Journal of Complexity 28 (2012) 524–538

Contents lists available at SciVerse ScienceDirect

Journal of Complexity

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

Adaptive parameter choice for one-sided finite difference schemes and its application in diabetes technology

V. Naumova *, S.V. Pereverzyev, S. Sivananthan

Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergstraße 69, A-4040 Linz, Austria

ARTICLE INFO

Article history:
Received 30 September 2011
Accepted 7 June 2012
Available online 26 June 2012

Keywords:
Numerical differentiation
One-sided finite difference schemes
Adaptive parameter choice
Balancing principle
Quasi-optimality criterion

ABSTRACT

In this paper we discuss the problem of approximation of the first derivative of a function at the endpoint of its definition interval. This problem is motivated by diabetes therapy management, where it is important to provide estimations of the future blood glucose trend from current and past measurements. A natural way to approach the problem is to use one-sided finite difference schemes for numerical differentiation, but, following this way, one should be aware that the values of the function to be differentiated are noisy and available only at given fixed points. Then (as we argue in the paper) the number of used point values is the only parameter to be employed for regularization of the above mentioned ill-posed problem of numerical differentiation. In this paper we present and theoretically justify an adaptive procedure for choosing such a parameter. We also demonstrate some illustrative tests, as well as the results of numerical experiments with simulated clinical data.

1. Problem formulation

In this paper we consider the problem of approximation of a derivative $y'(B)$ at the boundary point of some interval $[b, B]$ under the condition that at the given points

$$B = t_N > t_{N-1} > \cdots > t_1 \geq b \tag{1}$$

only noisy values $y_\delta(t_j)$ of $y(t_j)$ are available such that

$$|y(t_j) - y_\delta(t_j)| \leq \delta. \tag{2}$$

* Corresponding author.
E-mail addresses: valeriya.naumova@oeaw.ac.at (V. Naumova), sergei.pereverzyev@oeaw.ac.at (S.V. Pereverzyev), sivananthan.sampath@oeaw.ac.at (S. Sivananthan).

0885-064X/$ – see front matter
doi:10.1016/j.jco.2012.06.001

It should be noted that the problem in such formulation arises in various practical applications, for example, in the diabetes therapy management. More about this application will be given in the forthcoming section.

At the same time, it is noteworthy to mention that the problem of approximation of $y'(B)$ does not seem to have been discussed so intensively as the problem of the reconstruction of the derivative at a given interior point. For the latter one there is a substantial amount of publications, see, for example [1,14,16,22,24], just to mention a few.

At this point, it should be noted that the problem of the numerical differentiation is known to be ill-posed [6]. Hence, in order to approximate a value of the derivative in a stable way, regularization methods should be applied.

In the context of approximation $y'(\tau)$ at an interior point $\tau \in (b, B)$ the role of such a regularization method is usually played by a finite difference formula of the form

$$y'(\tau) \approx S_{n,h} y_\delta(\tau) = \sum_{j=-n}^{n} \frac{a_j^n}{h} y_\delta(\tau + jh), \tag{3}$$

where the value h of the step-size is used as a regularization parameter. Within this scheme it is presupposed that a step-size h can be freely chosen in dependence on a noise level δ (see, for example [14,16,24]). Note that such a scheme is not relevant for our problem, since the points (1) are assumed to be given and cannot be chosen at will. Moreover, formula (3) cannot be used for approximation of the derivative value at the boundary point, which is our main interest in this paper.

With this interest in mind, we consider another tool for numerical differentiation, known as one-sided finite difference formula, see, for example [23]. Similar to [23] we assume that the given sequence (1) is formed by points of several equidistant grids with step-sizes $h_1 > h_2 > \cdots > h_\nu$ such that

$$\{t_k\}_{k=1}^{N} = \bigcup_{s=1}^{\nu} \{t_{j,s}\}_{j=0}^{N_s-1}, \tag{4}$$

where $t_{j,s} = B - jh_s,\ j = 0, 1, \ldots, N_s - 1,\ N_s = \left[\frac{B-b}{h_s}\right]$, and $[a]$ denotes the integer part of a. For instance, if we are given to the grid with the step-size h_s, then the grids with step-sizes $h_i = ih_s,\ i = 1, 2, \ldots,$ also can be used.

Then the one-sided finite difference formula used for the approximation of $y'(B)$ can be written in the form

$$y'(B) \approx S_{n,h_s} y_\delta = \sum_{j=0}^{n} \frac{a_j^n}{h_s} y_\delta(B - jh_s), \quad n \le N_s. \tag{5}$$

From the viewpoint of the Information-Based Complexity (IBC) theory [29] a family of formulae (5) with a fixed array of coefficients $\{a_j^n\}_{j=0,\ n=1}^{n,\ N_s}$ can be seen as a sequence of algorithms of increasing computational cost, which is proportional either to the amount $(n + 1)$ of used standard information $\{y(t_{j,s})\}_{j=0}^{n}$ or, which is almost the same, to the order n of a formula. But these algorithms operate with noisy information $\{y_\delta(t_{j,s})\}_{j=0}^{n},\ n = 1, 2, \ldots, N_s$, and aim to approximate the value of the functional $Sy := y'(B)$, which is unbounded on the space $C[0, 1]$ of continuous functions. Due to this unboundedness, the considered approximation problem is ill-posed. In contrast to the well-posed problems with noisy information, which were mainly studied until now in the literature on IBC (see, e.g., [21]), in the considered case an uncontrolled increase of the order n of the (5) leads not only to unnecessary computational costs, but also to an amplification of the noise propagation error (see Example 1 below). Therefore, the goal is to use a value n of the complexity measure as a regularization parameter, and to find among a fixed sequence of algorithms (5) such one for which the noise propagation error is balanced with the error of approximation of Sy at given y from the corresponding noise-free information. The latter error usually decreases with increasing amount $(n+1)$ of information. However, this is not always the case for the considered problem (see Remark 1

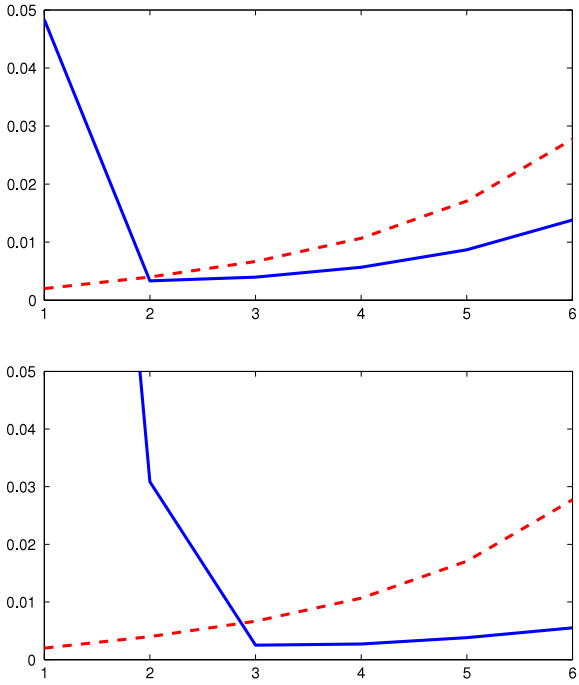

Fig. 1. The functions $\phi(n)$ (solid line) and $\psi(n)\frac{\delta}{h_s}$ (dotted line) for $h_s = 10^{-1}$, $\delta = 10^{-4}$ and $y \in C^3$ (upper panel) and $y \in C^4$ (lower panel).

and the illustrations of Assumption 3 by Figs. 1 and 2) that somehow differs our analysis from the known ones in the literature.

Thus, if in (5) the step-size h_s cannot be freely chosen, then the performance of the high order finite difference formula can be rather poor if the order n is not properly related with δ and h_s. This issue is important in some applications, where a poor performance of numerical differentiation can lead to dangerous implications. For example, from Tables 10 and 11 presented in the last section it can be seen that the use of the numerical differentiation formula with the finest step-size and a priori fixed (highest) order dramatically increases the percentage of erroneous predictions of the blood glucose (BG) concentration of diabetic patients compared to the formulae with proper parameter choice.

In this study, we present a new approach for selection of optimal parameter n, which is based on the balancing principle [14].

The article is organized in the following way. In Section 2, we present an adaptive parameter choice rule in (5) based on the balancing principle. We also provide and prove error estimates for this rule under some assumptions. In the same section, we present some illustrative tests supporting theoretical results. In Section 3, we conclude our study with a discussion and illustration of how the proposed approach can be used in the diabetes therapy management, in particular, for predicting blood glucose values from current and past BG measurements.

2. The regularized finite difference method

We return now to the discussion of the problem of estimating the value of the derivative at the boundary point B. Let $y : [b, B] \to \mathbb{R}$ be an r-times continuously differentiable function, $r \geq 2$. The difference between $y'(B)$ and the estimate (5) can be bounded as follows

$$|y'(B) - S_{n,h_s} y_\delta| \leq |y'(B) - S_{n,h_s} y| + |S_{n,h_s} y - S_{n,h_s} y_\delta|, \tag{6}$$

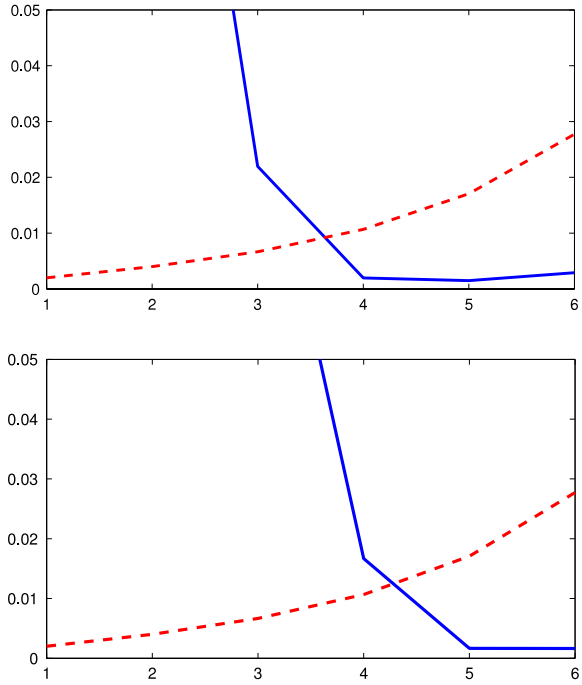

Fig. 2. The functions $\phi(n)$ (solid line) and $\psi(n)\frac{\delta}{h_s}$ (dotted line) for $h_s = 10^{-1}$, $\delta = 10^{-4}$ and $y \in C^5$ (upper panel) and $y \in C^6$ (lower panel).

where the first term on the right-hand side is the approximation error, whereas the second term is the noise propagation error. For the latter one we have a bound

$$|S_{n,h_s}y - S_{n,h_s}y_\delta| \le \psi(n)\frac{\delta}{h_s}; \tag{7}$$

here $\psi(n) = \sum_{j=0}^{n} |a_j^n|$, and for $n = 0$ we define $\psi(0) = 0$. Note that under the assumption (2) the bound (7) cannot be improved in its terms.

At this point it is worth observing that any formula of the form (5) can be seen as a linear bounded functional S_{n,h_s} defined on the space $C[0, 1]$ such that $\|S_{n,h_s}\| = \frac{\psi(n)}{h_s}$, and our aim is to use a family $\{S_{n,h_s}\}_{n=1}^{N_s}$ of such functionals for approximating an unbounded functional $Sy = y'(B)$. Thus, to guarantee a good approximation, the functionals from a chosen family $\{S_{n,h_s}\}$ should not be uniformly bounded. Even more, to the best of our knowledge, in the literature such formulae (5) are mainly considered for which $\|S_{n,h_s}\|$ increases with n.

Therefore, dealing with a fixed h_s, we make the following assumption on the function $\psi(n)$.

Assumption 1. The function $\psi(n)$ is an increasing function of n.

The following example illustrates this assumption.

Example 1. Following [9] we consider one-sided formulae of the form (5) with coefficients presented in Table 1. Note that if $y(t)$ is an algebraic polynomial of degree n on $[b, B]$, then for the formula S_{n,h_s} with coefficients from Table 1 one has $y'(B) = S_{n,h_s}y$.

In the sequel, we will use these formulae for numerical illustrations, since in practical applications the formulae of higher order are rarely used. The reason for this can be seen in the numerical test with the function (21) presented below. From Table 4 it follows that for some step-sizes (e.g. $h = 0.056$, $h = 0.039$) the use of higher order formulae even for analytic functions may lead

Table 1
Coefficients of the one-sided formulae of the form (5).

n	a_0^n	a_1^n	a_2^n	a_3^n	a_4^n	a_5^n	a_6^n
1	1	−1					
2	$\frac{3}{2}$	−2	$\frac{1}{2}$				
3	$\frac{11}{6}$	−3	$\frac{3}{2}$	$-\frac{1}{3}$			
4	$\frac{25}{12}$	−4	3	$-\frac{4}{3}$	$\frac{1}{4}$		
5	$\frac{137}{60}$	−5	5	$-\frac{10}{3}$	$\frac{5}{4}$	$-\frac{1}{5}$	
6	$\frac{49}{20}$	−6	$\frac{15}{2}$	$-\frac{20}{3}$	$\frac{15}{4}$	$-\frac{6}{5}$	$\frac{1}{6}$

Table 2
The values of the function $\psi(n)$ for corresponding n.

n	1	2	3	4	5	6
$\psi(n)$	2	4	6.67	10.67	17.07	27.73

to the reduction of the order of accuracy. Similar effect can be observed for functions of a finite smoothness (see, e.g. Table 5).

For the formulae with the coefficients from Table 1 the values of the function $\psi(n)$ are given in Table 2. It is clear that for the one-sided formulae with the coefficients from Table 1 Assumption 1 is satisfied.

Suppose we are given a finite set $\mathcal{N} = \{n_i\}_{i=1}^M$ of possible orders of formulae (5). Note that for the formulae from Example 1 we have $\mathcal{N} = \{1, 2, \ldots, 6\}$.

Then, for a fixed step-size h_s the corresponding set of approximate values of the derivative $y'(B)$ is given as $\{S_{n_i,h_s}y_\delta\}_{i=1}^M$.

The topic we are going to discuss now is how to choose $n \in \mathcal{N}$ to guarantee a good approximation of the derivative $y'(B)$ from the set $\{S_{n_i,h_s}y_\delta\}_{i=1}^M$. Here we propose a selection strategy that is only by a constant factor worse than the best possible error bound.

Note that the bound (7) for the noise propagation error does not depend on the function to be differentiated, while the first term in the bound (6) depends on the smoothness of this function, which is usually unknown. Therefore, one needs a strategy that allows an adaptation of the order $n \in \mathcal{N}$ to the unknown smoothness of y. We will formulate such a strategy under the following additional assumptions.

Assumption 2.

$$\psi(1)\frac{\delta}{h_s} < |y'(B) - S_{1,h_s}y|. \tag{8}$$

Note that if the condition (8) is not satisfied, there is no sense to choose $n > 1$, because the approximation error is dominated by the noise propagation error for any $n \in \mathcal{N}$.

Assumption 3. For given $y \in C^r$, $r \geq 2$ and $h_s > 0$ we assume that there is a continuous function $\phi(u) = \phi(u; h_s, y)$, $u \in [1, n_M]$, such that

(i) $|y'(B) - S_{n,h_s}y| \leq \phi(n; h_s, y)$, $n \in \mathcal{N}$;
(ii) if $n_{\min}$ is the smallest number such that

$$n_{\min} = \arg\min\{\phi(n), n \in \mathcal{N}\},$$

then the function $\phi(n)$ is non-increasing on $[1, n_{\min})$ and for any $n \in [n_{\min}, n_M]$

$$\phi(n) \leq \psi(n)\frac{\delta}{h_s}. \tag{9}$$

Any function satisfying Assumption 3 is called admissible. We denote by $\Phi(y, h_s)$ the set of all admissible functions.

If Assumption 3(ii) does not hold, then the noise level is low and negligible. For example, if (9) is not satisfied with $n = n_{\min}$, then one can ignore the noise propagation and $n = n_{\min}$ is the best choice for noisy data, as well as for noise-free.

The next example illustrates these assumptions.

Example 1 (*continuation*)**.** Consider the problem of estimating the first derivative of a function $y \in C^r[0, 1]$ at the boundary point $t = 1$. Similar to [18] to eliminate the influence of the other boundary point $t = 0$ we assume that $y \in C_0^r = \{y : y \in C^r[0, 1],\ y^{(i)}(0) = 0,\ i = 0, 1, \ldots, r-1\}$.

Then y can be represented by Taylor's theorem as follows

$$y(t) = \int_0^1 \frac{(t-\tau)_+^{r-1}}{(r-1)!} y^{(r)}(t)d\tau,$$

where $(a)_+ = \max\{a, 0\}$, and for a finite difference scheme (5) one can derive the bound

$$|y'(1) - S_{n,h_s}y| \le \|y^{(r)}\|_C \int_0^1 \left| \frac{(1-\tau)_+^{r-2}}{(r-2)!} - \sum_{j=0}^{n} \frac{a_j^n}{h_s} \frac{(1-jh_s-\tau)_+^{r-1}}{(r-1)!} \right| d\tau; \tag{10}$$

here $\|\cdot\|_C$ denotes the maximum norm in $C[0, 1]$.

In view of (10) an admissible function $\phi(n; h_s, y)$ can be taken of the form

$$\phi(n; h_s, y) = \|y^{(r)}\|_C h_s^{r-1} V_r(n, h_s), \tag{11}$$

where

$$V_r(n, h_s) = h_s^{1-r} \int_0^1 \left| \frac{(1-\tau)_+^{r-2}}{(r-2)!} - \sum_{j=0}^{n} \frac{a_j^n}{h_s} \frac{(1-jh_s-\tau)_+^{r-1}}{(r-1)!} \right| d\tau. \tag{12}$$

Using (12) one can easily calculate (symbolically and numerically) the values of $V_r(n, h_s)$. The graphs of admissible functions (11) for $h_s = 10^{-1}$ and $\|y^{(r)}\|_C = 10^{r-3}$, $r = 3, 4, 5, 6$, are presented in Figs. 1 and 2. On these figures, we also display the graphs of the function $\psi(n)\frac{\delta}{h_s}$ for $\delta = 10^{-4}$. Using these figures, one can conclude that for the considered noise level Assumption 3 is satisfied for any function $y \in C_0^r$ such that $\|y^{(r)}\|_C \le 10^{r-3}$, $r = 3, 4, 5, 6$. Note that in order to illustrate the behaviour of the functions in a more transparent way we keep the same scale for all graphs.

Remark 1. In [14,16] the authors consider the modulus of continuity of a continuous function y' as an admissible function. Note also that Assumption 3 is more general than the assumption that an admissible function is non-increasing, which has been used in [14], and it makes a difference in the analysis compared to [14,16].

From (7) and Assumption 3 the difference between $y'(B)$ and (5) can be bounded as follows

$$|y'(B) - S_{n,h_s}y_\delta| \le \phi(n) + \psi(n)\frac{\delta}{h_s}. \tag{13}$$

In view of (13) the quantity

$$e(y, \mathcal{N}, S_{n,h_s}) = \min_{n\in\mathcal{N}} \inf_{\phi\in\Phi(y,h_s)} \left\{ \phi(n) + \psi(n)\frac{\delta}{h_s} \right\} \tag{14}$$

is the best possible error bound that can be guaranteed for the approximation $y'(B)$ within the framework of the scheme S_{n,h_s} under Assumptions 1–3 and (2).

We now present a principle for the adaptive choice of $n_+ \in \mathcal{N}$ that allows us to reach the best possible error bound up to multiplier 6ρ, where

$$\rho = \rho(\mathcal{N}) = \max_{i=1,\ldots,M-1} \frac{\psi(n_{i+1})}{\psi(n_i)}.$$

In particular, from Table 2 in Example 1 one can see that $\rho(\mathcal{N}) = 2$ for $\mathcal{N} = \{1, \ldots, 6\}$.

As we will see such n_+ can be chosen knowing only noisy data values $y_\delta(t_j)$ for a fixed step-size h_s and without any a priori information concerning smoothness of $y \in C^r[b, B]$. To achieve the optimal order under Assumptions 1–3, we employ the balancing principle in the form presented in [14].

Consider the set

$$\mathcal{N}(S_{n,h_s}) = \left\{ n_i \in \mathcal{N} : |S_{n_i,h_s} y_\delta - S_{n_j,h_s} y_\delta| \le C\psi(n_j)\frac{\delta}{h_s},\ j = i+1, i+2, \dots, M \right\}, \tag{15}$$

where C is a tuning parameter. In particular, we use $C = 4$ for theoretical proofs and we adjust the value of C for the numerical experiments.

The order n_+ we are interested in is defined as

$$n_+ = \min\{\mathcal{N}(S_{n,h_s})\}. \tag{16}$$

We stress that an admissible function $\phi(n) \in \Phi(y, h_s)$, as well as any other information about the smoothness of the function y is not involved in the process of choosing n_+. Now we will formulate the main result of this section.

Theorem 1. *Let $y \in C^2[b, B]$. Then under Assumptions 1–3 we have*

$$|y'(B) - S_{n_+,h_s} y_\delta| \le 6\rho e(y, \mathcal{N}, S_{n,h_s}).$$

The proof of this theorem is a modification of the one provided in [14,15].

Proof. Let $\phi \in \Phi(y, h_s)$ be any admissible function and let us consider the numbers

$$n_* = n_*(\phi) = \min\left\{ n \in \mathcal{N} : \phi(n) \le \psi(n)\frac{\delta}{h_s} \right\},$$

$$n_{**} = n_{**}(\phi) = \arg\min\left\{ \phi(n_j) + \psi(n_j)\frac{\delta}{h_s},\ n_j \in \mathcal{N} \right\}.$$

Observe that from Assumption 3 it follows that $n_* \le n_{\min}$. Then

$$\psi(n_*)\frac{\delta}{h_s} \le \rho\left(\phi(n_{**}) + \psi(n_{**})\frac{\delta}{h_s}\right), \tag{17}$$

because either $n_{**} \le n_* - 1 < n_*$, so that by definition we have $\psi(n_* - 1)\frac{\delta}{h_s} < \phi(n_* - 1)$ and

$$\begin{aligned}\psi(n_*)\frac{\delta}{h_s} &= \frac{\psi(n_*)}{\psi(n_* - 1)}\psi(n_* - 1)\frac{\delta}{h_s} < \rho\phi(n_* - 1) \le \rho\phi(n_{**}) \\ &< \rho\left(\phi(n_{**}) + \psi(n_{**})\frac{\delta}{h_s}\right), \quad \rho > 1,\end{aligned}$$

or $n_* < n_{**}$, in which case

$$\psi(n_*)\frac{\delta}{h_s} < \psi(n_{**})\frac{\delta}{h_s} < \rho\left(\phi(n_{**}) + \psi(n_{**})\frac{\delta}{h_s}\right).$$

Moreover, note that for $n_j > n_*$, $n_j \in \mathcal{N}$ it follows that

$$\phi(n_j) < \psi(n_j)\frac{\delta}{h_s}. \tag{18}$$

In the case of $n_j > n_{\min}$ the inequality (18) follows from Assumption 3. If now for $n_j \le n_{\min}$ we assume the inequality $\phi(n_j) \ge \psi(n_j)\frac{\delta}{h_s}$, which is opposite to (18), then it would lead to

$$\phi(n_*) \ge \phi(n_j) \ge \psi(n_j)\frac{\delta}{h_s} > \psi(n_*)\frac{\delta}{h_s},$$

that is in contradiction with the definition of n_*.

Now we can also conclude that $n_* \geq n_+$. Indeed, for any $n_j > n_*$, $n_j \in \mathcal{N}$, we have

$$\begin{aligned}
|S_{n_*,h_s} y_\delta - S_{n_j,h_s} y_\delta| &\leq |y'(B) - S_{n_*,h_s} y_\delta| + |y'(B) - S_{n_j,h_s} y_\delta| \\
&\leq \phi(n_*) + \psi(n_*)\frac{\delta}{h_s} + \phi(n_j) + \psi(n_j)\frac{\delta}{h_s} \\
&\leq 2\psi(n_*)\frac{\delta}{h_s} + \phi(n_j) + \psi(n_j)\frac{\delta}{h_s} \\
&< 2\frac{\delta}{h_s}\psi(n_*) + 2\frac{\delta}{h_s}\psi(n_j) < 4\frac{\delta}{h_s}\psi(n_j).
\end{aligned}$$

It means that $n_* \in \mathcal{N}(S_{n,h_s})$ and

$$n_* \geq n_+ = \min\{\mathcal{N}(S_{n,h_s})\}.$$

Then using (17) one can finally obtain

$$\begin{aligned}
|y'(B) - S_{n_+,h_s} y_\delta| &\leq |y'(B) - S_{n_*,h_s} y_\delta| + |S_{n_*,h_s} y_\delta - S_{n_+,h_s} y_\delta| \\
&\leq \phi(n_*) + \psi(n_*)\frac{\delta}{h_s} + 4\psi(n_*)\frac{\delta}{h_s} \leq 6\psi(n_*)\frac{\delta}{h_s} \\
&\leq 6\rho\left(\phi(n_{**}) + \psi(n_{**})\frac{\delta}{h_s}\right) \\
&\leq 6\rho \min_{n\in\mathcal{N}}\left\{\phi(n) + \psi(n)\frac{\delta}{h_s}\right\}. \qquad (19)
\end{aligned}$$

This estimation holds true for an arbitrary admissible function $\phi \in \Phi(y, h_s)$. Therefore, we conclude that

$$|y'(B) - S_{n_+,h_s} y_\delta(B)| \leq 6\rho \min_{n\in\mathcal{N}} \inf_{\phi\in\Phi(y,h_s)}\left\{\phi(n) + \psi(n)\frac{\delta}{h_s}\right\}.$$

The proof is complete. □

2.1. Numerical tests

To illustrate the theoretical results of the previous section, we consider the estimation of the first derivative of a function $y \in C^r[0, 1]$, $r \geq 2$ at the boundary point $t = 1$ by means of a one-sided finite difference formula (5). The balancing principle (15), (16) is used as the criterion for choosing the optimal order $n_+ \in \mathcal{N}$.

To perform numerical experiments we consider the function of a finite smoothness

$$y(t) = |t|^7 + |t - 0.25|^7 + |t - 0.5|^7 + |t - 0.75|^7 + |t - 0.85|^7 \in C^6[0, 1], \qquad (20)$$

as well as analytic functions

$$y(t) = \frac{t^3}{3} - \frac{t^2}{2} \in C^\infty[0, 1], \qquad (21)$$

$$y(t) = e^t \in C^\infty[0, 1]. \qquad (22)$$

The numerical values $y_\delta(1 - jh_s)$ for the functions used in the tests are generated in the form $y_\delta(1 - jh_s) = y(1 - jh_s) + \delta\xi_j$, where $\delta \sim 10^{-5}$ and ξ_j are random values sampled uniformly in the interval $[-1, 1]$.

For our numerical experiments, we consider 8 equidistant grids with step-sizes

$$h_s = 1.5^{-s}, \quad s = 4, 5, \ldots, 11, \qquad (23)$$

and the one-sided finite difference formulae (5) with coefficients from Example 1.

Table 3
The results of the application of the choice rule (16).

Function	Step-size h_s	0.167	0.125	0.083	0.056	0.039	0.026	0.017	0.012
(20)	Order n_+	5	5	5	5	5	5	5	5
	Order of accuracy	10^{-1}	10^{-2}	10^{-3}	10^{-3}	10^{-4}	10^{-3}	10^{-3}	10^{-3}
(21)	Order n_+	5	4	5	5	5	5	5	5
	Order of accuracy	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-3}	10^{-4}	10^{-3}	10^{-3}
(22)	Order n_+	5	5	5	5	5	5	5	5
	Order of accuracy	10^{-4}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-2}	10^{-3}	10^{-2}

Table 4
The orders of the accuracy of the numerical differentiation formulae for the analytic function (21) and step-sizes (23).

	Order n					
Step-size h_s	1	2	3	4	5	6
0.167	10^{-2}	10^{-2}	10^{-4}	10^{-4}	10^{-4}	10^{-4}
0.125	10^{-2}	10^{-2}	10^{-4}	10^{-4}	10^{-4}	10^{-4}
0.083	10^{-2}	10^{-3}	10^{-4}	10^{-4}	10^{-4}	10^{-4}
0.056	10^{-2}	10^{-3}	10^{-4}	10^{-4}	10^{-4}	10^{-3}
0.039	10^{-2}	10^{-4}	10^{-3}	10^{-3}	10^{-3}	10^{-3}
0.026	10^{-2}	10^{-3}	10^{-3}	10^{-3}	10^{-4}	10^{-3}
0.017	10^{-3}	10^{-4}	10^{-4}	10^{-4}	10^{-3}	10^{-3}
0.012	10^{-3}	10^{-4}	10^{-3}	10^{-3}	10^{-3}	10^{-3}

Table 5
The orders of the accuracy of the numerical differentiation formulae for the function of a finite smoothness (20) and step-sizes (23).

	Order n					
Step-size h_s	1	2	3	4	5	6
0.167	10^{0}	10^{0}	10^{-1}	10^{-1}	10^{-1}	10^{-3}
0.125	10^{0}	10^{-1}	10^{-1}	10^{-1}	10^{-2}	10^{-3}
0.083	10^{0}	10^{-1}	10^{-1}	10^{-2}	10^{-3}	10^{-3}
0.056	10^{0}	10^{-1}	10^{-2}	10^{-3}	10^{-3}	10^{-3}
0.039	10^{-1}	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-3}
0.026	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-3}	10^{-3}
0.017	10^{-1}	10^{-2}	10^{-4}	10^{-4}	10^{-3}	10^{-3}
0.012	10^{-1}	10^{-2}	10^{-3}	10^{-3}	10^{-3}	10^{-3}

At first, for the fixed step-size h_s we find the optimal order $n_+ \in \mathcal{N}$ by means of the balancing principle (15), (16). As we have already mentioned in applications the value of the tuning parameter in (15) can be adjusted. The adjustment can be done as follows: we simulate data using some function, say (21), and find a value of C that leads to a good performance of the principle (16) on simulated data. This value of C is used in numerical applications. As a result of such an adjustment procedure, we have found $C = 0.0021$.

The results of the application of the choice rule (16) for the functions (20)–(22) and step-sizes (23) are displayed in Table 3.

In view of the results [16], one may expect that higher order finite difference schemes allow an order of accuracy, which is the best possible for a given noise level δ. Using Table 3, we can conclude that for a fixed step-size it is not the case. For example, in the case of the function (20), which belongs to $C^r[0, 1]$, $r = 6$, the best guaranteed order of accuracy in terms of the noise level is known to be $\delta^{\frac{r-1}{r}}$. For $r = 6$ and $\delta \sim 10^{-5}$ it gives the order of 10^{-4}, and this order is really achieved by the formula S_{n,h_s} for $n = n_+ = 5$, $h_s = 0.039$. At the same time, direct calculations show that for the considered function the formula S_{6,h_s} gives an accuracy only of order 10^{-3} for all h_s from (23) (see Table 5).

Table 6
The approximation errors produced by $S_{n_{+(h_+)},h_+}y_\delta$.

Function	(20)	(21)	(22)
Order of accuracy	10^{-4}	10^{-4}	10^{-4}
Order n_+	5	5	5
Step-size h_+	0.039	0.167	0.167

2.1.1. Combination of the balancing principle and the quasi-optimality criterion

Table 3 also shows that the accuracy of the formulae S_{n_+,h_s} depends on h_s, as one could expect. Therefore, in the situation, when one has some choice in the selection of the value of h_s, as in (23), for example, it is natural to utilize it. For this purpose, one can use a heuristic approach known as the quasi-optimality criterion [28,11].

Once $n = n_+ = n_+(h_s)$ has been specified as a function of h_s (see, for example, Table 3), the quasi-optimality criterion can be implemented as follows. For all available h_s one needs to calculate approximations $S_{n_{+(h_s)},h_s}y_\delta$ given by (5). Then one needs to calculate the absolute difference

$$\sigma(s) = |S_{n_{+(h_s)},h_s}y_\delta - S_{n_{+(h_{s-1})},h_{s-1}}y_\delta|$$

and finds

$$h_+ = h_p : \quad p = \arg\min\{\sigma(s), s = 5, \ldots, 11\}. \tag{24}$$

The approximation errors produced by $S_{n_{+(h_+)},h_+}y_\delta$ for considered functions are presented in Table 6.

As it can be seen from Tables 3 and 6, in all three cases the proposed combination of the quasi-optimality criterion and the balancing principle really allows the best choice of the values of h_s.

Remark 2. Note that to avoid ill-conditioning of numerical differentiation one may use standard regularization methods. As it has been mentioned in [24], this approach typically involves writing the derivative as the solution to an integral Volterra equation and then reducing the integral equation to a family of well-posed problems that depend on a regularization parameter. Then several parameter choice strategies can be used to find an optimal value of this parameter (see, e.g. [14] and references therein). Unfortunately, as it follows from the paper [8] (see Example 5 there) even for some analytic functions the order of accuracy $O\left(\delta^{\frac{1}{3}}\right)$ cannot be improved in this way. Therefore, to guarantee the accuracy of better order $O\left(\delta^{\frac{r-1}{r}}\right)$ for sufficiently smooth functions one should use another approach. One of them has been described above. We are grateful to an anonymous referee who inspired us to make this remark.

3. One-sided finite difference schemes in blood glucose prediction

In this section, we discuss a possibility to use the proposed approach in the diabetes therapy management, in particular, for blood glucose prediction.

Management of the diabetes mainly focuses on keeping BG concentration as close to normal range as possible without causing dangerous events, namely hypo- and hyperglycaemia, when BG lower than 70 (mg/dL) or higher than 180 (mg/dL) respectively. This can be achieved by balancing the amount of insulin injections, meals and physical exercises. Keeping in mind that the onset of rapid-acting insulin occurs within 10–15 (min), and the onset of meal responses on glucose levels occurs approximately within 5–10 (min), it is of great importance to know the future BG level at least 15 (min) ahead of time.

Recent progress in the diabetes therapy management is related to the so-called Continuous Glucose Monitoring (CGM) systems. In short, these devices provide ongoing monitoring of glucose level on an automated basis during the day, namely BG estimations each 5 or 10 (min). For mathematical details, see [17].

Thus, the promising concept in the diabetes therapy management is the prediction of the future BG evolution using CGM data [26]. The importance of such prediction has been shown by several applications [3,19].

Mathematically the problem can be formulated as follows. Assume that at the time moment $t = t_N = B$ we are given n estimates $y_N, y_{N-1}, y_{N-2}, \ldots, y_{N-n+1}$ of a patient's BG concentration sampled correspondingly at the time moments $t_N > t_{N-1} > t_{N-2} > \cdots > t_{N-n+1}$ within the sampling horizon $SH = t_N - t_{N-n+1}$. The goal is to construct a predictor that uses these past measurements to predict BG concentration $y_j = y(t_j)$ for m subsequent future time moments $\{t_j\}_{j=N+1}^{N+m}$ within the prediction horizon $PH = t_{N+m} - t_N$ such that $t_N < t_{N+1} < t_{N+2} < \cdots < t_{N+m}$.

There are several prediction techniques, and a variety of glucose predictors has been recently proposed; see, for example, [26,7,20,25,27]. In this section we discuss the predictors based on the numerical differentiation. Note that such approach has been used in several BG predictors. In [13] it is even mentioned that practically all BG predictions are based on projection ahead of the current glucose trend (or derivative). For example, one of such predictors is described in the patent [10], and another one has been recently discussed, in particular, in [26].

In the context of the diabetes therapy management, such predictors estimate the rate of change of BG concentration at the prediction moment $t = t_N = B$ from current and past measurements. Keeping in mind that the above mentioned rate is nothing but the value of the derivative $y'(B)$ of the function $y(t)$ describing the BG evolution, it is clear that only one-sided numerical differentiation scheme (5) can be used in this approach.

In this section, we assess the performance of the BG predictor that is based on the adaptive parameter choice rule (16), (24) implemented on the set of one-sided formulae from Example 1. It is remarkable that this BG predictor outperforms the predictors based on any of the formulae from the above mentioned set.

To illustrate this we use data set of 100 virtual subjects which are obtained from Padova/University of Virginia simulator [12]. For each in silico patient the CGM readings have been simulated and sampled with a frequency of 1 (min) during 3 days. We perform our illustrative tests with data of the same 10 virtual subjects that have been considered in [26]. Data of the first three subjects (Virtual ID 1, 2, and 3) have been randomly chosen from 100 traces generated by the simulator. The other seven data traces (Virtual ID 17, 18, 24, 33, 34, 42, and 47) have been chosen because they contain dangerous events, i.e. hypo- and hyperglycaemic periods.

Following [10], in our numerical experiments we consider the predictors, which are based on the linear extrapolation. To be more specific, a future BG concentration at any time moment $t \in [t_N, t_{N+m}]$ can be estimated from the given past BG observations $\mathbf{z} = ((t_{N-n+1}, y_{N-n+1}), \ldots, (t_N, y_N))$ as follows

$$y(t) = y'(B) \cdot t + y_N - y'(B) \cdot t_N, \tag{25}$$

where $y'(B) = y'(t_N)$ is estimated by means of the one-sided finite difference schemes from Example 1.

To quantify the clinical accuracy of considered predictors, we use the Prediction Error-Grid Analysis (PRED-EGA) [26], which has been designed especially for the blood glucose predictors. This assessment methodology records reference glucose estimates paired with the estimates predicted for the same moments and looks at two essential aspects of the clinical accuracy: rate error grid analysis (R-EGA) and point error grid analyses (P-EGA). The first one is measured, in fact, by values of time derivatives calculated at the assessment moments for reference as well as for assessed glucose profile, whereas the P-EGA account for accuracy, point-by-point, in comparison with a reference value. As a result, the PRED-EGA distinguishes Accurate (Acc.), Benign (Benign) and Erroneous (Error) predictions in hypoglycaemic (0–70 mg/dL), euglycaemic (70–180 mg/dL) and hyperglycaemic (180–450 mg/dL) ranges. This stratification is of great importance because consequences caused by a prediction error in the hypoglycaemic range are very different from ones in the euglycaemic range [4].

First of all, we assess the performance of considered predictors (25) in the so-called ideal situation, when simulated data used as predictor inputs and references in the PRED-EGA assumed to be noise-free. Note that even though the simulator can produce virtual CGM readings every 1 (min), in our tests we try to mimic the data flow from some widely used CGM sensors, when a new reading appears every 5 (min) and updates the prediction input. It means that a new predicted glucose profile is produced

Table 7
The performance assessment matrix given by the PRED-EGA for the predictors (25), (5), based on any of the formulae from Example 1, operating on simulated noise-free data with $PH = 15$ (min).

Patient Vir. ID	BG $\leq$ 70 (mg/dL) (%)			BG 70–180 (mg/dL) (%)			BG $\geq$ 180 (mg/dL) (%)		
	Acc.	Benign	Error	Acc.	Benign	Error	Acc.	Benign	Error
1	–	–	–	99.88	0.12	–	100	–	–
2	–	–	–	99.88	0.12	–	–	–	–
3	–	–	–	99.88	0.12	–	–	–	–
17	99.69	0.31	–	100	–	–	–	–	–
18	99.71	0.29	–	100	–	–	–	–	–
24	100	–	–	99.81	0.19	–	–	–	–
33	99.71	0.29	–	100	–	–	100	–	–
34	99.60	0.40	–	100	–	–	100	–	–
42	100	–	–	99.83	0.17	–	100	–	–
47	99.73	0.27	–	100	–	–	100	–	–
Avg.	**99.78**	**0.22**	–	**99.93**	**0.07**	–	**100**	–	–

Table 8
The performance assessment matrix given by the PRED-EGA for the BG predictors (25), (5), based on the adaptive parameter choice rule $n = n_+(h_+)$ implemented on the set of formulae from Example 1, operating on simulated noise-free data with $PH = 15$ (min).

Patient Vir. ID	BG $\leq$ 70 (mg/dL) (%)			BG 70–180 (mg/dL) (%)			BG $\geq$ 180 (mg/dL) (%)		
	Acc.	Benign	Error	Acc.	Benign	Error	Acc.	Benign	Error
1	–	–	–	100	–	–	100	–	–
2	–	–	–	100	–	–	–	–	–
3	–	–	–	100	–	–	–	–	–
17	100	–	–	100	–	–	–	–	–
18	100	–	–	100	–	–	–	–	–
24	100	–	–	100	–	–	–	–	–
33	100	–	–	100	–	–	100	–	–
34	100	–	–	100	–	–	100	–	–
42	100	–	–	100	–	–	100	–	–
47	100	–	–	100	–	–	100	–	–
Avg.	**100**	–	–	**100**	–	–	**100**	–	–

every $\Delta t = 5$ (min). Moreover, for our numerical experiments we consider 3 equidistant grids with step-sizes $h_s = 5, 10, 15$ (min).

Table 8 demonstrates the performance assessment matrix given by the PRED-EGA for 15 min ahead glucose predictions (25), where $t = t_{N+m}$, $t_{N+m} - t_N = 15$ min, and $y'(B)$ is estimated by (5) with $\delta = 0$, $n = n(h_s)$ and h_s chosen in accordance with (16), (24) (in the considered case we formally put $\delta = 1$ (mg/dL) in the definition of the set (15)).

At the same time, Table 7 demonstrates the assessment given by the PRED-EGA for predictions (25) with the same $t = t_{N+m} = t_N + 15$ (min), but with $y'(B)$ estimated by any of the one-sided formulae (5) from Example 1, where $\delta = 0$, $h_s = 5$ (min).

The comparison of both tables allows a conclusion that in the ideal situation of noise-free data the BG predictor (25), (5) that is based on the adaptive parameter choice rule $n = n_+(h_+)$ implemented on the set of one-sided formulae from Example 1 outperforms the predictors (25), (5) based on any of the formulae from the same set.

Remark 3. Note that statistical filtering techniques can also be used for constructing finite difference formulae of numerical differentiation. For example, in [2] a filtering procedure which allows a construction of central difference formulae of the form (3) has been presented and analysed. For constructing one-sided finite difference formulae (5), which are the objects of the present paper, a statistical filtering approach has been used in [10], where $y'(B)$ has been approximated by the

Table 9
The performance assessment matrix given by the PRED-EGA for the predictors (25), (26), based on the Savitzky–Golay derivative filter, which operate on simulated noise-free data with $PH = 15$ (min).

Patient Vir. ID	BG ≤ 70 (mg/dL) (%)			BG 70–180 (mg/dL) (%)			BG ≥ 180 (mg/dL) (%)		
	Acc.	Benign	Error	Acc.	Benign	Error	Acc.	Benign	Error
1	–	–	–	99.88	0.12	–	100	–	–
2	–	–	–	99.88	0.12	–	–	–	–
3	–	–	–	99.88	0.12	–	–	–	–
17	99.69	0.31	–	100	–	–	–	–	–
18	99.71	0.29	–	100	–	–	–	–	–
24	100	–	–	99.81	0.19	–	–	–	–
33	99.71	0.29	–	100	–	–	100	–	–
34	99.60	0.40	–	94.98	4.35	0.67	57.14	42.86	–
42	100	–	–	98.18	1.82	–	100	–	–
47	99.73	0.27	–	96.46	3.54	–	100	–	–
Avg.	**99.78**	**0.22**	–	**98.91**	**1.03**	**0.07**	**91.43**	**8.57**	–

Table 10
The performance assessment matrix given by the PRED-EGA for the predictors (25), (5), based on the formula $S_{6,5}$, operating on simulated noisy data with $PH = 15$ (min).

Patient Vir. ID	BG ≤ 70 (mg/dL) (%)			BG 70–180 (mg/dL) (%)			BG ≥ 180 (mg/dL) (%)		
	Acc.	Benign	Error	Acc.	Benign	Error	Acc.	Benign	Error
1	–	–	–	52.81	44.44	2.75	47.06	47.06	5.88
2	–	–	–	45.36	51.47	3.17	–	–	–
3	–	–	–	48.48	50.12	1.4	–	–	–
17	62.89	13.52	23.58	48.88	50	1.12	–	–	–
18	54.57	12.98	32.45	54.35	43.91	1.74	–	–	–
24	53.23	14.15	32.62	62.71	36.91	0.38	–	–	–
33	78.17	6.19	15.63	57.28	36.61	6.1	50	12.5	37.5
34	60.96	9.16	29.88	61.87	33.45	4.68	42.86	42.86	14.29
42	65.16	10.66	24.18	59.17	33.39	7.44	28.57	71.43	–
47	67.73	9.6	22.67	52.71	44.17	3.12	–	100	–
Avg.	**63.25**	**10.89**	**25.86**	**54.36**	**42.45**	**3.19**	**33.7**	**54.77**	**11.53**

Savitzky–Golay derivative filter of order 7 for a sampling frequency of $h_s = 5$ (min). More precisely, the following formula of the form (5) has been suggested in [10] for estimating $y'(B) = y'(t_N)$ in (25):

$$y'(B) \approx \frac{3}{140}[y_\delta(B) - y_\delta(B - 6h_s)] + \frac{2}{140}[y_\delta(B - h_s) - y_\delta(B - 5h_s)] + \frac{1}{140}[y_\delta(B - 2h_s) - y_\delta(B - 4h_s)]. \tag{26}$$

In Table 9 we present the performance assessment matrix given by the PRED-EGA for 15 min ahead predictions (25), (26) with $\delta = 0$. By comparing Tables 7–9 one can conclude that, for given data, the implementation of filtering approach (26) suggested in [10] is not superior to the formulae from Example 1.

To assess the effect of the noise on the choice of the best order $n = n_+(h_s)$ and the step-size $h = h_+$ for the finite difference schemes from Example 1, we use a predictor input formed by adding white noise to the simulated readings.

The performance assessment matrices for BG predictions (25) with $PH = 15$ (min) have been calculated for the case when simulated CGM readings have been corrupted by random white noise with an SD of 6 (mg/dL) and then used as predictor inputs. We would like to stress that the assessment has been done with respect to the references given as simulated noise-free CGM readings.

The performance of the predictors (25), (5) with the adaptive parameter choice rule $n = n_+(h_+)$ is presented in Table 11 and compared with the performance of the predictors (25), (5) based on the

Table 11
The performance assessment matrix given by the PRED-EGA for the BG predictors (25), (5), based on the adaptive parameter choice rule $n = n_+(h_+)$ implemented on the set of formulae from Example 1, operating on simulated noisy data with $PH = 15$ (min).

Patient Vir. ID	BG $\leq$ 70 (mg/dL) (%)			BG 70–180 (mg/dL) (%)			BG $\geq$ 180 (mg/dL) (%)		
	Acc.	Benign	Error	Acc.	Benign	Error	Acc.	Benign	Error
1	–	–	–	94.17	5.34	0.49	100	–	–
2	–	–	–	92.63	7.37	–	–	–	–
3	–	–	–	94.17	5.83	–	–	–	–
17	95.9	–	4.1	94.85	5.15	–	–	–	–
18	91.12	–	8.88	97.22	2.78	–	–	–	–
24	92.62	–	7.38	97.48	2.52	–	–	–	–
33	97.93	–	2.07	91.51	5.66	2.83	100	–	–
34	90	–	10	96.4	2.74	0.86	85.71	14.29	–
42	96.31	–	3.69	94.92	3.22	1.86	71.43	28.57	–
47	97.06	0.53	2.41	94.64	5.15	0.21	100	–	–
Avg.	**94.42**	**0.08**	**5.5**	**94.8**	**4.58**	**0.62**	**91.43**	**8.57**	–

one-sided finite difference formula of the order $n = 6$, displayed in Table 10. We have chosen the formula with fixed order $n = 6$ for the comparison, because for $\Delta t = 5$ (min) the sampling horizon $SH = 30 = 6\Delta t$ (min) has been suggested in [10,30] as the optimal one for BG prediction.

Comparing Tables 7, 8, 10 and 11, we can conclude that the BG predictor (25), (5) with the adaptive parameter choice rule $n = n_+(h_+)$ implemented on the set of formulae from Example 1 outperforms all predictors (25), (5) based on any of the formulae from the given set.

Moreover, the numerical tests have shown that the constructed BG predictor (25), (5) is able to predict more accurately the occurrences of dangerous events, i.e. hypo- and hyperglycaemic periods. Potentially, one may consider the combination of the proposed adaptive BG predictor with a predictor based on more advanced techniques, such as mentioned in [26], for example.

Acknowledgments

This research has been performed in the course of the project "DIAdvisor" [5] funded by the European Commission within 7th Framework Programme. The authors gratefully acknowledge the support of the "DIAdvisor" — consortium.

References

[1] R.S. Anderssen, F.R. de Hoog, Finite difference method for the numerical differentiation of non-exact data, Computing 33 (1984) 259–267.
[2] B. Anderssen, F. de Hoog, M. Hegland, A stable finite difference ansatz for higher order differentiation of non-exact data, Bull. Aust. Math. Soc. 58 (1998) 223–232.
[3] B. Buckingham, H.P. Chase, E. Dassau, E. Cobry, P. Clinton, V. Gage, K. Caswell, J. Wilkinson, F. Cameron, H. Lee, B.W. Bequette, F.J. Doyle III, Prevention of nocturnal hypoglycemia using predictive alarm algorithms and insulin pump suspension, Diabetes Care 33 (2010) 1013–1018.
[4] W.L. Clarke, S. Anderson, L. Farhy, M. Breton, L. Gonder-Frederick, D. Cox, B. Kovatchev, Evaluating the clinical accuracy of two continuous glucose sensors using Continuous glucose–error grid analysis, Diabetes Care 28 (2005) 2412–2417.
[5] Diadvisor: personal glucose predictive diabetes advisor, 2008.
[6] H. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, in: Mathematics and Its Applications, vol. 375, Kluwer Academic Publishers, Dordrecht, Boston, London, 1996.
[7] M. Eren, A. Cinar, L. Quinn, D. Smith, Estimation of future glucose concentrations with subject-specific recursive linear models, Diabetes Technol. Ther. 11 (2009) 243–253.
[8] J. Flemming, B. Hofmann, P. Mathé, Sharp converse results for the regularization error using distance functions, Inverse Problems 27 (2011) 025006. 18 pp.
[9] B. Fornberg, Generation of finite difference formulas on arbitrarily spaced grids, Math. Comp. 51 (1988) 699–706.
[10] A. Hayes, J.J. Mastrototaro, S.B. Moberg, J.C. Mueller Jr., H.B. Clark, M.C.V. Tolle, G.L. Williams, B. Wu, G.M. Steil, Algorithm sensor augmented bolus estimator for semi-closed loop infusion system, US Patent Application Publication, US2009/0234213 (2009).
[11] S. Kindermann, A. Neubauer, On the convergence of the quasi-optimality criterion for (iterated) Tikhonov regularization, Inverse Probl. Imaging 2 (2008) 291–299.

[12] B. Kovatchev, M. Breton, C.D. Man, C. Cobelli, In silico preclinical trials: a proof of concept in closed-loop control of type 1 diabetes, J. Diabetes Sci. Technol. 3 (2009) 44–55.
[13] B. Kovatchev, W. Clarke, Peculiarities of the continuous glucose monitoring data stream and their impact on developing closed-loop control technology, J. Diabetes Sci. Technol. 2 (2008) 158–163.
[14] S. Lu, S.V. Pereverzev, Numerical differentiation from a viewpoint of regularization theory, Math. Comp. 75 (2006) 1853–1870.
[15] P. Mathé, The Lepskii principle revisited, Inverse Problems 22 (2006) 030011. 6 pp.
[16] P. Mathé, S.V. Pereverzev, The use of higher order finite difference schemes is not dangerous, J. Complexity 25 (2009) 3–10.
[17] V. Naumova, S.V. Pereverzyev, S. Sampath, Extrapolation in variable RKHSs with application to the blood glucose reading, Inverse Problems 27 (2011) 075010. 13 pp.
[18] S.M. Nikolskiy, Quadrature formulae, NAUKA, Moscow, 1974.
[19] C. Palerm, B.W. Bequette, Hypoglycemia detection and prediction using continuous glucose monitoring a study on hypoglycemic clamp data, J. Diabetes Sci. Technol. 1 (2007) 624–629.
[20] S. Pappada, B. Cameron, P. Rosman, Development of neural network for prediction of glucose concentration in type 1 diabetes patients, J. Diabetes Sci. Technol. 2 (2008) 792-–801.
[21] L. Plaskota, Noisy Information and Computational Complexity, Cambridge University Press, Cambridge, 1996.
[22] R. Qu, A new approach to numerical differentiation and integration, Mathl. Comput. Modelling 24 (1996) 55–68.
[23] K. Rahul, S.N. Bhattacharyya, One-sided finite-difference approximations suitable for use with Richardson extrapolation, J. Comput. Phys. 219 (2006) 13–20.
[24] A. Ramm, A. Smirnova, On stable numerical differentiation, Math. Comp. 70 (2001) 1131–1153.
[25] J. Reifman, S. Rajaraman, A. Gribok, W.K. Ward, Predictive monitoring for improved management of glucose levels, J. Diabetes Sci. Technol. 1 (2007) 478–486.
[26] S. Sivananthan, V. Naumova, C.D. Man, A. Facchinetti, E. Renard, C. Cobelli, S. Pereverzyev, Assessment of blood glucose predictors: the prediction-error grid analysis, Diabetes Technol. Ther. 13 (2011) 787–796.
[27] G. Sparacino, F. Zanderigo, S. Corazza, A. Maran, Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series, IEEE Trans. Biomed. Eng. 54 (2007) 931–937.
[28] A.N. Tikhonov, V.B. Glasko, Use of the Regularization Methods in Non-linear Problems, in: USSR Comput. Math. Phys., vol. 5, 1965.
[29] J.F. Traub, G.W. Wasilkowski, H. Wozniakowski, Information-based Complexity, Academic Press, Boston, San Diego, and New York, 1988.
[30] F. Zanderigo, G. Sparacino, B. Kovatchev, C. Cobelli, Glucose prediction algorithms from continuous monitoring data: assessment of accuracy via continuous glucose error-grid analysis, Diabetes Technol. Ther. 1 (2007) 645–651.