

# On Flux Coupling Analysis of Metabolic Subsystems

Sayed-Amir Marashi<sup>a,b,\*</sup>, Laszlo David<sup>a,c,d</sup>, Alexander Bockmayr<sup>a,c,\*</sup>

<sup>a</sup>*FB Mathematik und Informatik, Freie Universität Berlin, Arnimallee 6, D-14195 Berlin, Germany*

<sup>b</sup>*International Max Planck Research School for Computational Biology and Scientific Computing (IMPRS-CBSC), Max Planck Institute for Molecular Genetics, Ihnestr. 63-73, D-14195 Berlin, Germany*

<sup>c</sup>*DFG-Research Center Matheon, Berlin, Germany*

<sup>d</sup>*Berlin Mathematical School (BMS), Berlin, Germany*

---

## Abstract

Genome-scale metabolic networks are useful tools for achieving a system-level understanding of metabolism. However, due to their large size, analysis of such networks may be difficult and algorithms can be very slow. Therefore, some authors have suggested to analyze subsystems instead of the original genome-scale models. Flux coupling analysis (FCA) is a well-known method for detecting functionally related reactions in metabolic networks. In this paper, we study how flux coupling relations may change if we analyze a subsystem instead of the original network. We show mathematically that a pair of fully, partially or directionally coupled reactions may be detected as uncoupled in certain subsystems. Interestingly, this behavior is the opposite of the flux coupling changes that may occur due to missing reactions, or equivalently, deletion of reactions. Computational experiments suggest that the analysis of plastid (but not mitochondrial) subsystems may significantly influence the results of FCA. Therefore, the results of FCA for subsystems, especially plastid subsystems, should be interpreted with care.

### *Keywords:*

FCA, Elementary modes, Boundary reactions, Uptake reactions, Exchange fluxes, Subnetworks, Subsystems.

---

\*To whom correspondence should be addressed.

*Email addresses:* marashi@molgen.mpg.de (Sayed-Amir Marashi), Laszlo.David@fu-berlin.de (Laszlo David), Alexander.Bockmayr@fu-berlin.de (Alexander Bockmayr)

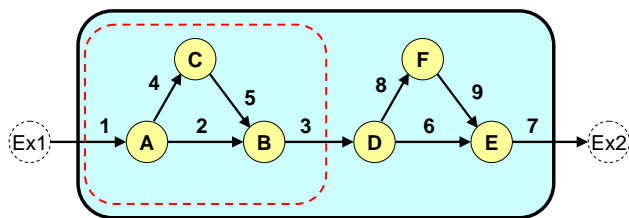


Figure 1: A metabolic network with nine reactions and six internal metabolites. The system boundary is shown as a solid black line. A subsystem of the original network can be selected by including only three metabolites (A–C) and five reactions (1–5). The new boundary is shown as dashed line. Note that reaction 3 is a boundary reaction in the selected subsystem, while it is an internal reaction in the original network.

---

## 1. Background

Genome-scale metabolic networks are useful models for the analysis of metabolism at the systems level (Oberhardt et al., 2009). However, due to the existence of hundreds to thousands of reactions in a genome-scale network model, it is not easy to analyze such networks. For this reason, some authors suggest to study only some interesting “subsystems” within these networks. A mathematical definition of a subsystem will be given in the next section. Informally speaking, a subsystem can be chosen by “cutting out” the uninteresting components, i.e., by redrawing the boundary of the network to include only a subset of reactions and metabolites (see Fig. 1). Metabolites within the system boundary will be called internal, while metabolites outside the system boundary are called external (e.g., in Fig. 1 the metabolites  $D$ ,  $E$ , and  $F$  are internal in the original network, while they become external in the subsystem). Reactions that convert only internal metabolites to each other are called internal. By choosing a new boundary, some of the internal reactions may become exchange reactions of the subsystem (e.g., reaction 3 in Fig. 1). Other reactions may become external (like reactions 6, 7, 8, and 9 in Fig. 1). External reactions convert only external metabolites to each other.

Subsystems may appear naturally in biological networks. For example, due to compartmentalization of eukaryotic cells, organelle networks can be considered as subsystems within the whole-cell network. Many authors prefer to study natural subnetworks, e.g., metabolic networks of eukaryotic or-

ganelles in isolation (Poolman et al., 2003; Vo et al., 2004). Similarly, some authors have studied other subsystems in isolation (e.g. see Teusink et al., 2009). A related approach is to split a genome-scale metabolic network into smaller subsystems and to study them in isolation (Schilling and Palsson, 2000; Schilling et al., 2002; Schuster et al., 2002; Schwarz et al., 2005; Verwoerd, 2010, 2011).

As mentioned, in the analysis of metabolic networks, time and memory requirements are critical issues (Terzer and Stelling, 2008; David et al., 2011; Jevremovic et al., 2011). Analysis of metabolic subsystems has the advantage that considerable savings can be achieved in computation time and the memory requirements. However, this “simplification” of the network may sometimes result in wrong conclusions. For example, instead of the original elementary modes of the network, one may obtain “pathway fragments” in a subsystem (Imielinski and Belta, 2008), which may not be part of any steady-state flux distribution in the original network (Kaleta et al., 2009). Additionally, interdependencies between pairs of fluxes, which can be determined by flux coupling analysis (FCA), can be missed when subsystems are analyzed. It has been previously shown that FCA of subsystems results in smaller sets of flux-coupled reactions compared to FCA of genome-scale networks (Burgard et al., 2004).

In the present study, we focus on subsystem-based vs. genome-scale flux coupling analysis. We first formally introduce the concepts used in this study. Then, we mathematically explore the effect of subsystem selection on FCA.

## 2. Formal Definitions

### 2.1. Mathematical preliminaries

We use  $[n]$  as an abbreviation for the set  $\{1, \dots, n\}$ . For an  $m \times n$  matrix  $M$ ,  $P \subseteq [m]$  and  $Q \subseteq [n]$ , we denote by  $M_{P,Q}$  the submatrix of  $M$  induced by the rows in  $P$  and the columns in  $Q$ . The  $m \times n$  zero matrix  $\mathbf{0}$  is defined as the matrix where all elements are equal to zero. The  $i$ -th unit vector  $\mathbf{e}_i \in \mathbb{R}^n$  is the column vector  $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)^T$ , where only the  $i$ -th element is 1. Here,  $\cdot^T$  denotes transposition.

### 2.2. Concepts related to metabolic networks

In a metabolic network with  $n$  reactions and  $m$  internal metabolites, the stoichiometric matrix  $S$  is an  $m \times n$  matrix where element  $S_{ij}$  is the stoichiometric coefficient of metabolite  $i$  in reaction  $j$ . Reactions can be

reversible or irreversible. In this paper, a *metabolic network*  $N = (S, Irr)$  is characterized by its stoichiometric matrix  $S$  and the set of irreversible reactions  $Irr$ .

Reactions that produce and consume internal metabolites only are called internal. Exchange or boundary reactions are reactions that produce (resp. consume) internal metabolites, but consume (resp. produce) external metabolites. They have only positive (resp. negative) entries in the stoichiometric matrix  $S$ . An irreversible boundary reaction producing internal metabolites is called an *uptake reaction*. Reactions that both produce and consume only external metabolites are considered external to the metabolic network and not included in the stoichiometric matrix.

We adopt the definition of subsystems suggested in Imielinski and Belta (2008). Suppose that in a network  $N = (S, Irr)$  the set of metabolites is equal to  $[m]$ . Without loss of generality, suppose that the stoichiometric matrix  $S^*$  of a network  $N^* = (S^*, Irr)$  is obtained by deleting from  $S$  the set of rows  $\{r + 1, \dots, m\}$ , with  $1 \leq r < m$ . We refer to the second network  $N^*$  as the *subsystem* and to the first network  $N$  as the *extension* or the *extended network*. By selecting a subsystem, it is possible to get a set of zero columns,  $Z$ . For simplicity, we redefine the stoichiometric matrix of the subsystem as  $S^* = S_{[r];[n] \setminus Z}$ , which is an  $r \times (n - |Z|)$  matrix in which only the non-zero columns are kept.  $Irr^* = Irr \setminus Z$  is the subset of  $Irr$  including the irreversible reactions in the subsystem.

If the concentrations of the internal metabolites do not change, we say that the *steady-state condition* holds. A vector  $v$  satisfying the equation  $S \cdot v = 0$  is called a (steady-state) flux distribution. The (steady-state) flux cone of the network is defined as  $C = \{v \in \mathbb{R}^n \mid S \cdot v = 0, v_i \geq 0 \text{ for all } i \in Irr\}$ . A flux vector  $e \in C$  is called an *elementary mode* (EM) (Schuster and Hilgetag, 1994; Schuster et al., 2000) if there is no vector  $v \in C \setminus \{0\}$  such that  $supp(v) \subsetneq supp(e)$ . Here,  $supp(v) = \{i \in [n] \mid v_i \neq 0\}$  denotes the support of a vector  $v \in \mathbb{R}^n$ , i.e., the set of indices of non-zero components. Thus, each EM represents a minimal set of reactions that can work together in steady-state.

In a metabolic network, some reactions may be unable to carry a non-zero flux at steady-state. Such reactions are called *blocked* (Burgard et al., 2004). For a pair of unblocked reactions, there may exist certain dependencies between the fluxes through these reactions, known as flux coupling relations Burgard et al. (2004):

- If for all  $v \in C$ ,  $v_i \neq 0$  implies  $v_j \neq 0$  then  $i$  is said to be *directionally coupled* to  $j$ , or  $i \longrightarrow j$ .
- If for all  $v \in C$ ,  $v_i \neq 0$  implies  $v_j \neq 0$  and vice versa, then  $i$  is said to be *partially coupled* to  $j$ , or  $i \longleftrightarrow j$ .
- If for all  $v \in C$ ,  $v_i \neq 0$  implies  $v_j \neq 0$  and vice versa, and additionally  $v_j/v_i = \lambda$  is a constant value, then  $i$  is said to be *fully coupled* to  $j$ , or  $i \iff j$ .

If two reactions  $i$  and  $j$  are not coupled, they are called *uncoupled*. There are two types of uncoupling relations (Marashi and Bockmayr, 2011):

- if  $i$  and  $j$  are uncoupled and additionally there exists an EM  $e \in C$  such that  $e_i \neq 0$  and  $e_j \neq 0$ , then  $i$  and  $j$  will be called *sometimes coupled* ( $i \xrightarrow{S.C.} j$ ).
- if  $i$  and  $j$  are uncoupled and additionally, for all EMs  $e \in C$ ,  $e_i = 0$  or  $e_j = 0$ , then  $i$  and  $j$  will be called *mutually exclusive* ( $i \xrightarrow{M.E.} j$ ).

### 3. Results and Discussion

In this section, we first observe how selecting a subsystem (or equivalently, redrawing network boundaries) can affect the results of flux coupling analysis. We then mathematically study the impact of subsystem selection on flux coupling relations. Afterwards, we briefly compare the effect of deleting reactions (cf. Marashi and Bockmayr, 2011) with subsystem selection. Finally, we study the relevance of our results for the analysis of real metabolic networks.

#### 3.1. Impact of redrawing the network boundaries on FCA

Fig. 2A shows a metabolic network. In Supplementary file 1, the corresponding METATOOL input file is presented. Using the METATOOL software (Pfeiffer et al., 1999; von Kamp and Schuster, 2006), only one EM (which includes all reactions) is found in this network. Therefore, reactions  $i$  and  $j$  (like other reactions) are fully coupled.

In Fig. 2B-E and Supplementary file 1, a number of scenarios are shown, where different subsystems are selected. When the subsystem in Fig. 2B is chosen, metabolites A and B are assumed as external. In this subsystem,

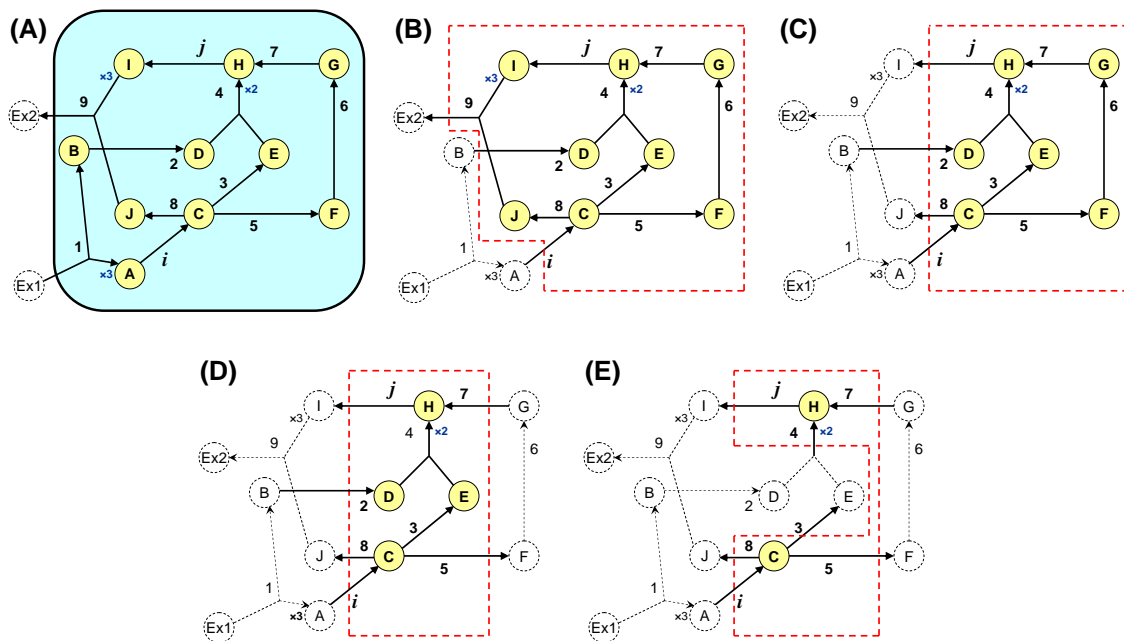


Figure 2: Flux coupling between reactions  $i$  and  $j$  depends on subsystem selection. (A) In the original network, the network boundary is shown as solid line. In this network,  $i \leftrightarrow j$ . However, depending on the selected subsystem boundaries (dashed lines), different flux coupling relations are observed: (B)  $i \leftrightarrow j$ ; (C)  $j \rightarrow i$ ; (D)  $i \xleftrightarrow{S.C.} j$ ; (E)  $i \xleftrightarrow{M.E.} j$ . See the text and Supplementary file 1 for more details.

two EMs are found, and reactions  $i$  and  $j$  are partially coupled to each other. Choosing the subsystem in Fig. 2C results in a system with three EMs, where  $j$  is directionally coupled to  $i$ . If the network boundaries are redrawn to include the subsystem in Fig. 2D,  $i$  and  $j$  become uncoupled, or more precisely, sometimes coupled. Finally, keeping only metabolites C and H in the subsystem will result in Fig. 2E, where  $i$  and  $j$  are mutually exclusive.

Fig. 3A shows another small example network. In this network no reaction consumes metabolite D. As a result, reactions 3 and 4 are blocked under steady-state conditions and reactions  $i$  and  $j$  are mutually exclusive. However, if we select the subsystem shown in Fig. 3B, metabolite D is considered as an external metabolite. In this case, reactions 3 and 4 are not blocked. Therefore,  $i \xleftrightarrow{S.C.} j$  in this subsystem. In Supplementary file 2, the METATOOL input files of the two networks are presented.

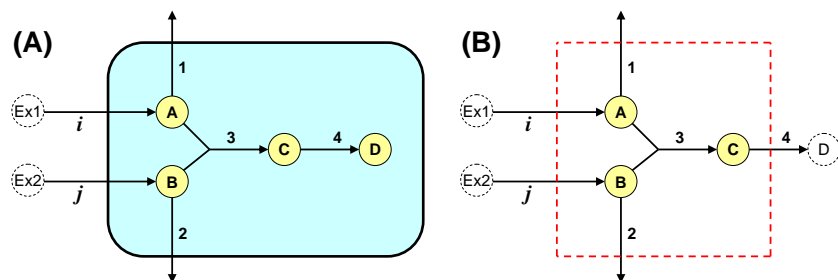


Figure 3: (A): A metabolic network with six reactions and four internal metabolites. The system boundary is shown as a solid black line. In this network, we have  $i \xleftrightarrow{M.E.} j$ ; (B): A subsystem of the network is selected. The new boundary is shown as dashed line. In this subsystem, we have  $i \xleftrightarrow{S.C.} j$ . See the text and Supplementary file 2 for more details. This figure can be seen also as a schematic representation for understanding Theorem 2:  $N^*$  is the network in (B);  $\Lambda = \{4\}$ ;  $\text{supp}(d) = \{i, j, 3, 4\}$ ;  $\text{supp}(f) = \{i, 1\}$ ;  $\text{supp}(g) = \{j, 2\}$ ; and the extension  $N$ , is shown in (A).

### 3.2. Mathematical analysis of flux coupling relations in subsystems

In the previous section, we observed that it is possible to have certain changes in flux coupling relation of two reactions, depending on the selection of the subsystem. The following theorems summarize these changes<sup>1</sup>.

**Theorem 1.** *For any two different unblocked uptake reactions  $i$  and  $j$  in a metabolic network  $N^*$ , the following holds:*

1. *There exists an extension  $N$  in which  $i \iff j$ .*
2. *If  $i$  and  $j$  are not fully coupled in  $N^*$ , there exists an extension  $N$  in which  $i \longleftrightarrow j$  and not  $i \iff j$ .*
3. *If  $i$  and  $j$  are uncoupled in  $N^*$ , there exists an extension  $N$  in which  $j \rightarrow i$  (or  $i \rightarrow j$ ) holds.*
4. *If  $i$  and  $j$  are mutually exclusive in  $N^*$ , there exists an extension  $N$  in which  $i$  and  $j$  are sometimes coupled.*

**Theorem 2.** *Let  $i$  and  $j$  be two different unblocked uptake reactions with  $i \xleftrightarrow{S.C.} j$  in a metabolic network  $N^*$ . Suppose there exists a subset  $\Lambda$  of boundary reactions that satisfies the following three conditions: (a) for each elementary mode  $d$ , with  $d_i, d_j \neq 0$ , there exists  $r \in \Lambda$  such that  $d_r \neq 0$ ; (b)*

<sup>1</sup>Formal proofs are given in Appendix A.

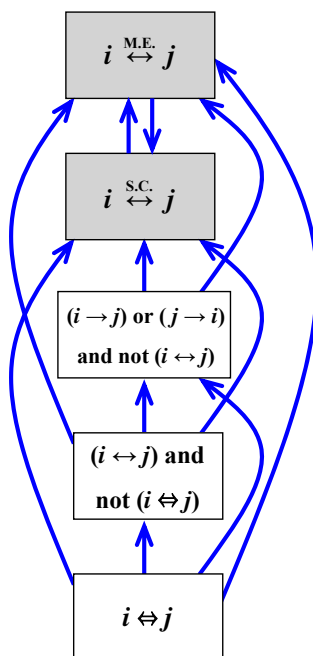


Figure 4: Possible changes in flux coupling relations due to subsystem selection. An arrow from relation A to relation B indicates that while relation A holds in the extended network, relation B may be observed in the subsystem. Gray and white boxes indicate uncoupling and coupling relations, respectively.

there exists an EM  $f$  in  $N^*$  with  $f_i \neq 0, f_j = 0$  such that  $f_r = 0$  for all  $r \in \Lambda$ ; and (c) there exists an EM  $g$  in  $N^*$  with  $g_i = 0, g_j \neq 0$ , such that  $g_r = 0$  for all  $r \in \Lambda$ . Then, there exists an extension  $N$  in which  $i \xrightarrow{M.E.} j$ .

We refer to Fig. 3 for an illustrative example.

The possible changes in flux coupling relations due to subsystems selection (Theorems 1 and 2) are summarized in Fig. 4.

Theorems 1 and 2 discuss the flux coupling relation changes only for a pair of boundary reactions  $i$  and  $j$ . However, these subsystem-induced changes are the only possible ones, independently of whether the reactions  $i$  and  $j$  are boundary or internal reactions. This is shown in Theorem 3.

**Theorem 3.** Consider a metabolic subsystem  $N^*$  with an extension  $N$ . All possible changes in flux coupling relations that may occur for two unblocked reactions  $i, j$  are the following:

1.  $i \iff j$  in  $N$  vs. any other flux (un)coupling relation in  $N^*$ ,



2.  $i \longleftrightarrow j$  in  $N$  vs.  $i \longrightarrow j$ ,  $i \xleftrightarrow{S.C.} j$ , or  $i \xleftrightarrow{M.E.} j$  in  $N^*$ ,
3.  $i \longrightarrow j$  in  $N$  vs.  $i \xleftrightarrow{S.C.} j$  or  $i \xleftrightarrow{M.E.} j$  in  $N^*$ ,
4.  $i \xleftrightarrow{S.C.} j$  in  $N$  vs.  $i \xleftrightarrow{M.E.} j$  in  $N^*$ ,
5.  $i \xleftrightarrow{M.E.} j$  in  $N$  vs.  $i \xleftrightarrow{S.C.} j$  in  $N^*$ .

### 3.3. Subsystem selection vs. reaction deletion

In a previous study, we showed that flux coupling relations can also change if reactions are deleted (or missing) from the network (Marashi and Bockmayr, 2011). Deleting a reaction (or equivalently, absence of a reaction from a metabolic model) is equivalent to deletion/absence of a *column* from the stoichiometric matrix (Marashi and Bockmayr, 2011). It was shown that pairs of uncoupled reactions in the original network may become coupled in the resulting incomplete network (but not vice versa). Additionally, deletion of reactions results in a decrease in the number of EMs.

In the present paper, we show that restriction to metabolic subsystems has very different consequences. Choosing a subsystem is equivalent to deletion of a subset of *rows* from the stoichiometric matrix. In a subsystem, coupled reaction pairs in the original network may become uncoupled in the resulting subsystem (but not vice versa). Moreover, the number of EMs in a subsystem can be greater than the number of EMs in the original network (Schuster et al., 2002).

It is important to notice that subsystem selection and reaction deletion are very different concepts. When a reaction is not included in the selected subsystem, this is *not* equivalent to “deleting” the reaction from the network. Fig. 5 shows an example. One can see that deleting reaction 3 from the network in Fig. 5A blocks all other reactions in steady-state (Fig. 5B), while considering a subsystem excluding metabolite B and reaction 3 does not have the same effect (Fig. 5C). This is due to the fact that by deleting a reaction, the corresponding internal metabolites are not deleted, and therefore, they can become unbalanced (Dandekar et al., 2003). Consequently, such metabolites will become “dead-ends”, and the associated reactions will be blocked. Note that considering the above subsystem (by excluding metabolite B and reaction 3) is also different from deleting all reactions that are involved with metabolite B (Fig. 5D).

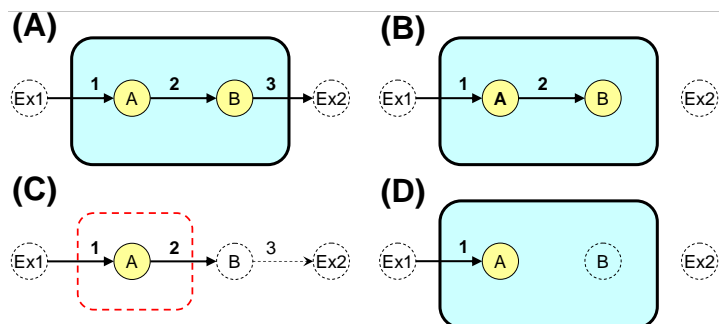


Figure 5: Reaction deletion vs. subsystem selection. (A) A small metabolic network with three reactions and two internal metabolites; (B) the same metabolic network, when reaction 3 is deleted (or missing); (C) the same metabolic network, when metabolite B (and reaction 3) are assumed to be external; (D) the same metabolic network, when reaction 2 and 3 (which produce and consume metabolite B) are deleted.

#### 4. Biological implications

It is believed that some organelles like mitochondria and plastids originated from free-living bacteria which became endosymbionts of the ancestral eukaryotic cells (Sagan, 1967; Gross and Bhattacharya, 2009). These organelles are enclosed by two membranes, which strictly control the inflow and outflow of metabolites and proteins. Therefore, one may expect the metabolic subsystems of plastids and mitochondria to be almost independent from the rest of the network. As mentioned in Section 1, some authors have studied these subsystems in isolation (e.g. Poolman et al., 2003; Vo et al., 2004; Urbanczik, 2007).

In contrast, one may argue that only a limited number of the proteins in mitochondria and plastids are encoded in their genomes. In fact, most of these proteins are encoded in the nuclear genome, synthesized in cytosol, and then transported to these organelles. There is a possibility that during evolution, some of these enzymatic functions are replaced by their cytosolic counterparts. As a result, the metabolism in these organelles might be highly interconnected to the cytosolic enzymatic activities. Therefore, the question arises: can we study the organelle subsystems in isolation, without (much) influencing the dependencies and couplings among the fluxes?

In order to answer this question, we studied ten organelles (see Section 6.1). Plastid subsystems were selected from the genome-scale metabolic networks of *Hordeum vulgare* (barley), *Arabidopsis thaliana*, *Chlamydomonas*

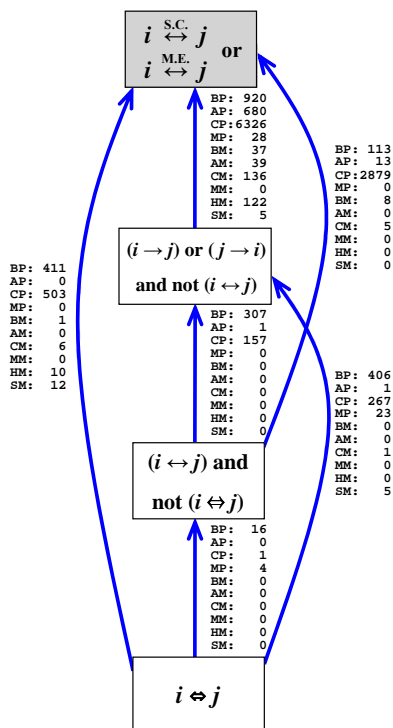


Figure 6: Observed changes in flux (un)coupling relations due to subsystem selection for the subsystems: AP: plastid of *A. thaliana*; AM: mitochondrion of *A. thaliana*; BP: plastid of barley; BM: mitochondrion of barley; CP: plastid of *C. reinhardtii*; CM: mitochondrion of *C. reinhardtii*; MP: plastid of maize; MM: mitochondrion of maize; HM: mitochondrion of human; and SM: mitochondrion of *S. cerevisiae*. An arrow from relation A to relation B indicates that while relation A holds in the extended network, relation B may be observed in the subsystem. Gray and white boxes indicate uncoupling and coupling relations, respectively.

*reinhardtii*, and *Zea mays* (maize). Moreover, six mitochondrial subsystems were selected from the genome-scale metabolic network of barley, *A. thaliana*, *C. reinhardtii*, maize, human (*Homo sapiens*) and baker's yeast (*Saccharomyces cerevisiae*). We considered all pairs of unblocked reactions in each subsystem, and computed pairwise flux coupling relations: (i) when the genome-scale network is analyzed; and (ii) when the isolated subsystem is analyzed. The results of cases (i) and (ii) are compared in Fig. 6.

Table 1 shows the frequencies of changes in flux coupling relations due to analysis of organelle subsystems instead of the complete network. The ratio of the changed flux coupling relations to the total coupling relations is

Model name	Organelle subsystem	subsystem size	Total number of coupling relations	Number of changes	Change ratio
Barley	plastid	139	9 591	2 173	22.7%
<i>A. thaliana</i>	plastid	112	6 216	695	11.2%
<i>C. reinhardtii</i>	plastid	512	131 328	10 133	7.7%
Maize	plastid	52	1 326	55	4.1%
Barley	mitochondrion	45	990	46	4.6%
<i>A. thaliana</i>	mitochondrion	51	2 250	39	1.5%
<i>C. reinhardtii</i>	mitochondrion	274	37 675	148	0.39%
Maize	mitochondrion	26	325	0	0%
Human	mitochondrion	484	116 886	132	0.11%
Yeast	mitochondrion	161	12 880	22	0.17%

Table 1: Frequencies of changes in flux coupling relations due to analysis of organelle subsystems instead of the complete network. Subsystem size is the number of reactions of a subsystem which are unblocked in the complete genome-scale network.

shown in this table. This ratio can be seen as a “measure” of the dependence of the metabolic subsystem on the network: very small ratio means that most of the flux coupling relations are not changed, and therefore, isolation of the subsystem does not much change the functional dependencies of the metabolic fluxes. On the other hand, an increased ratio means that a higher number of functional dependencies of the metabolic fluxes are changed due to analysis of the subsystem in isolation, and therefore, the subsystem is more dependent on the fluxes outside the subsystem.

From this table, one can see that the mitochondrial subsystems are relatively independent of the rest of the metabolic networks, with change ratios ranging between 0% to 4.6%. In the plastids, however, the change ratios are generally higher, 4.1% – 22.7%.

In order to show the statistical significance of the results, we studied the change ratio for random subsystems (see the Methods section). Briefly, in case of each plastid or mitochondrion subsystem, we randomly selected 1000 subsystems of similar size in the corresponding genome-scale metabolic network. Then, the actual plastid or mitochondrion change ratio was compared to the change ratio distribution of random subsystems. With this analysis, we found out that the change ratio for every plastid subsystem was highly significantly greater than the change ratios of random subsystems. However, with two exceptions (human and *A. thaliana*), the change ratio for every mitochondrial subsystem was highly significantly less than the change ratios

of random subsystems (see Supplementary file 3). This observation suggests that mitochondrial metabolic functions do not depend much on the “extra-organelle” metabolic fluxes. Thus, these subsystems can generally be studied in isolation without losing much information about the functional dependencies. On the other hand, the analysis of plastid subsystems in isolation may result in the loss of functional dependencies between metabolic fluxes.

The discrepancy between the level of independence in the plastids and mitochondria subsystems might be simply an artefact of metabolic network reconstruction. If this is not the case, one possible explanation for the difference between the plastids and mitochondria could be that during evolution, some pathways might have evolved in plants which increase the dependencies between the internal and external fluxes of the plastids. Suggesting a well-supported hypothesis is only possible when more comprehensive genome-scale metabolic networks of eukaryotes become available.

## 5. Conclusion and further research

In this paper, we show that the analysis of a subsystem instead of the complete network can cause the flux coupling relations to undergo certain changes. In particular, a pair of (fully, partially or directionally) coupled reactions may be detected as uncoupled in the chosen subsystem, but not vice versa. Interestingly, this behavior is the opposite of the flux coupling changes that may happen due to missing reactions, or equivalently, deletion of reactions.

Interdependencies of fluxes in subsystems are not always different from those in the original metabolic networks. For example, in Fig. 1, all the flux coupling relations in the selected subnetwork (between reaction pairs from  $\{1, 2, \dots, 5\}$ ) are the same as the corresponding flux coupling relations in the original network. In such a subnetwork, the set of EMs is equal to the set of EMs of the original network when projected to the subspace of reactions in the selected subnetwork (Wiback and Palsson, 2002). It is interesting to investigate the properties of such subnetworks.

With the analysis of real-world metabolic models, we observed that a non-negligible number of reaction pairs in plastids may have altered flux coupling relations when the plastids are studied in isolation. However, in the case of (most) mitochondrial subsystems, a relatively small fraction of the reaction pairs will have different flux coupling relations if these subsystems are studied in isolation. The reason for this behavior is yet to be discovered.

## 6. Methods

### 6.1. Datasets: Genome-scale network models and organelle subsystems

Six genome-scale metabolic network models are used in this study: the network of *Hordeum vulgare* (barley) (Grafahrend-Belau et al., 2009); *AraGEM*, the network of *Arabidopsis thaliana*, (de Oliveira Dal’Molin et al., 2010); *iRC1080*, the network of *Chlamydomonas reinhardtii* (Chang et al., 2011); *iRS1563*, the network of *Zea mays* (maize) (Saha et al., 2011); *Recon 1*, the network of *Homo sapiens* (human) (Duarte et al., 2007); and *iND750*, the network of *Saccharomyces cerevisiae* (baker’s yeast) (Duarte et al., 2004).

Ten organelle subsystems were considered: plastid subsystems from barley (140 reactions including 139 unblocked ones), *A. thaliana* (123 reactions including 112 unblocked ones), *C. reinhardtii* (657 reactions including 512 unblocked ones); and maize (114 reactions including 52 unblocked ones); and mitochondrial subsystems from barley (45 reactions, all unblocked), *A. thaliana* (58 reactions including 51 unblocked ones), *C. reinhardtii* (339 reactions including 274 unblocked ones), maize (53 reactions including 26 unblocked ones), human (600 reactions including 484 unblocked ones) and baker’s yeast (263 reactions including 161 unblocked ones). To select each organelle subsystem from the stoichiometric matrix of the genome-scale network, we considered a submatrix which includes all the rows corresponding to the metabolites in the organelle subsystem.

### 6.2. Flux coupling analysis

Flux coupling analysis was performed by the FFCA software (David et al., 2011). We considered all pairs of unblocked reactions in each subsystem, and computed pairwise flux coupling relations: (i) when the genome-scale network is analyzed; and (ii) when the isolated subsystem is analyzed.

We also studied the change ratio of randomly selected subsystems. The procedure was as follows. Consider an organelle subsystem with  $n$  reactions. In the beginning, a metabolite and its associated reactions were randomly selected and included in the subsystem. Then, in each step, a neighboring metabolite was added to the subsystem together with its associated reactions. This step is repeated until the number of reactions in the subsystem becomes equal to  $n \pm 5$ .

In case of each organelle, the change ratio of the organelle was compared to the distribution of the change ratios of 1000 random subsystems. We used

one-sided Wilcoxon signed rank test to determine the statistical significance of the change ratio of the organelle.

## Acknowledgements

We would like to thank Jason Papin (University of Virginia) for his helpful discussion on different versions of *C. reinhardtii* metabolic network.

Burgard, A. P., Nikolaev, E. V., Schilling, C. H., Maranas, C. D., 2004. Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome Research* 14, 301–312.

Chang, R. L., Ghamsari, L., Manichaikul, A., Hom, E. F. Y., Balaji, S., Fu, W., Shen, Y., Hao, T., Palsson, B. O., Salehi-Ashtiani, K., Papin, J. A., 2011. Metabolic network reconstruction of *Chlamydomonas* offers insight into light-driven algal metabolism. *Molecular Systems Biology* 7, 518.

Dandekar, T., Moldenhauer, F., Bulik, S., Bertram, H., Schuster, S., 2003. A method for classifying metabolites in topological pathway analyses based on minimization of pathway number. *BioSystems* 70, 255–270.

David, L., Marashi, S.-A., Larhlimi, A., Mieth, B., Bockmayr, A., 2011. FFCA: a feasibility-based method for flux coupling analysis of metabolic networks. *BMC Bioinformatics* 12, 236.

de Oliveira Dal’Molin, C. G., Quek, L.-E., Palfreyman, R. W., Brumbley, S. M., Nielsen, L. K., 2010. AraGEM, a genome-scale reconstruction of the primary metabolic network in Arabidopsis. *Plant Physiology* 152, 579–589.

Duarte, N. C., Becker, S. A., Jamshidi, N., Thiele, I., Mo, M. L., Vo, T. D., Srivas, R., Palsson, B. O., 2007. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proceedings of the National Academy of Sciences of the United States of America* 104, 1777–1782.

Duarte, N. C., Herrgård, M. J., Palsson, B. O., 2004. Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Research* 14, 1298–1309.

- Grafahrend-Belau, E., Schreiber, F., Koschützki, D., Junker, B. H., 2009. Flux balance analysis of barley seeds: A computational approach to study systemic properties of central metabolism. *Plant Physiology* 149, 585–598.
- Gross, J., Bhattacharya, D., 2009. Mitochondrial and plastid evolution in eukaryotes: an outsiders' perspective. *Nature Reviews Genetics* 10, 495–505.
- Imielinski, M., Belta, C., 2008. Exploiting the pathway structure of metabolism to reveal high-order epistasis. *BMC Systems Biology* 2, 40.
- Jevremovic, D., Boley, D., Sosa, C. P., 2011. Divide-and-conquer approach to the parallel computation of elementary flux modes in metabolic networks. in: *Proceedings of the IEEE International Symposium on Parallel and Distributed Processing Workshops and PhD Forum (IPDPSW)*, Shanghai, China, pp. 502–511.
- Kaleta, C., de Figueiredo, L. F., Schuster, S., 2009. Can the whole be less than the sum of its parts? Pathway analysis in genome-scale metabolic networks using elementary flux patterns. *Genome Research* 19, 1872–1883.
- Marashi, S.-A., Bockmayr, A., 2011. Flux coupling analysis of metabolic networks is sensitive to missing reactions. *BioSystems* 103, 57–66.
- Oberhardt, M. A., Palsson, B. O., Papin, J. A., 2009. Applications of genome-scale metabolic reconstructions. *Molecular Systems Biology* 5, 320.
- Pfeiffer, T., Sánchez-Valdenebro, I., Nuño, J. C., Montero, F., Schuster, S., 1999. METATOOL: for studying metabolic networks. *Bioinformatics* 15, 251–257.
- Poolman, M. G., Fell, D. A., Raines, C. A., 2003. Elementary modes analysis of photosynthate metabolism in the chloroplast stroma. *European Journal of Biochemistry* 270, 430–439.
- Sagan, L., 1967. On the origin of mitosing cells. *Journal of Theoretical Biology* 14, 225–274.
- Saha, R., Suthers, P. F., Maranas, C. D., 2011. *Zea mays* iRS1563: A comprehensive genome-scale metabolic reconstruction of maize metabolism. *PLoS ONE* 6, e21784.



- Schilling, C. H., Covert, M. W., Famili, I., Church, G. M., Edwards, J. S., Palsson, B. O., 2002. Genome-scale metabolic model of *Helicobacter pylori* 26695. *Journal of Bacteriology* 184, 4582–4593.
- Schilling, C. H., Palsson, B. O., 2000. Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis. *Journal of Theoretical Biology* 203, 249–283.
- Schuster, S., Fell, D. A., Dandekar, T., 2000. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nature Biotechnology* 18, 326–332.
- Schuster, S., Hilgetag, C., 1994. On elementary flux modes in biochemical reaction systems at steady state. *Journal of Biological Systems* 2, 165–182.
- Schuster, S., Pfeiffer, T., Moldenhauer, F., Koch, I., Dandekar, T., 2002. Exploring the pathway structure of metabolism: decomposition into sub-networks and application to *Mycoplasma pneumoniae*. *Bioinformatics* 18, 351–361.
- Schwarz, R., Musch, P., von Kamp, A., Engels, B., Schirmer, H., Schuster, S., Dandekar, T., 2005. YANA – a software tool for analyzing flux modes, gene-expression and enzyme activities. *BMC Bioinformatics* 6, 135.
- Terzer, M., Stelling, J., 2008. Large-scale computation of elementary flux modes with bit pattern trees. *Bioinformatics* 24, 2229–2235.
- Teusink, B., Wiersma, A., Jacobs, L., Notebaart, R. A., Smid, E. J., 2009. Understanding the adaptive growth strategy of *Lactobacillus plantarum* by in silico optimisation. *PLoS Computational Biology* 5, e1000410.
- Urbanczik, R., 2007. Enumerating constrained elementary flux vectors of metabolic networks. *IET Systems Biology* 1, 274–279.
- Verwoerd, W. S., 2010. Dissecting Metabolic Networks into Functional Subnets. In: Fung, G. (Ed.), *A Practical Guide to Bioinformatics Analysis*. iConcept Press, pp. 79–98.
- Verwoerd, W. S., 2011. A new computational method to split large biochemical networks into coherent subnets. *BMC Systems Biology* 5, 25.

Vo, T. D., Greenberg, H. J., Palsson, B. O., 2004. Reconstruction and functional characterization of the human mitochondrial metabolic network based on proteomic and biochemical data. *Journal of Biological Chemistry* 279, 39532–39540.

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

Wiback, S. J., Palsson, B. O., 2002. Extreme pathway analysis of human red blood cell metabolism. *Biophysical Journal* 83, 808–818.

## Appendix A. Theorems and Proofs

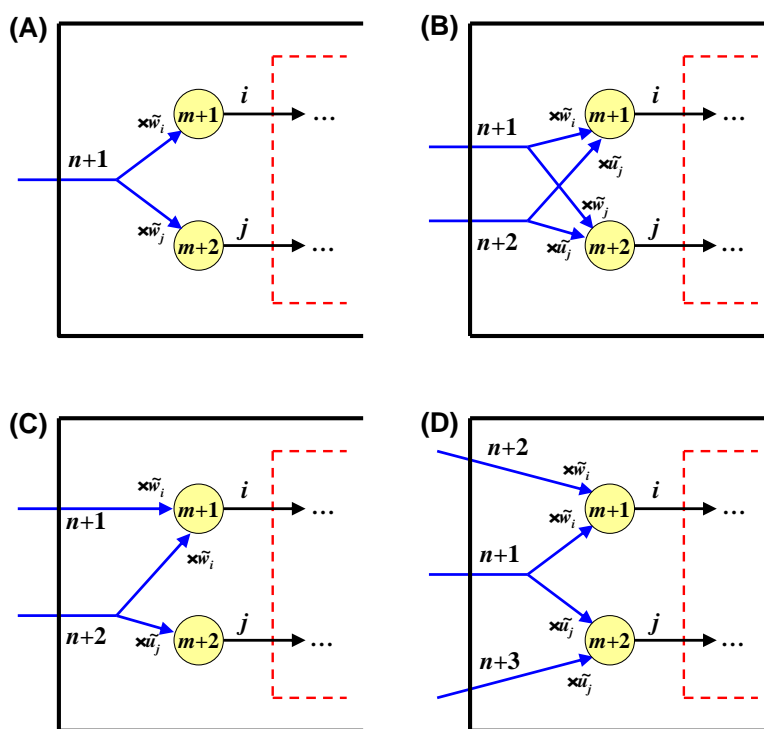


Figure A.7: (A-D): Schematic representations for understanding parts 1-4 of Theorem 1. The dashed line represents the boundary of the subsystem, while the solid line represents the boundary of the extension. In each case,  $i$  and  $j$  are uptake reactions in the subsystem, while they become internal reactions in the extension network.

## Proof of Theorem 1

Let  $C^*$  denote the flux cone of  $N^*$ . We extend the network by adding two distinct external metabolites  $m + 1$  and  $m + 2$  consumed by reactions  $i$  and  $j$ , respectively.

1. Since  $i$  and  $j$  are unblocked in  $N^*$ , there exists  $\tilde{u}, \tilde{v} \in C^*$  such that  $\tilde{u}_i > 0$  and  $\tilde{v}_j > 0$ . Let  $\tilde{w} = \tilde{u} + \tilde{v}$ . Since  $i$  and  $j$  are irreversible, we get  $\tilde{w}_i, \tilde{w}_j > 0$ . We construct an extension  $N$  by adding an irreversible reaction  $n + 1$  that produces  $m + 1$  and  $m + 2$  (see Fig. A.7A). The flux cone of  $N$  is defined as

$$C = \{v \in \mathbb{R}^{n+1} \mid S \cdot v = 0, v_i \geq 0, \text{ for all } i \in Irr\} \quad (\text{A.1})$$

where  $Irr = Irr^* \cup \{n + 1\}$ , and

$$S = \begin{pmatrix} S^* & \mathbf{0} \\ -\mathbf{e}_i^T & \tilde{w}_i \\ -\mathbf{e}_j^T & \tilde{w}_j \end{pmatrix}.$$

Note that  $\begin{pmatrix} \tilde{w} \\ 1 \end{pmatrix} \in C$ . Hence,  $i, j$  and  $n + 1$  are unblocked reactions in  $N$ . Additionally,  $Sv = 0$  implies that  $-v_i + \tilde{w}_i v_{n+1} = 0$  and  $-v_j + \tilde{w}_j v_{n+1} = 0$ . Thus, for all  $v \in C$ ,  $v_j > 0$  implies  $v_i > 0$  (and  $v_{n+1} > 0$ ) and vice versa, with  $v_i/v_j = (\tilde{w}_i v_{n+1})/(\tilde{w}_j v_{n+1}) \stackrel{\text{def}}{=} \lambda$  a constant. Therefore,  $i \iff j$ .

2. If  $i$  and  $j$  are unblocked and not fully coupled in  $N^*$ , there exist flux vectors  $\tilde{w}, \tilde{u} \in C^*$  with  $\tilde{w}_i, \tilde{w}_j, \tilde{u}_i, \tilde{u}_j > 0$  such that  $\tilde{w}_i/\tilde{w}_j \neq \tilde{u}_i/\tilde{u}_j$ . We construct an extension  $N$  by adding two irreversible reactions  $n + 1$  and  $n + 2$  which both produce  $m + 1$  and  $m + 2$ , but with different stoichiometric coefficients (see Fig. A.7B). The flux cone of  $N$  is defined as:

$$C = \{v \in \mathbb{R}^{n+2} \mid S \cdot v = 0, v_i \geq 0, \text{ for all } i \in Irr\} \quad (\text{A.2})$$

where  $Irr = Irr^* \cup \{n + 1, n + 2\}$ , and

$$S = \begin{pmatrix} S^* & \mathbf{0} & \mathbf{0} \\ -\mathbf{e}_i^T & \tilde{w}_i & \tilde{u}_i \\ -\mathbf{e}_j^T & \tilde{w}_j & \tilde{u}_j \end{pmatrix}.$$

First, note that  $\begin{pmatrix} \tilde{w} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \tilde{u} \\ 0 \\ 1 \end{pmatrix} \in C$ . Thus  $i, j, n+1$  and  $n+2$  are unblocked in  $N$ . Additionally, since  $\tilde{w}_i/\tilde{w}_j \neq \tilde{u}_i/\tilde{u}_j$ ,  $i$  and  $j$  are not fully coupled in  $N$ . Now we prove that for all  $v \in C$ ,  $v_i = 0$  implies  $v_j = 0$ . Since  $Sv = 0$ , we have  $-v_i + v_{n+1}\tilde{w}_i + v_{n+2}\tilde{u}_i = 0$  and  $-v_j + v_{n+1}\tilde{w}_j + v_{n+2}\tilde{u}_j = 0$ . If  $v_i = 0$ , we have  $v_{n+1}\tilde{w}_i + v_{n+2}\tilde{u}_i = 0$ . However,  $\tilde{w}_i, \tilde{u}_i > 0$  and  $v_{n+1}, v_{n+2} \geq 0$ , which implies  $v_{n+1} = v_{n+2} = 0$ . Hence,  $v_j = 0$ . Similarly, for all  $v \in C$ ,  $v_j = 0$  implies  $v_i = 0$ . Therefore,  $i \longleftrightarrow j$ .

3. If  $i$  and  $j$  are unblocked and uncoupled, there exist  $\tilde{w}, \tilde{u} \in C^*$  such that  $\tilde{w}_i > 0, \tilde{w}_j = 0$  and  $\tilde{u}_i = 0, \tilde{u}_j > 0$ . We construct an extension  $N$  by adding two irreversible reactions  $n+1$  and  $n+2$ . The first reaction only produces  $m+1$ , while the second reaction produces both  $m+1$  and  $m+2$  (see Fig. A.7C). The flux cone of  $N$  is defined as:

$$C = \{v \in \mathbb{R}^{n+2} \mid S \cdot v = 0, v_i \geq 0, \text{ for all } i \in \text{Irr}\} \quad (\text{A.3})$$

where  $\text{Irr} = \text{Irr}^* \cup \{n+1, n+2\}$ , and

$$S = \begin{pmatrix} S^* & \mathbf{0} & \mathbf{0} \\ -\mathbf{e}_i^T & \tilde{w}_i & \tilde{w}_i \\ -\mathbf{e}_j^T & 0 & \tilde{u}_j \end{pmatrix}$$

Note, that  $\begin{pmatrix} \tilde{w} + \tilde{u} \\ 0 \\ 1 \end{pmatrix} \in C$ . Therefore,  $i, j$  and  $n+2$  are unblocked.

Moreover, for all  $v \in C$ , we have  $-v_i + \tilde{w}_i(v_{n+1} + v_{n+2}) = 0$  and  $-v_j + \tilde{u}_j v_{n+2} = 0$ . From these two equations, it can be seen that  $v_j > 0$

implies  $v_i > 0$ . Therefore,  $j \longrightarrow i$ . However, since  $\begin{pmatrix} \tilde{w} \\ 1 \\ 0 \end{pmatrix} \in C$ ,  $i \longrightarrow j$

cannot hold.

4. If  $i$  and  $j$  are unblocked and uncoupled in  $N^*$ , there exist  $\tilde{w}, \tilde{u} \in C^*$  such that  $\tilde{w}_i > 0, \tilde{w}_j = 0$  and  $\tilde{u}_i = 0, \tilde{u}_j > 0$ .

We extend the network by adding three new reactions  $n+1, n+2$  and  $n+3$  to  $N^*$ , as shown in Fig. A.7D. The extended network,  $N$ , has the following flux cone:

$$C = \{v \in \mathbb{R}^{n+3} \mid S \cdot v = 0, v_i \geq 0, \text{ for all } i \in Irr\} \quad (\text{A.4})$$

where  $Irr = Irr^* \cup \{n+1, n+2, n+3\}$ , and

$$S = \begin{pmatrix} S^* & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\mathbf{e}_i^T & \tilde{w}_i & \tilde{w}_i & 0 \\ -\mathbf{e}_j^T & \tilde{u}_j & 0 & \tilde{u}_j \end{pmatrix} \quad (\text{A.5})$$

Note that,  $\begin{pmatrix} \tilde{w} + \tilde{u} \\ 1 \\ 0 \\ 0 \end{pmatrix} \in C$ . As a result, reactions  $i$ ,  $j$  and  $n+1$  are

unblocked. It can be easily seen that for all  $v \in C$ , we have  $-v_i + \tilde{w}_i(v_{n+1} + v_{n+2}) = 0$  and  $-v_j + \tilde{u}_j(v_{n+1} + v_{n+3}) = 0$ . Therefore, for all  $v \in C$ ,  $v_{n+1} > 0$  implies that  $v_i, v_j > 0$ . Thus, there exists an EM,  $f$ ,

with  $f_i, f_j > 0$ . Additionally,  $\begin{pmatrix} \tilde{w} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \tilde{u} \\ 0 \\ 0 \\ 1 \end{pmatrix} \in C$ . Therefore, there exists

two EMs, say  $g, h$ , with  $g_i > 0, g_j = 0$  and  $h_i = 0, h_j > 0$ . We conclude that  $i$  and  $j$  are sometimes coupled.

## Proof of Theorem 2

We denote by  $C^*$  the flux cone of  $N^*$ . We extend the network by adding a new metabolite for each reaction in  $\Lambda = \{r_1, \dots, r_{|\Lambda|}\}$ . The extended network,  $N$ , has the flux cone

$$C = \{v \in \mathbb{R}^n \mid S \cdot v = 0, v_i \geq 0, \text{ for all } i \in Irr^*\}, \quad (\text{A.6})$$

where

$$S = \begin{pmatrix} S^* \\ -\mathbf{e}_{r_1}^T \\ \dots \\ -\mathbf{e}_{r_{|\Lambda|}}^T \end{pmatrix}.$$

Therefore,  $C = C^* \cap \{v \in \mathbb{R}^n \mid v_r = 0 \text{ for all } r \in \Lambda\}$ . Since  $f, g \in C$ ,  $i$  and  $j$  are uncoupled in  $N$ . Suppose there exists an elementary mode  $d$  in  $N$  with  $d_i, d_j \neq 0$ . One easily checks that  $d$  is also an elementary mode in  $N^*$ . By assumption (a), there exists  $r \in \Lambda$  such that  $d_r \neq 0$ . This implies  $d \notin C$ , which is a contradiction. We conclude  $i \overset{M.E.}{\longleftrightarrow} j$ .

### Proof of Theorem 3

The flux cone of the subsystem can be defined as

$$C^* = \{v \in \mathbb{R}^n \mid S^* \cdot v = 0, v_i \geq 0, \text{ for all } i \in Irr^*\} \quad (\text{A.7})$$

Let  $N$  be an extension of  $N^*$  obtained by adding  $q$  metabolites (and possibly,  $r$  reactions). Then,  $N$  has the following flux cone:

$$C = \{v \in \mathbb{R}^{n+r} \mid S \cdot v = 0, v_i \geq 0, \text{ for all } i \in Irr\} \quad (\text{A.8})$$

where

$$S = \begin{pmatrix} S^* & \mathbf{0} \\ M & P \end{pmatrix}$$

with  $M \in \mathbb{R}^{q \times n}$  and  $P \in \mathbb{R}^{q \times r}$ . Clearly, for all  $u \in \mathbb{R}^n$  and  $v \in \mathbb{R}^r$ ,  $\begin{pmatrix} u \\ v \end{pmatrix} \in C$  implies  $u \in C^*$ . Now, we show that the only possible changes in the flux coupling relations are the ones mentioned in the theorem:

- i) If  $i$  and  $j$  are uncoupled in  $N$ , then there exists  $\begin{pmatrix} u \\ v \end{pmatrix} \in C$  such that  $u_i = 0$  and  $u_j \neq 0$ . Also, there exists  $\begin{pmatrix} u' \\ v' \end{pmatrix} \in C$  such that  $u'_i \neq 0$  and  $u'_j = 0$ . Therefore,  $i$  and  $j$  cannot be directionally, partially, or fully coupled in  $N^*$ .
- ii) If  $i \rightarrow j$  (and not  $i \leftrightarrow j$ ) in  $N$ , then there exists  $\begin{pmatrix} u \\ v \end{pmatrix} \in C$  such that  $u_i = 0$  and  $u_j \neq 0$ . Therefore,  $i$  and  $j$  cannot be partially, or fully coupled in  $N^*$ .
- iii) If  $i \leftrightarrow j$  (and not  $i \iff j$ ) in  $N$ , then there exists  $\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} u' \\ v' \end{pmatrix} \in C$  such that  $u_i, u_j, u'_i, u'_j \neq 0$  and  $u_i/u_j \neq u'_i/u'_j$ . Therefore,  $i$  and  $j$  cannot be fully coupled in  $N^*$ .