ARTICLE

BIOTECHNOLOGY and BIOENGINEERING

# A Reliable Simulator for Dynamic Flux Balance Analysis

**K. Höffner, S. M. Harwood, P. I. Barton**

Process System Engineering Laboratory, Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue Cambridge, Massachusetts 02139; telephone: +1-617-253-6526; fax: +1-617-258-5042; e-mail: pib@mit.edu

**ABSTRACT:** Dynamic flux balance analysis (DFBA) provides a platform for detailed design, control and optimization of biochemical process technologies. It is a promising modeling framework that combines genome-scale metabolic network analysis with dynamic simulation of the extracellular environment. Dynamic flux balance analysis assumes that the intracellular species concentrations are in equilibrium with the extracellular environment. The resulting underdetermined stoichiometric model is solved under the assumption of a biochemical objective such as growth rate maximization. The model of the metabolism is coupled with the dynamic mass balance equations of the extracellular environment via expressions for the rates of substrate uptake and product excretion, which imposes additional constraints on the linear program (LP) defined by growth rate maximization of the metabolism. The linear program is embedded into the dynamic model of the bioreactor, and together with the additional constraints this provides an accurate model of the substrate consumption, product secretion, and biomass production during operation. A DFBA model consists of a system of ordinary differential equations for which the evaluation of the right-hand side requires not only function evaluations, but also the solution of one or more linear programs. The numerical tool presented here accurately and efficiently simulates large-scale dynamic flux balance models. The main advantages that this approach has over existing implementation are that the integration scheme has a variable step size, that the linear program only has to be solved when qualitative changes in the optimal flux distribution of the metabolic network occur, and that it can reliably simulate behavior near the boundary of the domain where the model is defined. This is illustrated through large-scale examples taken from the literature.

Biotechnol. Bioeng. 2013;110: 792–802.



**KEYWORDS:** numerical tools; system biology; computational methods; flux balance analysis; dynamic flux balance analysis

Correspondence to: P.I. Barton

Additional supporting information may be found in the online version of this article.
Received 25 May 2012; Revision received 21 September 2012;
Accepted 25 September 2012
Accepted manuscript online 10 October 2012;
Article first published online 18 October 2012 in Wiley Online Library (http://onlinelibrary.wiley.com/doi/10.1002/bit.24748/abstract)
DOI 10.1002/bit.24748

## Introduction

The objective of this article is to present a numerical integration scheme for large-scale systems of differential equations encountered in dynamic flux balance analysis (DFBA). DFBA is a subject of growing interest both from the metabolic engineering as well as the process systems engineering perspective. The motivation for this research is the increasing use of genome-scale DFBA models for analysis, control and optimization of biochemical processes. The need for fast, accurate, and reliable simulations becomes evident as the size and number of embedded metabolic network models increases. In addition to the increasing level of complexity of the models, recent research has extended the approach to analyze processes with two or more microbial species (Hanly and Henson, 2011; Zhuang et al., 2011).

### Flux Balance Analysis

Flux balance analysis (FBA) is a genome-scale constraint-based modeling approach for metabolic networks. It is a successful framework for metabolic engineering and analysis of metabolic networks (Orth et al., 2010; Palsson, 2006; Stephanopoulos et al., 1998). Applications of FBA for industrial and medical biotechnology have been recently reviewed by Milne et al. (2009).

Construction of FBA models involves four essential steps: defining the system; obtaining reaction stoichiometry; defining biologically relevant objective functions, and adding other biochemical constraints; and solving the resulting linear program (Raman and Chandra, 2009). Based on genomic analysis, a metabolism can be modeled as a network of reactions that must satisfy simple mass balance constraints. The network reconstruction determines the stoichiometry of the metabolism under the balanced growth assumption (Palsson, 2006; Stephanopoulos et al., 1998). However, this network is often underdetermined; the fluxes of the different substrates and metabolites can vary and still produce a solution that satisfies the mass balance constraints. Thus, it is assumed that the fluxes will be such that some cellular objective is maximized. For example, an

evolutionary argument can be made that a microorganism will maximize its growth rate if sufficient nutrients are provided (Edwards and Palsson, 2000; Kauffman et al., 2003). Other candidates for objective functions include (Vargas et al., 2011): maximization of ATP production (Majewski and Domach, 1990; Ramakrishna et al., 2001), minimization of metabolic adjustment (MoMa) (Segré et al., 2002), or minimization of ATP consumption (Pizarro et al., 2007; Raghunathan et al., 2006). Additional physico-chemical constraints, spatial or topological constraints, condition dependent environmental constraints, and regulatory constraints can be added if such information is available (Palsson, 2006). The result for FBA is a linear program given by the cellular objective under the stoichiometry and additional constraints.

A growing number of reconstructions for a variety of microorganisms are available at Feist et al. (2008) and at the BiGG online database (Schellenberger et al., 2010). The size of these models ranges from 240 metabolites and 263 reactions to 2,774 metabolites and 3,726 reactions.

## Dynamic Flux Balance Analysis

Dynamic flux balance analysis is an extension of FBA which has the advantage that it enables analysis of interactions between the metabolism and the environment. It generates dynamic predictions of substrate, biomass, and product concentrations for growth in batch or fed-batch cultures (Hjersted and Henson, 2009). DFBA provides a structured model of the biochemical process where the reaction pathways within the microorganism change depending on the environmental conditions, which is effectively represented by changes in the functional dependency on the substrates. In contrast, simplified unstructured models often used in process control represent the growth rate of the microorganism as a simple function of a substrate and have therefore only limited predictive capabilities. This type of model, for example Monod models, has the disadvantage that it is only valid for a small range of conditions (Zhuang et al., 2011) and therefore it is not suitable for batch operation with constantly changing conditions.

In addition to the interaction through the expressions for uptake and production, regulatory changes due to changes in the environment can also be included in the model formulation. For example, Covert et al. (2001) extended FBA models using dynamic FBA and Boolean logic to represent transcriptional regulatory constraints. Recently Vargas et al. (2011) presented dynamic models which incorporate regulatory information through dynamic constraints, time-dependent objective functions, and other non-steady-state phenomena.

### *Applications*

The number of DFBA applications has increased in recent years, with more applications based on genome-scale models. Small-scale (less than 200 fluxes per FBA model and one or two compartments) applications of DFBA include bacteria (Anesiadis et al., 2008; Lequeux et al., 2010; Mahadevan et al., 2002; Meadows et al., 2010; Varma and Palsson, 1994;), *Saccharomyces cerevisiae* (Lee et al., 2008; Pizarro et al., 2007; Sainz et al., 2003), plant (Luo et al., 2009), and animal (Luo et al., 2006). Genome-scale applications of DFBA include (Vargas et al., 2011): *S. cerevisiae* (Hjersted and Henson, 2009; Hjersted et al., 2007) and bacteria (Oddone et al., 2009; Zhuang et al., 2011). Table I summarizes the existing applications of DFBA

**Table I.** Overview of existing DFBA simulation studies (sorted by date).

| Refs. | FBA model | Met. | Fluxes | Method/solver |
|---|---|---|---|---|
| Varma and Palsson (1994) | Based on Majewski and Domach (1990) | 24 | 34 | SOA/— |
| Mahadevan et al. (2002) | Based on Schilling et al. (2000) | 3 | 4 | SOA/CPLEX |
| | | | | DOA/fmincon |
| Sainz et al. (2003) | – | 43 | 38 | SOA/— |
| Luo et al. (2006) (MDFBA) | – | 7 | 8 | DOA/fmincon |
| Hjersted and Henson (2006, 2009) | iGH99 | 98 | 82 | DA/CONOPT |
| Pizarro et al. (2007) | Based on Sainz et al. (2003) | 38 | 39 | SOA/— |
| Hjersted et al. (2007) | iND750 | 1,059 | 1,265 | DA/MOSEK |
| Anesiadis et al. (2008) | iJR904 | 625 | 931 | SOA/CPLEX |
| Lee et al. (2008) (idFBA) | – | — | 13 | SOA/— |
| Luo et al. (2009) (MDFBA) | – | 8 | 5 | DOA/fmincon |
| Oddone et al. (2009) | IL1403 | 422 | 621 | SOA/Mathematica |
| Lequeux et al. (2010) (MDFA) | — | 24 | 34 | polynomial fitting |
| Salimi et al. (2010) | iFS2007 | 679 | 712 | DA/— |
| | iFS431 | 603 | 621 | |
| Zhuang et al. (2011) | *G. sulfurreducens* | 541 | 522 | DA/LINDO |
| | *R. ferrireducens* | 790 | 762 | |
| Meadows et al. (2010) | Based on Varma and Palsson (1994) | 30 | 123 | ODE15S/linprog |
| Vargas et al. (2011) | idFV715(iFF708) | 590 | 1,181 | SOA/LINDO |
| Nolan and Lee (2011) (MDFA) | — | 150 | 136 | SOA/— |
| Hanly and Henson (2011) | iRJ904 | 625 | 931 | DA/MOSEK |
| Hanly et al. (2012) | iND750 | 1,059 | 1,265 | |

sorted by date. There are three different foci in application of DFBA which may be pursued independently or simultaneously, dynamic metabolic engineering, co-culture simulation, and optimal (fed-)batch control.

Dynamic metabolic engineering studies focus on gene insertion and deletion and their effect on the dynamic behavior (Hjersted et al., 2007). Related dynamic approaches such as Metabolic Adjustment Dynamic Flux Balance Analysis (MDFBA) (Luo et al., 2009), based on Minimization of Metabolic Adjustment (MoMa), and integrated DFBA (idFAB) (Lee et al., 2008) embed quadratic programs rather than linear programs. Discussion of these approaches is outside of the scope of this article.

Utilizing the complex interactions between several microorganisms is a promising avenue for industrial biotechnology. *In silicio* studies of such co-cultures allow for fast prescreening of possible compositions of the consortia and possible additional metabolic engineering approaches to improve the productivity of the consortia. The dynamic multi-species metabolic modeling (DMMM) framework of Zhuang et al. (2011) has been developed for co-culture simulation and several computational studies have been presented so far (Hanly and Henson, 2011; Salimi et al., 2010; Zhuang et al., 2011).

Optimal control of batch or fed-batch operation of bioreactors is important in application such as the production of renewable liquid fuels. Optimal control applications of genome-scale DFBA models have been limited due to the increased computation time for large models (Hjersted et al., 2007). Hjersted and Henson (2006) studied fed-batch optimization of a bioreactor with a small-scale model of yeast metabolism. The optimization problem is temporally discretized and the resulting mathematical program with equality constraints (MPEC) was solved using CONOPT. The feed stream concentrations were modeled as piecewise constant control inputs with the same discretization as the states.

### *Implementation*

Two approaches to simulate DFBA models were presented in Mahadevan et al. (2002): static optimization approach (SOA) and dynamic optimization approach (DOA). SOA finds a solution of the LP and integrates the ODE with the fixed LP solution for some fixed time, then the LP is resolved with the values of the dynamics at the final time of this time horizon. This process is repeated until the end of the simulation is reached. A disadvantage of SOA is that one needs to balance between accurate results, that is, small step size and computation time. DOA involves optimization over the entire time period of interest to obtain time profiles of fluxes and metabolite levels. The dynamic optimization problem is transformed through collocation to a nonlinear programming (NLP) problem and the NLP problem is solved once (Mahadevan et al., 2002). DOA is limited to small-scale metabolic models due to the explosion of dimension that occurs through time discretization. More recently, a direct approach (DA) of including an LP solver in the right-hand side of the ODE has also been considered, which requires resolving the LP at every evaluation of the right-hand side of the ODE.

## Methods

This section details the proposed method for numerical integration of DFBA simulations. Theoretical treatment of dynamic systems with LPs embedded and a complete description of the proposed implementation can be found in Harwood et al. (2012). The goal of this work is to present a reliable integration scheme for dynamic flux balance simulation, which are a subclass of dynamic systems with LPs embedded considered in Harwood et al. (2012). In particular DFBA applications face challenges in terms of problem size, non-uniqueness of the solutions, and stiffness which can be overcome by the proposed method. First, flux balance models are formally introduced and the uniqueness of solutions of LPs is discussed. Then the dynamic states are introduced and their interconnection with the FBA model is formulated. Finally, the implementation as a system of differential algebraic equations (DAEs) is presented.

### Stoichiometric Model

The stoichiometry of the metabolic reaction network is derived from the balance equations for the metabolites. Let $\mathbf{v} \in \mathbb{R}^{n_v}$ be the vector of $n_v$ fluxes, called the *flux distribution*, and let $\mathbf{z} \in \mathbb{R}^{n_m}$ be the vector of $n_m$ metabolites. The mass balance equations for a metabolism can be expressed as $\dot{\mathbf{z}} = \mathbf{S}\mathbf{v}$, where the dotted variables denote time derivatives of the variables. It is assumed that the transients inside the metabolism are fast compared to the dynamics of extracellular environment and the metabolic network is at quasi steady-state (Stephanopoulos et al., 1998). Hence, the time derivatives of the metabolites are set to zero, which results in a set of linear algebraic equations for the flux distribution. In general the metabolic network has more fluxes than metabolites, which implies that the algebraic equations $\mathbf{S}\mathbf{v} = \mathbf{0}$ are underdetermined. The null space of $\mathbf{S}$ defines the subspace of allowable flux distributions for the metabolism, that is, there are an infinite number of flux distributions that satisfy this mechanistic model. This subspace is further constrained by other biochemical considerations, topological or environmental constraints or explicit expressions in terms of the extracellular concentrations. These additional constraints are expressed by inequality constraints of the form

$$\mathbf{v}^L \leq \mathbf{v} \leq \mathbf{v}^U \tag{1}$$

where $v_i^L, v_i^U \in \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ for all $i = 1, \ldots, n_v$. Some biochemical objective of the metabolism, expressed by the cost vector $\mathbf{c}$, is specified to optimize the

underdetermined set of equations under the inequality constraints (1). A FBA model is therefore a linear program that determines the optimal value and has the form

$$\begin{aligned} \mu = \max_{\mathbf{v}\in\mathbb{R}^{n_v}} \ & \mathbf{c}^{\mathrm{T}}\mathbf{v} \\ \text{s.t. } & \mathbf{S}\mathbf{v} = \mathbf{0} \\ & \mathbf{v}^L \le \mathbf{v} \le \mathbf{v}^U \end{aligned} \tag{2}$$

**Remark 2.1.** The proposed algorithm is formulated for LPs in standard form

$$\begin{aligned} \mu = \min_{\overline{\mathbf{v}}\in\mathbb{R}^{n_{\overline{v}}}} \ & \overline{\mathbf{c}}^{\mathrm{T}}\overline{\mathbf{v}} \\ \text{s.t. } & \mathbf{A}\overline{\mathbf{v}} = \mathbf{b} \\ & \overline{\mathbf{v}} \ge \mathbf{0} \end{aligned} \tag{3}$$

where $\overline{\mathbf{v}}$, $\overline{\mathbf{c}} \in \mathbb{R}^{n_{\overline{v}}}$, $\mathbf{b} \in \mathbb{R}^{n_c}$ with $n_c \ge n_m$ and $\mathbf{A} \in \mathbb{R}^{n_c \times n_{\overline{v}}}$. Note that maximizing $\mathbf{c}^{\mathrm{T}}\mathbf{v}$ is equivalent to minimizing $-\mathbf{c}^{\mathrm{T}}\mathbf{v}$. It is well known that any linear program can be rewritten in standard form Bertsimas and Tsitsiklis (1997). Specifically, let $\mathbf{s}$ denote the slack variables and let $\overline{\mathbf{v}} = [(\mathbf{v} - \mathbf{v}^L)^T, \mathbf{s}]$, then maximization of $\mathbf{c}^{\mathrm{T}}\mathbf{v}$ is equivalent to minimization of $\overline{\mathbf{c}}^{\mathrm{T}}\overline{\mathbf{v}}$ with $\overline{\mathbf{c}} = [-\mathbf{c}, \mathbf{0}]$. With slight abuse of notation we drop the bars over $\mathbf{c}$ and $\mathbf{v}$ and continue to use $\mathbf{c}$ and $\mathbf{v}$ for LPs in standard form as the FBA model. Furthermore, it is assumed that the matrix $\mathbf{A}$ has full row rank.

Note that the optimal cost of a linear program is unique, but in general the solution vector is not, a fact which is sometimes referred to as the *existence of alternate solutions*. Resolution of this non-uniqueness has been discussed in the context of DFBA (Mahadevan and Schilling, 2003). Also note that the optimal cost and solution depends continuously on the right-hand side of the equality constraints (Bertsimas and Tsitsiklis, 1997). This is reflected in the expression

$$\begin{aligned} \widehat{U}(\mathbf{b}) = \ & \arg\min_{\mathbf{v}\in\mathbb{R}^{n_v}} \ \mathbf{c}^{\mathrm{T}}\mathbf{v} \\ \text{s.t. } & \mathbf{A}\mathbf{v} = \mathbf{b} \\ & \mathbf{v} \ge \mathbf{0} \end{aligned} \tag{4}$$

where $\widehat{U}(\mathbf{b}) \subset \mathbb{R}^{n_v}$ is called *the solution set* for the right-hand side $\mathbf{b} \in \mathbb{R}^{n_c}$. It can be described as a (set-valued) function. Denote the set of all right-hand sides for which there exists a finite solution by $F = \{\mathbf{A}\mathbf{v} \in \mathbb{R}^m : \mathbf{v} \ge \mathbf{0}\}$ and let $\widehat{\mathcal{S}} = \{\widehat{U}(\mathbf{b}) : \mathbf{b} \in F\} \subset \mathcal{P}(\mathbb{R}^{n_v})$ be the set of all solution sets, where $\mathcal{P}(\mathbb{R}^{n_v})$ denotes the power set of $\mathbb{R}^{n_v}$. A unique solution can be specified by introducing a Lipschitz continuous map $\mathbf{g} : \mathcal{S} \to \mathbb{R}^{n_u}$, $n_u \le n_v$ such that $\mathbf{u}(\mathbf{b}) \equiv \mathbf{g}(\widehat{U}(\mathbf{b}))$ is unique for all $\mathbf{b}$, where $\mathcal{S} = \widehat{\mathcal{S}} \cup \{\{\mathbf{u}\} : \mathbf{u} \in V : V \in \widehat{\mathcal{S}}\}$. For example, $\mathbf{g}$ can be implemented as a secondary LP that maximizes the production of some desired metabolic product, such as ethanol, for a fixed optimal growth rate. In general, for each additional dimension of $\mathbf{u}$, an additional optimization has to be formulated to guarantee uniqueness of the solution. An implementation of such a hierarchical optimization without explicitly solving multiple LPs at every integration step is discussed in Implementation Section. Another implementation of the map $\mathbf{g}$ could be based on the geometry of the flux cone to resolve the non-uniqueness of the optimal flux distribution. This approach is limited by the size of models that can be considered (Smallbone and Simeonidis, 2009).

## Dynamic System

The dynamic mass balance equations of the relevant chemical species in the bioreactor are ordinary differential equations (ODEs). They are coupled with the flux balance models of the mircoorganisms via the growth rate, uptake and excretion kinetic expressions. The growth rate is the optimal cost of the embedded LP, the uptake and excretion kinetics are typically expressed as a Monod reaction in terms of the limiting substrates, that is, expressing the bounds in Equation (1) as functions of the process variables. Additional inhibition terms can also be included in the expressions. For the LP in standard form (3) this translates into functional dependence on the right-hand side of the equality constraints. The mass balance for the extracellular species and the biomass are described as a system of ordinary differential equations

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t),\ \mathbf{u}(t, \mathbf{x}(t))) \tag{5}$$

where the dynamic state vector $\mathbf{x}$ represents the concentrations of the chemical species inside the bioreactor, for example, glucose, oxygen and ethanol, and the initial conditions are given by $\mathbf{x}(t_0) = \mathbf{x}_0$. The right-hand side $\mathbf{f}$ depends on the solution $\mathbf{u}(t, \mathbf{x}(t)) \equiv \mathbf{g}(U(t, \mathbf{x}(t))$ of the linear program

$$\begin{aligned} U(t, \mathbf{x}(t)) = \arg\min_{\mathbf{v}\in\mathbb{R}^{n_v}} \ & \mathbf{c}^{\mathrm{T}}\mathbf{v} \\ \text{s.t. } & \mathbf{A}\mathbf{v} = \mathbf{b}(t, \mathbf{x}(t)) \\ & \mathbf{v} \ge \mathbf{0} \end{aligned} \tag{6}$$

with $\mathbf{b} : \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_c}$ with $U(t, \mathbf{x}(t)) = \widehat{U}(\mathbf{b}(t, \mathbf{x}(t)))$. We refer to (5) and (6) as a *dynamic system with a LP embedded*.

**Remark 2.2.** From a theory and application point of view it makes sense to require existence and uniqueness of the solution to (5). We assume that the technical conditions on the ODEs (5) and LP (6), detailed in Harwood et al. (2012), are satisfied such that a unique solution to the initial value problem exists over the time horizon of interest, for example, the duration of the batch.

## Implementation

One way to approximate the solution of Equation (5) numerically is to embed a linear program solver in the function evaluation subroutine for a numerical integrator,

referred to as the direct approach. However, this is a fairly naïve approach. First, compared to the method that follows, solving a LP is more computationally expensive. Second, the naïve formulation does not provide analytical derivative information of which most advanced numerical integrators can take advantage and are important for optimization. The combined effect of these two facts makes the numerical integration of (5) slow. Third, the solution of a LP such as (6) is a continuous, but not smooth function of $\mathbf{b}$. Consequently, the right-hand side of the ODE is nonsmooth thus violates the assumptions required for the use of implicit numerical integration algorithms. The main advantage in the proposed approach is gained by reformulating the system (5) as a hybrid DAE system using standard optimization theory and integration of this system with a suitable algorithm. Furthermore, integration near the boundary of the set $F$, where the embedded LP is feasible can cause numerical difficulties, since evaluation of the right-hand side of the ODE outside of $F$ is not possible. The reformulation proposed here provides a robust approach to integrate along such solutions.

It is well known that optimality conditions for linear programs are given by a set of algebraic equations and inequalities known as the Karush–Kuhn–Tucker (KKT) conditions (Bertsimas and Tsitsiklis, 1997). This fact also has been recognized in applications of DFBA (Hjersted and Henson, 2006), but only in the context of batch and fed-batch optimization where the natural bilevel optimization formulation is reformulated as a mathematical program with equilibrium constraints, which, as discussed before, is limited to small-scale FBA models. In order to describe the details of the implementation, basic concepts on solutions of LPs are recalled, and more details can be found for example in Bertsimas and Tsitsiklis (1997). The feasible set of an LP is a convex polyhedron. The vertices of this polyhedron are called *basic feasible solutions*. An optimal solution will occur at one of these vertices. This vertex can be described by a basis set. For a *nondegenerate solution*, this basis set will be $B = \{i : v_i \neq 0\}$. That is, it contains the indices $i$ of the nonzero variables $v_i$. $B$ will contain exactly $m$ (the number of equality constraints) elements. For a *degenerate solution*, $B$ will again contain $m$ elements, but one or more of these indices will correspond to a variable that equals zero. For ease of notation, $B(j)$ will denote the elements of $B = \{B(j) : j = 1, \ldots, m\}$. Define the *basic variables* as $\mathbf{v}_B = [v_{B(1)} \cdots v_{B(m)}]^{\mathrm{T}}$, and the *nonbasic variables* as $\mathbf{v}_{\mathrm{NB}} = \{v_i : i \notin B\}$. The nonbasic variables are also called the active variables (the inequality constraint $v_i \geq 0$ is "active"—holds with equality—for them). Similarly, the basis matrix $\mathbf{A}_B$ is a nonsingular square matrix formed from the columns of the constraint matrix $\mathbf{A}$. If $\mathbf{a}_i$ is the $i$th column of $\mathbf{A}$, then $\mathbf{A}_B = [\mathbf{a}_{B(1)} \cdots \mathbf{a}_{B(m)}]$.

The integration scheme presented here integrates the differential algebraic equations of the ODE and the algebraic KKT conditions. The code that has been developed to integrate systems of the form (5) efficiently and accurately is called DSL48LPR. It employs DAEPACK (Tolsma and Barton, 2000) component DSL48E, and it is a numerical integrator for a dynamical system with an embedded LP, where the dynamic states determine the right-hand side of the equality constraints. DSL48E can numerically integrate index-one DAEs and detect user-defined events (Park and Barton, 1996), which may represent a discrete change in the definition of the system behavior. In the most general sense, DSL48LPR solves (6), the embedded LP, once using the simplex algorithm (Bertsimas and Tsitsiklis, 1997) and records the basis set $B$. Then the solution set (while this basis is optimal) is obtained very simply. The basic variables of the solution set are given by $\mathbf{v}_B = \mathbf{A}_B^{-1}\mathbf{b}(t, \mathbf{x}(t))$ and the non-basic variables are simply zero. Thus one obtains a DAE system

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(t, \mathbf{x}(t),\ \mathbf{c}^{\mathrm{T}}\mathbf{v}(t)) \\ \mathbf{A}_B\mathbf{v}_B(t) &= \mathbf{b}(t,\ \mathbf{x}(t)) \\ \mathbf{v}_{\mathrm{NB}}(t) &= \mathbf{0} \end{aligned}$$

A basis set is optimal until it becomes infeasible; this is indicated by a basic variable crossing zero. Such an event can be detected accurately by DSL48E using the algorithm of Park and Barton (1996), integration is stopped at the event, and the basis set is updated by a LP solve before continuing integration. In the context of DFBA, a basis change occurs when a qualitative change in the flux distribution is enforced by changes in the extracellular environment. In contrast, the optimal flux distribution does not change qualitatively between basis changes, meaning that the optimal pathways vary in their weighting over time, but fluxes can only be added or subtracted at a basis change. Consequently, for many of the integration steps, a system of linear equations is solved, rather than a linear program (indeed, the DSL48E algorithm only factors the matrix very infrequently, on average every 10 time steps). This framework provides an efficient method for solving a system of ODEs with embedded LPs. The main advantages that this approach has over existing implementations are that the integration scheme has a variable step size and the linear program only has to be solved when qualitative changes in the optimal flux distribution of the metabolic network occur.

As already mentioned, the existence of alternate solutions can be resolved by formulating a secondary optimization problem. This essentially involves optimizing over the solution set of one linear program hopefully to obtain a uniquely valued function. To be precise, one possible formulation is to let $\mathbf{g} : \mathcal{S} \to \mathbb{R}^2$ be defined by

$$\begin{aligned} g_1(V) = \min_{\mathbf{v} \in \mathbb{R}^{n_v}}\ & \mathbf{c}^{\mathrm{T}}\mathbf{v} \\ \text{s.t. } & \mathbf{A}\mathbf{v} = \mathbf{b} \\ & \mathbf{v} \geq \mathbf{0} \end{aligned} \tag{7}$$

$$\begin{aligned} g_2(V) = \min_{\mathbf{v} \in \mathbb{R}^{n_v}}\ & \mathbf{c}_2^{\mathrm{T}}\mathbf{v} \\ \text{s.t. } & \mathbf{A}\mathbf{v} = \mathbf{b} \\ & \mathbf{c}^{\mathrm{T}}\mathbf{v} = \mathbf{c}^{\mathrm{T}}\mathbf{u} \\ & \mathbf{v} \geq \mathbf{0} \end{aligned} \tag{8}$$

where $\mathbf{u}$ is any element of $V$, the solution set, of the linear program (7), and $\mathbf{c}_2 \in \mathbb{R}^{n_v}$ is the cost vector of the secondary optimization. For example, $\mathbf{c}_2$ can be chosen such that $\mathbf{c}_2^{\mathrm{T}}\mathbf{v}$ describes the ethanol flux. The *reduced cost* of a nonbasic variable $u_j$ for an optimal basis $B$ is given by $c_j - \mathbf{c}_B^{\mathrm{T}}\mathbf{A}_B^{-1}\mathbf{a}_j$. It can be shown that if all reduced costs are zero, then the optimal solution is unique. This provides a simple test to determine if a specific flux is unique, and since the reduced cost does not depend on the right-hand side, the flux is also unique as long as the basis is optimal. LP solvers based on the simplex algorithm provide the reduced costs for an optimal basis. If all reduced costs are zero then the solution is unique. If a reduced cost is greater than zero then $\widehat{B} = \{j\} \cup B$ is a basis for the LP (8). The strong similarity between the two programs and their optimal bases can be exploited such that calculating an optimal solution point of (8) is reduced to calculating an optimal solution point of the lower-level LP (6) using a specific basis $B^*$. That is, $B^*$ is specifically chosen as a basis describing an optimal point for both LP (6) and (8), and so effectively only one basis set and set of primal variables must be calculated to determine the solutions of both linear programs required to define the function $\mathbf{g} \circ \widehat{U}$. Furthermore, this process can be repeated to add another objective leading to a hierarchical optimization in which possibly all fluxes can be determined uniquely (see Harwood et al., 2012 for details).

One challenge that is shared by SOA and the direct approach, discussed in Introduction Section, is that near the boundary of the set of state variable values for which the embedded LP has a solution, also referred to as the *viability domain*, evaluation of the right-hand side of the ODE is not possible if the integrator attempts to evaluate it outside the viability domain (e.g., if finite differences are used for derivative approximation). The approach presented here circumvents this situation by "locking" the model in a fixed mode, described by expressing the optimal solution in terms of an optimal basis. In this mode the viability domain is extended for numerical computation and evaluation outside of the domain and will not cause errors, since the algorithm does not try to solve an infeasible LP, but only uses a basis that is not optimal.

The current implementation uses CPLEX and includes a C program that uses the CPLEX Callable Library to construct the standard form LPs that are embedded in the system and subsequently solves them at events. However, the user can easily replace this with a call to their preferred LP solver. Additional user-defined and possibly discontinuous functions can also be added to the right-hand side of the equality constraints. The functions have to satisfy the condition described in Harwood et al. (2012). Examples of the type of functions are provided in Results Section. A sample implementation based on the first case study is included in the Supplementary Information A.3.

## Results

Three case studies are presented to illustrate the proposed algorithm. The formulation is based on the dynamic multispecies metabolic modeling (DMMM) framework (Zhuang et al., 2011). Since a limiting factor in previous DFBA simulations has been the size of the FBA model, two of the largest FBA models are considered. In all case studies the dynamic mass balance equations of the extracellular environment of the batch or fed-batch bio-reactor have the general form

$$\dot{y}_j(t) = \mu_j(t, \mathbf{x}(t))y_j(t) - D(t)y_j(t), \quad j = 1, \ldots, N \tag{9}$$

$$\dot{s}_i(t) = \sum_{j=1}^{N} \nu_i^j(t, \mathbf{x}(t))y_j(t) + D(t)F_i(t) - D(t)s_i(t), i = 1, \ldots, M \tag{10}$$

$$\dot{V}(t) = \sum_{i=1}^{M} F_i(t) \tag{11}$$

where $N$ is the number of biological species and $M$ is the number of extracellular chemical species. The variable $y_j$ is the biomass concentration of the $j$th metabolism, $s_i$ is the extracellular concentration of the $i$th chemical species, $F_i$ is the feed rate of chemical species $i$, $D = \sum_{i=1}^{M} F_i/V$ is the dilution rate, $\nu_i^j$ is the uptake or production rate of chemical species $i$ due to species $j$ and $V$ is the volume of the bio-reactor. The functions $\nu_i^j$ can have components of $\mathbf{g}$ as arguments, for example when the flux of a metabolic product is determined through secondary optimization. The complete dynamic state vector is denoted by $\mathbf{x} = [\mathbf{y}^{\mathrm{T}}\ \mathbf{s}^{\mathrm{T}}\ V]^{\mathrm{T}}$. Furthermore, we assume that an initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$ is given such that all flux balance models have at least one feasible solution which continues to exists for some time. The growth rates $\mu_j(t, \mathbf{x}(t))$ are determined by the flux balance models:

$$\begin{aligned} \mu_j(t, \mathbf{x}(t)) = \min_{\mathbf{v}^j \in \mathbb{R}^{n_j}} \ & \mathbf{c}^{j\mathrm{T}}\mathbf{v}^j, \quad j = 1, \ldots, N \\ \text{s.t. } & \mathbf{A}^j\mathbf{v}^j = \mathbf{b}^j(t, \mathbf{x}(t)) \\ & \mathbf{v}^j \geq \mathbf{0} \end{aligned} \tag{12}$$

Note that the $N$ linear programs can be written as one linear program of the form

$$\begin{aligned} \mu(t, \mathbf{x}(t)) = \sum_{j=1}^{N} \mu_j(t, \mathbf{x}(t)) = \min_{\substack{\mathbf{v}^j \in \mathbb{R}^{n_j} \\ j \in \{1, \ldots, N\}}} & \sum_{j=1}^{N} \mathbf{c}^{j\mathrm{T}}\mathbf{v}^j \\ \text{s.t. } \mathbf{A}^j\mathbf{v}^j = \mathbf{b}^j(t, \mathbf{x}(t)) & \\ \mathbf{v}^j \geq \mathbf{0}, \quad j = 1, \ldots, N & \end{aligned} \tag{13}$$

Furthermore, the map $\mathbf{g}$ can be implemented in this case as $\mathbf{g} \circ \widehat{U} = [\mu_1 \cdots \mu_N]^{\mathrm{T}}$. Hence, this formulation is of the assumed form (5) and (6).

### *E. coli* Batch Fermentation

In the first case study, we consider the metabolic network model of the *E. coli* K-12 bacterium, iJR904 (Reed et al., 2003), which is available online (Schellenberger et al., 2010). This study is based on the study presented in Hanly and Henson (2011). The model consists of 625 unique metabolites in two compartments, it has 931 intracellular, 144 exchange fluxes and an additional flux representing the biomass generation as growth rate $\mu$. The extracellular metabolites are glucose ($g$), xylose ($z$), and ethanol ($e$). Hence, $\mathbf{s} = [g\, z\, e]^{\mathrm{T}}$. The uptake kinetics for glucose, xylose, and oxygen are given by the Michaelis–Menten kinetics

$$v_g(t) = v_{g,\max} \frac{g(t)}{K_g + g(t)} \frac{1}{1 + \frac{e(t)}{K_{ie}}} \tag{14}$$

$$v_z(t) = v_{z,\max} \frac{z(t)}{K_z + z(t)} \frac{1}{1 + \frac{e(t)}{K_{ie}}} \frac{1}{1 + \frac{g(t)}{K_{ig}}} \tag{15}$$

$$v_o(t) = v_{o,\max} \frac{o(t)}{K_o + o(t)} \tag{16}$$

where $v_{g,\max}$, $v_{z,\max}$, $v_{o,\max}$, $K_g$, $K_z$, $K_{ie}$, $K_{ig}$, and $K_o$ are known constant parameters. The additional inhibition factor dependent on the ethanol concentration accounts for reduced glucose consumption in a high ethanol concentration environment. Similarly, the xylose uptake rate is limited by high glucose and ethanol concentrations. It is assumed that the oxygen concentration in the reactor is controlled and no mass balance for oxygen is considered, therefore $o(t)$ is a user-defined piecewise constant function. The ethanol exchange flux $v_e$ is determined by implementing the secondary optimization as discussed above and was unique through the whole batch cultivation. The functions $v_i^j = v_i$ in (10) are $v_1 = -v_g$, $v_2 = -v_z$, and $v_3 = v_e$. The metabolic network model is connected to the extracellular environment through the exchange fluxes for glucose, xylose and oxygen by adding Equations (14)–(16) to the equality constraints. The total number of constraints is then 764, accounting also for multiplicity of the metabolites between compartments. Note that the *E. coli* model has linearly dependent constraints, and since DSL48LPR assumes that the stoichiometry matrix has full row rank, linearly dependent equations are deleted. This reduced the number of equality constraints to 749. In standard form, the flux balance model has 749 constraints and 2,150 variables, which include additional slack variables for the inequality constraints.

The simulation represents a batch operation of the bioreactor under anaerobic conditions on glucose and xylose media. The time evolution of the dynamic states and a representative selection of the LP variables are shown in Figure 1. First glucose, as the preferred carbon source is consumed. After approximately 18 h glucose has been depleted, hence the embedded LP has to be resolved to find an alternative flux distribution that maximizes the growth, that is, the optimal basis changes. The new flux distribution is such that xylose becomes the main carbon source. The final batch time is reached when the glucose concentration is less than 0.1 g/L and the xylose concentration is less than 1.5 g/L (Hanly and Henson, 2011), which are minimal requirements for anaerobic growth. The vertical lines in Figure 1 indicate basis changes, which coincide with the switch from glucose to xylose as carbon source, and the final batch time when minimal growth conditions are no longer met. Note that apparent discontinuity in the fluxes is in fact a rapid transient induced by the basis change, indicating that the model is extremely stiff. All numerical parameter values including the initial conditions are according to Hanly and Henson (2011). A summary of all model parameters is included in the Supplementary Information A.2.

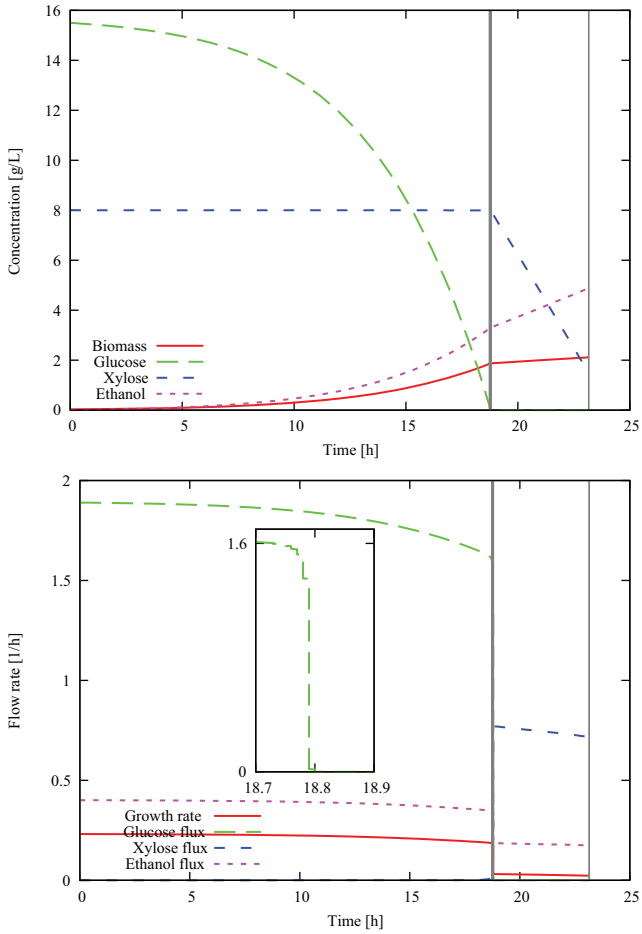


**Figure 1.** *E. coli* simulation: species concentration in bioreactor under anaerobic conditions (**top**); growth rates and exchange fluxes corresponding to the species in the bioreactor. Magnification of the glucose flux close to depletion of glucose shows stiff behavior of the DFBA model (**bottom**). Vertical dotted lines indicate changes of the optimal basis.

### *Performance of the Algorithm*

One difficulty encountered in the naïve approach of integration DFBA models is that when the value of the dynamic state approaches the boundary of the set of state values where a solution of the FBA model exists, integration is prone to failure. DSL48LPR is able to detect such events, as basis changes, and can integrate along the boundaries of these sets as long as a solution exists. In the example, the glucose flux rapidly goes to zero as the glucose concentration goes to zero. Any continuation of this trajectory beyond zero glucose flux results in a linear program that is infeasible. Advanced stiff integrators can recover from this situation by reducing the integration steps size only if some arbitrary, and at this point meaningless, solution to the LP is provided. Furthermore, this can drastically increase the computation time. Table II shows a comparison of the direct method with different Matlab solvers for ordinary differential equations (ode23, ode23s, and ode15s) and the implementation in DSL48LPR. For the first three integrators the LP solution had to be set to zero at infeasible points to complete the simulation, in contrast DSL48LPR completes the simulation faster and without attempting to solve an infeasible LP during the integration.

Furthermore, we have implemented a DFBA simulation with the most up-to-date *E. coli model* iJO1366 (Orth et al., 2011), assuming the same parameter for the uptake kinetics as described above. The computation time is doubled compared to the iJR904 model. This was expected since the computation time for DFBA models is mainly dependent on the size of the flux distribution and the number of fluxes in the iJO1366 model is roughly double compared to the iJR904 model (1,075 vs. 2,583 fluxes).

## *S. cerevisiae* Batch Fermentation

Batch operation of a bioreactor with *S. cerevisiae* under aerobic growth on glucose is simulated as a second example. This example illustrates how the hierarchical optimization is formulated to determine components of the flux distribution uniquely. Also, discontinuities in the right-hand side of the linear program are implemented. The main substrate during the fermentation is glucose, but *S. cerevisiae* can also consume ethanol under aerobic conditions. Following Hanly and Henson (2011), the ethanol exchange flux is assumed to be unconstrained, since ethanol diffuses freely through the plasma membrane. The ethanol uptake is only indirectly limited by the oxygen uptake. In order to model this behavior, the oxygen uptake rate has to be increase when glucose is not available. This is implemented by increasing the upper bound of the oxygen uptake rate from 1.5 to 8.0 mmol/gdw/h (see Hanly and Henson, 2011 for details). The extracellular chemical species for this simulation are glucose ($g$) and ethanol ($e$). The genome-scale network reconstruction of the *S. cerevisiae* metabolism iND750 (Duarte et al., 2004) was considered based on previous studies by Hjersted et al. (2007). The FBA model has 645 unique metabolites in eight compartments and 1,265 intracellular and exchange fluxes and an additional flux representing the biomass generation of growth rate $\mu$. The uptake kinetics for glucose is of the form (14). All numerical parameter values are according to Hjersted et al. (2007). The embedded LP in standard form has 984 constraints and 2,528 variables.

The map **g** is implemented by applying the procedure outlined in Implementation Section to determine the ethanol fluxes uniquely. In order to determine the range of ethanol production predicted by the model, the secondary optimization is first set to maximize the ethanol production. This determines the upper bound of ethanol production as a function of time. Then, a second simulation in which minimization of ethanol is the objective in the hierarchical optimization determines the lower bound of ethanol production. It can be observed that the ethanol flux is unique as long as glucose is consumed. The step change which occurs at the switch from glucose to ethanol as substrate introduces a discontinuity on the right-hand side of the equality constraints of the LPs, since this is also a user-defined event the timing of the event can be accurately calculated by DSL48LPR and integration restarted. The time evolution of the range of ethanol concentration predicted by the model is indicated by the gray shaded area in Figure 2. In addition to a representative selection of the flux distribution, the range of the ethanol flux, also referred to as the production envelope (Mahadevan and Schilling, 2003), is shown in the bottom half of Figure 2.

**Table II.** Comparison of different integrator performances for the *E. coli* batch fermentation example.

| | ODE23 (s) | ODE23S (s) | ODE15S (s) | DSL48LPR (s) |
|---|---|---|---|---|
| Integration interval $[0, t_{v_g=0}]$ | 3.429 | 9.266 | 2.508 | 0.858 |
| Integration interval $[0, t_f]$ | 360.94[a] | 26.682[a] | 10.511[a] | 3.077 |

[a]LP solution was set to zero at infeasible points.

## Co-Culture Batch Fermentation

As mentioned in the introduction, analysis, and synthesis of consortia of metabolisms will most likely play a vital role in the development of future biofuels. For example, two important intermediates in the conversion of ligneous biomass to ethanol are glucose and xylose (Hanly and Henson, 2011). In this example a co-culture of noncompeting *E. coli* and *S. cerevisiae* grown on glucose and xylose media is simulated. The *E. coli* strain ZSC113 (Eiteman et al., 2008), which is modified such that it does not process glucose, is considered, since wild-type *E. coli* competes with *S. cerevisiae* for glucose. The glucose uptake for the *E. coli* model iJR904 was constrained to zero to simulate the genetic modification that prohibits the glucose consumption. The

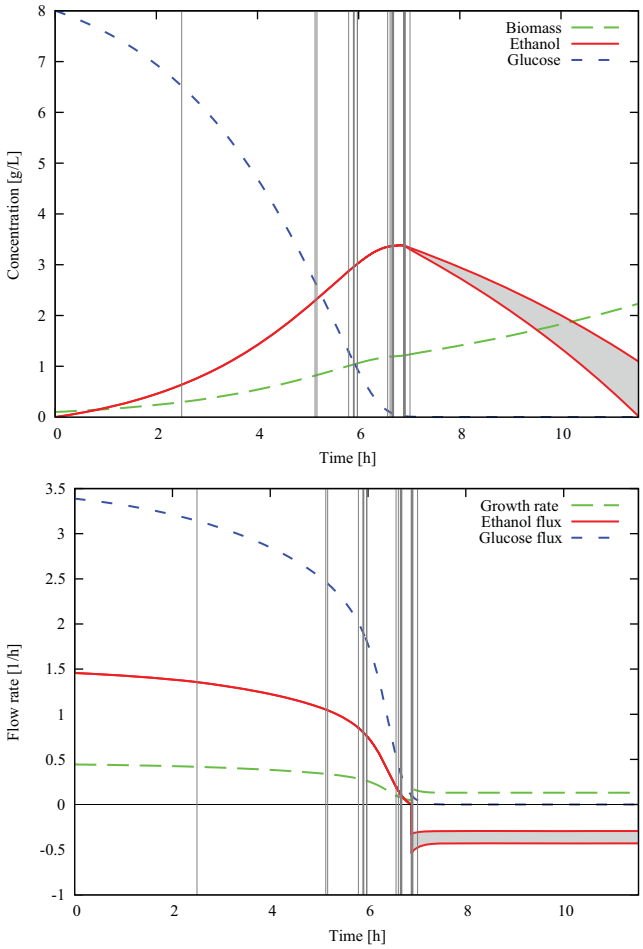


**Figure 2.** *S. cerevisiae* simulation: species concentration in bioreactor aerobic conditions (**top**); Growth rates, exchange fluxes exchange fluxes corresponding to the species in the bioreactor (**bottom**). The ethanol production envelope shown in the gray area between maximal and minimal ethanol flux (solid red lines). Vertical dotted lines indicate changes of the optimal basis.

only additional modification to the original iJR904 model was to set the lower bound of the non-growth associated maintenance (NGAM) to 10.5 mmol/gdw/h. This adjustment gives better prediction for non-optimal growth conditions with the uptake parameter values based on optimal growth conditions (see Hanly et al., 2012 for details). The wild-type *S. cerevisiae* does not process xylose, hence the iDN750 model with the modification described in *E. coli* Batch Fermentation Section is considered. To accurately predict anaerobic growth, the iDN750 model had to be changed as detailed in the Supplementary Information. For both metabolisms a secondary objective of maximal ethanol production is implemented. The dynamic equations for this model are of the form (9) where $y_1$ and $y_2$ denotes the *E. coli* and *S. cerevisiae* biomass concentration, respectively. The extracellular chemical species are glucose ($g$), xylose ($z$), and ethanol ($e$). Hence, the functions $v_i^j$ in (10) are $v_1^1 = -v_g$, $v_1^2 = v_2^1 = 0$, $v_2^2 = -v_z$, $v_3^1 = v_e^1$, and $v_3^2 = v_e^2$. The two metabolisms interact only indirectly through the common product ethanol which inhibits the substrate uptake for both metabolism. The kinetic parameters for both models presented in Hanly and Henson (2011) are used here. The combined embedded LP representing both microorganism has 1,738 constraints and 4,682 variables in standard form. For this simulation study, a switch from aerobic ($o = 0.24$ mmol/L) to anaerobic condition is implemented and the ratio between *E. coli* and *S. cerevisiae* inoculum could be varied while the total inoculum was set to 0.1 g/L. The switch time and the inoculum ratio are considered as free process parameters. The optimal combination of inoculum ratio and switch time is determined by repeated simulation for various combinations. Figure 3 shows a two-dimensional parameter scan of the ethanol production, defined as the final ethanol concentration divided by the batch time, with respect to the aerobic–anaerobic switching time and the *S. cerevisiae* inoculum. Graphical inspection puts the optimal combination of switching time and *S. cerevisiae* inoculum at 5.2 h and 0.085 g/L. The final batch time for all simulations is reached when the minimal growth conditions for the individual metabolism is reached. The *E. coli* strain has minimal required xylose flux of 3.75 mmol/gdw/h, this was implemented as a user-defined condition to terminate the simulation. The minimal glucose concentration was set to 1.5 g/L. The time plots of the dynamic states and a representative selection of the fluxes are shown in Figure 4.

## Conclusions

An efficient and reliable numerical integrator for dynamic flux balance analysis models has been presented. By reformulating the dynamic flux balance model as hybrid differential algebraic equations, the embedded flux balance

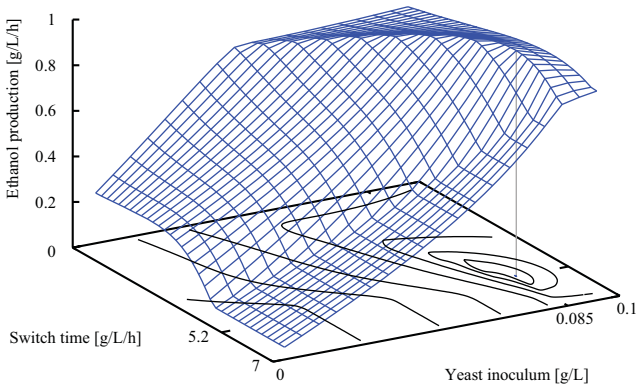


**Figure 3.** Co-culture simulation: Ethanol production as a function of switching time $t_f$ and *S. cerevisiae* inoculum $y_2(0)$. Every grid points represents a single DFBA simulation. The maximal ethanol production at (5.2 h, 0.085 g/L). Contour lines for $\frac{e(y_2(0),t_f)}{t_f} = [0.2, 0.4, 0.6, 0.8, 0.9, 0.975, 0.99]$.

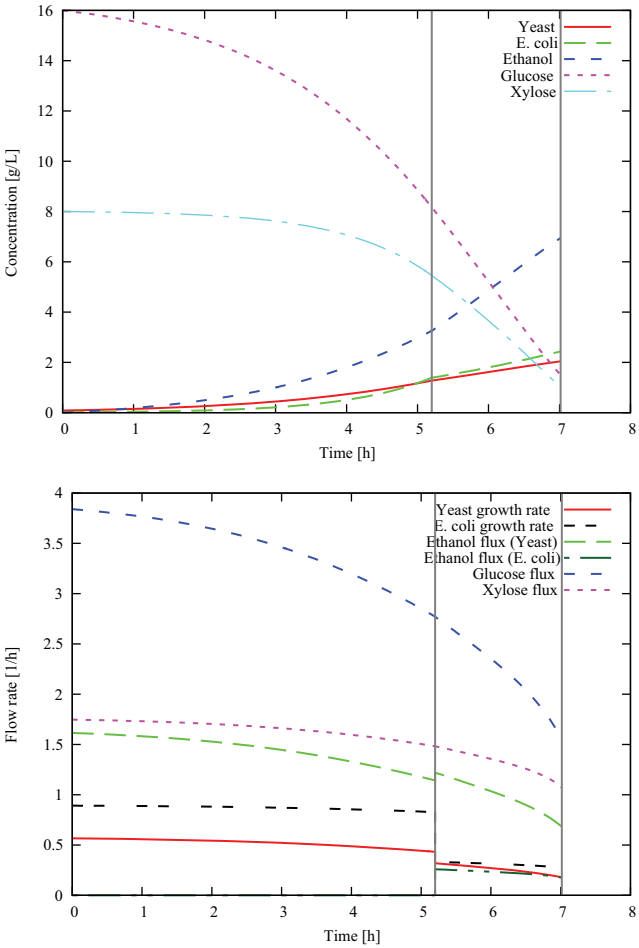


**Figure 4.** Co-culture simulation: species concentration in co-culture bioreactor for optimal combination of switching time and *S. cerevisiae* inoculum as shown in Figure 3 (**top**); growth rates and exchange fluxes (**bottom**). Vertical dotted lines indicate changes of the optimal basis, the basis changed at $t = 5.2$ h is forced by a step change in the oxygen concentration.

LP only has to be solved when the structure of the optimal flux distribution changes, which is a small number compared to the number of integration steps per simulation. Furthermore, hierarchical optimization was implemented to determine elements of the optimal flux distribution uniquely with minimal additional cost. Hierarchical optimization is required when the time evolution of some metabolic product should be predicted. The implementation makes use of the strong similarity between the primary and secondary optimization problem and their respective optimal basis. The results were illustrated by considering three genome-scale DFBA simulations from the literature, including one co-culture simulation.

Due to the nonsmooth nature of the solution of the embedded LP, the right-hand side of the dynamic equations is also nonsmooth. Hence, standard gradient-based optimization algorithms which assume smoothness of the data are not suitable. The formulation presented here allows for nonsmooth sensitivity analysis and nonsmooth dynamic optimization based on bundle solver methods and the numerical integrator DSL48LPR is well suited for such a task. Using the concept of the generalized Jacobian, a set-valued map, the dual problem to the embedded LP, and a nonsmooth bundle solver in combination with DSL48LPR an optimal solution can be computed. The implementation of nonsmooth sensitivity analysis for DFBA is the subject of a current research effort and a detailed study on this subject will appear elsewhere.

The first author thanks Michael Henson and his research group for their help on details of the flux balance models.

## References

Anesiadis N, Cluett WR, Mahadevan R. 2008. Dynamic metabolic engineering for increasing bioprocess productivity. Metab Eng 10(5):255–266.

Bertsimas D, Tsitsiklis J. 1997. Introduction to linear optimization. Athena scientific series in optimization and neural computation. Belmont, MA: Athena Scientific.

Covert MW, Schilling CH, Palsson BØ. 2001. Regulation of gene expression in flux balance models of metabolism. J Theor Biol 213(1):73–88.

Duarte NC, Herrgård MJ, Palsson BØ. 2004. Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. Genome Res 14(7):1298–1309.

Edwards JS, Palsson BØ. 2000. The Escherichia coli MG1655 in silico metabolic genotype: Its definition, characteristics, and capabilities. Proc Natl Acad Sci 97(10):5528–5533.

Eiteman M, Lee S, Altman E. 2008. A co-fermentation strategy to consume sugar mixtures effectively. J Biol Eng 2:1–8.

Feist A, Herrgård M, Thiele I, Reed J, Palsson B. 2008. Reconstruction of biochemical networks in microorganisms. Nat Rev Microbiol 7(2):129–143.

Hanly TJ, Henson MA. 2011. Dynamic flux balance modeling of microbial co-cultures for efficient batch fermentation of glucose and xylose mixtures. Biotechnol Bioeng 108(2):376–385.

Hanly TJ, Urello M, Henson MA. 2012. Dynamic flux balance modeling of *S. cerevisiae* and *E. coli* co-cultures for efficient consumption of glucose/xylose mixtures. Appl Microbiol Biotechnol 93:2529–2541.

Harwood SM, Höffner K, Barton PI. 2012. Solution of ordinary differential equations with linear programs embedded: The right-hand side case. SIAM J Sci Comput (submitted for publication).

Hjersted JL, Henson MA. 2006. Optimization of fed-batch *Saccharomyces cerevisiae* fermentation using dynamic flux balance models. Biotechnol Prog 22(5):1239–1248.

Hjersted JL, Henson MA. 2009. Steady-state and dynamic flux balance analysis of ethanol production by *Saccharomyces cerevisiae*. IET Syst Biol 3(3):167–179.

Hjersted JL, Henson MA, Mahadevan R. 2007. Genome-scale analysis of *Saccharomyces cerevisiae* metabolism and ethanol production in fed-batch culture. Biotechnol Bioeng 97(5):1190–1204.

Kauffman KJ, Prakash P, Edwards JS. 2003. Advances in flux balance analysis. Curr Opin Biotechnol 14(5):491–496.

Lee J, Gianchandani E, Eddy J, Papin J. 2008. Dynamic analysis of integrated signaling, metabolic, and regulatory networks. PLoS Comput Biol 4(5):e1000086.

Lequeux G, Beauprez J, Maertens J, Horen EV, Soetaert W, Vandamme E, Vanrolleghem PA. 2010. Dynamic metabolic flux analysis demonstrated on cultures where the limiting substrate is changed from carbon to nitrogen and vice versa. J Biomed Biotechnol 2010: http://www.hindawi.com/journals/jbb/2010/621645/cta/

Luo R, Liao S, Tao G, Li Y, Zeng S, Li Y, Luo Q. 2006. Dynamic analysis of optimality in myocardial energy metabolism under normal and ischemic conditions. Mol Syst Biol 2:0031.

Luo R, Wei H, Ye L, Wang K, Chen F, Luo L, Liu L, Li Y, Crabbe MJC, Jin L, Li Y, Zhong Y. 2009. Photosynthetic metabolism of C3 plants shows highly cooperative regulation under changing environments: A systems biological analysis. Proc Natl Acad Sci 106(3):847–852.

Mahadevan R, Schilling C. 2003. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab Eng 5(4): 264–276.

Mahadevan R, Edwards JS, Doyle FJ. 2002. Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. Biophys J 83(3):1331–1340.

Majewski RA, Domach MM. 1990. Simple constrained-optimization view of acetate overflow in *E. coli*. Biotechnol Bioeng 35(7):732–738.

Meadows AL, Karnik R, Lam H, Forestell S, Snedecor B. 2010. Application of dynamic flux balance analysis to an industrial *Escherichia coli* fermentation. Metab Eng 12(2):150–160.

Milne CB, Kim P-J, Eddy JA, Price ND. 2009. Accomplishments in genome-scale in silico modeling for industrial and medical biotechnology. Biotechnol J 4(12):1653–1670.

Nolan RP, Lee K. 2011. Dynamic model of CHO cell metabolism. Metab Eng 13(1):108–124.

Oddone GM, Mills DA, Block DE. 2009. A dynamic, genome-scale flux model of lactococcus lactis to increase specific recombinant protein expression. Metab Eng 11(6):367–381.

Orth JD, Thiele I, Palsson BØ. 2010. What is flux balance analysis? Nat Biotechnol 28(3):245–248.

Orth J, Conrad T, Na J, Lerman J, Nam H, Feist A, Palsson B. 2011. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. Mol Syst Biol 7:535.

Palsson BØ. 2006. Systems biology: Properties of reconstructed networks. New York, NY: Cambridge University Press.

Park T, Barton PI. 1996. State event location in differential-algebraic models. ACM Trans Model Comput Simul 6(2):137–165.

Pizarro F, Varela C, Martabit C, Bruno C, Pérez-Correa JR, Agosin E. 2007. Coupling kinetic expressions and metabolic networks for predicting wine fermentations. Biotechnol Bioeng 98(5):986–998.

Raghunathan A, Pérez-Correa J, Agosin E, Biegler L. 2006. Parameter estimation in metabolic flux balance models for batch fermentation formulation; solution using differential variational inequalities (DVIs). Ann Oper Res 148:251–270.

Ramakrishna R, Edwards JS, McCulloch A, Palsson BØ. 2001. Flux-balance analysis of mitochondrial energy metabolism: Consequences of systemic stoichiometric constraints. Am J Physiol Regul Integr Comp Physiol 280(3):R695–R704.

Raman K, Chandra N. 2009. Flux balance analysis of biological systems: Applications and challenges. Brief Bioinform 10(4):435–449.

Reed J, Vo T, Schilling C, Palsson BØ. 2003. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). Genome Biol 4(9):R54.

Sainz J, Pizarro F, Pérez-Correa JR, Agosin E. 2003. Modeling of yeast metabolism and process dynamics in batch fermentation. Biotechnol Bioeng 81(7):818–828.

Salimi F, Zhuang K, Mahadevan R. 2010. Genome-scale metabolic modeling of a clostridial co-culture for consolidated bioprocessing. Biotechnol J 5(7):726–738.

Schellenberger J, Park J, Conrad T, Palsson BØ. 2010. BiGG: a biochemical genetic and genomic knowledgebase of large scale metabolic reconstructions. BMC Bioinform 11(1):213.

Schilling C, Edwards J, Letscher D, Palsson B. 2000. Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems. Biotechnol Bioeng 71(4):286–306.

Segré D, Vitkup D, Church GM. 2002. Analysis of optimality in natural and perturbed metabolic networks. Proc Natl Acad Sci USA 99(23):15112–15117.

Smallbone K, Simeonidis E. 2009. Flux balance analysis: A geometric perspective. J Theor Biol 258:311–315.

Stephanopoulos GN, Aristidou AA, Nielsen J. 1998. Metabolic engineering: Principles and methodologies. Waltham, MA: Academic Press.

Tolsma J, Barton PI. 2000. DAEPACK: An open modeling environment for legacy models. Indus Eng Chem Res 39(6):1826–1839.

Vargas F, Pizarro F, Pérez-Correa J, Agosin E. 2011. Expanding a dynamic flux balance model of yeast fermentation to genome-scale. BMC Syst Biol 5(1):75.

Varma A, Palsson BØ. 1994. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. Appl Environ Microbiol 60(10): 3724.

Zhuang K, Izallalen M, Mouser P, Richter H, Risso C, Mahadevan R, Lovley D. 2011. Genome-scale dynamic modeling of the competition between rhodoferax and geobacter in anoxic subsurface environments. ISME J 5:305–316.