BMC Systems Biology

BioMed Central

Research article

Open Access

Hybrid optimization for ^{13}C metabolic flux analysis using systems parametrized by compactification

Tae Hoon Yang*[1,2], Oliver Frick[2] and Elmar Heinzle[2]

Address: [1]James Graham Brown Cancer Center & Department of Surgery, 2210 S. Brook St. Rm 342, Belknap Research Building, University of Louisville, Louisville, KY 40208, USA and [2]Biochemical Engineering Institute, Saarland University, Campus A1.5, D-66123 Saarbrücken, Germany

Email: Tae Hoon Yang* - th.yang@louisville.edu; Oliver Frick - o.frick@mx.uni-saarland.de; Elmar Heinzle - e.heinzle@mx.uni-saarland.de

* Corresponding author

Published: 26 March 2008

BMC Systems Biology 2008, **2**:29 doi:10.1186/1752-0509-2-29

Received: 8 October 2007

Accepted: 26 March 2008

This article is available from: http://www.biomedcentral.com/1752-0509/2/29

© 2008 Yang et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The importance and power of isotope-based metabolic flux analysis and its contribution to understanding the metabolic network is increasingly recognized. Its application is, however, still limited partly due to computational inefficiency. ^{13}C metabolic flux analysis aims to compute *in vivo* metabolic fluxes in terms of metabolite balancing extended by carbon isotopomer balances and involves a nonlinear least-squares problem. To solve the problem more efficiently, improved numerical optimization techniques are necessary.

Results: For flux computation, we developed a gradient-based hybrid optimization algorithm. Here, independent flux variables were compactified into [0, 1)-ranged variables using a single transformation rule. The compactified parameters could be discriminated between non-identifiable and identifiable variables after model linearization. The developed hybrid algorithm was applied to the central metabolism of *Bacillus subtilis* with only succinate and glutamate as carbon sources. This creates difficulties caused by symmetry of succinate leading to limited introduction of ^{13}C labeling information into the system. The algorithm was found to be superior to its parent algorithms and to global optimization methods both in accuracy and speed. The hybrid optimization with tolerance adjustment quickly converged to the minimum with close to zero deviation and exactly re-estimated flux variables. In the metabolic network studied, some fluxes were found to be either non-identifiable or nonlinearly correlated. The non-identifiable fluxes could correctly be predicted *a priori* using the model identification method applied, whereas the nonlinear flux correlation was revealed only by identification runs using different starting values *a posteriori*.

Conclusion: This fast, robust and accurate optimization method is useful for high-throughput metabolic flux analysis, *a posteriori* identification of possible parameter correlations, and also for Monte Carlo simulations to obtain statistical qualities for flux estimates. In this way, it contributes to future quantitative studies of central metabolic networks in the framework of systems biology.

Background

In recent years, metabolic flux analysis (MFA) has become an important tool for quantifying metabolic pathways which is essential for in-depth understanding of biological systems. Among the developed tools, ^{13}C flux analysis utilizing ^{13}C labeling patterns of metabolic products that result from feeding ^{13}C-labeled substrates provides

detailed information about intracellular pathway fluxes *in vivo* [1-3]

^{13}C-based MFA requires carbon flux modeling through the metabolic network, which describes the mathematical relationship between unknown fluxes and the available measurement data set. It requires modeling two connected equation systems, which describe reaction stoichiometry between metabolites and between carbon isotopomers, respectively. Using the model, fluxes can be computed from the measurements by solving a nonlinear least-squares problem (NLSP). The stoichiometric network is a linear equation system of material balances given for the reactions between metabolites, i.e., $\mathbf{S} \cdot \nu = \mathbf{0}$ with $\mathbf{S}$ denoting the stoichiometric matrix and ν the fluxes. The fluxes consist of non-measured intracellular fluxes ν_u and measured effluxes ν_m, i.e., $\nu = (\nu_u, \nu_m)^T$. Typically, realistic models are underdetermined, that is, the rank of $\mathbf{S}$ is smaller than the number of entries in ν_u [1]. This difference between $rank(\mathbf{S})$ and the number of entries of ν_u equals the number of fluxes that have to be chosen as the design parameters Θ (independent flux variables) and are required when parametrizing the network such that $\nu = \mathbf{F}_{flux}(\Theta)$.

Typically, modeling carbon isotopomer networks involves an equation system containing a few to several hundred variables to balance reactions between isotopomers [1,3,4]. The equation system of carbon isotopomer reactions is an implicit function $\mathbf{F}_{carbon}(\mathbf{x}, \nu) = \mathbf{0}$, bilinear but square with respect to carbon isotopomer fractions ($\mathbf{x}$) and nonlinear with respect to fluxes, where $\mathbf{x}$ consists of non-measured $\mathbf{x}_u$ and measured $\mathbf{x}_m$. Among the available tools, the cumomer concept developed by Wiechert *et al*. (1999) provides explicit solutions for carbon isotopomer fractions ($\mathbf{x}$) by transforming the bilinear system into a cascade of linear systems [5]. In addition to this, the explicit partial derivatives of the cumomer network with respect to fluxes are obtainable, which is useful for gradient-based optimization algorithms. When ^{13}C labeling information is applied, intracellular fluxes are determined by means of numerical optimization that seeks a constrained minimum of independent flux variables (Θ) to an objective function, e.g.,

$$\min_{\boldsymbol{\Theta} \in \Re^m} f(\boldsymbol{\Theta}) = \frac{1}{2} \left(\boldsymbol{\eta} - \mathbb{F}(\boldsymbol{\Theta}) \right)^T \cdot \boldsymbol{\Sigma}_\eta^{-1} \cdot \left(\boldsymbol{\eta} - \mathbb{F}(\boldsymbol{\Theta}) \right) \text{subject to } \mathbf{v}(\boldsymbol{\Theta}) \geq \mathbf{0}. \quad (1)$$

Here $f(\Theta)$ denotes the objective function to be minimized with respect to Θ and $\mathbb{F}$ (Θ) signifies the model function corresponding to the measured data set $\eta = (\eta_1, \eta_2, ..., \eta_n)^T$ consisting of measured ^{13}C labeling data ($\mathbf{x}_m$). Such data are typically received as mass isotopomer distributions measured by mass spectrometry (MS) and/or fractional carbon labeling measured by nuclear magnetic resonance (NMR) techniques, and effluxes (ν_m). The measurement error ε is typically assumed to have a normal distribution such that $\varepsilon \in N(\mathbf{0}, \Sigma_\eta)$, where Σ_η is the covariance matrix of measurements.

For ^{13}C-based MFA, the applied algorithms for numerical flux estimation are mainly gradient-based local optimization [6,7] or gradient-free global optimization [8-11] such as simulated annealing (SA) or genetic algorithms (GAs). Also, a hybrid technique of global-local optimization has been applied [12]. Such algorithms are described in detail elsewhere [13-15]. The stochastic global optimization methods can be inefficient due to the time required to obtain the so-called asymptotic convergence or reachability in high dimensional parameter spaces [11,16-21]. Moreover, such algorithms of random nature may fail to find the global solution unless the number of samplings tends to infinity, which is practically impossible [21,22]. In comparison, the gradient-based local optimizations have a much higher convergence speed, but the solution quality depends heavily on starting points [14,19]. Further to this, reaching the global optimum is ensured only for convex problems, whereas it is nontrivial to determine the convexity of general nonlinear problems with nonlinear constraints [23]. Thus, one may obtain solutions that are not necessarily global [14] and that might vary depending on starting points.

In this regard, a robust method is highly desirable which guarantees a convergence with speed and accuracy for the nonlinear flux estimation problem. In the present work, we developed a method for efficient parametrization of metabolic network by compactification. Additionally, a mathematical method is suggested for model identification that solves *a priori* flux identifiability problems regarding ^{13}C labeling experiments in terms of model linearization. On this basis, an optimization algorithm was developed that hybridizes two gradient-based optimization tools. The developed approaches were evaluated using the central metabolism of *Bacillus subtilis* with succinate and glutamate feeding as the only carbon sources. We examined whether global flux solutions are obtainable using the gradient-based deterministic algorithm.

Results

In this section, we describe the mathematical and computational procedures developed and their application to the nonlinear problem of metabolic flux estimation in a realistic metabolic network of *Bacillus subtilis*. The model represents an inherent difficulty arising from only succinate and glutamate feeding as carbon sources. Succinate is a symmetric molecule having only two distinguishable carbon atoms. On the other hand, glutamate is also con-

verted into succinate. Due to this, the information that is introduced by the ^{13}C-labels of the substrates is severely limited.

Parametrization of stoichiometric network

To formulate a non-linear least squares problem (NLSP), the system of interest has to be parametrized. The stoichiometric network, which is a rank-deficient linear system ($\mathbf{S} \cdot \nu = \mathbf{0}$), was parametrized using the following three steps.

Step (i)

The stoichiometric matrix $\mathbf{S}$ is transformed into its reduced row echelon form $\mathbf{S}_{RRE}$ using Gauss-Jordan elimination with partial pivoting. Subsequently, each row and column of $\mathbf{S}_{RRE}$ is analyzed to identify the dependent and independent variables. The first non-zero element in each row of $\mathbf{S}_{RRE}$ that is always 1 is called 'leading 1'. If the i^{th} column of $\mathbf{S}_{RRE}$ contains only zeros and one leading 1, then the i^{th} element of ν is dependent, otherwise it is independent. The number of fluxes selected as independent equals the degrees of freedom in the stoichiometric network, that is, the difference between the number of variables and the rank of $\mathbf{S}$. Once this step has been completed, we will see that certain intracellular fluxes and all effluxes (ν_m) are chosen as independent.

Step (ii)

Physiologically, metabolic fluxes are constrained such that $\mathbf{0} \le \nu < \infty$, if reversible reactions are considered as two independent reactions. For the effluxes, the corresponding measurements are available. Hence, the effluxes can be bounded, e.g., mean value $\pm \sqrt{\chi^2_{1,\phi}}$ × standard deviation, where $\chi^2_{1,\phi}$ denotes the inverse of χ^2-cumulative distribution function at a certain confidence level of ϕ, and the resulting range covers $100 \times \phi$ % of possible experimental observations.

The intracellular fluxes selected as independent can be compactified using a single rule such that:

$$\varphi_i = \frac{\nu_i}{\alpha + \nu_i} \quad \text{with} \quad 0 \le \nu_i < \infty \wedge \alpha > 0 \Rightarrow \varphi_i \in [0,\ 1). \tag{2}$$

The above compactified flux variables φ, the [0, 1)-fluxes, are a bijection of [0, ∞) domain onto [0, 1) range: if $\nu_i \to 0$, $\varphi_i \to 0$ and if $\nu_i \to \infty$, $\varphi_i \to 1$ for a certain real positive number of the parameter scaling constant α. These [0, 1)-fluxes can potentially elevate the output sensitivity and, thus, the convergence speed. The output sensitivities of a carbon flux system with respect to fluxes and [0, 1)-fluxes are $\partial \mathbf{x}_m / \partial \nu$ and $\partial \mathbf{x}_m / \partial \varphi$, respectively, where the latter equals $(\partial \mathbf{x}_m / \partial \nu) \cdot (\partial \nu / \partial \varphi)$ by the chain rule. Differentiating ν_i with respect to φ_i results in

$$\frac{d\nu_i}{d\varphi_i} = \frac{\alpha}{(1-\varphi_i)^2} \tag{3}$$

from (2). If $\alpha > (1 - \varphi_i)^2$ holds and, thus, $d\nu_i / d\varphi_i > 1$, the sensitivity given by $\partial \mathbf{x}_m / \partial \varphi = (\partial \mathbf{x}_m / \partial \nu) \cdot (\partial \nu / \partial \varphi)$ increases. Particularly, a higher sensitivity can always be obtained by setting $\alpha \ge 1$ due to finite values of fluxes, that is, $0 \le \nu_i < \infty$ or $0 \le \varphi_i < 1$. Hence, setting the parameter scaling constant $\alpha \ge 1$ is more preferable for numerical optimization than $\alpha > 0$. Moreover, the mapping such as (2) has proven to yield an extremely low curvature of ^{13}C labeling in the parameter space and is advantageous for model linearization [24]. The constant α can be adjusted during the optimization. This is described in detail in the Appendix.

Step (iii)

By symbolically solving the equation system consisting of stoichiometric balances and flux constraints such as (2) for the dependent fluxes (ν_{depend}), we get explicit expressions for all dependent fluxes such that $\nu_{depend} = \mathbf{F}_{flux}(\Theta)$ with $\Theta = (\varphi_1, \varphi_2, \ldots, \nu_{m1}, \nu_{m2}, \ldots)^T$.

Flux identifiability by model linearization

One important question to be answered prior to the numerical flux computation is the *a priori* parameter identifiability. To this end, the so-called Buchberger's algorithm [25] can be applied to compute the Gröbner basis [26]. However, this polynomial algebraic method can be very time consuming and, thus, is restricted to small systems, corresponding to the fact that the Gröbner basis can be extremely large for a metabolic carbon labeling system [27,28].

In previous works [28,29], the Jacobian matrix of the measurement model has been utilized for model identification: the matrix determinant or rank were applied to the flux identifiability analysis in carbon labeling systems. However, these approaches do not tell how to sort out identifiable flux variables in case not all fluxes are identifiable.

As an alternative, the first order optimality condition and null space analysis can be applied to model identification. For this purpose, the linear algebra approach is now extended to answer the question of the *a priori* flux identifiability problem regarding ^{13}C isotopomer analysis, which begins with model linearization. At the solution of the nonlinear problem (1), the gradient ∇f is expected to be 0 by the first order optimality condition, as in:

$$\nabla f(\boldsymbol{\Theta}) = \mathbf{J}(\boldsymbol{\Theta})^T \cdot \boldsymbol{\Sigma}_\eta^{-1} \cdot \left(\mathbb{F}(\boldsymbol{\Theta}) - \boldsymbol{\eta} \right) = \mathbf{0}. \quad (4)$$

Here, $\mathbf{J}(\hat{\mathbf{Q}})$ represents the Jacobian matrix that equals $\partial \mathbb{F} / \partial \Theta$ evaluated at $\hat{\mathbf{Q}}$. Assuming that the parameter estimate $\hat{\mathbf{Q}}$ is close to its true value $\hat{\mathbf{Q}}$, the above equation can linearly be approximated in the neighborhood of $\hat{\mathbf{Q}}$ by the first-order Taylor series expansion such that:

$$\mathbf{J}(\boldsymbol{\Theta})^T \cdot \boldsymbol{\Sigma}_\eta^{-1} \cdot \left(\mathbb{F}(\boldsymbol{\Theta}) + \mathbf{J}(\boldsymbol{\Theta}) \cdot \Delta\boldsymbol{\Theta} - \boldsymbol{\eta} \right) = \mathbf{0}. \quad (5)$$

Solving the above equation for $\Delta\hat{\boldsymbol{\Theta}}$ gives the particular solution of (1) as follows:

$$\Delta\hat{\boldsymbol{\Theta}} = \left(\mathbf{J}(\hat{\boldsymbol{\Theta}})^T \cdot \boldsymbol{\Sigma}_\eta^{-1} \cdot \mathbf{J}(\hat{\boldsymbol{\Theta}}) \right)^{-1} \cdot \mathbf{J}(\hat{\boldsymbol{\Theta}})^T \cdot \boldsymbol{\Sigma}_\eta^{-1} \cdot \left(\boldsymbol{\eta} - \mathbb{F}(\hat{\boldsymbol{\Theta}}) \right). \quad (6)$$

Assuming that the above linearization is a good approximation in the vicinity of solution, the *a priori* identification problem can be answered using linear system theory, similarly as applied for stoichiometric metabolite balancing [30]. The theory is now extended to the linearized carbon labeling system: the general solution of (1) can be formulated as a linear combination of the particular solution (6) with its homogeneous part, i.e.,

$$\Delta\hat{\boldsymbol{\Theta}} = \left(\mathbf{J}(\hat{\boldsymbol{\Theta}})^T \cdot \boldsymbol{\Sigma}_\eta^{-1} \cdot \mathbf{J}(\hat{\boldsymbol{\Theta}}) \right)^{-1} \cdot \mathbf{J}(\hat{\boldsymbol{\Theta}})^T \cdot \boldsymbol{\Sigma}_\eta^{-1} \cdot \left(\boldsymbol{\eta} - \mathbb{F}(\hat{\boldsymbol{\Theta}}) \right) + null(\mathbf{J}(\hat{\boldsymbol{\Theta}})) \cdot \boldsymbol{\beta}. \quad (7)$$

Here, $null(\mathbf{J}(\hat{\mathbf{Q}}))$ is the null space of $\mathbf{J}(\hat{\mathbf{Q}})$ and β is a vector containing arbitrary non-zero values. The i^{th} element of the search step $\Delta\hat{\boldsymbol{\Theta}}$ whose corresponding row of $null(\mathbf{J}(\hat{\mathbf{Q}}))$ consists of zeros can be determined uniquely from the measurement η. Thus, the unique solution of the corresponding i^{th} design parameter can be computed numerically.

Hybrid optimization algorithm

Due to the parametrization introduced above, the flux estimation problem (1) can now be formulated such that:

$$\min_{\boldsymbol{\Theta}} f(\boldsymbol{\Theta}) \quad \text{subject to} \quad \begin{cases} \mathbf{v}_{\text{depend}}(\boldsymbol{\Theta}) \geq \mathbf{0} \\ 0 \leq \varphi_i < 1 \\ \max(0, -\sqrt{\chi^2_{1,\phi}}\boldsymbol{\sigma}_{\text{m}}) \leq \mathbf{v}_{\text{m}} \leq \sqrt{\chi^2_{1,\phi}}\boldsymbol{\sigma}_{\text{m}} \end{cases}, \quad (8)$$

where $f(\Theta)$ denotes the objective function specified in (1) and $\Theta = (\varphi_1, \varphi_2, ..., \varphi_n, v_{\text{m}1}, v_{\text{m}2}, ...)^T$ the design parameter vector consisting of the [0, 1)-fluxes φ_i. The effluxes v_{m} comprise substrate uptake as well as product and biomass formation with known standard deviation σ_m, and v_{depend} corresponds to all unknown intracellular fluxes.

To solve the above constrained NLSP, we developed a logical algorithm (Figure 1) that interactively hybridizes two gradient-based local optimization methods, that is, the sequential quadratic programming (SQP) [31] and the subspace trust-region method based on the interior-reflective Newton method (STRiN) [32]. The developed method performs a series of sub-optimization trials by interactively switching between SQP and STRiN using the following features.

Analytical gradient and Hessian

Typically, a gradient-based optimization problem is solved more accurately and with higher efficiency when analytical gradients are provided compared to numerical gradients, which are calculated by finite-difference approximation. By providing the analytical Jacobian matrix consisting of $\partial \mathbf{x}_{\text{m}}/\partial\varphi$ and $\partial v_{\text{m}}/\partial\Theta$, the gradient of the objective function formulated in (A.1) of the Appendix can be calculated. The analytical Hessian (A.2) is calculated when STRiN is used. In comparison to this, SQP updates the Hessian using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) formula based on linearization of the gradient [14].

Adjustment of tolerance and parameter scaling constant

Usually, the tolerances placed on parameters or on the objective function which are utilized as termination criteria for numerical optimization cannot clearly be defined in advance. They depend on the errors associated with user-supplied data as well as output sensitivities. Thus, tolerances are set rather empirically. Here, we suggest a method to adjust tolerance values and find a proper value by repeatedly restarting optimization trials. The tolerance adjustment by the restart is implemented in the developed algorithm as shown in Figure 1-E. Due to this, the complete process consists of a series of sub-optimization trials. At each k^{th} trials, tolerance values are adjusted.

Further to this, the parameter scaling constant α is also adjusted during the optimization. As mentioned in the previous section, $\alpha \geq 1$ is the preferable choice due to the finite values of metabolic fluxes. In practice, α is regulated to render the model's Jacobian matrix into better condition. This is obviously advantageous for the gradient-based method, which typically involves matrix inversion to compute the search direction as shown by (6). The exact procedure how to adjust tolerance and the parameter scaling constant is described in the Appendix.

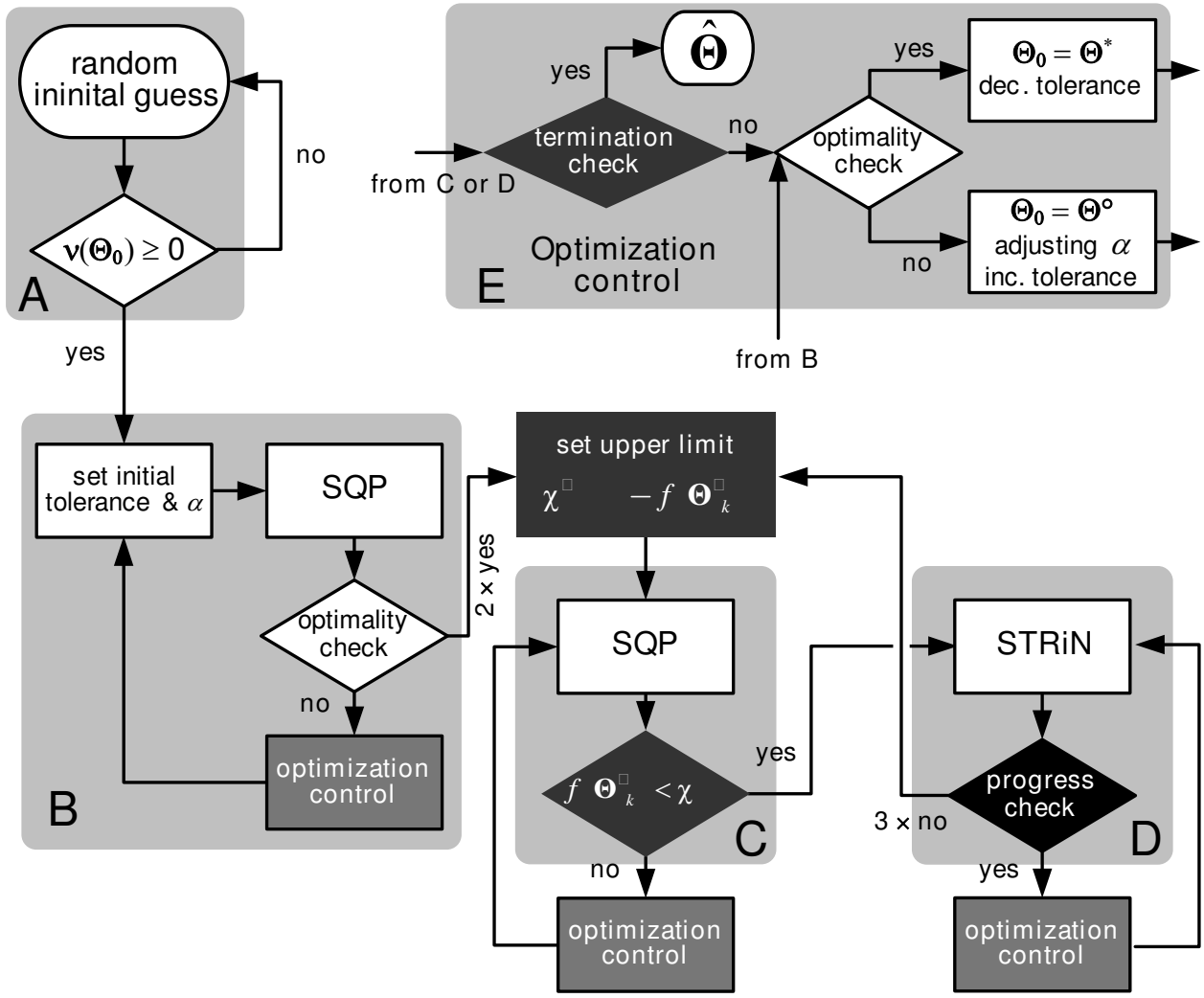

Figure 1

Developed hybrid optimization algorithm with tolerance adjustment consisting of the features: initialization within the feasible region (A); initial optimization using the SQP (B); interactive hybrid process using SQP (C); STRiN (D); and optimization control algorithm (E). $f(\Theta^*_k)$: objective function value at the current local minimizer Θ^*_k; χ^2_{UL}: upper limit of $f(\Theta^*_k)$ to invoke STRiN (if $f\Theta^*_k) < \chi^2_{UL}$; α: parameter scaling constant; Θ_0: initial guess; Θ^*: local minimizer from a successful sub-optimization; Θ°: iterate recorded for the smallest function value up to the current optimization trial; $\hat{\mathbf{Q}}$: ultimate minimizer.

Hybridization

SQP represents the state-of-the-art for solving constrained nonlinear optimization problems. By introducing the Lagrangian function, SQP solves a series of quadratic programming sub-problems, each involving the minimization of a quadratic approximation of the objective function subject to a linear approximation of the constraints. It shows its strength when solving problems with significant nonlinearity and from remote starting points [14,33].

STRiN solves a nonlinear large problem by linear approximation and the method of preconditioned conjugate gradients. This algorithm is based on the trust-region algorithm, is computationally inexpensive, and provides rapid convergence in the vicinity of the solution. A downside of STRiN is that it accepts any direction of negative

curvature, even when this direction gives an insignificant reduction in the objective function [14]

The key concept of the developed algorithm is hybridizing the merits of these two optimization algorithms, i.e., the loose starting-point-dependency of SQP and the convergence speed of STRiN in the solution vicinity. During the optimization trails with the tolerance adjustment, we get a feasible point that may be a more suitable initial guess for the next trial. Accordingly, we initiate a few trials under "relaxed" tolerance conditions using a robust algorithm that is less dependent on the quality of the initial points. Subsequently, the local minima obtained can be tested by another algorithm whether it gives a significant improvement with a rapid convergence.

In this context, the hybridization is carried out as shown in the flow chart in Figure 1. The algorithm is forced to initiate within the feasible region by generating an arbitrary starting point Θ_0 subject to the inequality constraint $\nu_{depend}(\Theta_0) \geq \mathbf{0}$ (Figure 1-A). This can be done by generating random numbers of design parameters within their bounds given by (8). For cases in which the random generation is expensive, e.g., when a design parameter has a very narrow range that satisfies the constraints, a deterministic method might be employed to get a starting point in the physically possible region. This is discussed in the Appendix.

As soon as a feasible set of initial values is obtained, the optimization starts using the SQP algorithm under relaxed tolerance criteria (Figure 1-B). After a few successful trials, an upper limit (χ^2_{UL}) that schedules the initiation of STRiN is placed, e.g., half of the current objective function value, $\chi^2_{UL} = 1/2\ f(\Theta^*_k)$, where Θ^*_k is the local minimizer isolated from the k^{th} sub-optimization trial. Afterwards, if the objective value $f(\Theta^*_k)$ resulting from the current SQP trial (Figure 1-C) is smaller than χ^2_{UL}, the STRiN optimization is activated (Figure 1-D); otherwise, the SQP is repeated. During the STRiN optimization, the progress of trials is monitored to prevent the non-reducing problem associated with the STRiN algorithm mentioned above. As a criterion of the progress check, the relative improvement of $|f(\Theta^*_k) - f(\Theta^*_{k-1})|/f(\Theta^*_k)$ can be measured at each STRiN termination. If the improvement is insignificant, i.e., less than a user-specified value (e.g., 0.05), the STRiN receives penalty. In case a few successive trials show insignificant improvement, the algorithm sets a new upperlimit χ^2_{UL} and returns to the SQP optimization. For the restart, the tolerance of the previous SQP trial is reutilized. This prevents too large changes in tolerance, which can be caused by the STRiN trials with insignificant improvement.

As a termination criterion of the hybrid optimization, the changes in local optima are measured (Figure 1-C), e.g., the absolute value of the slope resulting from the linear regression of $y = (f(\Theta^*_{k-4}), f(\Theta^*_{k-3}),...f(\Theta^*_k))^T$ with respect to $x = (1, 2,..., 5)^T$. If the absolute slope is less than a user-specified small value for a few further trials, the optimization can be terminated ultimately.

Application to Bacillus subtilis *metabolic network*

To evaluate the developed methods, a metabolic network of the wild type *B.subtilis* was constructed as shown in Figure 2 based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database specified for the strain. The network is composed of catabolic reactions of the central metabolism incorporating glycolysis/gluconeogenesis, pentose phosphate pathway, TCA cycle, C3/C4 inter-conversion and anabolic reactions. All effluxes including the biomass yield Y_{XS} in Figure 2 were assumed to be measured experimentally from a batch-cultivation on succinate and glutamate as carbon sources. All flux values specified in Figure 2 are generated from arbitrary values of $\Theta_{default}$ and normalized by the glutamate uptake rate. Note that each anabolic flux given in Figure 2 is the product of Y_{XS} (biomass production [$g_{DW}\ L^{-1}\ h^{-1}$] normalized by glutamate uptake [mM h^{-1}]) and a value that specifies the precursor requirement for growth (mmol precursor per g biomass) adopted from literature data [34]. For bidirectional reactions, the fluxes in the gluconeogenetic direction were declared as forward.

Parametrization

The metabolic network in Figure 2 has 27 intracellular fluxes (10 bidirectional and 17 unidirectional fluxes), 5 effluxes, and 10 anabolic fluxes expressed in terms of Y_{XS}. For 15 intracellular metabolites defined in the network, 15 flux balances were set up that are linearly independent of each other (Appendix). The stoichiometric matrix **S** was obtained by symbolic differentiation of the balances with respect to the whole fluxes including Y_{XS}. When considering that the flux of glutamate uptake is unity due to the normalization, the network owns 17 degrees of freedom. Accordingly, 12 intracellular fluxes ν_{2r}, ν_{3r}, ν_{4r}, ν_{5r}, ν_{6r}, ν_{7r}, ν_{8r}, ν_{14}, ν_{14r}, ν_{15r}, ν_{16}, ν_{17r}, 4 effluxes, and Y_{XS} were recognized as independent variables from the reduced row echelon form $\mathbf{S}_{RRE}$ of **S**. These independent intracellular fluxes were transformed into [0, 1)-fluxes as given by (2). Note that the counterpart of a backward flux of a bidirectional reaction can also be selected as an independent variable. This just depends on how to arrange the entries in the flux vector. For instance, when ν_{2r} is followed by ν_2 in the flux vector, ν_2 will be recognized as an independent variable instead of ν_{2r}.

Solving the equation system of the flux balances (Appendix) and the [0, 1)-flux equations for the dependent fluxes

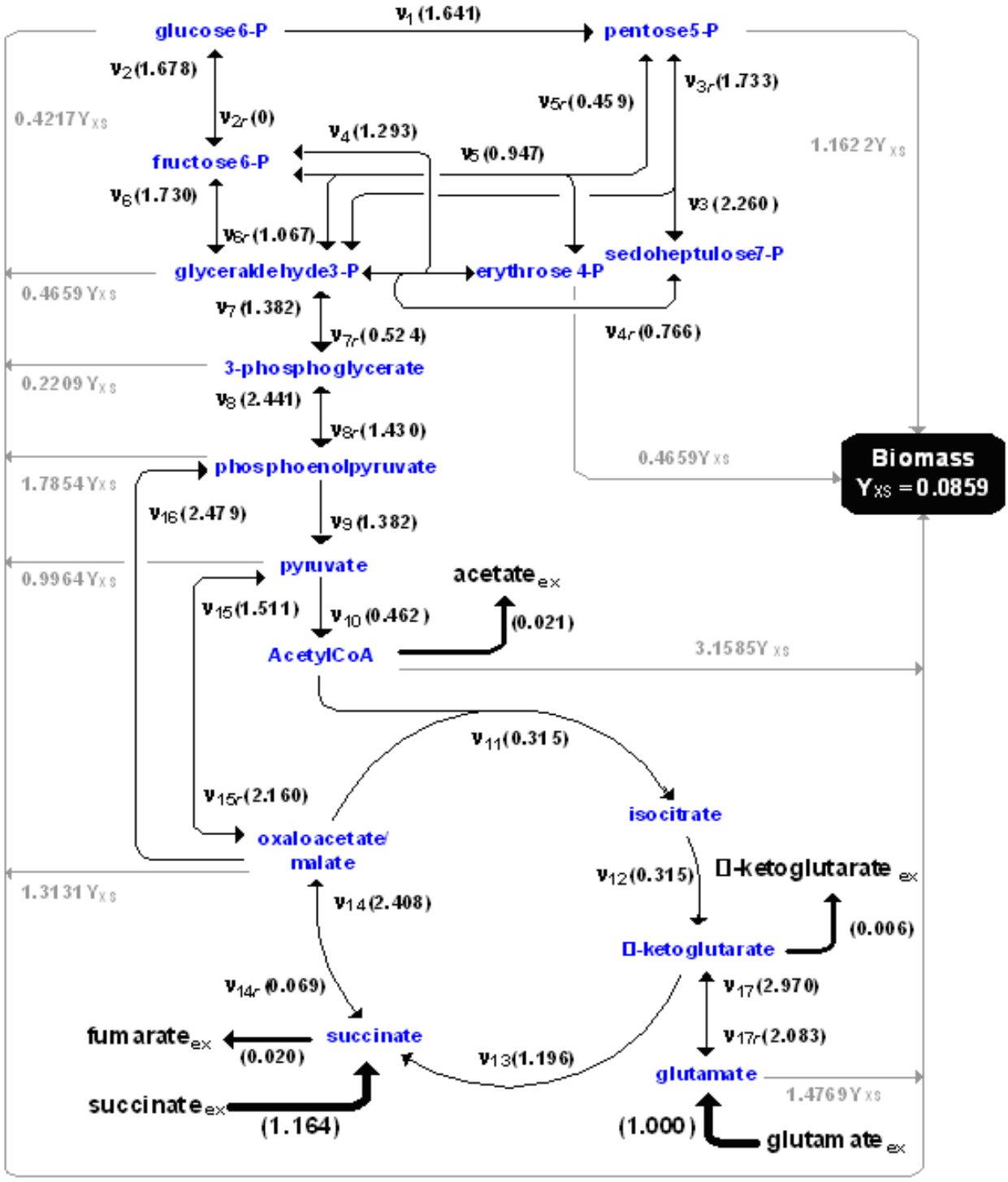

Figure 2
Metabolic network of the central metabolism of *Bacillus subtilis* utilizing glutamate and succinate as co-substrates. All flux values denoted in parentheses were generated by obeying the given stoichiometry. Effluxes and biomass yield were measured experimentally. The symbol 'ν' indicates the flux, the subscript '***r***' the reverse flux of the bidirectional flux pair, the subscript '**ex**' extracellular pools of substrates and products and Y_{XS} the biomass yield in g(biomass)/mmol(glutamate). All flux values are normalized by the glutamate uptake rate.

gives their explicit analytical expressions with respect to design parameters, i.e., $v_i = n_{flux}(\Theta)$ with 19 design parameters of $\Theta = (\varphi_{2r}, \varphi_{3r}, \varphi_{4r}, \varphi_{5r}, \varphi_{6r}, \varphi_{7r}, \varphi_{8r}, \varphi_{14}, \varphi_{14r}, \varphi_{15r}, \varphi_{16}, \varphi_{17r}$, succinate$_{ex}$, acetate$_{ex}$, α-ketoglutarate$_{ex}$, fumarate$_{ex}$, Y_{XS}, $pCO_{2,1}$, $pCO_{2,2})^T$. The parameters $pCO_{2,1}$ and $pCO_{2,2}$ estimate the CO_2 labeling pattern for each ^{13}C labeling experiment: CO_2 has only one carbon and is incorporated into other metabolites by carboxylation. Thus, its labeling state can simply be determined from other ^{13}C labeling data as an additional parameter of NLSP in case it is not measured.

For flux re-estimation studies, the default values of the design parameter were $\Theta_{default}$ = (0, 0.6341, 0.4337, 0.3148, 0.5162, 0.3436, 0.5884, 0.7066, 0.0643, 0.6835, 0.7125, 0.6756, 1.164, 0.021, 0.006, 0.02, 0.0859, 0.5601, 0.5601) at α = 1, which results in the flux values given in Figure 2.

Model identification

For the identifiability studies, the main question was whether Θ has one unique solution set that can be estimated from the available mass isotopomer measurements in terms of the gradient-based optimization. The effluxes (succinate$_{ex}$, acetate$_{ex}$, α-ketoglutarate$_{ex}$, fumarate$_{ex}$) and Y_{XS} are excluded for the identifiability analysis because they can be estimated as long as the corresponding measurements are available. To examine various flux scenarios, 200 stochastic simulations were performed for each of the 630 possible ^{13}C-experimental designs. To this end, uniform random numbers between 0 and 0.95 were created for Θ. At each simulation the null space of $\mathbf{J}(\Theta)$ was calculated.

According to the stochastic simulations, none of the designs had an empty null space and the entry of $null(\mathbf{J}(\Theta))$ corresponding to φ_{2r} was always non-zero disregarding the flux state and designs. Most designs, including the design considered here, yielded an identical null space of:

$$null(\mathbf{J}(\Theta)) = (1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)^T. \quad (9)$$

Accordingly, the design parameter φ_{2r} is not expected to have a unique solution, i.e., it has infinitively many solutions. Except v_2 and v_{2r}, other fluxes were not functions of φ_{2r}. This means that the bidirectional glucose 6-P isomerase cannot be determined from any ^{13}C substrates or dual substrate combinations. Only the net flux of the reaction can be calculated from the stoichiometry. Therefore, one can either regard the glucose 6-P isomerase as a unidirectional reaction or set φ_2 as an arbitrary constant. This renders the carbon flux network to have an empty null space of $\mathbf{J}(\Theta)$, i.e., full rank for $\mathbf{J}(\Theta)$, and all other design parameters are theoretically expected to have unique solutions. This theoretical expectation is examined later using the developed algorithm.

Hybrid optimization with tolerance adjustment

To evaluate the efficiency of the developed optimization algorithm, the flux values calculated in advance were re-estimated numerically from the carbon mass isotopomer distributions (MDVs) of output metabolites matching the default flux values ($v_{default}$). To this end, the default flux values, depicted in Figure 2, were calculated by the flux function $\mathbf{F}_{flux}(\Theta_{default})$ obtained by parametrization. Subsequently, the ^{13}C labeling data of output metabolites were calculated from $v_{default}$ and $\Theta_{default}$. The default fluxes were re-estimated by solving the constrained NLSP (8) whose inputs (η) were the effluxes, Y_{XS}, and the MDVs.

Prior to testing the hybrid optimization with tolerance adjustment (HATA), we examined whether the tolerance adjustment is beneficial for optimization. This was checked by performing the SQP optimization by providing the gradients for the objective function and for the flux inequality constraints ($\nabla c = (-\partial v_{depend}/\partial\Theta)^T$) analytically. As shown in Figure 3-A, the objective function value $f(\Theta^*)$ of each optimization trial decreased with respect to tolerance adjusted at each optimization restart. It was observed that $f(\Theta^*_k) < f(\Theta^*_{k-1})$ always holds when starting the k^{th} trial from the $(k-1)^{th}$ local minimizer Θ^*_{k-1}. Restarting the failed $(k-1)^{th}$ trial from the feasible iterate Θ° recorded for the smallest function value up to the current trial and increasing the tolerance was observed to give the same result, i.e., $f(\Theta^*_k) < f(\Theta^\circ)$. The efficiency of tolerance adjustment was further compared to the SQP optimization carried out at a constant tolerance of 1×10^{-20} (Figure 3-B). The SQP using the tolerance adjustment was observed to be more efficient in accuracy but more time-consuming than the case without adjusting. The SQP without adjustment reached a local minimum of around 10^{-7} much more rapidly but did not result in any further improvement, whereas the SQP with adjustment made slower but continuous progression. This gives an idea that the tolerance adjustment strategy might be useful to escape from possible local stationary regions and to achieve a lower minimum.

Using the tolerance adjustment, the hybrid algorithm consisting of SQP and STRiN (HATA, Figure 1) was compared with its parent algorithms and two global optimization methods. All optimizations except the genetic algorithm (GA) were initiated from an identical starting point for the numerical flux re-estimation. The GA applied does not need an external initial value set. At each initiation of the HATA trials, α was updated by choosing an integer between 1 and 10 that yields the best-conditioned Jacobian matrix of the model as mentioned previously. As shown in Figure 4-A, HATA accomplished the re-

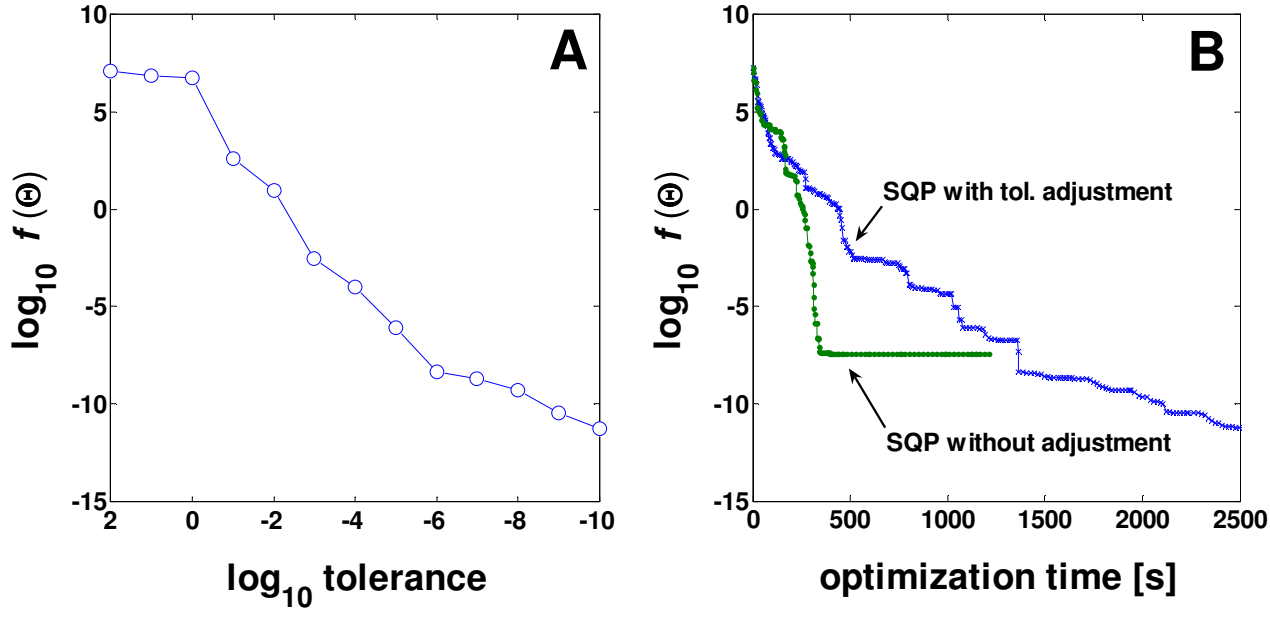

Figure 3
Decrease of the objective function at each termination of SQP sub-optimization using tolerance adjustment (**A**) and its comparison with SQP carried out at a constant tolerance during optimization (**B**).

estimation with the best efficiency regarding its accuracy and speed. It took about 300 seconds until the objective function became 10^{-16}. The SQP optimization with analytical gradient (SQP ∇_{user}) yielded the next satisfactory result. In comparison, the SQP optimization using the numerical gradient obtained by the finite difference (SQP ∇_{finite}) resulted in a poorer progress and was much slower and less accurate than HATA or SQP ∇_{user}. During the optimization using SQP ∇_{finite}, we observed a discrepancy between the gradient obtained by the finite difference and the analytical approach (A.1). The inaccuracy of the finite difference seemed to cause the poorer result of SQP ∇_{finite}.

The worst algorithm among the local methods was the STRiN using analytical gradient and Hessian (STRiN ∇_{user} $\mathbf{H}_{user}$). It was rapid at the beginning but improved the objective only from 10^7 to 10^5. Afterwards, the optimization crashed with the carbon labeling system becoming completely singular. As previously mentioned, this seems to be due to the STRiN limitation of accepting any direction of negative curvature even if it does not give significant reduction in the objective function. Also, the global optimization methods tested, which were the GA and the SA, were notably time-inefficient compared to the local methods. For 1500 seconds of optimization time, the objective function was very slowly decreased by two orders of magnitude only. SA decreased $f(\Theta)$ only to 1×10^3 after 5.1 hours and GA to 2×10^2 after 6.7 hours. Thus, global optimization is expected to demand much greater time to get the equivalent result as HATA.

HATA optimization from different starting points

The flux re-estimation was also performed from 200 different starting points created randomly and by simultaneously generating uniform random numbers for the [0, 1)-flux $\varphi_{2r} = \nu_{2r}/(\alpha + \nu_{2r})$. This was to examine whether the quality and speed of convergence has any starting-point-dependency and whether φ_{2r}, which *a priori* expected to have infinite solutions, has any influence on optimization results.

All 200 optimization runs using the hybrid algorithm were terminated successfully. On average, HATA required 5.3 ± 1.6 min to converge to a weighted objective function value smaller than 1×10^{-10}, which equals the non-weighted sum of squares less than 2×10^{-18}. As shown in Figure 4-A, 99.5 % of the flux re-estimations were completed within 10 min and about 76 % in less than 6 min.

Further to this, the flux re-estimates resulting from the 200 random simulations were compared with the true flux solutions calculated from $\Theta_{default}$ in advance. As shown in Figure 5-A, all fluxes except the flux pairs of transketolase 1 (TK1; ν_3, ν_{3r}) and transaldolase (TA; ν_4, ν_{4r}) could be re-estimated correctly. When plotting the flux re-estimates versus the true flux values given in Figure 5, the data

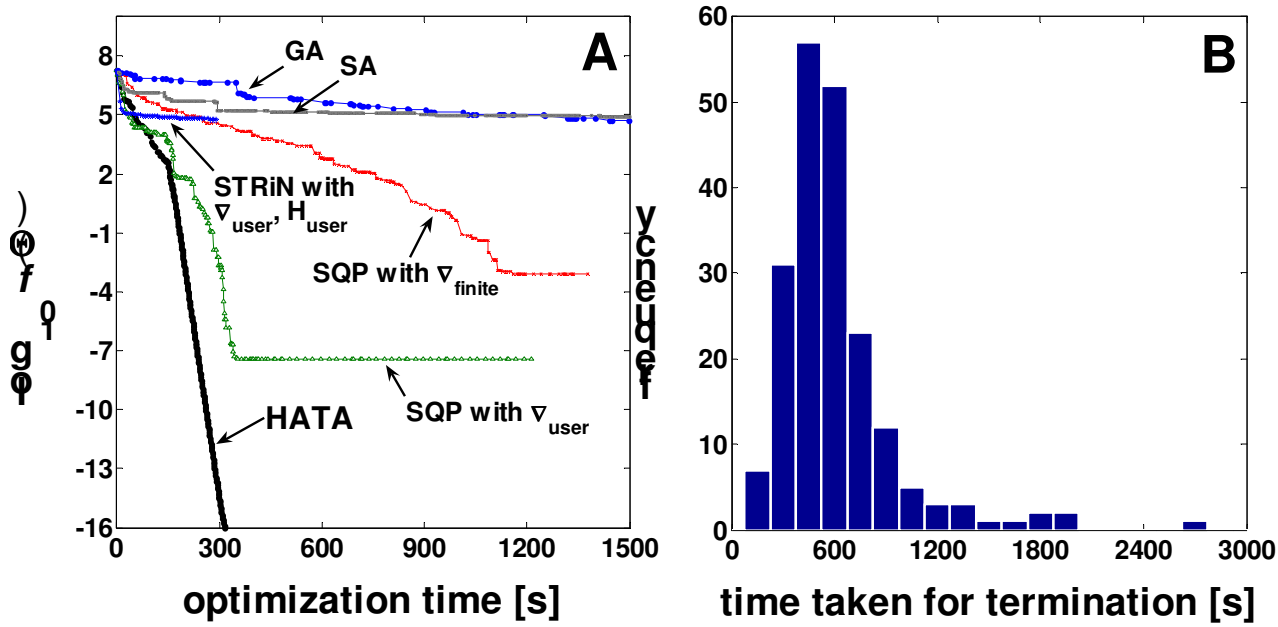

Figure 4
Comparison of the hybrid optimization with tolerance adjustment (HATA) with its parent algorithms of the SQP with user-supplied analytical gradient (SQP ∇_{user}) and the STRiN with user-supplied analytical gradient and Hessian (STRiN ∇_{user} $\mathbf{H}_{user}$) as well as with other algorithms, such as SQP with numerical gradient by finite differentiation (SQP ∇_{finite}), genetic algorithm (GA), and simulated annealing (SA) (**A**). All algorithms were initiated from an identical starting point. The objective function value at the i^{th} iterate is registered only if $f(\Theta_i) \leq f(\Theta_{i-1})$. Time efficiency of the HATA represented by histogram plot of time taken for termination of 200 runs of optimization using different starting points (**B**).

points except ν_3, ν_{3r}, ν_4, and ν_{4r} lie exactly on a line with a slope of 1 (1:1 line). In comparison to this, only the net fluxes of the TK1 (ν_3 - ν_{3r}) and TA (ν_4 - ν_{4r}) reactions could be recalculated properly: the net fluxes could be calculated from stoichiometric relations with other flux re-estimates. The same was observed for the net flux of glucose 6-P isomerase, of which [0, 1)-flux φ_{2r} had an arbitrary value at each simulation run.

By the null space investigation (9) based on (7), the design parameters φ_{3r} and φ_{4r} were not expected to be correlated. In accordance with this, the fluxes could rarely be re-fitted in case one or both of these parameters were set arbitrarily. This indicates that the parameters φ_{3r} and φ_{4r} have a certain significant correlation that must be fulfilled at termination to yield unique solutions for other fluxes. As expected, the flux estimates of TK1 and those of TA were found to be nonlinearly correlated with each other as depicted in Figure 5-B for ν_{3r} and ν_{4r}. The true solution, $(\varphi_{3r}, \varphi_{4r}) = (0.6341, 0.4337)$ also lies on the curve. At the termination, it is obvious that the output sensitivities of $\partial \mathbf{x}_m/\partial \varphi_{3r}$ and $\partial \mathbf{x}_m/\partial \varphi_{4r}$ are low: the objective function reached a very small value of nearly zero ($f(\hat{\mathbf{Q}}) < 1 \times 10^{-10}$) and we get arbitrary but correlated values for φ_{3r} and φ_{4r}. Due to this, we cannot practically get φ_{3r} and φ_{4r} estimated correctly and, accordingly, ν_3, ν_{3r}, ν_4, and ν_{4r}, while the corresponding net fluxes can always be calculated from other correct flux estimates in terms of stoichiometry.

We observed that a unique estimation of all four fluxes would be possible only if the mass isotopomers of sedoheptulose 7-phosphate were additionally measured. It was observed that providing these mass isotopomers renders the output sensitivities of $\partial \mathbf{x}_m/\partial \varphi_{3r}$ and $\partial \mathbf{x}_m/\partial \varphi_{4r}$ much less dependent on starting points. In contrast, the output sensitivities vary strongly and even become zero when the mass isotopomers of sedoheptulose 7-phosphate are not involved in the objective.

Further to this, the fluxes were not found to be influenced by the glucose 6-phosphate isomerase fluxes represented by φ_{2r}. Disregarding the random values of φ_{2r}, the fluxes except ν_3, ν_{3r}, ν_4, and ν_{4r} converged to their true values. Also φ_{3r} and φ_{4r} did not show any correlation with φ_{2r} as depicted in Figure 5-C and Figure 5-D, respectively.

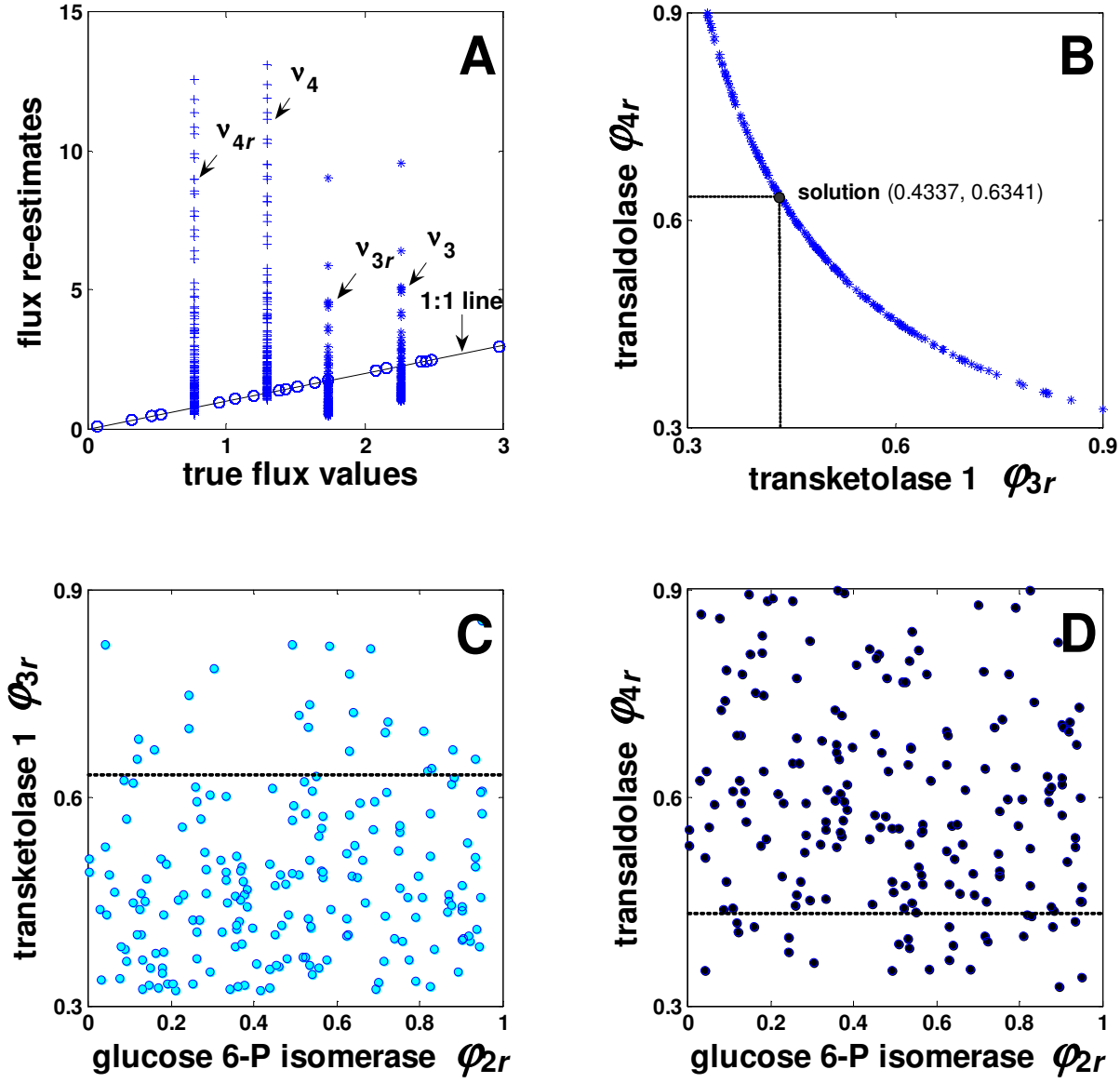

Figure 5
Flux re-estimation initiated from 200 random starting points. The fluxes which were successfully re-estimated lie on the 1:1 line when plotted against true flux values (A). Correlation analysis by plotting the transaldolase (TA) [0, 1)-flux re-estimates (φ_{4r}) versus those of transketolase 1 (TK1) (φ_{3r}) (B) and the [0, 1)-flux re-estimate of TK1 (C) as well as of TA (D) versus φ_{2r} of glucose 6-P isomerase. The dotted line indicates the solution of φ_{3r} = 0.4337 and φ_{3r} = 0.6341, respectively.

Hence, we can conclude that the observed nonlinear correlation between φ_{3r} and φ_{4r} becomes visible solely from using the different starting points for the optimization. In addition, including φ_{2r} as an additional design parameter gives the same result for the fluxes, whereas φ_{2r} results in randomly distributed values without any obvious pattern. This shows that the flux identification carried out by the model linearization and null space investigation is appropriate for its practical use.

Discussion

The tradeoff between robustness and convergence speed is a central issue in numerical optimization [14]. The developed method fulfills both criteria and enables exact met-

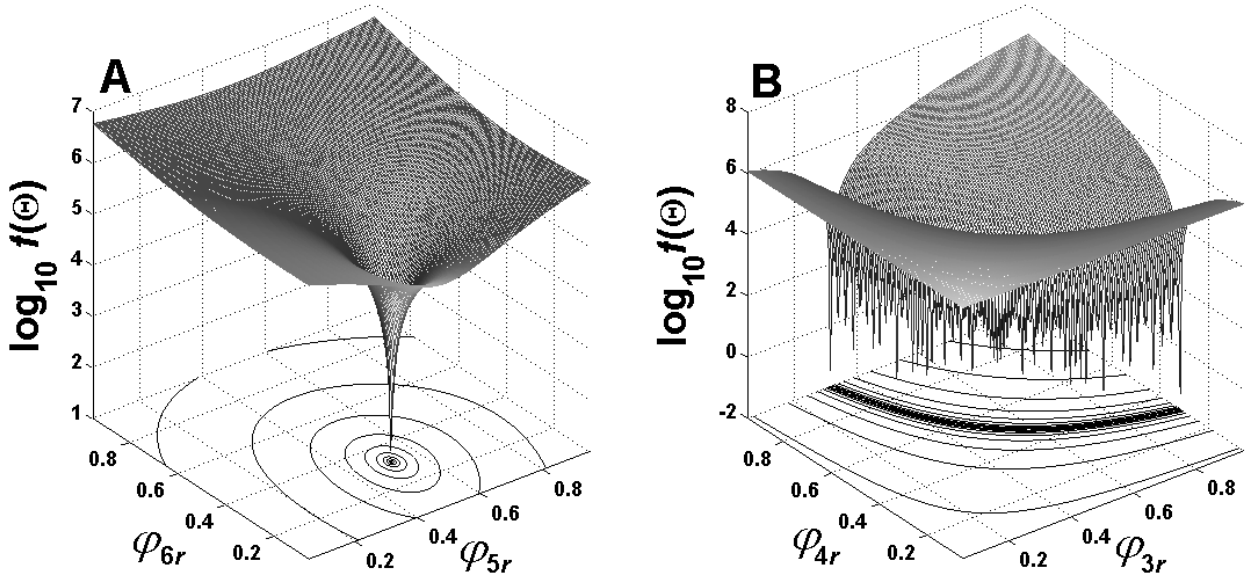

Figure 6
Behaviors of the objective function in the parameter space of non-correlated φ_{5r} and φ_{6r} (**A**) as well as correlated φ_{3r} and φ_{4r} (**B**). Two parameters were varied from 0.05 to 0.95 with a step of 0.005 while other parameters were fixed at the default solution $\Theta_{default}$.

abolic flux estimation. This was accomplished on the basis of parametrizing the metabolic flux network by compactification and hybridizing the merits of two different gradient-based algorithms. Interestingly, HATA, which hybridizes the worst local method with SQP ∇_{user} and utilizes the tolerance adjustment strategy, has proven to be superior to other approaches. This may be by virtue of providing STRiN with a more appropriate starting point advanced by SQP. Thus, because STRiN is computationally inexpensive and rapid in the vicinity of solution, it achieves a fast convergence. Furthermore, the interactive switching between the algorithms prevents STRiN from accepting an unfavorable search direction without significant objective improvement.

Further to this, we observed that the parametrization by the compactification of independent [0, ∞)-fluxes into [0, 1)-variables is advantageous compared to the non-compactified case. When the optimization was performed by selecting the [0, ∞)-fluxes (v_{2r}, v_{3r}, v_{4r}, v_{5r}, v_{6r}, v_{7r}, v_{8r}, v_{14}, v_{14r}, v_{15r}, v_{16}, v_{17r} with upper bounds of 200) directly as design parameters, the convergence took drastically longer (27 ± 45 min) than the cases using the compactified fluxes. Moreover, optimization runs often failed to converge or terminated at suboptimal points. In comparison to this, introducing [0, 1)-fluxes improved both the robustness and speed of convergence as demonstrated above. The same advantage is probably achieved when using the [0, 1)-compactifications such as exchange fluxes or flux partitioning ratios applied elsewhere [9,24,35]. Our compactification approach allows a straight parametrization because independent fluxes can easily be recognized from the stoichiometric matrix and compactified using a single rule of (2). Moreover, it is straightforwardly differentiable when considering that compactified [0, 1)-fluxes are continuous and smooth in the [0, ∞)-flux space and *vice versa*.

The parameter scaling constant α introduced during the compactification provides a way to implement numerical flux estimation with efficiency. In particular, the parameter scaling constant α is updated during optimization to render the model's Jacobian matrix into better condition. This is advantageous for the gradient-based methods for computing the search direction because it typically involves the inverse of the Jacobian matrix as given by (6). We examined whether different α values affect the efficiency of optimization. The α value was observed to affect the convergence speed but not the accuracy of the optimization. When starting from the same initial guess applied to the optimization trial in Figure 4, the fastest convergence speed was obtained when α is fixed at 10 (about 200 s for $f(\Theta) < 10^{-14}$). The speed efficiency decreased if α is smaller or larger than 10. A trial with α adjustment ini-

tiated at α = 1 automatically updated α to 10 after a few sub-optimization trials and took about 300 s to achieve $f(\Theta) < 10^{-14}$. Although it was slower than when α was fixed at 10, the proposed strategy with α adjustment recognized the optimal value. Among the 200 HATA optimization runs with α adjusting between 1 and 10 initialized with α = 1, 112 cases updated α to 10 after a few to several sub-optimization trials while α remained 1 in 78 cases. Moreover, the time taken for the optimization runs using different starting points was not necessarily shorter when α was updated to 10. This indicates that the most adequate α value for the optimization may depend on starting conditions.

We could find unique flux solutions except the fluxes recognized *a priori* as non-identifiable or those that were *a posteriori* found to have a significant nonlinear correlation. Such nonlinear parameter correlation can be identified only by the tedious process of executing flux estimation using different sets of starting values *a posteriori* [36]. The non-correlated [0, 1)-fluxes of which true values could be re-estimated, e.g., φ_{5r} and φ_{6r}, give a unique minimum of $f(\Theta)$, which is obviously achieved disregarding starting points (Figure 6-A). In comparison to this, the nonlinearly correlated [0, 1)-fluxes φ_{3r} and φ_{4r} result in an infinite number of minima that meet the correlation shown in Figure 5-B depending on starting points (Figure 6-B). Because the objective function value is nearly zero at each minimum ($f(\hat{\mathbf{Q}}) < 1 \times 10^{-10}$ with a weighting factor of 10^8), the system can be considered to have an infinite number of global minima when parameters are nonlinearly correlated. In this case, a global optimization method can be extremely expensive for identifying such unknown parameter correlations due to its time-inefficiency. As a consequence, a fast and reliable numerical method of flux estimation is highly desirable to repeat optimization runs at different starting points.

One interest of numerical optimization is whether the global solution can be obtained by the algorithm applied. We observed that the local minimizers from 200 random simulations agreed with each other and also with the true flux values, the global solution. Depending on experimental noise or metabolic state, the problem may contain a few to several local saddle points. In this case, one can consider a further hybridization with a global optimization approach. Global optimization algorithms are inefficient in terms of convergence speed. On the other hand, gradient-based algorithms can converge rapidly but lack a global perspective for non-convex problems. Also in other cases, the combination of global and local search procedures has proven to offer the advantages of both methods while offsetting their disadvantages [17,37-39]. Such hybridization with a global optimization algorithm is useful for non-convex problems containing several local minima and can be done in many ways [17,39,40]. For instance, the local minima obtained from a gradient-based method can be used to update the individuals in the population to prepare the next generation when combining with the GA [17,40,39,41]. Because we could prove the speed and accuracy of the developed optimization method, we consider it very useful for investigating metabolic fluxes and for developing efficient algorithms combined with a global optimization method.

In practice, the quality of the flux fitting can be judged by the χ^2-test [42,43] when using noisy measurement data. In our case, the upper critical χ^2 limit is 206 at 95 % confidence level at the given degree of freedom of 174 (188 ^{13}C labeling data + 5 effluxes - 19 design parameters). Hence, if the weighted objective value (1) is smaller than the critical value of 103 (this is due to the factor 1/2 in the objective function), the fit can be regarded as acceptable. This value can also be used as a termination criterion for the optimization control in Figure 1-E. As shown in Figure 4-A, the value was achieved at 150 s by HATA, at 170 s by SQP with ∇_{user}, and at 750 s of optimization time by SQP with ∇_{finite}. In comparison to this, GA or SA was expected to achieve the acceptable $f(\Theta)$ after 7-8 hours or more (by extrapolation of $\log_{10} f(\Theta)$with respect to optimization time). Hence, an efficient Monte Carlo simulation cannot be carried out by the global optimization approaches.

Further, we compared HATA with SQP with ∇_{user} by conducting 100 stochastic simulations under identical conditions. At each simulation run, normal random numbers were generated for the experimental data set (^{13}C mass isotopomer fractions and extracellular fluxes) by assuming a unit standard deviation of 1×10^{-3} and uniform random numbers for the starting points of design parameters. Subsequently, parameter re-estimation was performed in terms of HATA and SQP with ∇_{user}, respectively. Each optimization run was stopped if $f(\Theta)$ becomes smaller than the critical χ^2 limit of 103 or does not give any further improvement. On average, $f(\Theta)$ at termination was 102 ± 26 for HATA and 167 ± 567 for SQP with ∇_{user}, and the time taken for the optimization was 353 ± 216 s for HATA and 1234 ± 1356 s for SQP with ∇_{user}. The results support that the developed hybrid algorithm using tolerance adjustment will be more efficient and robust when computing metabolic fluxes from noisy experimental data in a real-case application. In contrast, the efficiency of the SQP algorithm seems to be affected by initial conditions (starting points and experimental data set) when considering the large variation in $f(\Theta)$ as well as in the computation time. In this regard, HATA will also achieve the most efficient optimization in the practical sense whenever

repeated optimization runs are necessary. A further reduction in computation time can be obtained by utilizing a minimal set of isotopomer balances described by Antoniewicz et al. (2007) [3].

Conclusion

We developed and examined a developed hybrid algorithm for numerical ^{13}C flux estimation using a metabolic network model parametrized using a single compactification rule. We could prove its practical usefulness using the central metabolic network of *Bacillus subtilis*, which is a prokaryotic model organism and an industrially relevant strain. Based on the parametrization by compactification and model identification, the hybrid algorithm with tolerance adjustment allowed a fast and robust flux computation from the ^{13}C labeling data created from ^{13}C-labeled succinate and glutamate feeding. Such fast optimization was also found to be essential for the *a posteriori* identification of possible parameter correlations and also for the Monte Carlo approach to obtain statistical qualities for metabolic flux estimates. Therefore, this work represents an important contribution to the quantitative study of metabolic networks in the framework of systems biology.

Methods

For the flux re-estimation simulations of the central metabolic network of *B. subtilis*, the following inputs and outputs were applied.

^{13}C Input Substrates

The choice of input tracer substrates was made by means of the *D*-optimality criterion [44]. Since parallel experiments using different ^{13}C input labels yield more information as theoretically shown for the respirometric flux analysis utilizing CO_2 labeling measurement [4]. Therefore, experimental designs for two parallel cultivations were also considered here. Among the investigated experimental designs (630 possible designs from commercially available ^{13}C glutamate and ^{13}C succinate species; refer to Additional file 1!), a design combining data from a ^{13}C labeling experiment using non-labeled glutamate and [2,3-$^{13}C_2$] succinate with those from an experiment using [1,2-$^{13}C_2$] glutamate and [1,4-$^{13}C_2$] succinate was expected to provide the richest information at a reasonable expense. This design was selected to study flux identifiability and to examine the developed hybrid optimization.

^{13}C Output Metabolites

For simulation studies, we assumed that the mass isotopomers of proteinogenic amino acids such as alanine, valine, serine, threonine, glycine, aspartate, isoleucine, leucine, phenylalanine, tyrosine, arginine, histidine, and glutamate are measurable. In addition, the mass isotopomers of ribose 5-phosphate from RNA hydrolysate and hexose from carbohydrate hydrolysate are also assumed to be measurable. This totally gives a mass isotopomer data size of 188 for the parallel experiment. The measurements of mass isotopomer distributions were assumed to have a unit standard deviation of 1×10^{-4}.

Software implementation

All simulations were carried out on a PC equipped with a CPU of 2.4 GHz. The hybrid optimization algorithm depicted in Figure 1 was implemented in MATLAB (The Mathworks Inc., Natick, MA). Numerical optimization was carried out using the Optimization Toolbox (Version 3.0.4) and the Genetic Algorithm and Direct Search Toolbox (Version 2.1) of MATLAB. The symbolic operations required for parametrization and for computing partial derivatives were conducted using the Symbolic Math Toolbox (Version 3.1.4) of MATLAB. During simulations, random numbers were generated using the Statistics Toolbox (Version 5.2) of MATLAB. The random numbers for [0, 1)-fluxes (φ_i) were bounded such that the relative values of the corresponding dependent fluxes (ν_{depend}) lie between 0 and 20.

An example of the hybrid optimization for the problem described in the current work was coded using MATLAB (Hybridoptimizer.m) and supplemented (see Additional file 2). The method can be incorporated with any isotopomeric models [2,3,5,8] that are analytically differentiable.

List of abbreviations used

BFGS: Broyden-Fletcher-Goldfarb-Shanno; **GA**: genetic algorithm; **HATA**: hybrid algorithm with tolerance adjustment; **KEGG**: Kyoto Encyclopedia of Genes and Genomes; **MDV**: mass isotopomer distribution vector; **MFA**; metabolic flux analysis; **NLSP**: nonlinear least-squares problem; **SA**: simulated annealing; **SQP**: sequential quadratic programming **SQP**$\nabla_{\mathbf{user}}$: SQP optimization with analytical gradient; **SQP**$\nabla_{\mathbf{finite}}$: SQP optimization using the numerical gradient obtained by finite difference; **STRiN**: subspace trust-region method based on the interior-reflective Newton method; **STRiN**$\nabla_{\mathbf{user}}\mathbf{H}_{\mathbf{user}}$: STRiN using analytical gradient and Hessian; **TA**: transaldolase; **TK**: transketolase

Authors' contributions

TH Yang conceived the study, developed the concepts, implemented simulation, and drafted the manuscript. O. Frick prepared the data required for modeling of the stoichiometric and isotopomer network of *B. subtilis*. E. Heinzle supervised the work and was involved in writing the manuscript. All authors read and approved the final manuscript.

Appendix

Gradient and Hessian

At each k^{th} iteration of (1), the model function $\mathbb{F}$ is evaluated at the k^{th} iterate Θ_k to get the values corresponding to $\eta = (\nu_m, \mathbf{x}_m)^T$. This is done by generating a certain flux state by $\nu_k = \mathbf{F}_{\text{flux}}(\Theta_k)$ and, subsequently, solving $\mathbf{F}_{\text{carbon}}(\mathbf{x}_k, \nu_k) = \mathbf{0}$ for $\mathbf{x}_k$ at the k^{th} flux state ν_k. Since $\mathbb{F}$ is differentiable with respect to Θ_k, the gradient ∇f can be computed analytically by:

$$\nabla f(\boldsymbol{\Theta}_k) = \frac{\partial \mathbb{F}(\boldsymbol{\Theta}_k)^T}{\partial \boldsymbol{\Theta}} \cdot \boldsymbol{\Sigma}_\eta^{-1} \cdot \left(\mathbb{F}(\boldsymbol{\Theta}_k) - \boldsymbol{\eta} \right). \qquad (A.1)$$

Further to this, the Hessian matrix **H** that specifies the curvature of the search surface can be obtained by linearization [14,15], i.e.,

$$\mathbf{H}_k = \frac{\partial \mathbb{F}(\boldsymbol{\Theta}_k)^T}{\partial \boldsymbol{\Theta}} \cdot \boldsymbol{\Sigma}_\eta^{-1} \cdot \frac{\partial \mathbb{F}(\boldsymbol{\Theta}_k)^T}{\partial \boldsymbol{\Theta}}. \qquad (A.2)$$

In practice, the constrained problem (1) is formulated as the Lagrangian function, a linear combination of the objective function and the constraints [14].

Adjustment of tolerance and parameter scaling constant

The tolerance adjustment by the restart in Figure 1-E requires a series of sub-optimization trials. The first optimization trial starts with "sufficiently relaxed" tolerances both placed on parameters and objective function, e.g., of 100. When the k^{th} trial is acceptable (e.g., first order optimality conditions are satisfied to the specified tolerance or changes in parameters or function are smaller than the specified tolerance), the $(k + 1)^{\text{th}}$ trial restarts with a 10-fold decreased tolerance, where the k^{th} local minimizer Θ^*_k is used for the new starting point. The previous trial provides a more adequate starting point for the next, and it can be advantageous for local search methods. This is because the efficiency of local methods depends heavily on the quality of starting points [14,19]. In case the current trial fails (e.g., by exceeding the maximum number of iterations allowed in the estimation process), the tolerance is increased and the next trial starts from the iterate Θ° recorded for the smallest function value up to the current trial in the feasible region (Figure 1-E). The tolerance at which an optimization trial fails is registered. In case optimization trials successively fail, which is rarely the case, the step of tolerance increase is set smaller such as 5-, 10-, 50-, or 100-fold, etc., of the registered value. Correspondingly, the tolerance decrease is also made by this reduced step until the current tolerance matches the registered. This allows finding an appropriate tolerance value for NLSP.

At the beginning and at every failed trial, α is re-adjusted by referencing the condition number of the model's Jacobian matrix. When the new starting point for the k^{th} trial is Θ° that is isolated from the previous $(k - n)^{\text{th}}$ trial with α_{k-n}, the [0, 1)-fluxes have to be rescaled in accordance with the new scaling constant α_k. Since $\nu = \alpha_{k-n}\varphi_{k-n}/(1 - \varphi_{k-n})$ from (2), substituting ν in $\varphi_k = \nu/(\alpha_k + \nu)$ results in:

$$\varphi_k = \frac{\alpha_{k-n}\varphi_{k-n}/(1-\varphi_{k-n})}{\alpha_k + \alpha_{k-n}\varphi_{k-n}/(1-\varphi_{k-n})}. \qquad (A.3)$$

Deterministic Method of Feasible Starting Point Generation

When the random generation of starting points (see Figure 1-A!) is difficult, the feasible starting points can be found deterministically by minimizing a suitable objective. For instance, one can minimize the following objective:

$$\min_{\boldsymbol{\Theta}} \; 2^{N_{\text{neg.flux}}} + \frac{1}{\left| \log 10 \sum_{i=1}^{n} |b_{\text{flux}}| \right|} \quad \text{subject to} \quad \begin{cases} 2^{N_{\text{neg.flux}}} = 1 \\ 0 \le \varphi_i < 1 \\ \max(0, -\sqrt{\chi^2_{1,\phi}}\boldsymbol{\sigma}_{\text{efflux}}) \le \mathbf{v}_{\text{efflux}} \le \sqrt{\chi^2_{1,\phi}}\boldsymbol{\sigma}_{\text{efflux}} \end{cases} \qquad (A.4)$$

Here, $N_{\text{neg.flux}}$ signifies the number of negative-valued fluxes, b_{flux} the steady-state stoichiometric flux balances set up for metabolites (see the next section!) evaluated at the current iterate Θ_k, and n the number of the balances. By the first term in the objective and the first constraint, the minimization is directed such that all fluxes become non-negative. The second term is to prevent flux values calculated that may not obey stoichiometry due to the limited floating-point accuracy (round-off error). This approach is useful for preparing initial values while preserving random nature for Monte Carlo simulations. In practice, the upper bound of φ_i can be set to a value smaller than 1 that gives sufficiently large flux values so that φ_i has a closed interval.

Stationary State Stoichiometric Balances are set up around 15 metabolites specified in Figure 2 as follows. The values multiplied with Y_{XS} equals the strain specific precursor demand for growth (mmol precursor per gram biomass).

glucose 6-P: $\nu_2 - (\nu_1 + \nu_{2r} + 0.4217Y_{XS}) = 0$

fructose 6-P: $(\nu_{2r} + \nu_4 + \nu_5 + \nu_6) - (\nu_2 + \nu_{4r} + \nu_{5r} + \nu_{6r}) = 0$

glyceraldehyde 3-P: $(\nu_3 + \nu_{4r} + \nu_5 + 2\nu_{6r} + \nu_7) - (\nu_{3r} + \nu_4 + \nu_{5r} + 2\nu_6 + \nu_{7r} + 0.4659Y_{XS}) = 0$

3-phosphoglycerate: $(\nu_{7r} + \nu_8) - (\nu_7 + \nu_{8r} + 0.2209Y_{XS}) = 0$

phosphoenolpyruvate: $(\nu_{16} + \nu_{8r}) - (\nu_8 + \nu_9 + 1.7854Y_{XS}) = 0$

pyruvate: $(\nu_9 + \nu_{15}) - (\nu_{10} + \nu_{15r} + 0.9964Y_{XS}) = 0$

pentose 5-P: $(\nu_1 + 2\nu_{3r} + \nu_{5r}) - (2\nu_3 + \nu_5 + 1.1622Y_{XS}) = 0$

erythrose 4-P: $(\nu_4 + \nu_{5r}) - (\nu_{4r} + \nu_5 + 0.4659Y_{XS}) = 0$

sedoheptulose 7-P: $(\nu_3 + \nu_{4r}) - (\nu_{3r} + \nu_4) = 0$

acetyl-CoA: $\nu_{10} - (\nu_{11} + \text{acetate}_{ex} + 3.1585Y_{XS}) = 0$

isocitrate: $\nu_{11} - \nu_{12} = 0$

α-ketoglutarate: $(\nu_{12} + \nu_{17}) - (\nu_{13} + \nu_{17r} + \alpha\text{-ketoglutarate}_{ex}) = 0$

glutamate: $(\nu_{17r} + \text{glutamate}_{ex}) - (\nu_{17} + 1.4769Y_{XS}) = 0$

succinate: $(\nu_{13} + \nu_{14r} + \text{succinate}_{ex}) - (\nu_{14} + \text{fumarate}_{ex}) = 0$

oxaloacetate: $(\nu_{14} + \nu_{15r}) - (\nu_{11} + \nu_{14r} + \nu_{15} + \nu_{16} + 1.3131Y_{XS}) = 0$

Additional material

Additional file 1

Optimal Experimental Design. the relative information index expected from the tracer designs consisting of different ^{13}C-glutamate and ^{13}C-succniate for single and two parallel experiments.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1752-0509-2-29-S1.doc]

Additional file 2

Hybridoptimizer. an example of the hybrid optimization for the problem described in the current manuscript coded using MATLAB.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1752-0509-2-29-S2.exe]

Acknowledgements

O. Frick thanks DFG (Deutsche Forschungsgemeinschaft), project He 3092/6 for the support of his Ph.D. We wish to thank Andrew Marsh at the University of Louisville's James Graham Brown Cancer Center for editing.

References

1. Wiechert W: **13C metabolic flux analysis.** *Metab Eng* 2001, **3(3):**195-206.
2. Yang TH, Wittmann C, Heinzle E: **Metabolic network simulation using logical loop algorithm and Jacobian matrix.** *Metab Eng* 2004, **6(4):**256-267.
3. Antoniewicz MR, Kelleher JK, Stephanopoulos G: **Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions.** *Metab Eng* 2007, **9(1):**68-86.
4. Yang TH, Heinzle E, Wittmann C: **Theoretical aspects of 13C metabolic flux analysis with sole quantification of carbon dioxide labeling.** *Comput Biol Chem* 2005, **29(2):**121-133.
5. Wiechert W, Mollney M, Isermann N, Wurzel M, de Graaf AA: **Bidirectional reaction steps in metabolic networks: III. Explicit solution and analysis of isotopomer labeling systems.** *Biotechnol Bioeng* 1999, **66(2):**69-85.
6. Wiechert W, Siefke C, de Graaf A, Marx A: **Bidirectional reaction steps in metabolic networks: II. Flux estimation and statistical analysis.** *Biotechnol Bioeng* 1997, **55:**118-135.
7. Wittmann C, Heinzle E: **Genealogy profiling through strain improvement by using metabolic network analysis: metabolic flux genealogy of several generations of lysine-producing Corynebacteria.** *Appl Environ Microbiol* 2002, **68(12):**5843-5859.
8. Arauzo-Bravo MJ, Shimizu K: **An improved method for statistical analysis of metabolic flux analysis using isotopomer mapping matrices with analytical expressions.** *J Biotechnol* 2003, **105(1-2):**117-133.
9. Dauner M, Bailey JE, Sauer U: **Metabolic flux analysis with a comprehensive isotopomer model in Bacillus subtilis.** *Biotechnol Bioeng* 2001, **76(2):**144-156.
10. Forbes NS, Clark DS, Blanch HW: **Using isotopomer path tracing to quantify metabolic fluxes in pathway models containing reversible reactions.** *Biotechnol Bioeng* 2001, **74(3):**196-211.
11. Schmidt K, Nielsen J, Villadsen J: **Quantitative analysis of metabolic fluxes in Escherichia coli, using two-dimensional NMR spectroscopy and complete isotopomer models.** *J Biotechnol* 1999, **71(1-3):**175-189.
12. Zhao J, Shimizu K: **Metabolic flux analysis of Escherichia coli K12 grown on 13C-labeled acetate and glucose using GC-MS and powerful flux calculation method.** *J Biotechnol* 2003, **101(2):**101-117.
13. Floudas CA, Pardalos PM: **Recent advances in global optimization.** In *Princeton series in computer science* Princeton, N.J. , Princeton University Press; 1992:x, 633 p..
14. Nocedal J, Wright SJ: **Numerical optimization.** In *Springer series in operations research* New York , Springer; 1999:xx, 636 p..
15. Press WH: **Numerical recipes in C: the art of scientific computing.** 2nd edition. Cambridge ; New York , Cambridge University Press; 1992:xxvi, 994 p..
16. Brackin P, Colton SC: **Using genetic algorithms to set target values for engineering characteristics in the house of quality.** In *J Comput Inf Sci Eng Volume* 2. Issue 2 ASME; 2002:106-114.
17. Kelner V, Capitanescu F, Léonard O, Wehenkel L: **An hybrid optimization technique coupling an evolutionary and a local search algorithm .** *J Comput Appl Math* 2007 in press.
18. Lambert TW, Hittle DC: **Optimization of autonomous village electrification systems by simulated annealing.** *Solar Energy* 2000, **68(1):**121-132.
19. Mendes P, Kell D: **Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation.** *Bioinformatics (Oxford, England)* 1998, **14(10):**869-883.
20. Moles CG, Mendes P, Banga JR: **Parameter estimation in biochemical pathways: a comparison of global optimization methods.** *Genome research* 2003, **13(11):**2467-2474.
21. Xu P: **A hybrid global optimization method: The multi-dimensional case.** *J Comput Appl Math* 2003, **155(2):**423-446.
22. Long CE, Polisetty PK, Gatzke EP: **Nonlinear model predictive control using deterministic global optimization.** *Journal of Process Control* 2006, **16(6):**635-643.
23. Nash SG, Sofer A: **Linear and Nonlinear Programming.** New York , McGraw-Hill; 1996.
24. Wiechert W, de Graaf A: **Bidirectional reaction steps in metabolic networks: I. Modeling and simulation of carbon isotope labeling experiments.** *Biotechnol Bioeng* 1997, **55:**102-117.
25. Buchberger B: **An algorithmical criterion for the solvability of algebraic system of equation.** *Aequationes Mathematicae* 1988, **4:**45-50.
26. Saccomani MP: **Some results on parameter identification of nonlinear systems.** *Cardiovascular Engineering: An International Journal* 2004, **4:**95-102.
27. Wiechert W: **Algebraic methods for the analysis of redundancy and identifiability in metabolic 13C labelling systems.** In *Bioinformatics: From nucleic acids and proteins to cell metabolism* Edited by: Lessel U. Weinheim , Verlag Chemie; 1995:169-184.
28. van Winden WA, Heijnen JJ, Verheijen PJ, Grievink J: **A priori analysis of metabolic flux identifiability from 13C-labeling data.** *Biotechnol Bioeng* 2001, **74(6):**505-516.

29. Isermann N, Wiechert W: **Metabolic isotopomer labeling systems. Part II: structural flux identifiability analysis.** *Math Biosci* 2003, **183(2):**175-214.
30. Klamt S, Schuster S, Gilles ED: **Calculability analysis in underdetermined metabolic networks illustrated by a model of the central metabolism in purple nonsulfur bacteria.** *Biotechnol Bioeng* 2002, **77(7):**734-751.
31. Boggs PT, Tolle JW: **Sequential quadratic programming.** *Acta Numerica* 1995, **4:**1-51.
32. Coleman TF, Li Y: **On the convergence of reflective Newton methods for large-scale nonlinear minimization subject to bounds.** Ithaca, NY , Cornell Theory Center, Cornell University; 1992:36 p..
33. Schittowski K: **NLQPL: A FORTRAN-subroutine solving constrained nonlinear programming problems.** *Annals of Operations Research* 1985, **5:**485-500.
34. Sauer U, Hatzimanikatis V, Hohmann HP, Manneberg M, van Loon AP, Bailey JE: **Physiology and metabolic fluxes of wild-type and riboflavin-producing Bacillus subtilis.** *Appl Environ Microbiol* 1996, **62(10):**3687-3696.
35. Wittmann C, Heinzle E: **Application of MALDI-TOF MS to lysine-producing Corynebacterium glutamicum: a novel approach for metabolic flux analysis.** *Eur J Biochem* 2001, **268(8):**2441-2455.
36. Hill MC, Osterby O: **Determining extreme parameter correlation in ground water models.** *Ground water* 2003, **41(4):**420-430.
37. Balasubramanian P, Bettina SJ, Pushpavanam S, Balaraman KS: **Kinetic parameter estimation in hydrocracking using a combination of genetic algorithm and sequential quadratic programming.** *Ind Eng Chem Res* 2003, **42(20):**4723-4731.
38. Klepeis JL, Pieja MJ, Floudas CA: **Hybrid global optimization algorithms for protein structure prediction: alternating hybrids.** *Biophysical journal* 2003, **84(2 Pt 1):**869-882.
39. Xu YG, Li GR, Wu ZP: **A novel hybrid genetic algorithm using local optimizer based on heuristic pattern move.** *Applied Artificial Intelligence* 2001, **15:**601-631.
40. Tahk MJ, Woo HW, Park MS: **A hybrid optimization method of evolutionary and gradient search.** *Engineering Optimization* 2007, **39(1):**87-104.
41. Mahinthakumar GK, Mohamed S: **Hybrid genetic algorithm - Local seach methods for solving groundwater source identification inverse problems.** *J Water Resour Plng and Mgmt* 2005, **131:**45-57.
42. Antoniewicz MR, Kraynie DF, Laffend LA, Gonzalez-Lergier J, Kelleher JK, Stephanopoulos G: **Metabolic flux analysis in a nonstationary system: fed-batch fermentation of a high yielding strain of E. coli producing 1,3-propanediol.** *Metab Eng* 2007/04/03 edition. 2007, **9(3):**277-292.
43. Massart DL: **Handbook of chemometrics and qualimetrics.** In *Data handling in science and technology; v 20* Amsterdam; New York , Elsevier; 1997.
44. Mollney M, Wiechert W, Kownatzki D, de Graaf AA: **Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments.** *Biotechnol Bioeng* 1999, **66(2):**86-103.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

BioMedcentral