbb{R}^n \mid Nr = 0, r_i \geq 0 \text{ for } i \in \text{Irrev} \}.
\]

The EMs are, up to scaling, the support-minimal non-zero points of \( F \), where the support of a vector \( x \) is defined as the index set \( \{ i \in N : x_i \neq 0 \} \).

For reasons of representation, we will initially restrict ourselves to the MCS problem for a given set of target reactions which will be generalized in a later step. The duality theory, which enables direct computation of MCSs as EMs of a dual network, is presented in Section 2.2. Implementation details and computational results for benchmark problems are provided in Sections 2.3 and 3.1. Generalizations together with an application example are described in the Section 3.2.

2 METHODS

2.1 Basic concepts

A metabolic reaction network consists of \( m \) metabolites and \( n \) biochemical reactions for which we denote the index sets by \( M \) and \( N \), respectively. We model such a network via the \( m \times n \) stoichiometric matrix \( N \) where each column encodes one reaction by storing the respective stoichiometric coefficients of the metabolites consumed and produced. Each reaction may operate with a certain rate and the full set of \( n \) reaction rates is collected in the (flux) vector \( r \). The reactions usually divide into a set of irreversible and a set of irreversible reactions, denoted by \( \text{Rev} \) and \( \text{Irrev} = N^\prime \setminus \text{Rev} \), respectively. Irreversible reactions can only proceed in the forward direction, hence, \( r_i \geq 0 \) for all reactions \( i \in \text{Irrev} \).

Given such a metabolic network we say that it is in steady state if each metabolite is equally consumed and produced by the reactions. The set of flux vectors keeping the network in steady state is described as the following convex polyhedral cone, also called flux cone:

\[
F = \{ r \in \mathbb{R}^n \mid Nr = 0, r_i \geq 0 \text{ for } i \in \text{Irrev} \}.
\]

The EMs are, up to scaling, the support-minimal non-zero points of \( F \), where the support of a vector \( x \) is defined as the index set \( \{ i \in N : x_i \neq 0 \} \).

For reasons of representation, we will initially restrict ourselves to the MCS problem for a given set of target reactions which will be generalized in a later step. The duality theory, which enables direct computation of MCSs as EMs of a dual network, is presented in Section 2.2. Implementation details and computational results for benchmark problems are provided in Sections 2.3 and 3.1. Generalizations together with an application example are described in the Section 3.2.
Thus, they are determined by the binary support pattern of which we will make use. Note that EMs correspond to the extreme rays of the steady-state flux cone of all reactions are irreversible. Generally this is not true. However, one may split the reversible reactions into its two irreversible parts to make use of algorithms known for the extreme ray generation problem. A more detailed discussion about the computation and geometric interpretation of EMs can be found in Gagneur and Klamm (2004).

One is often interested in combinations of reaction knock-outs (cuts) that block the operation of certain target reactions in steady state. If we denote the set of target reactions by \( \theta \), the set of all rows. We abbreviate the polyhedron \( \{ r \in R^d | |r|_0 = 0, r_j \geq 0 \} \) by \( C \). An MCS is an inclusion-wise smallest cut set, i.e. no subset of an MCS \( C \) fulfills the latter condition. Note that the restriction of \( R^d \) is not a loss of generality as each reversible reaction can be split into two irreversible parts.

It will be convenient to denote the target reaction set by a target reaction vector \( t \in \{ 0,1 \}^d \) which is a column vector indicating the targets by 1.

### 2.2 Computing MCSs as elementary modes in a dual network

In the following, we derive a method that detects all MCSs without the necessity to generate the EMs. The idea is to use the same polyhedral methods as for calculating EMs but for a dual stoichiometric network. It is obtained by polynomial duality. First, we review the basic concept and notation.

Let \( A \), \( b \), and \( c \) be \( m \times n \), \( c \times n \), and \( n \times n \) real matrices, and \( \mathbf{P} \) the respective index sets of their rows, \( \mathbf{N} \) the index set of columns (and thus also of the variables) and let \( \mathbf{R}^d, \mathbf{R}^e \), and \( \mathbf{R}^c \) be \( \mathbf{R}^d, \mathbf{R}^e \), and \( \mathbf{R}^c \). For \( \mathbf{J} \subseteq \mathbf{M} \), we denote by \( \mathbf{A}_J \), the submatrix of \( \mathbf{A} \) consisting of the rows corresponding to the index set \( \mathbf{J} \) and all columns. Accordingly, \( \mathbf{J} \) describes the submatrix with columns from \( \mathbf{J} \subseteq \mathbf{N} \) and all rows. We abbreviate the polyhedron \( \mathbf{S} = \{ \mathbf{x} \in \mathbf{R}^d | \mathbf{Ax} = \mathbf{b}, \mathbf{Cx} = \mathbf{d} \} \) by \( \mathbf{S} \).

In particular, we can relax the set of MCSs to the EMs of a certain system.

#### Lemma 1

Each MCS \( C \) for given stoichiometric matrix \( \mathbf{N} \in \mathbf{R}^{d \times n} \) and target reaction vector \( t \in \{ 0,1 \}^d \) can be identified with an EM of the system.

\[
\begin{align*}
\mathbf{Mr} & = \mathbf{0} \\
\mathbf{r} & \geq \mathbf{0} \\
\mathbf{tr} & \geq \mathbf{1} \\
\mathbf{r} & \in \mathbf{R}^d \\
\end{align*}
\]

where \( \mathbf{I} \) denotes the \( n \times n \)-dimensional identity matrix.

Proof: First, note that the system (1) is obviously infeasible, which justifies the consideration of its EMs. To prove the claim, we have to show that any MCS \( C \) corresponds to an EM of the system (1), and no other MCS \( C' \) corresponds to the same subsystem. To this end consider the system

\[
\begin{align*}
\mathbf{Mr} & = \mathbf{0} \\
\mathbf{r} & \geq \mathbf{0} \\
\mathbf{tr} & \geq \mathbf{1} \\
\mathbf{r} & \in \mathbf{R}^d \\
\end{align*}
\]

Since \( C \) is an EMs, this system is an inconsistent subsystem of (1). Otherwise, it is not true that \( r_j = 0 \) \( \forall j \in \mathbf{T} \) and thus \( C \) is not a cut set.

But clearly (2) is not necessarily irreducible. Thus, it remains to argue that (2) has an EM that contains all equalities \( r_j = 0 \), \( i \in \mathbf{C} \). As \( C \) is an EM, we know that \( \{ r | r_j = 0, r_j \in \mathbf{0} \} \) is a cut set of equalities \( \mathbf{r} \in \mathbf{R}^d \) \( \forall \mathbf{r} \in \mathbf{T} \). Thus, there exist minimal sets \( \mathbf{J} \subseteq \mathbf{M} \) and \( \mathbf{J} \subseteq \mathbf{N} \) such that \( \{ r | r_j = 0, r_j \in \mathbf{0} \} \) contains all \( \mathbf{r} \in \mathbf{R}^d \). Hence, the system

\[
\begin{align*}
\mathbf{Mr} & = \mathbf{0} \\
\mathbf{r} & \geq \mathbf{0} \\
\mathbf{tr} & \geq \mathbf{1} \\
\mathbf{r} & \in \mathbf{R}^d \\
\end{align*}
\]

is irreducible. By construction, all EMs relate to distinct EMs.

The relation between MCSs and EMs of (1) is clearly not one-to-one but one MCS corresponds to many EMs. The choice of \( \mathbf{J} \) and \( \mathbf{K} \), \( \mathbf{K} \), is a priori, not necessarily unique, but imposing preferences on indices (use e.g. a lexicographic order on the set \( \mathbf{J} \)) yields uniqueness. Furthermore, some EM correspond to supersets of MCSs. For example, \( t^2 \geq 1 \) is always an EM whereas blocking all target reactions \( \mathbf{J} \) is always a cut set but not necessarily a minimal one. Therefore, we compute EMs by finding all EMs of (1) and select those minimal with respect to the variables fixed to 0.

In Parker and Ryan (1996), the computation of all EMs of an inconsistent system \( \{ x \leq b, c(x) = d, x \geq 0 \} \) is reduced to enumerating all vertices of a certain polyhedron. It is shown that the support of each vertex of this polyhedron identifies exactly the index set of one EM of the original system. The auxiliary polyhedron is obtained by means of the Farkas Lemma, which characterizes the feasibility of an arbitrary system of linear equality and inequality systems [for detailed illustration see Schrijver (1986)]. For completeness, we state here one version of the famous result.

**Theorem 2** (Farkas Lemma (Farkas, 1902)). For matrices \( A, B, C, D \) and vectors \( a, b, c, d \) of matching dimensions exactly one of the following statements is true.

\[
\begin{align*}
&\text{there exist } x, y, z, s, t \text{ such that } \\
&Ax + By = a, \\
&Cs + Ct' = c, \\
&Dx + Ey = e, \\
&x \geq 0, \\
&y \geq 0, \\
&z \geq 0, \\
&s \geq 0, \\
&t' \geq 0, \\
&\text{or } \\
&\text{there exist } x, w, z, s, t \text{ such that } \\
&Ax + By = a, \\
&Cs + Ct' = c, \\
&x \geq 0, \\
&z \geq 0, \\
&s \geq 0, \\
&t' \geq 0.
\end{align*}
\]

Since system (1) is inconsistent, we can apply the Theorem. In our case, the matrices \( \{ A, B \} \) correspond to the vector 1. Matrix \( C \) contains the columns corresponding to reversible reactions, and \( B \) the columns referring to reversible reactions. Accordingly, \( a = \mathbf{0} \), and \( y = \mathbf{0} \). The right-hand side \( a \) is \( \mathbf{0} \). We also split the vector \( s \) of the dual system into \( s \) (associated with \( \mathbf{N} \)) and \( x \) (associated with \( \mathbf{I} \)). In the next row, \( \mathbf{D} \) corresponds to the vector \( t' \) (note that we had to extend \( t' \) with \( s \) to \( t' \)). Again, \( C \) refers to the irreversible and \( D \) to the reversible part. By definition of \( t' \), \( D = \mathbf{0} \). The right-hand side of this equation is \( b = \mathbf{1} \). Since \( C \) and \( D \) are (row) vectors in our particular case, the vector \( w \) reduces to a scalar. Thus, we know that the following system is consistent

\[
\begin{align*}
\begin{bmatrix} z \end{bmatrix} & = \mathbf{0} \quad (3) \\
\begin{bmatrix} u \end{bmatrix} & = \mathbf{0} \\
\begin{bmatrix} x \end{bmatrix} & = \mathbf{0}
\end{bmatrix}
\end{align*}
\]

where \( \mathbf{1} \) is the identity matrix and \( \mathbf{0} \) the zero matrix. The standard form of a state-space description of a reaction network is given \( z > \mathbf{0}, w \geq 0 \), which leads to the dual stoichiometric network with stoichiometric matrix \( N_{\text{stoch}} \).

\[
\begin{align*}
\begin{bmatrix} z \end{bmatrix} & = \mathbf{0} \quad (4) \\
\begin{bmatrix} u \end{bmatrix} & = \mathbf{0} \\
\begin{bmatrix} x \end{bmatrix} & = \mathbf{0}
\end{bmatrix}
\end{align*}
\]
is clear that all relevant solutions of (4) must satisfy
\[ w \in \text{minimal support of} \quad \text{with} \quad z \]
correspond to irreversible reactions, whereas the example at the end of this section). Furthermore, the variables \( w \) and \( z \) are associated with reversible reactions.

As proved in Glasson and Ryan (1990) and Parker and Ryan (1996), the IISs of (1) correspond to the support of those extreme rays of the cone (4) with \( w > 0 \). Thus, the extreme rays that determine the MCSs are those with minimal support of \( w \) and a strictly positive \( w \). From the Farkas Lemma, it is clear that all relevant solutions of (4) must satisfy \( w > 0 \) since system (1) without the inequality \( \mathbf{r}^\top \mathbf{v} \geq 1 \) is feasible.

We also remark that the dual of the dual network (4) yields:
\[
\begin{align*}
N_r &= 0 \\
\mathbf{r}^\top \mathbf{c} &= 0 \\
\mathbf{c}^\top \mathbf{v} &= \mathbf{k}^\top \\
\mathbf{k}^\top &> 0 \\
\end{align*}
\]
The fact that only particular EMs of the dual network are relevant in this procedure suggests a particular design of the algorithm to compute them which is briefly outlined in the following section.

2.3 Implementation details

Generally, the enumeration of EMs can be a demanding problem due to a possible combinatorial explosion during the generation of new candidate modes. As described above, we are interested in those EMs of the dual system (3), which are minimal with respect to reactions \( \mathbf{v} \) and fulfill \( w > 0 \).

For small systems, it is possible to calculate all EMs and then to select those that fulfill these properties. For larger systems, this turns out to be impractical. Therefore, we devised an enumeration scheme which effectively calculates the desired subset of EMs. Details of this implementation can be found in the Supplementary Material. Briefly, certain redundancies and particular structures of the dual system (3) can be exploited to simplify the procedure and to reduce the computational effort. For example, reactions \( \mathbf{v} \) can be seen as parallel reactions (isozymes) to the backward direction of the reversible reactions \( \mathbf{v}^\top \mathbf{r} \geq 1 \) is feasible.

Fig. 2. Example of a metabolic network and its dual illustrating the correspondence of their EMs and MCSs.
The procedure of computing MCSs as EMs in a dual network is described in detail. For the former, Metatool (von Kamp and Schuster, 2008) was used, while for the latter, API functions of CellNetAnalyzer were utilized (Klamt and von Kamp, 2011). In the dual network, Metatool is used for processing the rows and Matlab scripts have been written to handle intermediate modes and the selection of support-minimal and superseded vectors (Supplementary Material). All in both methods show similar computation times. However, during row iterations of the dual network method, we observed that the number of intermediate modes can considerably exceed the number of final MCSs, which may lead to high memory requirements, at least as far as our prototypical implementation is concerned. Therefore, in future work we will address a more dedicated implementation of the dual network approach.

### 3 RESULTS

#### 3.1 Computational results

In order to test our implementation of the dual approach described above, we applied it to the benchmark problem of calculating MCSs in a model of the central metabolism of E.coli (Haus et al., 2008; note that we did not consider multifunctional enzymes leading to slightly different numbers of MCSs). Concretely, the task is to compute MCSs that disable growth under different substrate uptake scenarios. To increase the performance of EM computation, the metabolic network (N) is compressed beforehand as usual and the same reduction is also used before dualizing the system. After calculating the (primal) MCSs as EMs of the dual network, they can be readily mapped back to the original reactions.

Computation times are shown in Table 1. For the conventional MCS computation procedure, the calculation time splits into the computation of EMs and minimal hitting set computation (cf. Fig. 1).

The former was done with Metatool (von Kamp and Schuster, 2006) while for the latter, API functions of CellNetAnalyzer were used (Klamt and von Kamp, 2011). In the dual network, Metatool is used for processing the rows and Matlab scripts have been written for filtering intermediate modes and for selection of support-minimal and superseded vectors (Supplementary Material). All in both methods show similar computation times. However, during row iterations of the dual network method, we observed that the number of intermediate modes can considerably exceed the number of final MCSs, which can lead to high memory requirements, at least as far as our prototypical implementation is concerned. Therefore, in future work we will address a more dedicated implementation of the dual network method, e.g. reduce memory demand which should also improve its runtime performance.

Furthermore, although the dual approach did not outperform the conventional method to compute the MCSs in the example, other approaches show similar computation times. However, during row iterations of the dual network method, we observed that the number of intermediate modes can considerably exceed the number of final MCSs, which can lead to high memory requirements, at least as far as our prototypical implementation is concerned. Therefore, in future work we will address a more dedicated implementation of the dual network method, e.g. reduce memory demand which should also improve its runtime performance.

### 3.2 Generalizations

The procedure of computing MCSs as EMs in a dual network outlined above can be further generalized to increase the scope of applications and to allow for a more precise specification of intervention problems. First of all, it is straightforward to relax the notion of target reactions to target flux vectors which need to be blocked. In case of target flux vectors, the goal is to block all flux vectors where at least one of the target reactions is active. In the primal system (1), this was expressed by the vector t which occurs after dualization as column in the dual stoichiometric matrix (3).

Another design strategy relevant e.g. for metabolic engineering is to block all flux vectors where the yield \(\gamma^{P/S} = r_P/r_S\) (\(r_P\): product excretion rate; \(r_S\): substrate uptake rate) of a certain product falls below a given threshold h, i.e. where \(r_P/r_S < h\). These target flux vectors can be specified by \(h r_S - r_P \geq 1\) and the target reaction vector \(t\) is thus constructed with zeros except for the \(-1\) for the product excretion and \(h\) for the substrate uptake reaction.

However, for some applications we need more inhomogeneous constraints to specify target flux vectors properly. To this end, we move away from the flux cone to the flux polyhedron of a metabolic network.

**Definition 2** (Target flux polyhedron). Given a stoichiometric matrix \(N \in \mathbb{R}^{m \times n}\) and a matrix \(T \in \mathbb{R}^{m \times m}\) that models inhomogeneous constraints on the reaction rates in connection with certain lower and upper bounds \(B = ([b^L_j, b^U_j]) | 1 \leq j \leq n\), then the (target) flux polyhedron is given by

\[
P = \{ r \in \mathbb{R}^n | Nr = 0, b^L \leq Tr \leq b^U, r_{\text{row}j} \geq 0 \}.
\]

The double inequalities can be written as two separate systems which read \(Tr \leq b^U\) and \(Tr \geq b^L\). For ease of notation, we will from now on put all constraints into one matrix \(T\) and the combined right-hand side is denoted by \(b\).

\[
P = \{ r \in \mathbb{R}^n | Nr = 0, Tr \leq b, r_{\text{row}j} \geq 0 \}.
\]

The original formulation with the target reaction vector \(t\) can be easily embedded within this description with \(T = -T\) and \(b = -t\).

Matrix \(T\) allows us to express complex sets of (target) flux vectors in a flexible way, not only by specifying target reactions but also by other homogenous or inhomogeneous constraints, e.g. on boundaries or certain ratios of reaction rates. Moreover, by moving a row from \(N\) to \(T\) and providing suitable lower and upper bounds, the steady-state assumption on certain metabolites can be relaxed to certain boundaries of accumulation/degradation rates of the corresponding metabolite. This generalization also covers the reaction simplex as introduced by Horn and Jackson (1972), more appropriately called concentration polyhedra, see Clarke (1980).

Example 1. For illustration of potential application scenarios, we return to the example network of Figure 2 and assume the system...
to be in steady state as usual. Additionally, we limit the rate of reaction $r_1$ within the range $[-2,10]$ which is expressed by $\mathbf{1}_1 \leq \mathbf{b}$, with:

$$\mathbf{r} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ 10 \end{pmatrix}$$  \hspace{1cm} (5)

We can fix certain reaction rates to a single value (which could represent, for example, the non-growth associated ATP maintenance demand) in combination with boundaries for others, e.g. $r_5 = 5$, $0 \leq r_3 \leq 3$, $0 \leq r_4 \leq 5$ which can be encoded by:

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{r} = \begin{pmatrix} 0 \\ 3 \\ -3 \\ 5 \end{pmatrix}$$  \hspace{1cm} (6)

Another possibility is to drop steady-state assumptions for certain (e.g. external) metabolites and to consider accumulation or/and degradation for these metabolites within certain boundaries. For example, accumulation of metabolite $A$ at a rate between 1 and 2 degradation for these metabolites within certain boundaries. For example, accumulation of metabolite $A$ at a rate between 1 and 2 and steady state for $B$ is modeled by moving the row of $A$ from $\mathbf{N}$ to $\mathbf{T}$ and defining a suitable $\mathbf{b}$:

$$\begin{pmatrix} 1 & -1 & 1 & 0 \\ 1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{r} = \begin{pmatrix} -2 \\ 4 \\ 3 \end{pmatrix}$$  \hspace{1cm} (7)

It is well known that polyhedra can alternatively be represented by an outer description as intersection of hyperplanes (as it is the case for $\mathcal{P}$) or by an inner description, as a Minkowski sum $\mathcal{P} = \text{conv} \mathcal{V} + \text{cone} \mathcal{E}$, using the vertices $\mathcal{V}$ and the extreme rays $\mathcal{E}$ of $\mathcal{P}$. The flux polyhedron $\mathcal{P}$, therefore, decomposes into a cone and a polytopal part, whose vertices characterize the extremal operating states of the network for the prescribed ranges of reaction rates or their linear combinations. Each vertex $\mathbf{v} \in \mathcal{V}$ is determined by some combination of simultaneously tight constraints of the inequality description. The vertices hence show whether the ranges required are in fact limiting, and how their limiting effects are intertwined.

Example 2 (Example 1 cont.). System (5) describes a limit on reaction rate 1. The rays of this flux polyhedron are all EMs in Figure 2 that do not contain $r_1$, as this is limited by 10 and cannot be applied with arbitrary values. Additionally, the polyhedron consists of four vertices to be read row-wise starting with $r_1$ up to $r_5$:

$$\begin{pmatrix} P_1 \\ P_2 \\ 20002 \\ 20002 \end{pmatrix}$$

While $P_2$ attains the upper bound on the reaction rate, $P_3$ and $P_4$ fulfill the lower bound with equality. Thus, we know which minimal combination of reactions yields which reaction rate of $r_1$, $P_1$ satisfies the sign constraints of $r_2$, $r_3$, $r_4$ and $r_5$ with equality.

With the (target) flux polyhedron $\mathcal{P}$ at hand, we wish again to find minimal knock-out sets that block all its flux vectors. The generalized definition of MCSs for $\mathcal{P}$ is straightforward.

DEFINITION 3 (MCSs of flux polyhedra). Let $\mathcal{P}$ be a flux polyhedron. A cat set $C$ is a set of reactions such that no non-zero point $\mathbf{r}$ in $\mathcal{P}$ satisfies $\mathbf{r}_i = 0$ for $i \in C$. A minimal cut set is an inclusion-wise smallest cut set.

Note that setting $r_i = 0$ for $i \in C$ in flux polyhedra may lead to an empty solution space while in the homogeneous case at least the trivial solution $\mathbf{r} = \mathbf{0}$ always remains.

In accordance with the case of a flux cone, by applying again Theorem 2 one can compute the MCSs of a given (target) flux polyhedron $\mathcal{P}$ as the vertices of the following dual polyhedron:

$$\begin{pmatrix} \mathbf{N}^\top \mathbf{1} - \mathbf{T}^\top \mathbf{1} \\ \mathbf{v} \\ \mathbf{z} \\ \mathbf{w} \end{pmatrix} = \mathbf{0}$$  \hspace{1cm} (8)

with $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ the identity matrix and $\mathbf{T}^\top \mathbf{1}$ the identity matrix filled with $\theta$ rows for reversible reactions. Again the MCSs correspond to those vertices with $q > 0$, i.e. with $\mathbf{b}^\top \mathbf{w} < 0$, and minimal support of $\mathbf{v}$. As before, we determine the vertices by computing the EMs (rays) for $q > 0$ and selecting afterwards the ones with support in $\mathbf{v}$ and minimal support in $\mathbf{v}$.

This generalization of MCSs in flux polyhedra yields a highly general modeling framework. Returning to scenario (6) for the network in Figure 2, it becomes clear that the capacity constraints on reaction rates $r_3$, $r_4$ and $r_5$ imply different MCSs blocking all (steady state) flux vectors obeying the inhomogeneous constraints (6): it is now sufficient to remove either $r_3$ or $r_4$ because then the constraint for $r_5$ cannot be fulfilled anymore. This does not obviously hold when neglecting the inhomogeneous constraints.

Combinations of inhomogeneous constraints can conveniently be integrated in the dual system (8) but are difficult to treat via the conventional way, i.e. by defining a set of target EMs and then computing the minimal hitting sets: whether an EM is a target EM or not is not a local property anymore; in general, only certain (conic) combinations of EMs fulfill a given set of inhomogeneous constraints. The following example illustrates the use of the dual description for a realistic application scenario where the classical way of MCSs computation cannot directly be used.

Example 3. As in Table 1, we consider again the stoichiometric model of $E$.coli’s central metabolism with the intervention goal to block growth. For sake of simplicity, we focus on growth on succinate. In addition to the scenario in Table 1, it is reasonable to assume that there is a maximal substrate uptake rate [a realistic value is $10 \text{mmol} \text{gDW h}^{-1}$] and that we want to block all flux vectors with a minimum growth rate of $\mu > 0.001$. Furthermore, non-growth-associated ATP maintenance demand was not considered in the previous example but obviously consumes resources (ATP) that can then not be used for biomass synthesis. We therefore fix the rate $\text{maintATP}$ of the ATP maintenance demand to a specific value and consider three cases: first, we set $\text{maintATP} = 0$. This directly corresponds to the scenario considered in Table 1 and leads to the same set of MCSs. Second, we set $\text{maintATP} = 8$ which is a typical value used in other studies (Suthers et al., 2009). Third, to simulate a higher ATP maintenance demand (caused e.g. by environmental stress) we set $\text{maintATP} = 30$. When moving from the first over the second to the third scenario, we expect the MCSs to get smaller since more substrate must be directed to ATP synthesis rendering the system less robust for growth (i.e. less knock-outs suffice to inhibit growth). In fact, as can be seen in Figure 3, the size distribution of MCSs shows a shift from larger to smaller MCSs when increasing the
We presented a novel method that allows the computation of ATP maintenance demand. Hence, the cells become more vulnerable against random failures or targeted removal of certain combinations of reactions.

Fig.3. Size distribution of MCSs (blocking growth of E.coli when growing on succinate) for different ATP maintenance demands.

ATP maintenance demand. Hence, the cells become more vulnerable against random failures or targeted removal of certain combinations of reactions.

4 DISCUSSION AND CONCLUSION

We presented a novel method that allows the computation of MCSs without knowing beforehand or producing simultaneously the EMs of a stoichiometric network. Our computational approach is based on a duality framework, which completes the picture of dual relationships between EMs and MCSs (Fig. 1). Interestingly, both EMs and MCSs can be computed with the same algorithm, which enumerates extreme rays of a polyhedral cone. In addition to computational aspects, this duality offers a new perspective on function and dysfunction in biochemical networks: minimal functional units imply minimal failure modes and vice versa and their role can be interchanged in a dual network. It depends on the application which path of calculation the most effective one is.

As a further extension, we proposed a generalization of the classical steady-state analysis in metabolic networks that permits to model arbitrary combinations of homogeneous and inhomogeneous conditions on reaction rates or even accumulation/exhaustion of certain metabolites. As computations can still be performed by the same method, this expansion is a promising tool to open the concept to computational aspects, this duality offers a new perspective on function and dysfunction in biochemical networks: minimal functional units imply minimal failure modes and vice versa and their role can be interchanged in a dual network. It depends on the application which path of calculation the most effective one is.

As a further extension, we proposed a generalization of the classical steady-state analysis in metabolic networks that permits to model arbitrary combinations of homogeneous and inhomogeneous conditions on reaction rates or even accumulation/exhaustion of certain metabolites. As computations can still be performed by the same method, this expansion is a promising tool to open the concept to a wider class of applications.

Funding: German Federal Ministry of Education and Research (FORSYS-Center MaCS and ERASysBio); Ministry of Education and Research of Saxony-Anhalt (Research Center ‘Dynamic Systems’) in part.

Conflict of Interest: none declared.

REFERENCES


