Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

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

V. Naumova, S. Pereverzyev, S. Sampath

RICAM-Report 2011-18

Adaptive Parameter Choice for One-Sided Finite Difference Schemes and its Application in Diabetes Technology

Valeriya Naumova, Sergei V. Pereverzyev, Sivananthan Sampath

Abstract

In this paper we discuss the problem of an adaptive parameter choice in one-sided finite difference schemes for the numerical differentiation in case when noisy values of the function to be differentiated are available only at the given points. 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. Here we show, how the proposed approach can be used for this purpose and 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

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

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 given interior point. For

the latter one there is a substantial amount of publications, see, for example [1, 2, 3, 4, 5], 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 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 [2, 3, 5]). 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 [7]. Similar to [7] 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 $t_{j,s} = B - jh_s, \quad j = 0, 1, \cdots, N_s - 1, \quad N_s = [\frac{B-b}{h_s}]$, and $[a]$ denotes the integer part of a.

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

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

Note that the formula (5) can be seen as a two-parameter regularization method, since we still have freedom in choosing n and h_s. At this point, it is worth to mention that (3) could also be considered as a two-parameter

regularization method. But as it has been pointed out in [3] the use of the high order finite difference schemes allows for optimal order reconstruction of the derivative (at any given interior point of the domain) provided that the choice of the step-size h is made properly. This is because the order n has a role similar to the qualification of regularization [4], and hence higher order finite difference schemes may be used for low order smoothness. Therefore, in the approach [3] the choice of n for numerical differentiation is only constrained by the computational limitations. On the other hand, if the step-size h can not be freely chosen then the performance of high order finite difference scheme can be rather poor, if the order n is not properly related with h. In this study we present a new approach for selection of optimal parameter n, which is based on the balancing principle [2].

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 (BG) values from current and past BG measurements.

2. 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

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

where the first term in 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

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

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.

In the sequel, without loss of generality, 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 [8] 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 the 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.

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

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

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

Table 2: The values of the functions $\psi(n)$ for corresponding n

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 the Example 1 we have $\mathcal{N} = \{1, 2, \ldots, 6\}$.

Then, for a fixed step-size h_s the corresponding set of the 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 function $\phi(n) = \phi(n, h_s, y)$ such that*

(i) $|y'(B) - S_{n,h_s} y| \leq \phi(n, h_s, y),\ n \in \mathcal{N}$;

(ii) $\phi(n)$ *is continuous on* $[1, n_M]$;

(iii) 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 the Assumption 3 (*iii*) does not hold then noise level is low and neglegeable. 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.

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 [9] to eliminate the influence of the other boundary point $t = 0$ we assume that $y \in C_0^r = \{y : \ y \in C^r[0,1], \ y^{(i)}(0) = 0, \ i = 0, 1, \cdots, r-1\}$.*

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 in 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 the Figures 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 considered noise level the 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 more transparent way we keep the same scale for all graphs.

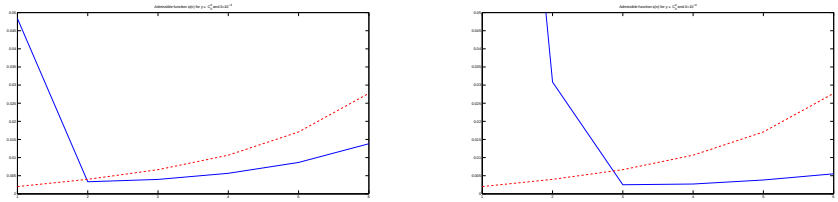

Figure 1: *The functions $\phi(n)$ (solid line) and $\psi(n)\frac{\delta}{h_s}$ (dotted line) for $h_s = 10^{-1}, \delta = 10^{-4}$ and $y \in C^3$ (left panel) and $y \in C^4$ (right panel).*

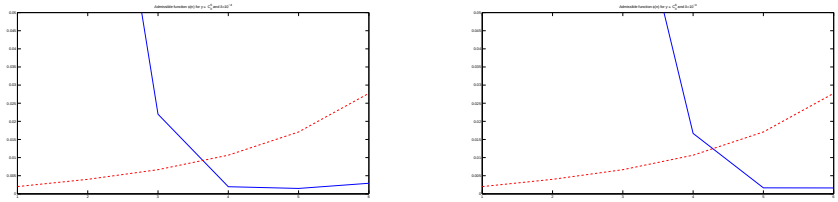

Figure 2: *The functions $\phi(n)$ (solid line) and $\psi(n)\frac{\delta}{h_s}$ (dotted line) for $h_s = 10^{-1}, \delta = 10^{-4}$ and $y \in C^5$ (left panel) and $y \in C^6$ (right panel).*

Remark 1. Note that the Assumption 3 is more general than the assumption that an admissible function is non-increasing, which has been used in [2].

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

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

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 the 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

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

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

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 the Assumptions 1–3 we employ the balancing principle in the form presented in [2].

Consider the set

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

C is a tuning parameter. In particular, we use $C = 4$ for theoretical proofs and we adjust $C-$value 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 are 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 the 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 [2, 10].

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

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

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

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

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

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

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

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

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

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

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

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

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

The proof is complete. □

2.1. Numerical tests

To illustrate the theoretical results of the previous section, we test the rule of estimating 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 final smoothness

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

as well as analytic functions

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

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

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^{-2}	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 3: The results of the application of the choice rule (16)

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

$$h_s = 1.5^{-s},\ s = 4, 5, \ldots, 11, \tag{23}$$

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

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. Then found $C-$value 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 [3] 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 the Table 3 we can conclude that for a fixed step-size it is not the case. For example, in case of the function (20), which belongs to $C^r[0, 1],\ r = 6$, the best guaranteed order of accuracy in terms of 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).

2.1.1. Combination of the balancing principle and 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 h_s-value, 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 [11], [12].

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 find

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

The approximation errors produced by $S_{n_{+(h_+)},h_+}y_\delta$ for considered functions are presented in Table 4.

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

Table 4: The approximation errors produced by $S_{n_{+(h_+)},h_+}y_\delta$

As it can be seen from Tables 3 and 4, in all three cases the proposed combination of the quasi-optimality criterion and the balancing principle really allows the best choice of h_s-value.

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 hyperglycemia, 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 [13].

Thus, the promising concept in the diabetes therapy management is the prediction of the future BG evolution using CGM data [14]. The importance of such prediction has been shown by several applications [15, 16].

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 $t_N > t_{N-1} > t_{N-2} > \ldots > t_{N-n+1}$ 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 $t_N < t_{N+1} < t_{N+2} < \ldots < t_{N+m}$.

There are several prediction techniques, and a variety of glucose predictors has been recently proposed, see, for example, [17, 18, 19, 20, 14]. In this section we discuss the predictors based on the numerical differentiation. Note that such approach has been used in several BG predictors. In [21] 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 [22], and another one has been recently discussed, in particular, in [14].

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 schemes (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 the one-sided formulae from the 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 [23]. 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 [14]. 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 hyperglycemic periods.

Following [22], 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 $y'(B)$ is estimated by means of the one-sided finite difference schemes from the Example 1.

To quantify the clinical accuracy of considered predictors we use the Prediction Error-Grid Analysis (PRED-EGA) [14], 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 look 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 hypoglycemic (0–70 mg/dL), euglycemic (70–180 mg/dL) and hyperglycemic (180–450 mg/dL) ranges. This stratification is of great importance because consequences caused by a prediction error in the hypoglycemic range are very different from ones in the euglycemic range [24].

First of all, we assess the performance of considered predictors (25) in the so-called ideal situation, when simulated data used as predictor inputs

Patient	BG $\leq$ 70 (mg/dL) (%)			BG 70-180 (mg/dL) (%)			BG $\geq$ 180 (mg/dL) (%)		
Vir. ID	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 5: The performance assessment matrix given by the PRED-EGA for the predictors (25), (5), based on any of the formulae from the Example 1, operating on simulated noise-free data with $PH = 15$ (min)

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 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 6 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).

At the same time, Table 5 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 the 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 the one-sided formulae from the Example 1 outperforms the predictors (25), (5) based on any of the formulae from the same set.

Patient	BG $\leq$ 70 (mg/dL) (%)			BG 70-180 (mg/dL) (%)			BG $\geq$ 180 (mg/dL) (%)		
Vir. ID	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.	**99.78**	**0.22**	–	**99.93**	**0.07**	–	**100**	–	–

Table 6: 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 the formulae from the Example 1, operating on simulated noise-free data with $PH = 15$ (min)

A more realistic situation can be modeled by adding white noise to the simulated readings and then by using them as predictor inputs. In this way we can assess the effect of the noise on the choice of the best order $n = n_+(h_s)$ and the step-size $h = h_+$ of the finite difference schemes from the Example 1.

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 8 and compared with the performance of the predictors (25), (5) based on the one-sided finite difference formula of the order $n = 6$, displayed in Table 7. 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 [25, 22] as the optimal one for BG prediction.

Comparing Tables 5–8, we can conclude that the BG predictor (25), (5) with the adaptive parameter choice rule $n = n_+(h_+)$ implemented on the

Patient	BG ≤ 70 (mg/dL) (%)			BG 70-180 (mg/dL) (%)			BG ≥ 180 (mg/dL) (%)		
Vir. ID	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.	**63.24**	**10.89**	**25.86**	**54.36**	**42.45**	**3.19**	**33.7**	**54.77**	**11.53**

Table 7: 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	BG ≤ 70 (mg/dL) (%)			BG 70-180 (mg/dL) (%)			BG ≥ 180 (mg/dL) (%)		
Vir. ID	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**	**5.50**	**94.80**	**4.58**	**0.62**	**91.43**	**8.57**	–

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 the formulae from the Example 1, operating on simulated noisy data with $PH = 15$ (min)

set of the formulae from the 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 hyperglycemic periods. Potentially, one may consider the combination of proposed adaptive BG predictor with a predictor based on more advanced techniques, such as mentioned in [14], for example.

Acknowledgment

This research has been performed in the course of the project "DIAdvisor" [26] funded by the European Commission within 7-th 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 (3–4) (1984) 259–267.

[2] S. Lu, S. V. Pereverzev, Numerical differentiation from a viewpoint of regularization theory, Math. Comp. 75 (256) (2006) 1853–1870.

[3] P. Mathé, S. V. Pereverzev, The use of higher order finite difference schemes is not dangerous, J. Complexity 25 (1) (2009) 3–10.

[4] R. Qu, A new approach to numerical differentiation and integration, Mathl. Comput. Modelling 24 (10) (1996) 55–68.

[5] A. Ramm, A. Smirnova, On stable numerical differentiation, Math. Comp. 70 (235) (2001) 1131–1153.

[6] H. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Vol. 375 of Mathematics and Its Applications, Kluwer Academic Publishers, Dordrecht, Boston, London, 1996.

[7] K. Rahul, S. N. Bhattacharyya, One-sided finite-difference approximations suitable for use with Richardson extrapolation, J. Computational Physics 219 (1) (2006) 13–20.

[8] B. Fornberg, Generation of finite difference formulas on arbitrarily spaced grids, Math. Comp. 51 (184) (1988) 699–706.

[9] S. M. Nikolskiy, Quadrature formulae, NAUKA, Moscow, 1974.

[10] P. Mathé, The Lepskii principle revisited, Inverse Problems 22 (3) (2006) 030011. 6 pp.

[11] A. N. Tikhonov, V. B. Glasko, Use of the regularization methods in non-linear problems, Vol. 5, USSR Comput. Math. Phys., 1965.

[12] S. Kindermann, A. Neubauer, On the convergence of the quasi-optimality criterion for (iterated) Tikhonov regularization, Inverse Problems and Imaging 2 (2) (2008) 291–299.

[13] V. Naumova, S. V. Pereverzyev, S. Sampath, Extrapolation in variable RKHSs with application to the blood glucose reading, Inverse Problems 27 (7) (2011) 075010. 13 pp.

[14] 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 (8) (2011) 787–796.

[15] 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. D. III, Prevention of nocturnal hypoglycemia using predictive alarm algorithms and insulin pump suspension, Diabetes Care 33 (5) (2010) 1013–1018.

[16] C. Palerm, B. W. Bequette, Hypoglycemia detection and prediction using continuous glucose monitoring a study on hypoglycemic clamp data, J Diabetes Sci Technol. 1 (5) (2007) 624–629.

[17] M. Eren, A. Cinar, L. Quinn, D. Smith, Estimation of future glucose concentrations with subject-specific recursive linear models, Diabetes Technol Ther 11 (4) (2009) 243–253.

[18] 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 (5) (2008) 792–801.

[19] J. Reifman, S. Rajaraman, A. Gribok, W. K. Ward, Predictive monitoring for improved management of glucose levels, J Diabetes Sci Technol. 1 (4) (2007) 478–486.

[20] 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 (5) (2007) 931–937.

[21] 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 (1) (2008) 158–163.

[22] A. Hayes, J. J. Mastrototaro, S. B. Moberg, J. C. M. 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.

[23] 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 (1) (2009) 44–55.

[24] 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 (10) (2005) 2412–2417.

[25] 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 (5) (2007) 645–651.

[26] Diadvisor: personal glucose predictive diabetes advisor (2008).
URL `http://www.diadvisor.eu`