



# Enzyme allocation problems in kinetic metabolic networks: Optimal solutions are elementary flux modes



Stefan Müller<sup>a,b,\*</sup>, Georg Regensburger<sup>a</sup>, Ralf Steuer<sup>c</sup>

<sup>a</sup> Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstraße 69, 4040 Linz, Austria

<sup>b</sup> CzechGlobe – Global Change Research Center, Academy of Sciences of the Czech Republic, Belidla 986/4a, 603 00 Brno, Czech Republic

<sup>c</sup> Institute for Theoretical Biology, Humboldt University Berlin, Invalidenstraße 43, 10115 Berlin, Germany

## HIGHLIGHTS

- Optimization principles are useful tools to study metabolic network organization.
- A metabolic optimization problem with general kinetics is formulated and studied.
- A rigorous proof is given that optimal solutions are elementary flux modes.
- The proof uses the theory of oriented matroids.
- The result helps to understand metabolic switches and the occurrence of low-yield pathways.

## ARTICLE INFO

### Article history:

Received 26 July 2013

Received in revised form

5 November 2013

Accepted 21 November 2013

Available online 1 December 2013

### Keywords:

Metabolic optimization

Enzyme kinetics

Oriented matroid

Elementary vector

Conformal sum

## ABSTRACT

The survival and proliferation of cells and organisms require a highly coordinated allocation of cellular resources to ensure the efficient synthesis of cellular components. In particular, the total enzymatic capacity for cellular metabolism is limited by finite resources that are shared between all enzymes, such as cytosolic space, energy expenditure for amino-acid synthesis, or micro-nutrients. While extensive work has been done to study constrained optimization problems based only on stoichiometric information, mathematical results that characterize the optimal flux in kinetic metabolic networks are still scarce. Here, we study constrained enzyme allocation problems with general kinetics, using the theory of oriented matroids. We give a rigorous proof for the fact that optimal solutions of the non-linear optimization problem are elementary flux modes. This finding has significant consequences for our understanding of optimality in metabolic networks as well as for the identification of metabolic switches and the computation of optimal flux distributions in kinetic metabolic networks.

© 2013 Elsevier Ltd. All rights reserved.

## 1. Introduction

Living organisms are under constant evolutionary pressure to survive and reproduce in complex environments. As a direct consequence, cellular pathways are often assumed to be highly adapted to their respective tasks, given the biochemical and biophysical constraints of their environment. Optimality principles have proven to be powerful methods to study and understand the large-scale organization of metabolic pathways (Berkhout et al., 2012; Heinrich and Schuster, 1996; Schuetz et al., 2012; Steuer and

Junker, 2008; Molenaar et al., 2009). A variety of recent computational techniques, such as flux-balance analysis (FBA), seek to identify metabolic flux distributions that maximize given objective functions, such as ATP regeneration or biomass yield, under a set of linear constraints. As one of their prime merits, FBA and related stoichiometric methods, including the generalization to time-dependent metabolism (Mahadevan et al., 2002; Antoniewicz, 2013), only require knowledge of the stoichiometry of a metabolic network – data that are available for an increasing number of organisms in the form of large-scale metabolic reconstructions (Oberhardt et al., 2009; Orth et al., 2010).

However, despite their explanatory and predictive success, constraint-based stoichiometric methods also have inherent limits. Specifically, FBA and related methods typically maximize stoichiometric yield. That is, the value of a designated output flux is maximized, given a set of limiting input fluxes. As emphasized in a number of recent studies, the assumption of maximal

\* Corresponding author at: Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstraße 69, 4040 Linz, Austria. Tel.: +43 1 51581 6553.

E-mail addresses: [stefan.mueller@ricam.oeaw.ac.at](mailto:stefan.mueller@ricam.oeaw.ac.at) (S. Müller), [georg.regensburger@ricam.oeaw.ac.at](mailto:georg.regensburger@ricam.oeaw.ac.at) (G. Regensburger), [ralf.steuer@hu-berlin.de](mailto:ralf.steuer@hu-berlin.de) (R. Steuer).

stoichiometric yield is not necessarily a universal principle of metabolic network function (Schuster et al., 2008, 2011; Goel et al., 2012; Molenaar et al., 2009). Quite on the contrary, examples of seemingly suboptimal metabolic behavior, at least from a stoichiometric perspective, are well-known for many decades. Among the most prominent instances are the Warburg and the Crabtree effect (Warburg et al., 1924; Crabtree, 1928; Hsu and Sabatini, 2008). Under certain circumstances, cells utilize a fermentative metabolism rather than aerobic respiration to regenerate ATP, despite the presence of oxygen and despite its significantly lower stoichiometric yield of ATP per amount of glucose consumed.

To account for such seemingly suboptimal behavior, several modifications and extensions of FBA have been developed recently. Conventional FBA is augmented with additional principles concerning limited cytosolic volume (Beg et al., 2007; Vazquez et al., 2008, 2010; Vazquez and Oltvai, 2011; Shlomi et al., 2011), membrane occupancy (Zhuang et al., 2011), and other, more general capacity constraints (Schuster et al., 2011; Goelzer et al., 2011). Each of these extensions allows for new insights into suboptimal stoichiometric behavior, and additional constraints often also induce the utilization of pathways with lower stoichiometric yield. However, none of the modifications of FBA addresses a metabolic network as a genuine dynamical system with particular kinetics that depend on a number of parameters. The neglect of the dynamical nature is a direct consequence of the extensive data requirements for parametrizing enzymatic reaction rates. Correspondingly, and despite its importance to understand metabolic optimality, only few mathematically rigorous results are currently available that allow to characterize solutions of constrained non-linear optimization problems arising from kinetic metabolic networks.

In this work, we formulate and study constrained enzyme allocation problems in metabolic networks with general kinetics. In particular, we are interested in enzyme distributions that maximize a designated output flux, given a limited total enzymatic capacity. We show that the optimal distributions of metabolic fluxes differ from solutions obtained by FBA and related stoichiometric methods. Most importantly, we give a rigorous proof for the fact that optimal flux distributions are elementary flux modes. Therein, we make use of results from the theory of oriented matroids that were hitherto only scarcely applied to metabolic networks (Beard et al., 2004), but offer great potential to unify and advance metabolic network analysis, as mentioned in Gagneur and Klamt (2004) and Müller and Bockmayr (2013). Our finding has significant consequences for the understanding of metabolic optimality as well as for the efficient computation of optimal fluxes in kinetic metabolic networks.

The paper is organized as follows. In Section 2, we introduce kinetic metabolic networks and state the enzyme allocation problem of interest. In Section 3, we illustrate our mathematical results and the ideas underlying our proofs by a conceptual example of a minimal metabolic network. In Section 4, we address the connection between metabolic network analysis and the theory of oriented matroids. In particular, we reformulate the optimization problem and show that, if the enzyme allocation problem has an optimal solution, then there is an optimal solution which is an elementary flux mode. Finally, we provide a discussion of our results in the context of metabolic optimization problems.

## 2. Problem statement

After introducing the necessary mathematical notation, we define kinetic metabolic networks and elementary flux modes, and state the metabolic optimization problem that we investigate in the following.

### 2.1. Mathematical notation

We denote the positive real numbers by  $\mathbb{R}_{>}$  and the non-negative real numbers by  $\mathbb{R}_{\geq}$ . For a finite index set  $I$ , we write  $\mathbb{R}^I$  for the real vector space of vectors  $x = (x_i)_{i \in I}$  with  $x_i \in \mathbb{R}$ , and  $\mathbb{R}_{>}^I$  and  $\mathbb{R}_{\geq}^I$  for the corresponding subsets. Given  $x \in \mathbb{R}^I$ , we write  $x > 0$  if  $x \in \mathbb{R}_{>}^I$  and  $x \geq 0$  if  $x \in \mathbb{R}_{\geq}^I$ . We denote the support of a vector  $x \in \mathbb{R}^I$  by  $\text{supp}(x) = \{i \in I \mid x_i \neq 0\}$ . For  $x, y \in \mathbb{R}^I$ , we denote the component-wise (or Hadamard) product by  $x \circ y \in \mathbb{R}^I$ , that is,  $(x \circ y)_i = x_i y_i$ .

### 2.2. Kinetic metabolic networks

A metabolic network  $(S, \mathcal{R}, N)$  consists of a set  $S$  of internal metabolites, a set  $\mathcal{R}$  of reactions, and the stoichiometric matrix  $N \in \mathbb{R}^{S \times \mathcal{R}}$ , which contains the net stoichiometric coefficients for each metabolite  $s \in S$  in each reaction  $r \in \mathcal{R}$ . The set of reactions is the disjoint union of the sets of reversible and irreversible reactions,  $\mathcal{R}_{\leftrightarrow}$  and  $\mathcal{R}_{\rightarrow}$ , respectively.

In the following, we assume that each reaction can be catalyzed by an enzyme. Let  $x \in \mathbb{R}_{\geq}^S$  denote the vector of metabolite concentrations,  $c \in \mathbb{R}_{\geq}^{\mathcal{R}}$  the vector of enzyme concentrations, and  $p \in \mathbb{R}^P$  a vector of parameters such as turnover numbers, equilibrium constants, and Michaelis–Menten constants. We write the vector of rate functions  $v : \mathbb{R}_{\geq}^S \times \mathbb{R}_{\geq}^{\mathcal{R}} \times \mathbb{R}^P \rightarrow \mathbb{R}^{\mathcal{R}}$  as

$$v(x; c, p) = c \circ \kappa(x, p)$$

with a function  $\kappa : \mathbb{R}_{\geq}^S \times \mathbb{R}^P \rightarrow \mathbb{R}^{\mathcal{R}}$ . In other words, each reaction rate  $v_r$  is the product of the corresponding enzyme concentration  $c_r$  with a particular kinetics  $\kappa_r$ .

A kinetic metabolic network  $(S, \mathcal{R}, N, v)$  is a metabolic network  $(S, \mathcal{R}, N)$  together with rate functions  $v$  as defined above. The dynamics of  $(S, \mathcal{R}, N, v)$  is determined by the ODEs

$$\frac{dx}{dt} = N v(x; c, p).$$

A steady state  $\bar{x} \in \mathbb{R}_{\geq}^S$  and the corresponding steady-state flux  $\bar{v} = v(\bar{x}; c, p) \in \mathbb{R}^{\mathcal{R}}$  are determined by

$$0 = N \bar{v}.$$

### 2.3. Elementary flux modes

A flux mode is a non-zero steady-state flux  $f \in \mathbb{R}^{\mathcal{R}}$  with non-negative components for all irreversible reactions. In other words, a flux mode is a non-zero element of the flux cone

$$C = \{f \in \mathbb{R}^{\mathcal{R}} \mid Nf = 0 \text{ and } f_r \geq 0 \text{ for all } r \in \mathcal{R}_{\rightarrow}\}.$$

An elementary flux mode (EFM)  $e \in \mathbb{R}^{\mathcal{R}}$  is a flux mode with minimal support:

$$\begin{aligned} f \in C \text{ with } f \neq 0 \text{ and } \text{supp}(f) \subseteq \text{supp}(e) \\ \Rightarrow \text{supp}(f) = \text{supp}(e). \end{aligned} \quad (\text{efm})$$

In fact,  $\text{supp}(f) = \text{supp}(e)$  further implies  $f = \lambda e$  with  $\lambda > 0$ . Otherwise, one can construct another flux mode from  $e$  and  $f$  with smaller support. As a consequence, there can be only finitely many EFMs (up to multiplication with a positive scalar). For references on elementary flux modes and related computational issues, see (Schuster and Hilgetag, 1994; Schuster et al., 2002; Klamt and Stelling, 2003; Gagneur and Klamt, 2004; Wagner and Urbanczik, 2005; Larhlami and Bockmayr, 2009).

### 2.4. Enzyme allocation problem

We are now in a position to state the optimization problem that we study in this work.

Let  $(S, \mathcal{R}, N, v)$  be a kinetic metabolic network with flux cone  $C$ . Fix a reaction  $r_* \in \mathcal{R}$ , positive weights  $w \in \mathbb{R}_{>0}^S$ , a subset  $X \subseteq \mathbb{R}_{>0}^S$ , and parameters  $p \in \mathbb{R}^p$ . Maximize the component  $\bar{v}_{r_*} > 0$  of the steady-state flux  $\bar{v} = c \circ \kappa(\bar{x}, p)$  by varying the steady-state metabolite concentrations  $\bar{x} \in X$  and the enzyme concentrations  $c \in \mathbb{R}_{>0}^{\mathcal{R}}$ . Thereby, fix the weighted sum of enzyme concentrations and require the steady-state flux to be a flux mode:

$$\max_{\bar{x} \in X, c \in \mathbb{R}_{>0}^{\mathcal{R}}} \bar{v}_{r_*} \quad (1a)$$

subject to

$$\sum_{r \in \mathcal{R}} w_r c_r = c^{\text{tot}}, \quad (1b)$$

$$\bar{v} \in C, \bar{v}_{r_*} > 0. \quad (1c)$$

Note that the constraint  $\bar{v} \in C$  implies  $0 = N\bar{v} = N(c \circ \kappa(\bar{x}, p))$  which is non-linear, in general.

Further, note that the enzyme allocation problem may be unfeasible, in particular, the constraint (1c) may be unsatisfiable. In this case, the flux cone and the kinetics are “incompatible”. Moreover, even if the problem is feasible, the maximum may not be attained at finite metabolite concentrations. On the other hand, if the problem is feasible, the kinetics is continuous in  $\bar{x}$ , and  $X$  is compact (bounded and closed), then the maximum is attained.

Problem (1) defines a very general metabolic optimization problem, in which the set of enzyme concentrations is adjusted in order to maximize a specific metabolic flux within the steady-state flux vector. In this respect, the weighted sum (1b) may encode different enzymatic constraints, such as limited cellular or membrane surface space, limited nitrogen or transition metal availability, as well as other constraints for the abundance of certain enzymes. In each case, the weight factors denote the fraction of the resource used per unit enzyme. Likewise, the flux component  $\bar{v}_{r_*}$  may stand for diverse metabolic processes, ranging from the synthesis rate of a particular product within a specific pathway to the rate of overall cellular growth. The optimization problem seeks to identify the maximum value of  $\bar{v}_{r_*}$ , the associated enzyme and steady-state metabolite concentrations,  $c$  and  $\bar{x}$ , as well as the corresponding flux  $\bar{v}$ .

### 3. A conceptual example

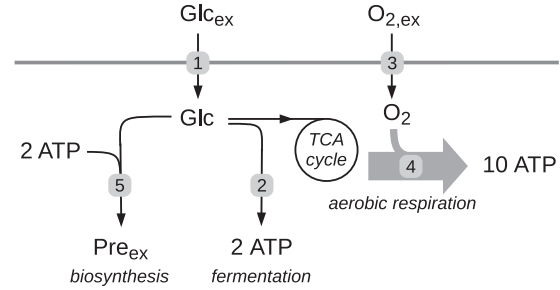
In order to illustrate our mathematical results and the ideas in its proofs, we first study a minimal metabolic network for the production of a precursor molecule from glucose. In particular, we consider two alternative pathways: fermentation (low yield) and respiration (high yield), cf. Fig. 1. The actual metabolic network  $(S, \mathcal{R}, N)$  is further simplified and involves the internal metabolites  $S = \{\text{Glc}, \text{O}_2, \text{ATP}\}$ , the set of reactions  $\mathcal{R}$  consisting of

- 1:  $\text{Glc}_{\text{ex}} \rightleftharpoons \text{Glc}$
- 2:  $\text{Glc} \rightleftharpoons 2 \text{ ATP}$
- 3:  $\text{O}_{2,\text{ex}} \rightleftharpoons \text{O}_2$
- 4:  $\text{Glc} + \text{O}_2 \rightleftharpoons 10 \text{ ATP}$
- 5:  $\text{Glc} + 2 \text{ ATP} \rightarrow \text{Pre}_{\text{ex}}$ ,

and the resulting stoichiometric matrix

$$N = \begin{array}{c} \text{Glc} \\ \text{O}_2 \\ \text{ATP} \end{array} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ +1 & -1 & 0 & -1 & -1 \\ 0 & 0 & +1 & -1 & 0 \\ 0 & +2 & 0 & +10 & -2 \end{pmatrix}.$$

The external substrates/products  $\text{Glc}_{\text{ex}}, \text{O}_{2,\text{ex}}, \text{Pre}_{\text{ex}}$  do not appear in  $(S, \mathcal{R}, N)$ , but their constant concentrations can enter the rate functions as parameters. In a kinetic metabolic network



**Fig. 1.** A minimal metabolic network. The pathway consists of two uptake reactions (1,3), two intracellular conversions (2,4), and the formation of a precursor molecule (5). The resulting EFMs are  $e^1 = (2, 1, 0, 0, 1)$  and  $e^2 = (\frac{6}{5}, 0, \frac{1}{5}, \frac{1}{5}, 1)$  with overall reactions  $2 \text{ Glc}_{\text{ex}} \rightarrow 1 \text{ Pre}_{\text{ex}}$  (fermentation) and  $6 \text{ Glc}_{\text{ex}} + 1 \text{ O}_{2,\text{ex}} \rightarrow 5 \text{ Pre}_{\text{ex}}$  (respiration).

$(S, \mathcal{R}, N, v)$ , the rate of reaction  $r$  is given as

$$v_r = c_r K_r,$$

that is, as a product of the corresponding enzyme concentration  $c_r$  and a particular kinetics  $K_r$ . In this example, we use the kinetics

$$\begin{aligned} K_1 &= k_1([\text{Glc}_{\text{ex}}] - K_1[\text{Glc}]) \\ K_2 &= k_2([\text{Glc}] - K_2[\text{ATP}]) \\ K_3 &= k_3([\text{O}_{2,\text{ex}}] - K_3[\text{O}_2]) \\ K_4 &= k_4([\text{Glc}][\text{O}_2] - K_4[\text{ATP}]) \\ K_5 &= k_5[\text{Glc}][\text{ATP}], \end{aligned}$$

however, our main result does not depend on the kinetics. In vector notation, we write

$$v(x; c, p) = c \circ \kappa(x, p),$$

thereby introducing the concentrations  $x$  of the internal metabolites and the parameters  $p$ ,

$$\begin{aligned} x &= ([\text{Glc}], [\text{O}_2], [\text{ATP}])^T, \\ p &= ([\text{Glc}_{\text{ex}}], [\text{O}_{2,\text{ex}}]; k_1, k_2, k_3, k_4, k_5; K_1, K_2, K_3, K_4)^T. \end{aligned}$$

The dynamics of the network is governed by the ODEs

$$\frac{dx}{dt} = N v(x; c, p).$$

A steady state  $\bar{x}$  and the corresponding steady-state flux  $\bar{v} = v(\bar{x}; c, p)$  are determined by

$$0 = N\bar{v}.$$

The goal is to maximize the production rate of the precursor  $\text{Pre}_{\text{ex}}$ , that is, the component  $\bar{v}_5$  of the steady-state flux  $\bar{v} = c \circ \kappa(\bar{x}, p)$ , by varying the steady-state metabolite concentrations  $\bar{x} \in \mathbb{R}_{>0}^3$  and the enzyme concentrations  $c \in \mathbb{R}_{>0}^5$ . Thereby, the (weighted) sum of enzyme concentrations is fixed and the steady-state flux must be a flux mode:

$$\max_{\bar{x}, c} \bar{v}_5 \quad (2a)$$

subject to

$$\sum_{r=1}^5 c_r = c^{\text{tot}}, \quad (2b)$$

$$N\bar{v} = 0, \bar{v}_5 > 0. \quad (2c)$$

For convenience, we use equal weights in the sum constraint.

To simplify the problem, we consider a restriction on the steady-state metabolite concentrations  $\bar{x}$ , in particular, we require  $\kappa(\bar{x}, p) > 0$ . In chemical terms, we assume the thermodynamic feasibility of a situation where all reactions can proceed from left to right. Since  $\bar{v} = c \circ \kappa$ , this implies  $\bar{v} \geq 0$ . That is, every component of the steady-state flux is non-negative; in fact, it is zero if and only if the corresponding enzyme concentration is zero.

Every feasible steady-state flux  $\bar{v}$  is a flux mode. In particular,  $\bar{v} = \bar{v}_5 f$  where  $f$  is a flux mode with  $f \geq 0$  and  $f_5 = 1$ . From  $\bar{v} = c \circ \kappa = \bar{v}_5 f$ , we further obtain

$$c_r = \bar{v}_5 \frac{f_r}{\kappa_r}.$$

Now, we can rewrite the constraint on the enzyme concentrations as

$$c^{\text{tot}} = \sum_{r=1}^5 c_r = \bar{v}_5 \sum_{r=1}^5 \frac{f_r}{\kappa_r}.$$

Instead of maximizing  $\bar{v}_5$ , we can minimize  $c^{\text{tot}}/\bar{v}_5 = \sum_{r=1}^5 f_r/\kappa_r(\bar{x}, p)$  by varying the steady-state metabolite concentrations  $\bar{x}$  and the flux mode  $f$ . Hence, the enzyme allocation problem (2) with the restriction  $\kappa(\bar{x}, p) > 0$  is equivalent to

$$\min_{\bar{x}, f} \sum_{r=1}^5 \frac{f_r}{\kappa_r(\bar{x}, p)} \quad (3a)$$

subject to

$$\kappa(\bar{x}, p) > 0, f \geq 0, \quad (3b)$$

$$Nf = 0, f_5 = 1. \quad (3c)$$

As shown in Section 4.2, every flux mode  $f \geq 0$  is a non-negative linear combination of elementary flux modes (EFMs)  $e \geq 0$ . In fact, there are two such EFMs,

$$e^1 = (2, 1, 0, 0, 1)^T \\ e^2 = (\frac{6}{5}, 0, \frac{1}{5}, \frac{1}{5}, 1)^T,$$

representing fermentation and respiration, and hence

$$f = \alpha_1 e^1 + \alpha_2 e^2 \quad \text{with} \quad \alpha_1, \alpha_2 \geq 0.$$

Note that we have scaled the EFMs  $e$  such that  $e_5 = 1$ . The condition  $f_5 = 1$  implies  $\alpha_1 + \alpha_2 = 1$ . As a result, we obtain another equivalent formulation of the restricted enzyme allocation problem:

$$\min_{\bar{x}, \alpha_1, \alpha_2} \sum_{r=1}^5 \frac{\alpha_1 e_r^1 + \alpha_2 e_r^2}{\kappa_r(\bar{x}, p)} \quad (4a)$$

subject to

$$\kappa(\bar{x}, p) > 0, \quad (4b)$$

$$\alpha_1 + \alpha_2 = 1. \quad (4c)$$

We observe that the objective function is linear in  $\alpha_1$  and  $\alpha_2$ :

$$\alpha_1 \underbrace{\sum_{r=1}^5 \frac{e_r^1}{\kappa_r(\bar{x}, p)}}_{g_1} + \alpha_2 \underbrace{\sum_{r=1}^5 \frac{e_r^2}{\kappa_r(\bar{x}, p)}}_{g_2}$$

with

$$g_1(\bar{x}, p) = \frac{2}{\kappa_1(\bar{x}, p)} + \frac{1}{\kappa_2(\bar{x}, p)} + \frac{1}{\kappa_5(\bar{x}, p)} \\ g_2(\bar{x}, p) = \frac{\frac{6}{5}}{\kappa_1(\bar{x}, p)} + \frac{\frac{1}{5}}{\kappa_3(\bar{x}, p)} + \frac{\frac{1}{5}}{\kappa_4(\bar{x}, p)} + \frac{1}{\kappa_5(\bar{x}, p)}.$$

Clearly,  $g_1(\bar{x}, p) > 0$  and  $g_2(\bar{x}, p) > 0$ , since  $\kappa(\bar{x}, p) > 0$ . Assume that the minima of  $g_1$  and  $g_2$  are attained at  $\bar{x}^1$  and  $\bar{x}^2$ , respectively. That is,  $\hat{g}_1 = \min_{\bar{x}} g_1(\bar{x}, p) = g_1(\bar{x}^1, p)$  and  $\hat{g}_2 = \min_{\bar{x}} g_2(\bar{x}, p) = g_2(\bar{x}^2, p)$ . If  $\hat{g}_1 < \hat{g}_2$ , then the objective function attains its minimum  $\hat{g}_1$  at  $\bar{x} = \bar{x}^1$ ,  $\alpha_1 = 1$  and  $\alpha_2 = 0$ , that is, for  $f = e^1$ . To see this, assume  $\alpha_2 > 0$ ; then, for all  $\bar{x}$ ,

$$\alpha_1 g_1(\bar{x}, p) + \alpha_2 g_2(\bar{x}, p) \geq \alpha_1 \hat{g}_1 + \alpha_2 \hat{g}_2 > \alpha_1 \hat{g}_1 + \alpha_2 \hat{g}_1 = (\alpha_1 + \alpha_2) \hat{g}_1 = \hat{g}_1.$$

Conversely, if  $\hat{g}_1 > \hat{g}_2$ , the minimum is attained for  $f = e^2$ , and finally, if  $\hat{g}_1 = \hat{g}_2$ , both  $f = e^1$  and  $f = e^2$  are optimal. In the degenerate case where  $\hat{g}_1 = g_1(\bar{x}^0, p) = g_2(\bar{x}^0, p) = \hat{g}_2$  at the same minimum point  $\bar{x}^0$ , any  $f = \alpha_1 e^1 + \alpha_2 e^2$  (with  $\alpha_1, \alpha_2 \geq 0$  and  $\alpha_1 + \alpha_2 = 1$ ) is optimal.

We can summarize our result as follows: generically, the steady-state flux  $\bar{v}$  related to an optimal solution of the restricted enzyme allocation problem (3) is an EFM. The same holds for all appropriate restrictions and hence for the full enzyme allocation problem (2).

For variable external substrate concentrations  $[\text{Glc}_{\text{ex}}]$  and  $[\text{O}_{2,\text{ex}}]$ , we are interested in which EFM is optimal and when a switch between EFMs occurs. To this end, we determine the optimal solution for each EFM. The optimization problem restricted to EFM  $e^1$  is equivalent to

$$\min_{\bar{x}} g_1(\bar{x}, p)$$

subject to

$$\kappa_1(\bar{x}, p) > 0, \quad \kappa_2(\bar{x}, p) > 0, \quad \kappa_5(\bar{x}, p) > 0.$$

In EFM  $e^1$ , reactions 3 and 4 do not carry any flux, that is,  $e_3^1 = e_4^1 = 0$ . Hence, the corresponding enzyme concentrations are zero, that is,  $c_3 = c_4 = 0$ , and there are no constraints involving  $\kappa_3(\bar{x}, p)$  and  $\kappa_4(\bar{x}, p)$ . From the optimal metabolite concentrations  $\bar{x}$ , we determine the optimal enzyme concentrations  $c_1$ ,  $c_2$ , and  $c_5$  as

$$c_r = \bar{v}_5 \frac{e_r^1}{\kappa_r} = c^{\text{tot}} \frac{\frac{e_r^1}{\kappa_r(\bar{x}, p)}}{\sum_{r=1,2,5} \frac{e_r^1}{\kappa_r(\bar{x}, p)}}.$$

Explicitly, the optimization problem for EFM  $e^1$  amounts to

$$\min_{[\text{Glc}], [\text{ATP}]} \left( \frac{2}{k_1([\text{Glc}_{\text{ex}}] - K_1[\text{Glc}])} + \frac{1}{k_2([\text{Glc}] - K_2[\text{ATP}])} + \frac{1}{k_5[\text{Glc}][\text{ATP}]} \right)$$

subject to

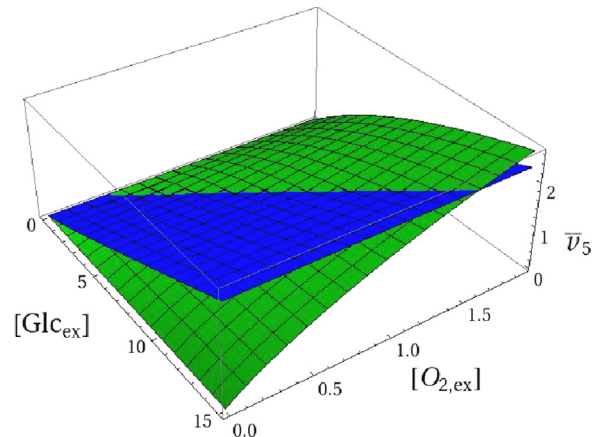
$$[\text{Glc}_{\text{ex}}] - K_1[\text{Glc}] > 0, \quad [\text{Glc}] - K_2[\text{ATP}] > 0, \quad [\text{Glc}][\text{ATP}] > 0,$$

where we omit the bar over the steady-state metabolite concentrations. From the optimal metabolite concentrations  $[\text{Glc}]$ ,  $[\text{ATP}]$ , we determine the optimal enzyme concentrations  $c_1$ ,  $c_2$ , and  $c_5$ . For example,

$$c_1 = c^{\text{tot}} \frac{\frac{2}{k_1([\text{Glc}_{\text{ex}}] - K_1[\text{Glc}])}}{\frac{2}{k_1([\text{Glc}_{\text{ex}}] - K_1[\text{Glc}])} + \frac{1}{k_2([\text{Glc}] - K_2[\text{ATP}])} + \frac{1}{k_5[\text{Glc}][\text{ATP}]}}.$$

The optimization problem restricted to EFM  $e^2$  is treated analogously.

Finally, we vary  $[\text{Glc}_{\text{ex}}]$  and  $[\text{O}_{2,\text{ex}}]$ , solve the restricted optimization problems for EFMs  $e^1$  and  $e^2$ , and compare the resulting



**Fig. 2.** Maximum production rate  $\bar{v}_5$  as a function of external substrate concentrations  $[\text{Glc}_{\text{ex}}]$  and  $[\text{O}_{2,\text{ex}}]$  plotted for EFM  $e^1$  (fermentation, blue) and EFM  $e^2$  (respiration, green). Parameters:  $k_1 = k_2 = k_3 = k_4 = k_5 = 1$ ,  $K_1 = K_3 = 1$ ,  $K_2 = K_4 = 0.1$ . (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)



maximum values of  $\bar{v}_5$ , cf. Fig. 2. Clearly, the optimal solution of the enzyme allocation problem switches between EFM  $e^1$  and  $e^2$  which involves a discontinuous change of enzyme and metabolite concentrations, cf. Fig. 3, where we fix  $[O_{2,ex}] = 1$  and vary  $[Glc_{ex}]$ .

#### 4. Mathematical results

We reformulate the enzyme allocation problem (1) and characterize its solutions. To this end, we employ concepts from the theory of oriented matroids like elementary vectors, sign vectors, and conformal sums.

Realizable oriented matroids arise from vector subspaces. Essentially, a realizable oriented matroid is the set of sign vectors of a subspace or, equivalently, all sign vectors with minimal support. Abstract oriented matroids can be characterized by axiom systems for (co-)vectors (satisfied by the sign vectors of a subspace), (co-)circuits (satisfied by the sign vectors with minimal support), or, equivalently, chirotopes. For an introduction to oriented matroids, we refer to the survey (Richter-Gebert and Ziegler, 1997), the textbooks (Bachem and Kern, 1992) and (Ziegler, 1995, Chapters 6 and 7), and the encyclopedic treatment (Björner et al., 1999).

In applications to metabolic network analysis, the involved oriented matroids are realizable. For example, the sign vector of a thermodynamically feasible steady-state flux must be orthogonal to all (internal) circuits (Beard et al., 2004). In the original proof, the circuit axioms for oriented matroids are used explicitly; however, the result also follows from basic facts about the orthogonality of sign vectors of subspaces (Ziegler, 1995, Chapter 6); alternatively, it can be proved using linear programming duality (Müller, 2012; Noor et al., 2012). We note that oriented matroids also appear in the study of directed hypergraph and Petri net models of biochemical reactions (Oliveira et al., 2001) and in the theory of chemical reaction networks with generalized mass action kinetics (Müller and Regensburger, 2012).

##### 4.1. Elementary vectors

An elementary vector (EV)  $e \in \mathbb{R}^{\mathcal{R}}$  of a vector subspace  $S \subseteq \mathbb{R}^{\mathcal{R}}$  is a non-zero vector with minimal support (Rockafellar, 1969):

$$\begin{aligned} f \in S \text{ with } f \neq 0 \text{ and } \text{supp}(f) \subseteq \text{supp}(e) \\ \Rightarrow \text{supp}(f) = \text{supp}(e). \end{aligned} \quad (\text{ev})$$

It is easy to see that EFM are exactly those EVs of  $\ker(N)$  that are flux modes. To our knowledge, this fact has not been clarified before.

**Lemma 1.** Let  $(S, \mathcal{R}, N)$  be a metabolic network and  $e \in \mathbb{R}^{\mathcal{R}}$ . The following statements are equivalent:

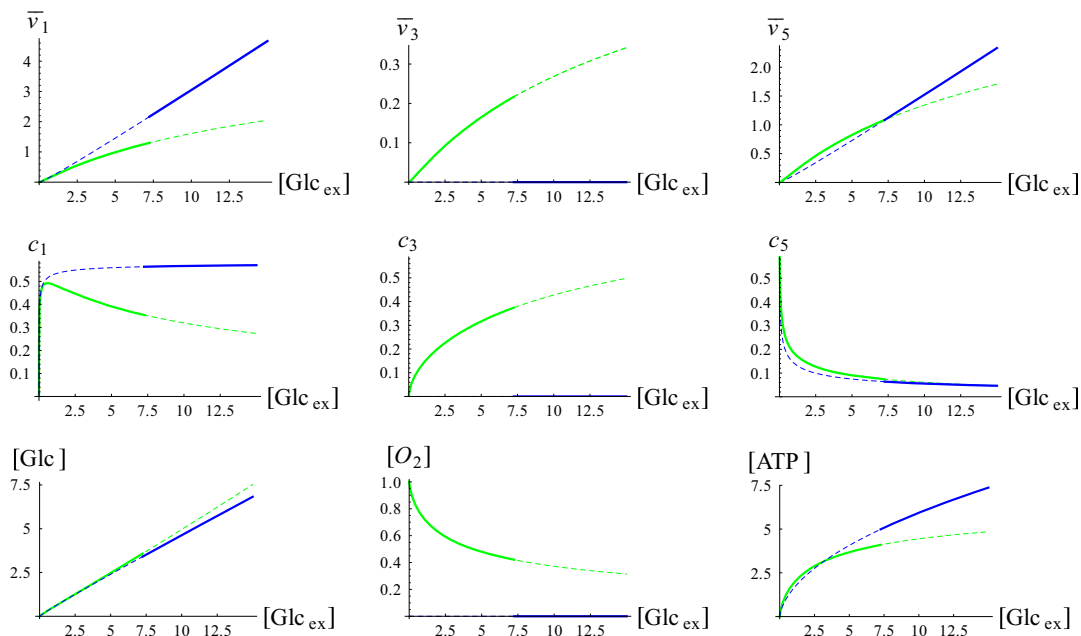
- (i)  $e$  is an EFM.
- (ii)  $e$  is an EV of  $\ker(N)$  and a flux mode.

**Proof.** (i)  $\Rightarrow$  (ii): We have to show that EFM  $e$  is an EV of  $\ker(N)$ . Suppose (ev) is violated, that is, there exists  $f \in \ker(N)$  with  $f \neq 0$  and  $\text{supp}(f) \subset \text{supp}(e)$ . If  $f_r \geq 0$  for all  $r \in \mathcal{R}_+$ , then  $f \in C$  in contradiction to (efm). Otherwise, consider  $f' = e + \lambda f$  with the largest scalar  $\lambda > 0$  such that  $f'_r \geq 0$  for all  $r \in \mathcal{R}_+$ . Then,  $f' \in C$  with  $f' \neq 0$  and  $\text{supp}(f') \subset \text{supp}(e)$  in contradiction to (efm). (ii)  $\Rightarrow$  (i): Let  $e$  be a flux mode and an EV of  $\ker(N)$ . Clearly, (ev) implies (efm), since  $f \in C$  implies  $f \in \ker(N)$ .  $\square$

##### 4.2. Sign vectors and conformal sums

We define the sign vector  $\sigma(x) \in \{-, 0, +\}^I$  of a vector  $x \in \mathbb{R}^I$  by applying the sign function component-wise. The relations  $0 < -$  and  $0 < +$  induce a partial order on  $\{-, 0, +\}^I$ : we write  $X \leq Y$  for  $X, Y \in \{-, 0, +\}^I$ , if the inequality holds component-wise. For  $x, y \in \mathbb{R}^I$ , we say that  $x$  conforms to  $y$ , if  $\sigma(x) \leq \sigma(y)$ . Analogously, for  $x \in \mathbb{R}^I$  and  $X \in \{-, 0, +\}^I$ , we say that  $x$  conforms to  $X$ , if  $\sigma(x) \leq X$ .

The following fundamental result about vectors and EVs will be rephrased for flux modes and EFM. For a proof, see (Rockafellar,



**Fig. 3.** Optimal fluxes  $\bar{v}_1, \bar{v}_3, \bar{v}_5$ , enzyme concentrations  $c_1, c_3, c_5$ , and internal metabolite concentrations  $[Glc], [O_2], [ATP]$  as functions of external substrate concentration  $[Glc_{ex}]$  plotted for EFM  $e^1$  (fermentation, blue) and EFM  $e^2$  (respiration, green). With increasing  $[Glc_{ex}]$ , the optimal solution switches from respiration (thick green lines) to fermentation (thick blue lines). Parameters:  $[O_{2,ex}] = 1$ ,  $k_1 = k_2 = k_3 = k_4 = k_5 = 1$ ,  $K_1 = K_3 = 1$ ,  $K_2 = K_4 = 0.1$ . (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

1969, Theorem 1; Bachem and Kern, 1992, Proposition 5.35; Ziegler, 1995, Lemma 6.7).

**Theorem 2.** Let  $S \subseteq \mathbb{R}^R$  be a subspace. Then every vector  $f \in S$  is a conformal sum of EVs. That is, there exists a finite set  $E$  of EVs conforming to  $f$  such that

$$f = \sum_{e \in E} e.$$

The set  $E$  can be chosen such that every  $e \in E$  has a component which is non-zero in  $e$ , but zero in all other elements of  $E$ . Hence,  $|E| \leq \dim(S)$  and  $|E| \leq |\text{supp}(f)|$ .

It is easy to see that every flux mode is the conformal sum of EFMs. For later use, we present a slightly rephrased version of this result. We note that, if  $e$  is an EFM, then any element of the ray

$$\{\lambda e \mid \lambda > 0\}$$

is an EFM. Hence, we may refer to one representative EFM on each ray.

**Corollary 3.** Let  $(S, \mathcal{R}, N)$  be a metabolic network,  $\tau \in \{-, 0, +\}^R$  be a sign vector and  $E_\tau$  be a set of representative EFMs conforming to  $\tau$ . Then, every flux mode  $f \in C$  conforming to  $\tau$  is a non-negative linear combination of elements of  $E_\tau$ :

$$f = \sum_{e \in E_\tau} \alpha_e e \quad \text{with} \quad \alpha_e \geq 0.$$

**Proof.** Clearly,  $f \in C$  implies  $f \in \ker(N)$ . By Theorem 2,  $f$  is the conformal sum of EVs of  $\ker(N)$ . However, for an EV  $e \in \ker(N)$  to conform to  $f \in C$ , it is required that  $e \in C$ . Hence, by Lemma 1,  $e$  is an EFM, which can be written as a positive scalar multiple of a representative EFM.  $\square$

#### 4.3. Problem reformulation

We start with the formal statement of an intuitive argument. Consider a feasible solution of the enzyme allocation problem (1) and the corresponding steady-state flux: if a reaction does not carry any flux, then the corresponding optimal enzyme concentration is zero. We add an appropriate constraint to the enzyme allocation problem and obtain an equivalent optimization problem.

**Lemma 4.** Let  $(S, \mathcal{R}, N, v)$  be a kinetic metabolic network. The enzyme allocation problem (1) is equivalent to the following optimization problem:

$$\max_{\bar{x} \in X, c \in \mathbb{R}_{\geq}^R} \bar{v}_{r_*} \quad (5a)$$

subject to

$$c_r = 0 \quad \text{if} \quad \bar{v}_r = 0, \quad (5b)$$

$$\sum_{r \in \mathcal{R}} w_r c_r = c^{\text{tot}}, \quad (5c)$$

$$\bar{v} \in C, \bar{v}_{r_*} > 0. \quad (5d)$$

**Proof.** For every feasible solution  $(\bar{x}, c)$  of (1) with objective function  $\bar{v}_{r_*}$ , we construct a feasible solution  $(\bar{x}, c')$  of (5) with objective function  $\bar{v}'_{r_*} \geq \bar{v}_{r_*}$ : Let  $S_{\bar{v}} = \text{supp}(\bar{v})$ . Using  $\lambda = c^{\text{tot}} / (\sum_{r \in S_{\bar{v}}} w_r c_r) \geq 1$ , we set

$$c'_r = \begin{cases} \lambda c_r & \text{if } r \in S_{\bar{v}}, \\ 0 & \text{if } r \notin S_{\bar{v}}. \end{cases}$$

Clearly,

$$\sum_{r \in \mathcal{R}} w_r c'_r = \sum_{r \in S_{\bar{v}}} w_r c'_r = \lambda \sum_{r \in S_{\bar{v}}} w_r c_r = c^{\text{tot}}.$$

Further,  $\bar{v}' = c' \circ \kappa = \lambda (c \circ \kappa) = \lambda \bar{v}$  implies  $N \bar{v}' = 0$  and  $\sigma(\bar{v}') = \sigma(\bar{v})$ , that is,  $\bar{v}' \in C$ . Hence,  $(\bar{x}, c')$  fulfills constraints (5b), (5c) and (5d) and  $\bar{v}'_{r_*} = \lambda \bar{v}_{r_*} \geq \bar{v}_{r_*}$ .  $\square$

As a consequence, variation over enzyme concentrations can be replaced by variation over flux modes.

**Lemma 5.** Let  $(S, \mathcal{R}, N, v)$  be a kinetic metabolic network. The enzyme allocation problem (5) is equivalent to the following optimization problem over  $\bar{x} \in X$  and  $f \in C$ :

$$\min_{\bar{x} \in X, f \in C} \sum_{r \in \text{supp}(\kappa)} \frac{w_r f_r}{\kappa_r(\bar{x}, p)} \quad (6a)$$

subject to

$$\sigma(f) \leq \sigma(\kappa), f_{r_*} = 1. \quad (6b)$$

Let  $(\bar{x}, c)$  and  $(\bar{x}, f)$  be corresponding feasible solutions of (5) and (6), respectively. The product of the related objective functions amounts to

$$\bar{v}_{r_*} \sum_{r \in \text{supp}(\kappa)} \frac{w_r f_r}{\kappa_r} = c^{\text{tot}}. \quad (7)$$

**Proof.** We show that, for every feasible solution  $(\bar{x}, c)$  of (5), there exists a feasible solution  $(\bar{x}, f)$  of (6), and vice versa. Moreover, that the related objective functions fulfill Eq. (7).

Assume that  $(\bar{x}, c)$  is a feasible solution of (5). Define  $f = \bar{v} / \bar{v}_{r_*}$ . Clearly,  $f \in C$ ,  $f_{r_*} = 1$ , and  $\sigma(f) = \sigma(\bar{v}) = \sigma(c \circ \kappa) \leq \sigma(\kappa)$ . Hence,  $(\bar{x}, f)$  is a feasible solution of (6). Let  $S_\kappa = \text{supp}(\kappa)$ . Using  $\bar{v} = c \circ \kappa = \bar{v}_{r_*} f$  and hence  $c_r = \bar{v}_{r_*} f_r / \kappa_r$  for  $r \in S_\kappa$ , we rewrite the sum constraint and obtain the desired Eq. (7):

$$c^{\text{tot}} = \sum_{r \in \mathcal{R}} w_r c_r = \sum_{r \in S_\kappa} w_r c_r = \bar{v}_{r_*} \sum_{r \in S_\kappa} \frac{w_r f_r}{\kappa_r}.$$

Conversely, assume that  $(\bar{x}, f)$  is a feasible solution of (6). Since  $\sigma(f) \leq \sigma(\kappa)$ , we can define  $\bar{v}_{r_*} > 0$  by Eq. (7), and we set  $c_r = \bar{v}_{r_*} f_r / \kappa_r$  for  $r \in S_\kappa$  and  $c_r = 0$  for  $r \notin S_\kappa$ . Clearly,  $c_r = 0$  if  $\bar{v}_r = c_r \kappa_r = 0$ . Further,

$$\sum_{r \in \mathcal{R}} w_r c_r = \sum_{r \in S_\kappa} w_r c_r = \bar{v}_{r_*} \sum_{r \in S_\kappa} \frac{w_r f_r}{\kappa_r} = c^{\text{tot}}.$$

By definition,  $\bar{v}_r = c_r \kappa_r = \bar{v}_{r_*} f_r$  for  $r \in S_\kappa$ , and, since  $\sigma(f) \leq \sigma(\kappa)$ ,  $\bar{v}_r = c_r \kappa_r = 0$  and  $f_r = 0$  for  $r \notin S_\kappa$ . That is,  $\bar{v} = \bar{v}_{r_*} f \in C$ . Hence,  $(\bar{x}, c)$  is a feasible solution of (5).  $\square$

We note that the inequality constraints involving the kinetics may be unfeasible. For given flux mode  $f \in C$ , the existence of steady-state metabolite concentrations  $\bar{x} \in \mathbb{R}_{\geq}^S$  such that  $\sigma(f) \leq \sigma(\kappa(\bar{x}, p))$  is equivalent to the existence of chemical potentials  $\mu \in \mathbb{R}^S$  such that  $\sigma(f) \leq \sigma(-(\mu N)^T)$ . Whereas conventional FBA has to be augmented with thermodynamic constraints (Beard et al., 2004; Müller and Bockmayr, 2013), they are incorporated in the definition of a metabolic network with known kinetics.

#### 4.4. Main results

The next statement characterizes optimal solutions of the enzyme allocation problem for fixed metabolite concentrations. Its proof involves the result on conformal sums obtained in Section 4.2.

**Proposition 6.** Let  $(S, \mathcal{R}, N, v)$  be a kinetic metabolic network. Consider the enzyme allocation problem (1) for fixed  $\bar{x} \in X$ . If this restricted optimization problem is feasible, then it has an optimal solution for which the corresponding steady-state flux is an EFM.

**Proof.** By Lemmas 4 and 5, the enzyme allocation problem (1) is equivalent to optimization problem (6). We consider (6) for fixed  $\bar{x} \in X$  and assume that this restricted problem is feasible.

We write  $\kappa$  short for  $\kappa(\bar{x}, p)$  and introduce  $\tau = \sigma(\kappa)$  and  $S_\kappa = \text{supp}(\kappa)$ . In (6), we vary over  $f \in C$  such that  $\sigma(f) \leq \sigma(\kappa) = \tau$  and  $f_{r_*} = 1$ . By Corollary 3, every flux mode  $f \in C$  conforming to  $\tau$  is a non-negative linear combination of elements of  $E_\tau$ , which is a set of representative EFM's conforming to  $\tau$ . We assume the EFM's to be scaled by component  $r_*$  and divide the set  $E_\tau$  into two subsets,  $E_\tau = E_1 \cup E_0$ , such that  $e \in E_1$  implies  $e_{r_*} = 1$  and  $e \in E_0$  implies  $e_{r_*} = 0$ . We have

$$f = \sum_{e \in E_1} \alpha_e e + \sum_{e \in E_0} \beta_e e \quad \text{with} \quad \alpha_e, \beta_e \geq 0.$$

From  $f_{r_*} = 1$ , we obtain the constraint

$$1 = f_{r_*} = \sum_{e \in E_1} \alpha_e e_{r_*} + \sum_{e \in E_0} \beta_e e_{r_*} = \sum_{e \in E_1} \alpha_e.$$

Using the conformal sum for  $f$  in (6), we obtain an equivalent formulation of the restricted problem:

$$\min_{\alpha_e, \beta_e} \sum_{r \in S_\kappa} \frac{w_r (\sum_{e \in E_1} \alpha_e e_r + \sum_{e \in E_0} \beta_e e_r)}{\kappa_r} \quad (8a)$$

subject to

$$\sum_{e \in E_1} \alpha_e = 1. \quad (8b)$$

We observe that the objective function is linear in  $\alpha_e$  and  $\beta_e$ :

$$\sum_{e \in E_1} \alpha_e \underbrace{\sum_{r \in S_\kappa} \frac{w_r e_r}{\kappa_r}}_{g_e} + \sum_{e \in E_0} \beta_e \underbrace{\sum_{r \in S_\kappa} \frac{w_r e_r}{\kappa_r}}_{g_e}$$

with

$$g_e = \sum_{r \in S_\kappa} \frac{w_r e_r}{\kappa_r} \quad \text{for} \quad e \in E_\tau.$$

Since all  $e \in E_\tau$  conform to  $\tau = \sigma(\kappa)$ , that is,  $\sigma(e) \leq \sigma(\kappa)$ , we have  $w_r e_r / \kappa_r \geq 0$  for all  $e \in E_\tau$  and  $r \in S_\kappa$ . Moreover, for all  $e \in E_\tau$ , there is  $r \in S_\kappa$  such that  $e_r \neq 0$  and hence  $w_r e_r / \kappa_r > 0$ . Consequently,  $g_e > 0$  for all  $e \in E_\tau$ .

Since there is no further restriction on  $\beta_e \geq 0$ , the minimum of the objective function is attained at  $\beta_e = 0$  for all  $e \in E_0$ . In other words, EFM's  $e \in E_0$  do not contribute to the optimal solution.

Let  $e' \in E_1$  be an EFM such that  $g_{e'} \leq g_e$  for all  $e \in E_1$ . Since  $\alpha_e \geq 0$  and  $\sum_{e \in E_1} \alpha_e = 1$ , we have

$$\sum_{e \in E_1} \alpha_e g_e \geq \sum_{e \in E_1} \alpha_e g_{e'} = g_{e'},$$

and the minimum of the objective function is attained at  $\alpha_{e'} = 1$  and  $\alpha_e = 0$  for all other  $e \in E_1$ , that is, for  $f = e'$ . To conclude, we consider a degenerate case: If there are several  $e \in E_{\min} \subseteq E_1$  for which  $g_e$  is minimal, then any  $f = \sum_{e \in E_{\min}} \alpha_e e$  (with  $\alpha_e \geq 0$  and  $\sum_{e \in E_{\min}} \alpha_e = 1$ ) is optimal.  $\square$

The following statement is the main result of this work.

**Theorem 7.** Let  $(S, \mathcal{R}, N, v)$  be a kinetic metabolic network. If the enzyme allocation problem (1) has an optimal solution, then it has an optimal solution for which the corresponding steady-state flux is an EFM.

**Proof.** Let an optimal solution of the enzyme allocation problem (1) be attained at  $\bar{x} \in X$ . Clearly, optimization problem (1) restricted to this particular  $\bar{x}$  is feasible. By Proposition 6, this restricted problem has an optimal solution for which the corresponding steady-state flux is an EFM.  $\square$

In applications, we use Theorem 7 to study the switching behavior of kinetic metabolic networks. Depending on external parameters, the optimal solution of the enzyme allocation problem may switch from one EFM to another, involving a discontinuous change of enzyme and metabolite concentrations. In a first

approach, one may vary the external parameters and determine the optimal solution for each EFM in order to find the optimal solution of the full problem. To this end, we transform the optimization problem restricted to an EFM.

**Corollary 8.** Let  $(S, \mathcal{R}, N, v)$  be a kinetic metabolic network. In the enzyme allocation problem (1), let the steady-state flux be restricted to  $\{\lambda e \mid \lambda > 0\}$ , where  $e \in C$  is an EFM with  $e_{r_*} = 1$ . Then, this restricted optimization problem is equivalent to the following optimization problem over  $\bar{x} \in X$ :

$$\min_{\bar{x} \in X} \sum_{r \in \text{supp}(e)} \frac{w_r e_r}{\kappa_r(\bar{x}, p)} \quad (9a)$$

subject to

$$\sigma(e) \leq \sigma(\kappa). \quad (9b)$$

The corresponding enzyme concentrations  $c \in \mathbb{R}_{\geq}^{\mathcal{R}}$  are given by

$$c_r = c^{\text{tot}} \frac{\frac{e_r}{\kappa_r(\bar{x}, p)}}{\sum_{s \in \text{supp}(e)} \frac{w_s e_s}{\kappa_s(\bar{x}, p)}}. \quad (10)$$

**Proof.** By Lemmas 4 and 5, the enzyme allocation problem (1) is equivalent to optimization problem (6). Hence, we consider (6) for fixed  $f = e$ . Clearly, this restricted problem is feasible if and only if optimization problem (9) is feasible. If  $\sigma(e) \leq \sigma(\kappa)$ , then  $\text{supp}(e) \subseteq \text{supp}(\kappa)$  and hence

$$\sum_{r \in \text{supp}(e)} \frac{w_r e_r}{\kappa_r(\bar{x}, p)} = \sum_{r \in \text{supp}(e)} \frac{w_r e_r}{\kappa_r(\bar{x}, p)}.$$

That is, the objective functions of the two optimization problems are identical.

By using  $c^{\text{tot}} \kappa = \lambda e$  with  $\lambda > 0$  and the constraint  $\sum_{r \in \mathcal{R}} w_r c_r = c^{\text{tot}}$ , we obtain  $\lambda = c^{\text{tot}} / (\sum_{r \in \text{supp}(e)} w_r e_r / \kappa_r)$  and hence  $c_r = \lambda \frac{e_r}{\kappa_r}$  for  $r \in \mathcal{R}$ .  $\square$

## 5. Discussion and outlook

In this work, we have studied a general metabolic optimization problem under enzymatic capacity constraints. Our analysis was motivated by the fact that the total enzymatic capacity of a metabolic network is limited by finite shared resources, such that an increase in the concentration of one enzyme necessitates a decrease in the concentration of one or more other enzymes. This scenario can be caused by molecular crowding, where enzymes compete for cytosolic space, limited membrane space, a finite availability of macro-nutrients, such as nitrogen or phosphorus, or micro-nutrients, such as transition metals, as well as a limited energy expenditure for amino-acid synthesis. In each of these instances, the global constraint can be formulated in terms of a weighted sum of enzyme concentrations, where the weight factors specify the fraction of the shared resource utilized per unit enzyme.

Recently, such global constraints have been incorporated into large-scale stoichiometric models of metabolism, most notably to explain the occurrence of low-yield pathways (Schuster et al., 2011; Shlomi et al., 2011; Vazquez and Oltvai, 2011; Zhuang et al., 2011). However, these works were restricted to linear stoichiometric optimization problems and did not consider enzyme concentrations within a kinetic description of a metabolic network. Here, we have addressed a non-linear kinetic optimization problem, where we assume that each reaction can be catalyzed by an enzyme, but allow for arbitrary enzyme kinetics. Most importantly, we have derived a rigorous proof that an optimal flux distribution under global enzymatic constraints is necessarily an elementary flux mode. This finding has significant consequences for our understanding of metabolic optimality and

metabolic switches as well as for the computational identification of optimal fluxes in kinetic metabolic networks.

In particular, our results allow us to efficiently compute and compare the optimal enzyme distributions for individual flux modes. In a recent work (Flamholz et al., 2013), the trade-off between energy yield and protein cost was studied for several alternative prokaryotic glycolytic pathways. To this end, different pathway designs were compared by fixing the output flux at a certain value and estimating the necessary protein investments. Clearly, such an approach does not exclude the possibility that a combined pathway might incur an even lower enzymatic cost. However, our result shows that it is indeed sufficient to compare the solutions for EFMs.

Notwithstanding its theoretical merits, the practical implications of our approach have to be studied further. Firstly, for genome-scale metabolic networks, an exhaustive evaluation of all EFMs is computationally infeasible. A natural next step is to design a combined linear/non-linear optimization algorithm for the identification of optimal EFMs. Secondly, while the enzymatic capacity constraint is applicable to a multitude of possible limitations, our approach may be extended to include *de novo* synthesis or uptake of limiting resources. We conjecture that the problem can still be formulated in a way that our overall conclusions remain valid if such more general models are considered. Thirdly, the effects of co-limitation in cellular metabolism and the simultaneous optimization of more than one objective function pose new challenges for theoretical analysis.

Finally, we note that experimentally observed flux distributions are not necessarily always EFMs. While some well-known metabolic switches indeed show an exclusive choice between alternative metabolic states, such as most instances of catabolite repression, it is known that alternative metabolic strategies sometimes operate simultaneously, such as a residual respiration in cancer cells (Moreno-Sanchez et al., 2007). Whether such a co-occurrence of metabolic strategies arises due to additional constraints or optimality principles, or to what extent these instances are sub-optimal adaptations, remains to be studied. We note that mixed strategies were also observed in recent computational studies using large-scale FBA models (Vazquez et al., 2010) as well as for integrated models of enzyme synthesis and metabolism to investigate shifts in growth strategies (Molenaar et al., 2009). Again, our analytical results provide a strong incentive to study the consequences of additional constraints or optimality principles on the solutions of metabolic optimization problems. We believe that only the further understanding of the properties of optimal flux distributions will allow us to investigate the fundamental trade-offs in cellular resource allocation.

## Acknowledgments

We acknowledge numerous helpful comments from two anonymous reviewers. During the revision process, we were informed by the authors of Wortel et al. (accepted for publication) that they have arrived at similar conclusions as presented in this work. The results in the two papers were derived independently.

**Funding Statement.** This work is an output of the project “Local Team and International Consortium for Computational Modeling of a Cyanobacterial Cell”, Reg. No. CZ.1.07/2.3.00/20.0256, financed by the European Union and supported by the Ministry of Education of the Czech Republic (SM, RS), as well as the project “Übergangsmetalle und phototrophes Wachstum: Ein neuer Ansatz der constraint-basierten Modellierung großer Stoffwechselnetzwerke” funded by the Einstein Stiftung Berlin (RS). The funding bodies had no role in study design, data analysis, and the decision to submit the manuscript.

## References

- Antoniewicz, M.R., 2013. Dynamic metabolic flux analysis—tools for probing transient states of metabolic networks. *Curr. Opin. Biotechnol.* 24 (6), 973–978.
- Bachem, A., Kern, W., 1992. *Linear Programming Duality, An Introduction to Oriented Matroids*. Springer-Verlag, Berlin.
- Beard, D.A., Babson, E., Curtis, E., Qian, H., 2004. Thermodynamic constraints for biochemical networks. *J. Theor. Biol.* 228 (3), 327–333.
- Beg, Q.K., Vazquez, A., Ernst, J., de Menezes, M.A., Bar-Joseph, Z., Barabási, A.-L., Oltvai, Z.N., 2007. Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proc. Natl. Acad. Sci. USA* 104 (31), 12663–12668.
- Berkhout, J., Bruggeman, F.J., Teusink, B., 2012. Optimality principles in the regulation of metabolic networks. *Metabolites* 2, 529–552.
- Björner, A., Las Vergnas, M., Sturmfels, B., White, N., Ziegler, G.M., 1999. *Oriented Matroids*, 2nd ed. *Encyclopedia of Mathematical Applications*, vol. 46. Cambridge University Press, Cambridge.
- Crabtree, H.G., 1928. The carbohydrate metabolism of certain pathological overgrowths. *Biochem. J.* 22 (5), 1289–1298.
- Flamholz, A., Noor, E., Bar-Even, A., Liebermeister, W., Milo, R., 2013. Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proc. Natl. Acad. Sci. USA* 110 (24), 10039–10044.
- Gagneur, J., Klamt, S., 2004. Computation of elementary modes: a unifying framework and the new binary approach. *BMC Bioinformatics* 5, 175.
- Goel, A., Wortel, M.T., Molenaar, D., Teusink, B., 2012. Metabolic shifts: a fitness perspective for microbial cell factories. *Biotechnol. Lett.* 34 (12), 2147–2160.
- Goelzer, A., Fromion, V., Scorletti, G., 2011. Cell design in bacteria as a convex optimization problem. *Automatica* 47, 1210–1218.
- Heinrich, R., Schuster, S., 1996. *The Regulation of Cellular Systems*. Chapman and Hall.
- Hsu, P.P., Sabatini, D.M., 2008. Cancer cell metabolism: Warburg and beyond. *Cell* 134 (5), 703–707.
- Klamt, S., Stelling, J., 2003. Two approaches for metabolic pathway analysis? *Trends Biotechnol.* 21 (2), 64–69.
- Larhlmi, A., Bockmayr, A., 2009. A new constraint-based description of the steady-state flux cone of metabolic networks. *Discret. Appl. Math.* 157 (10), 2257–2266.
- Mahadevan, R., Edwards, J.S., Doyle, F.J., 2002. Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* 83 (3), 1331–1340.
- Molenaar, D., van Berlo, R., de Ridder, D., Teusink, B., 2009. Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Syst. Biol.* 5, 323.
- Moreno-Sanchez, R., Rodriguez-Enriquez, S., Marin-Hernandez, A., Saavedra, E., 2007. Energy metabolism in tumor cells. *FEBS J.* 274 (6), 1393–1418.
- Müller, A.C., 2012. Thermodynamic constraints in metabolic networks (Master's thesis). Fachbereich Mathematik und Informatik, Freie Universität Berlin.
- Müller, A.C., Bockmayr, A., 2013. Fast thermodynamically constrained flux variability analysis. *Bioinformatics* 29, 903–909.
- Müller, S., Regensburger, G., 2012. Generalized mass action systems: complex balancing equilibria and sign vectors of the stoichiometric and kinetic-order subspaces. *SIAM J. Appl. Math.* 72, 1926–1947.
- Noor, E., Lewis, N., Milo, R., 2012. A proof for loop-law constraints in stoichiometric metabolic networks. *BMC Syst. Biol.* 6 (1), 140.
- Oberhardt, M., A. Palsson, B.O., Papin, J.A., 2009. Applications of genome-scale metabolic reconstructions. *Mol. Syst. Biol.* 5, 320.
- Oliveira, J.S., Bailey, C.G., Jones-Oliveira, J.B., Dixon, D.A., 2001. An algebraic-combinatorial model for the identification and mapping of biochemical pathways. *Bull. Math. Biol.* 63 (6), 1163–1196.
- Orth, J.D., Thiele, I., Palsson, B.O., 2010. What is flux balance analysis? *Nat. Biotechnol.* 28 (3), 245–248.
- Richter-Gebert, J., Ziegler, G.M., 1997. Oriented matroids, In: *Handbook of Discrete and Computational Geometry*. CRC, Boca Raton, FL, pp. 111–132.
- Rockafellar, R. T., 1969. The elementary vectors of a subspace of  $\mathbb{R}^N$ . In: *Combinatorial Mathematics and its Applications* (Conference Proceedings, University of North Carolina, Chapel Hill, N.C., 1967), Univ. North Carolina Press, Chapel Hill, N.C., pp. 104–127.
- Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M., Sauer, U., 2012. Multi-dimensional optimality of microbial metabolism. *Science* 336 (6081), 601–604.
- Schuster, S., de Figueiredo, L.F., Schroeter, A., Kaleta, C., 2011. Combining metabolic pathway analysis with evolutionary game theory: explaining the occurrence of low-yield pathways by an analytic optimization approach. *Biosystems* 105 (2), 147–153.
- Schuster, S., Hilgetag, C., 1994. On elementary flux modes in biochemical reaction systems at steady state. *J. Biol. Syst.* 2, 165–182.
- Schuster, S., Hilgetag, C., Woods, J.H., Fell, D.A., 2002. Reaction routes in biochemical reaction systems: algebraic properties, validated calculation procedure and example from nucleotide metabolism. *J. Math. Biol.* 45 (2), 153–181.
- Schuster, S., Pfeiffer, T., Fell, D.A., 2008. Is maximization of molar yield in metabolic networks favoured by evolution? *J. Theor. Biol.* 252 (3), 497–504.
- Shlomi, T., Benyamini, T., Gottlieb, E., Sharan, R., Ruppin, E., 2011. Genome-scale metabolic modeling elucidates the role of proliferative adaptation in causing the Warburg effect. *PLoS Comput. Biol.* 7 (3), e1002018.
- Steuer, R., Junker, B., 2008. Computational models of metabolism: stability and regulation in metabolic networks. In: Rice, S.A. (Ed.), *Advances in Chemical Physics*, vol. 142. Wiley.



- Vazquez, A., Beg, Q.K., Demenezes, M.A., Ernst, J., Bar-Joseph, Z., Barabási, A.-L., Boros, L. G., Oltvai, Z.N., 2008. Impact of the solvent capacity constraint on *E. coli* metabolism. *BMC. Syst. Biol.* 2, 7.
- Vazquez, A., Liu, J., Zhou, Y., Oltvai, Z.N., 2010. Catabolic efficiency of aerobic glycolysis: the Warburg effect revisited. *BMC. Syst. Biol.* 4, 58.
- Vazquez, A., Oltvai, Z.N., 2011. Molecular crowding defines a common origin for the Warburg effect in proliferating cells and the lactate threshold in muscle physiology. *PLoS One* 6 (4), e19538.
- Wagner, C., Urbanczik, R., 2005. The geometry of the flux cone of a metabolic network. *Biophys. J.* 89, 3837–3845.
- Warburg, O., Posener, K., Negelein, E., 1924. On the metabolism of carcinoma cells. *Biochem. Z.* 152, 309–344.
- Wortel, M. T., Peters, H., Hulshof, J., Teusink, B., Bruggeman, F. J., 2014. Metabolic states with maximal specific rate carry flux through an elementary flux mode, *FEBS J.*, accepted for publication.
- Zhuang, K., Vemuri, G.N., Mahadevan, R., 2011. Economics of membrane occupancy and respiro-fermentation. *Mol. Syst. Biol.* 7, 500.
- Ziegler, G.M., 1995. Lectures on Polytopes, Vol. 152 of Graduate Texts in Mathematics. Springer-Verlag, New York.