Computers and Chemical Engineering 71 (2014) 728–733

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

A constrained wavelet smoother for pathway identification tasks in systems biology

Sepideh Dolatshahi, Brani Vidakovic, Eberhard O. Voit*

The Wallace H. Coulter Department of Biomedical Engineering at Georgia Tech and Emory University, 313 Ferst Drive, Suite 4103, Atlanta, GA 30332-0535, USA

ARTICLE INFO

Article history:
Received 8 May 2014
Accepted 20 July 2014
Available online 30 July 2014

Keywords:
Biochemical systems theory (BST)
Dynamic flux estimation (DFE)
Metabolic pathway analysis
Parameter estimation
Smoothing
Wavelets

ABSTRACT

Metabolic time series data are being generated with increasing frequency, because they contain enormous information about the pathway from which the metabolites derive. This information is not directly evident, though, and must be extracted with advanced computational means. One typical step of this extraction is the estimation of slopes of the time courses from the data. Since the data are almost always noisy, and the noise is typically amplified in the slopes, this step can become a critical bottleneck. Several smoothers have been proposed in the literature for this purpose, but they all face the potential problem that smoothed time series data no longer correspond to a system that conserves mass throughout the measurement time period. To counteract this issue, we are proposing here a smoother that is based on wavelets and, through an iterative process, converges to a mass-conserving, smooth representation of the metabolic data. The degree of smoothness is user defined. We demonstrate the method with some didactic examples and with the analysis of actual measurements characterizing the glycolytic pathway in the dairy bacterium *Lactococcus lactis*. MATLAB code for the constrained smoother is available as a Supplement.

1. Introduction – motivation

Data characterizing genomic, proteomic or metabolic processes in the form of time series measurements contain very valuable, yet implicit information about the structure and dynamics of biological systems. In an effort to gain deeper insight into these systems, numerous recent articles have been addressing the extraction of this information and its integration into functional models, which may subsequently be utilized for explanation, prediction, manipulation, and optimization. The majority of methods for the extraction of information from time series employ techniques for minimizing the discrepancy between the measured data, *i.e.*, the time profiles, and the assumed model, which typically consists of a system of non-linear ordinary differential equations that are to be parameterized (Chou and Voit, 2009). The currently available estimation techniques include different regression, simulated annealing, or evolutionary optimization approaches, such as genetic algorithms, as well as various support algorithms for preprocessing the data. The latter algorithms are intended to reduce noise and smooth the time courses, which naturally are often quite ragged.

Dynamic flux estimation (DFE) (Goel et al., 2008) is a methodological framework for extracting information from time series measurements. It is distinct from all other methods, as it does not presume knowledge of an appropriate underlying model. DFE combines the tenets of stoichiometric (Gavalas, 1968; Heinrich and Schuster, 1996; Stephanopoulos et al., 1998) and flux balance analysis (FBA; *e.g.*, (Palsson, 2006)), which are genuinely static, with dynamic aspects of ODE modeling. DFE consists of two phases. The first phase is model-free and essentially assumption-free and includes steps of data preprocessing, time course smoothing, and slope estimation, and ultimately yields point-wise inferences of dynamic time series profiles for all fluxes in the system. Expressed differently, the first phase results in a numerical representation of flux values throughout the time period of the experiment. One must note, however, that this numerical representation does not directly reveal optimal, or even appropriate explicit functional forms that capture the dynamics of the fluxes in the system. The second phase addresses this issue. It consists of the mathematical characterization of the numerical flux profiles, based on an assumed format. Fig. 1 summarizes the phased approach of DFE and the expected outcomes of each step.

* Corresponding author. Tel.: +1 404 385 5057; fax: +1 404 894 4243.
E-mail addresses: sepid.dsh@gmail.com (S. Dolatshahi), brani@bme.gatech.edu (B. Vidakovic), eberhard.voit@bme.gatech.edu (E.O. Voit).

http://dx.doi.org/10.1016/j.compchemeng.2014.07.019
0098-1354/

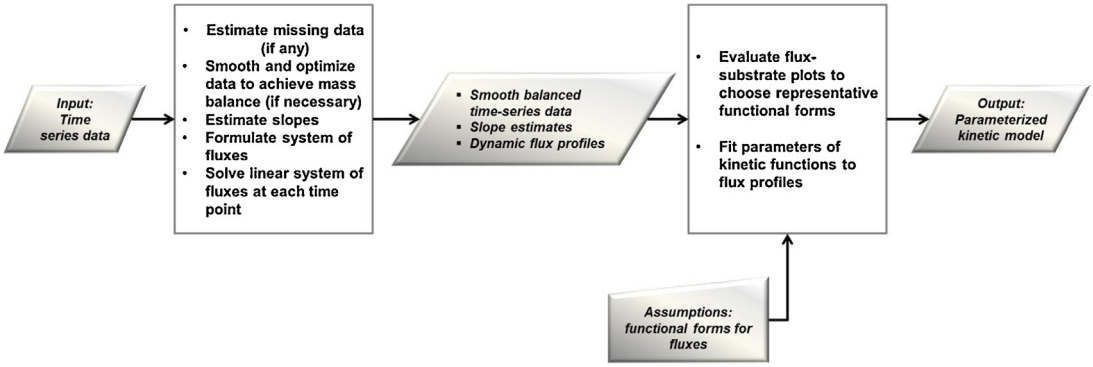

Fig. 1. Phases and steps of the dynamic flux estimation (DFE) technique utilized for metabolic time-series data.

While the performance of DFE can be excellent, it has the significant drawback that a direct application of the method requires a stoichiometric matrix that has full rank. This is seldom the case, because most pathway systems contain more fluxes than metabolites. Several auxiliary methods have been proposed that "fill the rank" with additional information ((Chou and Voit, 2012; Iwata et al., 2013; Voit et al., 2009); see also (Voit, 2013)) but none of them presents a perfect solution.

A secondary issue of a more computational nature is the need for using slopes of the time courses in the model-free phase. On the one hand, the use of slopes is very advantageous, because parameter values may be estimated without the integration of differential equations (Varah, 1982; Voit and Savageau, 1982a,b). Indeed, the integration of a system of differential equations is computationally expensive and prone to a host of technical challenges, associated with complicated error surfaces that can contain numerous local minima (Voit and Almeida, 2004). On the other hand, the slopes are rather sensitive to noise in the time courses, which renders it necessary to smooth and balance the data. Smoothing reduces noise, while balancing assures that there is no gain or loss of mass over time in a closed system.

Numerous methods have been proposed for smoothing time course data. They include splines, moving average algorithms, finite difference approximations, and various types of non-linear programming (Eilers, 2003; Vilela et al., 2007; Whittaker, 1923). These methods are time consuming and need to be performed interactively, or at least in a closely supervised manner. Furthermore, this type of smoothing process can lead to secondary issues. Especially important for the purposes of metabolic pathway analysis is the potential problem that the overall mass in a system may no longer be constant if the data are smoothed. To address these issues, we propose here an automated smoothing technique that takes as input any given data set and estimates and removes noise while at the same time satisfying the required mass balance within the system. The proposed approach is iterative and called constrained iterative wavelet-based smoother (CIWS).

2. Background and data

2.1. Multiresolution analysis using wavelets

The proposed smoothing technique is built upon the notion of multiresolution analysis (MRA) from wavelet theory which we will briefly explain here.

Wavelets are becoming a standard data analysis tool that is excellent for tasks of data compression as well as for denoising and smoothing. One of their advantages is that they are flexible as well as local, which means that they do not ignore desirable functional details. The reason is that the resolution in MRA can be adapted to the situation at hand.

Mathematically speaking, wavelets are orthogonal basis functions which span the space of all square-integrable functions ($L^2(\boldsymbol{R})$). Thus, any element in $L^2(\boldsymbol{R})$ may be represented as a possibly infinite linear combination of these basis functions. An important property of this linear representation is that it may be partitioned into orthogonal subspaces $W_j = span[\psi_{j,k}(x)]$, each of which captures a certain level of "detail" information. The key concept of orthogonal MRA is to partition a given function $f(x)$ into its components $f^{(j)}(x) \in W_j$. Here, the space W_j consists of functions with lower resolution than the ones in W_{j+1} which means that if some arbitrary function $g(x)$ is in W_j, then $g(2x)$ is in W_{j+1} (Strang, 1989).

For example, in the traditional wavelet representation

$$f(x) = \sum\nolimits_{k\in Z} C_{J_0,k}\phi_{J_0,k}(x) + \sum\nolimits_{j\geq j_0, k\in Z} d_{j,k}\psi_{j,k}(x)$$

the second sum contains the terms which capture the higher levels of detail (*i.e.*, $U_{j\geq j_0} W_j$, which is the union of all levels of detail greater than or equal to J_o). Choosing the appropriate coarsest resolution J_0 gives rise to different transforms. We can also just approximate

$$f(x) \cong \sum\nolimits_{k\in Z} c_{J_0,k}\phi_{J_0,k}(x).$$

The choice of J_0 provides us with the flexibility of selecting the desired level of detail, which is traded against the desired level of smoothness. In the above representation of $f(x)$, the functions

$$\phi_{J_0,k}(x) = 2^{j/2}\phi(2^j x - k)$$

and $\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$ are scaling and wavelet functions, which correspond to commonly called "smooth" and "detail" coefficients, respectively; j is the dilation/scale index, and k indicates shift or position (Vidakovic, 1999).

In wavelet decomposition, as mentioned before, the wavelet coefficients represent details, and if these are small, they can actually be removed without affecting the general trend of the data. In fact, wavelet transformations are known to be parsimonious in that they can be well described by a relatively small number of "energetic" wavelet coefficients.

Wavelet thresholding is the process of removing the wavelet coefficients that are smaller in magnitude than some threshold λ. The resulting signal, after the inverse wavelet transformation, is expected to have its noise removed or at least reduced. The characteristics of the data determine the magnitude of noise, and it is therefore useful to specify the threshold value magnitude of noise, and it is therefore useful to specify the threshold λ based on the variability of the data at hand. Different thresholding policies and threshold values are discussed in Section 4 in more detail.

All wavelet computations were performed in *WaveLab*, a MATLAB wavelet toolbox available from the website of Stanford University (Buckheit and Donoho, 1995). Sample MATLAB codes

using the functions available in the *WaveLab* toolbox are available in the Supplement B.

2.2. Description of data

The proposed smoothing method was tested with time series data of metabolite concentrations in the glycolytic pathway of the bacterium *Lactococcus lactis*. These data had been generated with carbon based *in vivo* nuclear magnetic resonance (^{13}C NMR) techniques.

The dataset, which is published in Carvalho et al. (2013), consists of time courses of glucose, lactate, UDP-glucose (uridine diphosphate glucose), minor amounts of ethanol and 2,3-butanediol, as well as intermediate metabolites including FBP (fructose 1,6-bisphosphate) and 3PGA (3-phosphoglyceric acid). These time courses were obtained from non-growing cells of *L. lactis*, following a bolus of 40 mM of [1-^{13}C] glucose, at 30 °C. The experiments were executed under anaerobic conditions, and pH was controlled at 6.5 and monitored online by *in vivo* ^{13}C NMR.

3. Constrained iterative wavelet-based smoother (CIWS)

3.1. Basic concepts of CIWS

During the model-free phase of DFE, noisy time courses are to be smoothed for later slope estimation and balanced in such a fashion that no material is gained or lost (Goel et al., 2008). The latter aspect is not a triviality because the model is based on the implicit assumption that all material is accounted for. Thus, if the (smoothed) data do not maintain mass conservation, the model structure is immediately at odds with the data, and the estimation process will introduce undesirable means of numerical compensation.

While several advanced algorithms for general smoothing have been developed (*e.g.*, Eilers, 2003; Vilela et al., 2007; Voit and Almeida, 2004; Whittaker, 1923), they have not ascertained the need to conserve mass. The task at hand here is therefore to smooth the time series $f_i(t)$, $i = 1,\ldots,n$, conditioned on the constraint that their sum remains constant in time. We approach this task by constructing a wavelet transform of each $f_i(t)$ in the following form

$$f_i(t) = \sum_{k\in Z} c^{(i)}_{J_0 k}\phi_{J_0,k}(t) + \sum_{j\geq J_0, k\in Z} d^{(i)}_{j,k}\psi_{j,k}(t).$$

Again, the functions ϕ_{J_0}, k and $\psi_{j,k}$ are scaling and wavelet functions respectively, and (j,k) is standard scale/shift wavelet indexing. The two sets of coefficients C_{J_0}, k, $d_{j,k}$ have a direct interpretation: the smooth and coarse coefficients C_{J_0}, k are responsible for trends and global features, while the detail coefficients $d_{j,k}$ describe mostly the noise in the decomposed time series $f_i(t)$. By thresholding the detail coefficients, that is, by setting to zero those coefficients small in magnitude, and inverting back to the domain of original data, the individual functions are smoothed. We will discuss later how small this magnitude is to be set. In the following, we will denote a smoothed version of f_i as f_i^*. In practical applications, wavelet decompositions of a given sampled function f are found by Mallat's algorithm (Mallat, 1989). This algorithm consists of data filtering by two filters h and g which are low pass and high pass wavelet filters. The form of filters is fully determined by the choice of the scaling and wavelet functions.

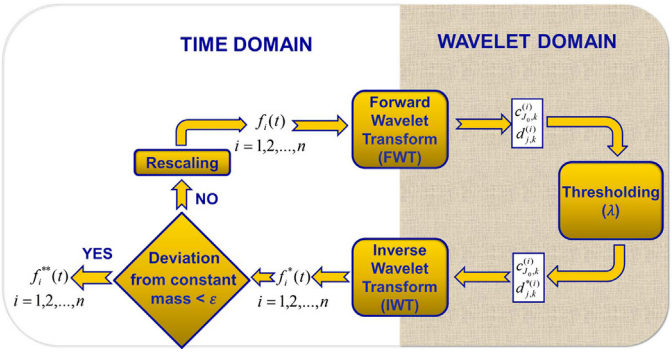

Fig. 2. Diagram of the constrained iterative wavelet smoothing (CIWS) technique.

Unless further precautions in the process of smoothing are introduced, the mass balance among all metabolites will likely be violated, leading to $\sum_i f_i^*(t) = g(t) \neq C$. We propose to balance the sum by rescaling each $f_i^*(t)$ to $f_i^{**}(t) = Cf_i^*(t)/g(t)$ and to repeat the process of wavelet-transforming, thresholding, back-inverting, and rescaling. While a rigorous proof of convergence is not possible in generality, simulation results with various representative examples (see later) suggest that this procedure generally leads to a set of smooth metabolic time trend functions, while ascertaining a constant mass balance over time. Indeed, as we will discuss later, the method leads to smooth time courses in our application. Fig. 2 depicts the concepts of this procedure.

This constrained iterative wavelet-based smoother (CIWS) consists of an iterated 3-step procedure.

Step 1 (. Construct wavelet transforms of each time series in terms of the scaling and wavelet functions ϕ and ψ.). For this task, we have a choice among an infinite set of basis functions. As a default, we use the Daubechies, Symmlet, and Coiflet families (Daubechies, 1992):

$$\begin{cases} f_i(t) = 1, 2, \ldots, m \\ \sum_i f_i(t) = C, \forall t \end{cases} \xrightarrow{\text{Wavelet transform}} f_i(t) = \sum_{k\in Z} c^{(i)}_{J_0,k}\phi_{J_0,k}(t) + \sum_{j\geq J_0, k\in Z} d^{(i)}_{j,k}\psi_{j,k}(t)$$

Step 2 (. Threshold the detail wavelet coefficients and invert back to smooth time–domain functions.). Detail coefficients describe the noise in the decomposed time series $f_i(t)$, while the smooth and coarse detail coefficients are responsible for trends and global features. Hard thresholding operates on the detail coefficients by setting $d^{*(i)}_{j,k} = \begin{cases} 0 & \text{if } |d^{(i)}_{j,k}| < \lambda \\ d^{(i)}_{j,k} & \text{if } |d^{(i)}_{j,k}| \geq \lambda \end{cases}$. The threshold λ is estimated from the standard deviation of the time series data, as outlined before. Once Step 2 is performed and the result inverted back to the time domain, one obtains time series $f_i^*(t)$, $i = 1, 2, \ldots, m$, which however are not mass conserving throughout time (*i.e.*, $\sum_i f_i^*(t_l = g(t) \neq C)$).

Step 3 (. Recover mass balance by appropriately rescaling each time series.). If the rescaled functions $f_i^{**}(t) = Cf_i^*(t)/g(t)$ are sufficiently smooth, then terminate the smoothing process. Otherwise return to Step 1. Here, sufficient smoothness is defined as $||\vec{g} - C|| < \varepsilon$, where $\varepsilon > 0$ is an acceptable error for deviations from mass conservation and $\vec{g} = (g(t_1), g(t_2), \ldots, g(t_N))$, $g(t_l) = \sum_i f_i^*(t_l) \neq C$.

3.2. Estimating the appropriate threshold and wavelet functions

We consider N noisy time points of a function $f_i(t)$ sampled at time points t_j, $j = 1,\ldots,N$:

$$f_i(t_j) = y_j + z_j, j = 1, \ldots, N$$

Here the noise is assumed to be white $z_j \sim WN(0,\sigma^2)$; expressed differently the noise is independent and identically distributed with

$$E\{Z_j\} = 0, \quad Var\{Z_j\} = \sigma^2.$$

Our goal is to estimate a smoothed vector of time points $f_i^{**} = [y_1, y_2, \ldots, y_N]$. As a first step it seems that we need to specify a class $\mathfrak{F}$ of sampled functions to which y_i, $i = 1,\ldots, m$ are supposed to belong. However, we often have no *a priori* knowledge about such a class $\mathfrak{F}$. Correspondingly, we assume no functional form for the manner with which the concentrations of metabolites change over time. Instead, determining the appropriate wavelet function, thresholding, and defining a threshold value are central to the performance and convergence of the proposed CIWS smoothing technique.

The process of thresholding wavelet coefficients may be divided into two steps. First, we need to choose an appropriate thresholding rule. In the step-by-step description of the CIWS algorithm, hard thresholding was introduced as the method of choice in principle, but other standard choices are available, including soft thresholding (Donoho and Johnstone, 1994), the non-negative garrote shrinkage (Breiman, 1995), and others (Vidakovic, 1999). The most prominent thresholding rules are represented in Eqs. (1)–(3) below:

$$T^{hard}(d, \lambda) = d\mathbf{1}(|d| > \lambda) \tag{1}$$

$$T^{soft}(d, \lambda) = (d - sgn(d)\lambda)\mathbf{1}(|d| > \lambda) \tag{2}$$

$$T^{garrote}(d, \lambda) = \left(d - \frac{\lambda^2}{d}\right)\mathbf{1}(|d| > \lambda) \tag{3}$$

where $\mathbf{1}(|d| > \lambda) = \begin{cases} 0 & |d| \le \lambda \\ 1 & |d| > \lambda \end{cases}$. Both hard and soft thresholding have advantages and disadvantages. Hard thresholding was discussed before as a viable option. Soft shrinkage tends to have a higher bias since it shrinks large coefficients. This tends to impair the convergence of our iterative method. The non-negative garrote shrinkage function was first introduced by Breiman (1995) in a different context and tends to improve the performance of hard thresholding slightly while preserving the convergence properties.

No matter what thresholding policy is chosen, a threshold value is to be determined. For each iteration of CIWS, the standard deviation of the remaining noise is estimated directly from the data. The threshold λ is defined as a linear function of the standard deviation of the sampled $f_i^*(t)$, because the noise is expected to decrease as the number of iterations increases. Thus, we set $\lambda = \gamma \cdot \sigma_i$ where σ_i is the standard deviation of the ith time series of concentrations. The coefficient γ still needs to be determined in order for the smoothed function to have the required and desired properties.

Because of implementation issues, the number of data points needs to be a power of two. When this is not the case in a given dataset, we artificially do a mirror-image extension of the last q data points so that the total length is a power of 2. This strategy facilitates smoothness at the boundaries of the observed dataset and avoids the so-called Gibbs effect associated with very short time series; it also yields more robust standard deviation estimates. Assuming that noise in the data is white with variance σ_i^2, the universal threshold of $\lambda_i = \sqrt{2\log n}\sigma_i$ was shown to remove noise with high probability, thus contributing to the visual quality of the reconstructed signals (Donoho and Johnstone, 1994).

3.3. Selecting an appropriate wavelet function

For any wavelet generating multiresolution analysis (MRA), there is a trade-off between the smoothness of the wavelet and scaling functions on the one hand, and the locality (that is, the accuracy of local representation) of the wavelets and wavelet representations on the other. Since a wavelet representation of discrete data must interpolate between the points with a shift of the scaling function, it is important that the scaling function generating MRA is smooth, because the wavelet decomposition represents a smooth physical process. However, a wavelet that is too smooth, due to a long wavelet filter, is not sufficiently local and modifies possibly important fluctuations far from the location that it wants to address. In the CIWS algorithm, the functions are iteratively rescaled to ensure a constant sum after each iterative smoothing step. This rescaling is driven point-wise, and for this reason it is important that wavelets retain sufficient locality. Indeed, wavelets that are too smooth and not sufficiently local may cause the CIWS algorithm to diverge. Details regarding the selection of appropriate wavelet functions among the standard families (Daubechies, Coiflet, Symmlet, and Pollen) are presented in Supplement A.

A traditional measure of smoothness of a function f is the Holder exponent α, which is defined through the following inequality: $(\forall x, y)\, \exists\, C \ge 0, \alpha \ge 0 : |f(x) - f(y)| \le C|x - y|^{\alpha}$. It was shown that the Holder exponent may be expressed in terms of a large enough vanishing moment M, namely as $\alpha \cong 0.2075\,\mathrm{M}$ (Vidakovic, 1999). Here, the kth moment of a wavelet function ψ is defined as $\int_{-\infty}^{\infty} x^k \psi(x)dx$, and the terminology that "a wavelet function has M vanishing moments" means that this integral is zero for $k = 0, 1,\ldots, M-1$. For the Daubechies family, for example, M vanishing moments are achieved by discrete wavelet filters that are of length 2 m.

One useful criterion for the choice of an adequate wavelet basis is entropy (Coifman and Wickerhauser, 1992): Among different candidate wavelet functions, the most appropriate is chosen by minimizing the maximum entropy among different time series of concentrations. Supplement A describes how this criterion facilitates the identification of the best wavelet functions for the purpose of smoothing. For our case of metabolic time series data, Coiflet 1 results in the lowest maximum entropy. This wavelet also well balances smoothness and locality and is differentiable. Supplement Table A.1 provides entropy values for our case study.

3.4. Avoidance of negative concentrations during back-conversion to the time domain

A typical problem with any wavelet-based estimation of non-negative functions and densities is the fact that the estimators may not be fully non-negative. Only the so-called Haar wavelet has a non-negative scaling function and induced kernel, and always results in a non-negative estimator; however, it is known that the convergence rates for strictly positive kernels are inferior to general kernel-based methods, where the kernels can be locally negative (Walter et al., 2001). Wavelet-based kernels (except for the Haar) are necessarily locally negative.

A common strategy to circumvent this issue is to fit and smooth log functions or square root functions and, once fitting and smoothing are accomplished, use an exponential transformation or square the results. For our case, this strategy poses a problem with controlling the mass balance. Thus, instead of a transformation, our solution uses a smooth wavelet and always curtails the negative components when iterating. Beyond satisfying our needs, this strategy could be tailored toward constraints other than mass conservation.

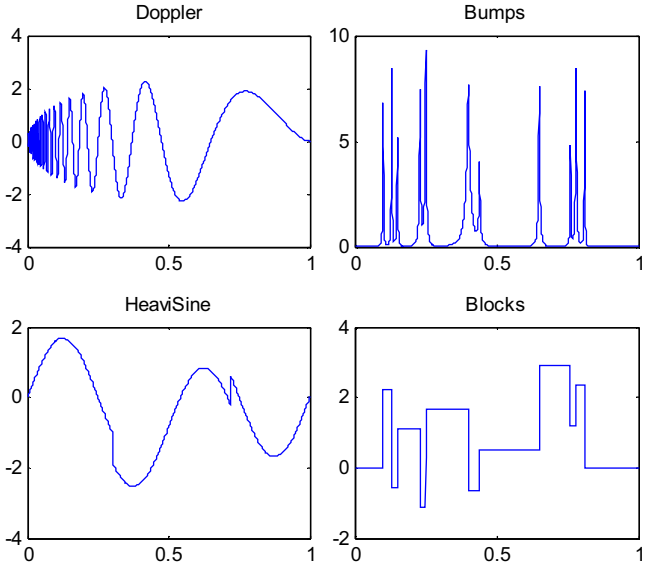

Fig. 3. Four of the set of five test functions, called *Doppler*, *Bumps*, *HeaviSine*, and *Blocks* without noise.

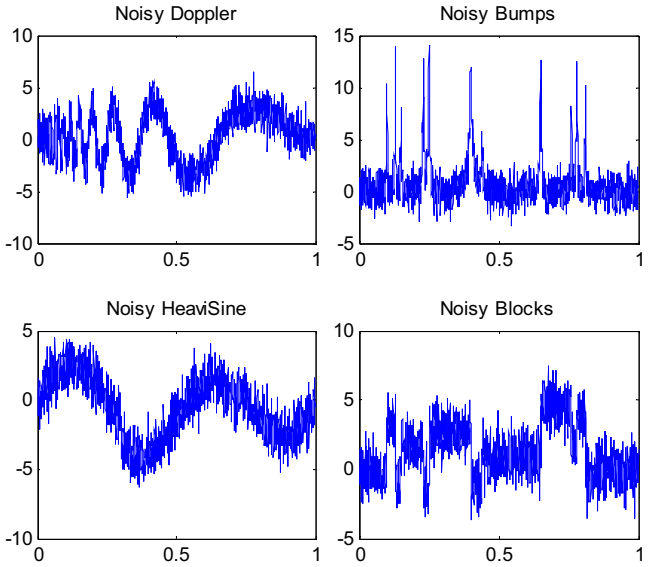

Fig. 4. Four of the set of five test functions called *Doppler*, *Bumps*, *HeaviSine*, and *Blocks* with additive white (Gaussian) noise of SNR = 15 dB.

4. Results

4.1. Convergence of the constrained smoother

For the metabolic time series in our case study, CIWS converges in fewer than 40 steps to smooth, mass-conserving time courses, and the degree of smoothness is adjustable to our specifications.

A formal, general proof of convergence of CIWS seems not possible, as it would require knowledge of the nature of the functions $f_i(t)$, constrained to some smoothness spaces, and the interplay between thresholding strategies and types of wavelets. The task might be feasible if the goal function, consisting of the residual error to be minimized, were convex. However due to the arbitrary nature of the functions to be smoothed and the wavelets, the task of analytically optimizing the goal function is intractable. Instead we assess convergence with Monte Carlo simulations and representative functional shapes. Specifically, we use synthetic data from a battery of standard test functions proposed by Donoho and Johnstone (1994), known as *Blocks, Bumps*, *Doppler* and *HeaviSine*. These functions had been selected because they portray significant spatial inhomogeneity and mimic functions arising in signal processing tasks, including imaging and NMR spectroscopy. An additional test function, *Mishmash*, is defined as:

$$Mishmash = C - (Blocks + Bumps + Doppler + Heavi\sin e).$$

Here C is the constant sum of all five functions. In our case, *Mishmash* reflects the need that the total mass remains constant over time.

For simulation purposes, the test functions are sampled at 2048 equally spaced time points within the interval [0,1]. Choosing the Coiflet 1 wavelet function, CIWS applied to this set converges after only two iterations for a smoothness parameter $\varepsilon = 0.1$, if the signal-to-noise ratio (SNR) is set to either 5 or 15 dB. Figs. 4–6 show four of these functions without noise, with added white (Gaussian) noise of SNR = 15 dB, and the output of the CIWS algorithm (Fig. 3).

The estimated functions in Fig. 5 together with the smoothed *Mishmash* function preserve the balance of energy C, and at the same time have noise mostly eliminated. Locally, the presence of some remaining noise is visible and it could be eliminated at the expense of losing resolution at high frequency features (as in *Doppler*).

4.2. Data analysis

CIWS was tested with different time series data of metabolites in the glycolytic pathway of the bacterium *L. lactis*, which had been measured under anaerobic conditions following a glucose pulse of 40 mM. The data were obtained with *in vivo* nuclear magnetic resonance (^{13}C NMR) techniques. The algorithm converged to mass-conserving time courses, whose smoothness was adjustable to our specifications.

Extensive simulations with different datasets demonstrated that CIWS is very time efficient and converges in quite a small number of iterations (between 2 and 60) if a reasonable value for the mass conservation error constant is chosen, such as $\varepsilon \sim 0.01$. As an illustration, Fig. 6 depicts the CIWS results for the glycolytic time series data from *L. lactis*. The wavelet function used in this case is *Coiflet 1*. It converges in 28 iterations for $\varepsilon = 0.1$. Decreasing ε to 0.001 does not result in a significant visual difference but increases the iterations until CIWS converges to 86. Note that CIWS retains

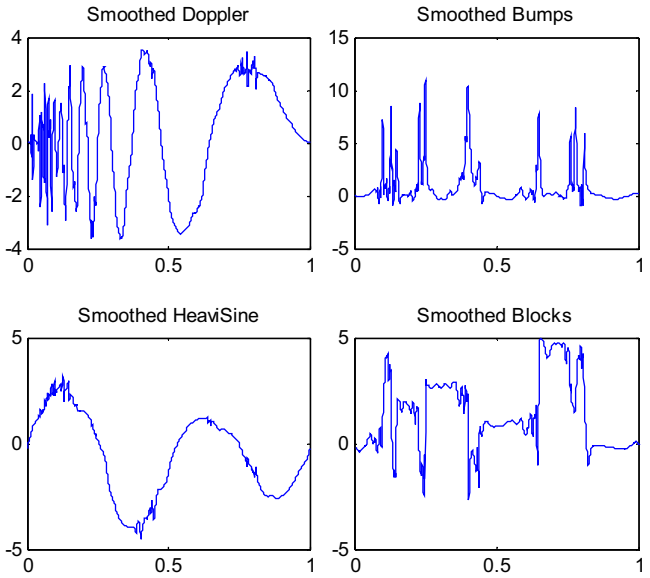

Fig. 5. Estimated test functions as output of the CIWS algorithm (compare to Fig. 4).

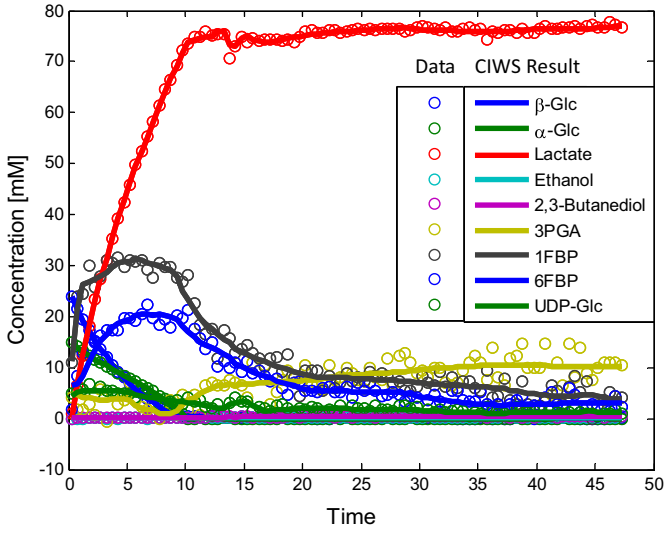

Fig. 6. Results of CIWS applied to one sample set of time series data characterizing the dynamics of the glycolytic pathway in *Lactococcus lactis* under anaerobic conditions and with an input glucose pulse of 40 Mm. Circles represent the measured time series data, while CIWS results are represented with lines of the corresponding color.

the observed drop in lactate around time 14, while smoothing the ascent in this metabolite $t \in (0, 10)$, as well as its more or less stationary phase $t \in (16, 47)$.

5. Discussion

In this article, we propose a novel constrained iterative wavelet-based smoothing method that permits noise reduction and smoothing, while assuring a mass conservation constraint. Unlike curve fitting, where the main emphasis is on matching the data as closely as possible, smoothing contains the somewhat vague concept of time course values that are expected to change relatively slowly from one time point to the next. This concept renders the performance criterion for smoothing techniques application dependent. For example, what might be considered as a good smoothing technique in one application where removing of outliers is of interest, might not be considered satisfactory in another application where one might be interested in local artifacts and features of a signal. This data-adaptive interplay between smoothness and local representation renders wavelets suitable tools for flexible smoothing tasks. For the dataset of our case study, the primary purpose is to smooth the data and prepare them for slope estimation while avoiding spurious spikes in the dynamics of the data.

The main novelty of the proposed CIWS smoothing technique is the property of mass conservation, which is practically important since it directly affects the consistency between experimental data and the fitted model.

A secondary advantage with respect to computation and implementation is the fact that CIWS does not require operator interaction or supervision and thus can be automated. Finally, the fast convergence properties of the wavelet transform techniques render this algorithm computationally very efficient. Mallat's cascade algorithm can be used in the implementation of each iteration, which results in computational complexity of order $O(N)$. This fact is particularly beneficial when have larger datasets with thousands of data points are to be smoothed. Such larger datasets, which are increasingly more prevalent due to modern high-throughput experimental techniques, will render the proposed CIWS smoothing method even more effective, because short signals decompose in only a few multiresolution spaces and render the thresholding ineffective due to the availability of limited number of scales.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.compchemeng.2014.07.019.

References

Breiman L. Better subset regression using the nonnegative garrote. Technometrics 1995;37:373–84.
Buckheit J, Donoho D. WaveLab and reproducible research. In: Antoniadis A, Oppenheim G, editors. Wavelets and statistics, vol. 103. New York: Springer; 1995. p. 55–81.
Carvalho AL, Turner DL, Fonseca LL, Solopova A, Catarino T, Kuipers OP, et al. Metabolic and transcriptional analysis of acid stress in *Lactococcus lactis*, with a focus on the kinetics of lactic acid pools. PLOS ONE 2013;8:14.
Chou I-C, Voit EO. Recent developments in parameter estimation and structure identification of biochemical and genomic systems. Math Biosci 2009;219:57–83.
Chou I-C, Voit EO. Estimation of dynamic flux profiles from metabolic time series data. BMC Syst Biol 2012;6:84.
Coifman RR, Wickerhauser MV. Entropy-based algorithms for best basis selection. IEEE Trans Inf Theory 1992;38:713–8.
Daubechies I. Ten lectures on wavelets. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1992.
Donoho DL, Johnstone IM. Ideal spatial adaptation by wavelet shrinkage. Biometrika 1994;81:425–55.
Eilers PHC. A perfect smoother. Anal Chem 2003;75:3631–6.
Gavalas GR. Nonlinear differential equations of chemically reacting systems. Berlin: Springer-Verlag; 1968.
Goel G, Chou I-C, Voit EO. System estimation from metabolic time-series data. Bioinformatics 2008;24:2505–11.
Heinrich R, Schuster S. The regulation of cellular systems. New York: Chapman & Hall; 1996.
Iwata M, Shiraishi F, Voit EO. Course but efficient identification of metabolic pathway systems. Int J Syst Biol 2013;4:57–72.
Mallat SG. A theory for multiresolution signal decomposition – the wavelet representation. IEEE Trans Pattern Anal Mach Intell 1989;11:674–93.
Palsson BØ. Systems biology: properties of reconstructed networks. New York: Cambridge University Press; 2006.
Stephanopoulos G, Aristidou AA, Nielsen JH. Metabolic engineering: principles and methodologies. San Diego: Academic Press; 1998.
Strang G. Wavelets and dilation equations – a brief introduction. SIAM Rev 1989;31:614–27.
Varah JM. A spline least squares method for numerical parameter estimation in differential equations. SIAM J Sci Stat Comput 1982;3:28–46.
Vidakovic B. Statistical modeling by wavelets. New York: Wiley; 1999.
Vilela M, Borges CC, Vinga S, Vasconcelos AT, Santos H, Voit EO, et al. Automated smoother for the numerical decoupling of dynamics models. BMC Bioinform 2007;8:305.
Voit EO. Characterizability of metabolic pathway systems from time series data. Math Biosci 2013;246:315–25.
Voit EO, Almeida J. Decoupling dynamical systems for pathway identification from metabolic profiles. Bioinformatics 2004;20:1670–81.
Voit EO, Goel G, Chou I-C, da Fonseca L. Estimation of metabolic pathway systems from different data sources. IET Syst Biol 2009;3:513–22.
Voit EO, Savageau MA. Power-law approach to modeling biological systems: III. Methods of analysis. J Ferment Technol 1982a;60:223–41.
Voit EO, Savageau MA. Power-law approach to modeling biological systems: II. Application to ethanol production. J Ferment Technol 1982b;60:229–32.
Walter GG, Shen X, Walter GG. Wavelets and other orthogonal systems. 2nd ed. Boca Raton: Chapman & Hall/CRC; 2001.
Whittaker E. On a new method of graduation. Edinburgh Mathematical Society; 1923. p. 63–75.