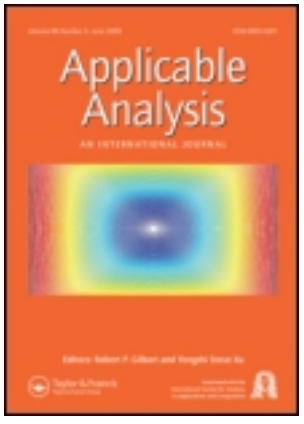

Applicable Analysis: An International Journal

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gapa20

Numerical differentiation by a Tikhonov regularization method based on the discrete cosine transform

Bang Hu [a] & Shuai Lu [a]

[a] School of Mathematical Sciences, Fudan University, Shanghai 200433, China

Version of record first published: 23 Sep 2011.

To cite this article: Bang Hu & Shuai Lu (2012): Numerical differentiation by a Tikhonov regularization method based on the discrete cosine transform, Applicable Analysis: An International Journal, 91:4, 719-736

To link to this article: http://dx.doi.org/10.1080/00036811.2011.598862

Applicable Analysis
Vol. 91, No. 4, April 2012, 719–736

Taylor & Francis
Taylor & Francis Group

Numerical differentiation by a Tikhonov regularization method based on the discrete cosine transform

Bang Hu and Shuai Lu*

School of Mathematical Sciences, Fudan University, Shanghai 200433, China

Communicated by B. Hon

(*Received 26 February 2011; final version received 15 June 2011*)

In this article, we discuss a classical ill-posed problem on numerical differentiation by a Tikhonov regularization method based on the discrete cosine transform (DCT). After implementing an eigenvalue decomposition of the second-order difference matrix in the penalty we obtain an explicit minimizer of the Tikhonov functional in a linear combination of DCT-2 components. The choice of the regularization parameter thus can be properly chosen after calling a single variable bounded nonlinear function minimization. Moreover, we propose two approaches to weaken the Gibbs phenomena appearing near the boundary of the reconstructed derivatives. Several numerical examples are provided to show the computational efficiency of our proposed algorithm. Finally, the extension to the multidimensional numerical differentiation leads the approach to a wider scope.

Keywords: (multidimensional) numerical differentiation; Tikhonov regularization; parameter choice rules; Gibbs phenomena

AMS Subject Classifications: 65D25; 65J20

1. Introduction

Numerical differentiation is a classical computational inverse problem where one needs to approximate the derivative $y'(t)$ of an unknown smooth function $y(t)$ from the noisy observation data $y^\delta(t)$ satisfying

$$\| y(t) - y^\delta(t) \|_{L^2(0,1)} \leq \delta \tag{1.1}$$

for $t \in [0, 1]$. Such a problem is well-known as a typical ill-posed problem [1] in the sense that a small perturbation in the observation may lead to large errors in the reconstructed derivatives. Since in real applications, observed data may always contain some noise, one has to deal with the instability which is induced by the imperceptible difference between the exact data and noisy one. In order to approximate the derivative in a stable way, different regularization methods have been considered in various context, for instance, [2–9].

*Corresponding author. Email: slu@fudan.edu.cn

ISSN 0003–6811 print/ISSN 1563–504X online

http://dx.doi.org/10.1080/00036811.2011.598862
http://www.tandfonline.com

To be more precise, when we say regularization methods in numerical differentiation, usually there are two categories of such methods. The first one contains those regularization methods which use a continuous information of the noisy data [2,5,6]. The forward finite difference formula, for example,

$$(y^{\delta}(t))' \approx \frac{y^{\delta}(t+h) - y^{\delta}(t)}{h}$$

is the simplest case where one tries to approximate the derivative of noisy observation $y^{\delta}(t)$ pointwisely. In this category, as it has been observed in [2,6], the stepsize h plays the role of the regularization parameter and has to be chosen properly. An adaptive choice of the stepsize h which yields an order optimal convergence rate without knowledge of the smoothness of the exact function $y(t)$ is shown in [5,10] based on the balancing principle [11].

In this article, we concentrate our interests on the second category of the regularization methods, which are those that use a discrete knowledge of noisy observation $y^{\delta}(t)$ at a fixed sampling points set $\{t_i\}_{i=0}^{n-1} \subset [0, 1]$ with $t_0 = 0$, $t_{n-1} = 1$ (see, e.g. [3–5,7–9]). One simple regularization method in this category is to construct the derivative $S'_{n-1}(t)$ of a natural cubic spline $S_{n-1}(t)$ by solving the minimization approach (cf. [7])

$$\frac{1}{n}\sum_{i=1}^{n}(y^{\delta}(t_i) - S_{n-1}(t_i))^2 + \alpha\|S''_{n-1}(t)\|^2 \to \min$$

where the value of the function $y(t)$ at the boundary points are exactly known, i.e. $y^{\delta}(t_0) = y(t_0)$ and $y^{\delta}(t_{n-1}) = y(t_{n-1})$. Such an approach is usually called the Tikhonov regularization where α near the penalty term is the regularization parameter whose choices are well-developed both in *a priori* and *a posteriori* ways [3,8] depending on the noise level δ. We note that similar approaches to find the derivative $\tilde{y}'(t)$ of a function $\tilde{y}(t)$ satisfying the minimization approach

$$\|y^{\delta} - \tilde{y}\|^2 + \alpha\|\tilde{y}''\|^2 \to \min \tag{1.2}$$

have been realized in different discretization forms, for instance, multiscale kernels [4], B-splines [5] and radial basis functions [9]. At the same time, in statistics and data analysis, (1.2) is usually called penalized least squares regression which is firstly introduced in [12] and also well-studied in [7]. In a particular case of evenly distributed sampling points, Cholesky decomposition [13] and discrete cosine transform [14] are shown to be relatively effective in realizing the minimization approach.

Motivated by the results in [14], the goal of this article is to discuss the numerical differentiation within the framework of DCT [15] in the form of Tikhonov regularization method (1.2) for evenly distributed sampling data. We show that by implementing the DCT it is not necessary to generate the matrix of the direct operator and the regularization parameter can be properly chosen by calling an explicit single variable bounded nonlinear function minimization. It is worth noting that a similar approach for one-dimensional numerical differentiation based on the fast Fourier transform can be found in [16].

This article is organized as follows. In Section 2, we present the mathematical formulation of the Tikhonov regularization based on the DCT. Different choices of

the regularization parameter α are discussed in Section 3. One can observe that all the mentioned criteria can be realized in an explicit minimization of a single variable bounded nonlinear function. Section 4 contains two simple approaches in order to weaken the Gibbs phenomena appearing near the boundary of the reconstructed derivatives. Finally, we show that our approach can be easily extended to the multidimensional numerical differentiation in Section 5.

2. Regularization scheme reformulation

Without loss of generality, let $y(t)$ be a smooth function on $[0,1]$ from which evenly distributed sampling data $y_i = y(\frac{i}{n-1})$ at $i = 0, 1, \ldots, n-1$ are produced. Here we use $\{t_i\}_{i=0}^{n-1}$ with $t_i = \frac{i}{n-1}$ denoting the equal-distance discrete sampling points in $[0,1]$. When we say a DCT, we mean a discrete linear operator DCT$(\cdot)$ which maps the discrete sampling data $\{y_i\}_{i=0}^{n-1}$ to a set of DCT coefficients $\{Y_m\}_{m=0}^{n-1}$ such that

$$Y_m = \mathrm{DCT}(\{y_i\}_{i=0}^{n-1}) = \sum_{i=0}^{n-1} y_i f_i(m), \quad m = 0, 1, \ldots, n-1, \tag{2.1}$$

where $f_i(m)$ is the forward DCT components. In the context, we concentrate our interests on the DCT-2 components [15] where

$$f_i(m) = c(m)\frac{\sqrt{2}}{\sqrt{n}}\cos\left(\frac{im\pi}{n} + \frac{m\pi}{2n}\right)$$

with $c(m) = 1/\sqrt{2}$ for $m = 0$ and $c(m) = 1$ otherwise. An inverse discrete cosine transform (IDCT) is defined in an analogous way such that

$$\tilde{y}(t) = \mathrm{IDCT}(\{Y_m\}_{m=0}^{n-1}) = \sum_{m=0}^{n-1} Y_m g_m(t), \tag{2.2}$$

with $g_m(t) = c(m)\frac{\sqrt{2}}{\sqrt{n}}\cos((t \cdot \frac{n-1}{n} + \frac{1}{2n})m\pi)$ defined as the (continuous) IDCT-2 components. $\tilde{y}(t)$ can be seen as a continuous approximation of $y(t)$ and one can easily observe that $\tilde{y}(t_i) = y(t_i)$ when we choose the same sampling points $\{t_i\}_{i=0}^{n-1}$. From a mathematical viewpoint, (2.1) and (2.2) equal the following minimization approach:

$$\min_{\tilde{y}(t) \in \mathrm{span}\{g_m(t)\}} \sum_{i=0}^{n-1} (y(t_i) - \tilde{y}(t_i))^2. \tag{2.3}$$

Once the reconstructed function $\tilde{y}(t)$ is obtained from the discrete sampling data $\{y_i\}_{i=0}^{n-1}$, one can use the derivative $\tilde{y}'(t)$ as an approximation of the original derivative $y'(t)$. For instance, one can introduce a reconstructed derivative transform RDT$(\cdot)$ to obtain $\tilde{y}'(t)$ in the following form:

$$\begin{aligned}\tilde{y}'(t) &= \mathrm{RDT}(\{Y_m\}_{m=0}^{n-1}) = \sum_{m=0}^{n-1} Y_m g_m'(t) \\ &= -\sum_{m=0}^{n-1} Y_m c(m)\frac{\sqrt{2}}{\sqrt{n}}\frac{(n-1)m\pi}{n}\sin\left(\left(t \cdot \frac{n-1}{n} + \frac{1}{2n}\right)m\pi\right).\end{aligned} \tag{2.4}$$

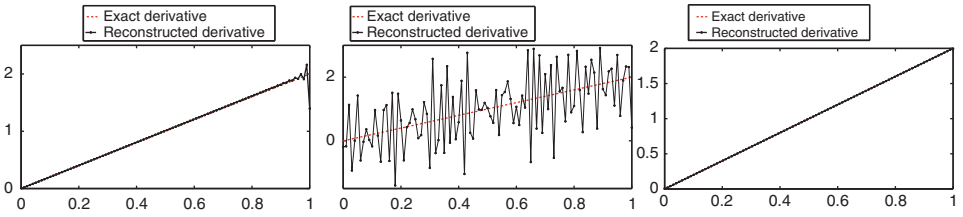

Figure 1. The reconstructed derivative based on the least squares method (2.3) and (2.4). The left graph is a direct reconstructed derivative vs. the exact derivative on the noise-free data where the Gibbs phenomena appears near the right boundary. The middle graph is the reconstructed derivative on the noisy observation with 1% noise. The right graph is the reconstructed derivative after weakening the Gibbbs phenomena near the right boundary for the noise-free data.

In the context, we call $g'_m(t)$ the derivative component of the original IDCT-2 component $g_m(t)$. The accuracy of the proposed approach can be observed in Example 2.1. As one can notice in the same example, such an idea (2.3) will fail when the sampling data $y^\delta = \{y^\delta(t_i)\}_{i=0}^{n-1}$ contains some observation noise δ satisfying $\sum_{i=0}^{n-1} |y^\delta(t_i) - y(t_i)|^2 \le \delta^2$.

Example 2.1 The unknown exact function is $y(t) = t^2$, i.e. $y'(t) = 2t$. The discrete sampling points are $\{t_i\}_{i=0}^{99}$ with $n = 100$ and $t_i = \frac{i}{99}$. The reconstructed derivative $\tilde{y}'(t)$ following the description above is shown in the left graph of Figure 1. Meanwhile, assume that the sampling data $\{y^\delta(t_i)\}_{i=0}^{99}$ contains 1% observation noise. The reconstructed derivative in the middle graph of Figure 1 contains no valuable information. The right graph of Figure 1 represents a corresponding reconstructed derivative after weakening the Gibbs phenomena. Detailed discussion can be found in Section 4.

A natural way to decrease the influence of the observation noise is to use the regularization method, for instance, in the form of (1.2). One then needs to seek for the minimizer $\tilde{y}_\alpha^\delta(t)$ of the discrete regularization functional (1.2) such that

$$\min_{\tilde{y}_\alpha^\delta(t) \in \operatorname{span}\{g_m(t)\}} \sum_{i=0}^{n-1} (y^\delta(t_i) - \tilde{y}_\alpha^\delta(t_i))^2 + \alpha \|D_n \tilde{y}_\alpha^\delta\|^2. \tag{2.5}$$

The matrix D_n in the additional penalty term stands for a second-order divided difference matrix, or in another name, a discrete form of the second derivative. $\|D_n \cdot\|$ here is the vector norm in R^n and $\tilde{y}_\alpha^\delta$ is the discrete vector of $[\tilde{y}_\alpha^\delta(t_0), \ldots, \tilde{y}_\alpha^\delta(t_{n-1})]^T$.

Discretization of the differential operators is well-established in computational ill-posed problems and we refer to [17] for a numerical realization in Matlab environment. By assuming the reflective boundary conditions (i.e. $y(t_{-1}) = y(t_{n-1})$ and $y(t_n) = y(t_0)$), we obtain an $n \times n$-matrix D_n

$$D_n = \begin{pmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -1 \end{pmatrix} \tag{2.6}$$

as the difference matrix representing the second-order differential operator in (1.2). Notice the current discrete differential operator is not in the standard form as in [17]. Once the difference matrix D_n is obtained ref. [18] shows that such a D_n yields an eigenvalue decomposition in the form of $D_n = U_n \Gamma_n U_n^{-1}$ where

$$\Gamma_n = \text{diag}\{\lambda_0, \ldots, \lambda_{n-1}\} \quad \text{with } \lambda_i = -2 + 2\cos(i\pi/n). \tag{2.7}$$

Moreover, the matrices U_n^{-1} and U_n are $n \times n$ type-2 DCT and IDCT matrices which are unitary. Such a property gives $U_n^{-1} = U_n^T$.

Recall that the Euler equation of (2.5) yields a discrete reconstruction $\tilde{y}_\alpha^\delta$ from the noisy sampling data y^δ such that

$$\tilde{y}_\alpha^\delta = (I_n + \alpha D_n^T D_n)^{-1} y^\delta \tag{2.8}$$

for a fixed regularization parameter α, we then can formulate a simplified continuous reconstructed function $\tilde{y}_\alpha^\delta(t)$ in the form of DCT and IDCT such that

$$\tilde{y}_\alpha^\delta(t) = \text{IDCT}(\Pi_n(\alpha)\text{DCT}(y^\delta)), \tag{2.9}$$

where $\Pi_n(\alpha) = (I_n + \alpha\Gamma_n^2)^{-1}$. The reconstructed derivative $\left(\tilde{y}_\alpha^\delta(t)\right)'$ can thus be obtained similarly in the form of (2.4). More precisely, once the regularization parameter α is properly chosen, one first calculates the weighted DCT coefficients $\{\text{WDCT}_m(y^\delta)\}_{m=0}^{n-1}$ such that

$$\text{WDCT}_m(y^\delta) = \frac{1}{1 + \alpha\lambda_m^2}\text{DCT}_m(y^\delta), \quad \text{for } m = 0, \ldots, n-1, \tag{2.10}$$

where $\text{DCT}_m(y^\delta)$ denotes the DCT coefficients of the noisy sampling data y^δ, i.e.

$$\begin{aligned}\text{DCT}_m(y^\delta) &= \text{DCT}(\{y^\delta(t_i)\}_{i=0}^{n-1}) \\ &= \sum_{i=0}^{n-1} y^\delta(t_i) f_i(m), \quad m = 0, 1, \ldots, n-1.\end{aligned} \tag{2.11}$$

$\left(\tilde{y}_\alpha^\delta(t)\right)'$ then is formulated in the following form:

$$\begin{aligned}\left(\tilde{y}_\alpha^\delta(t)\right)' &= \text{RDT}\left(\left\{\text{WDCT}_m(y^\delta)\right\}_{m=0}^{n-1}\right) = \sum_{m=0}^{n-1}\text{WDCT}_m(y^\delta)\, g_m'(t) \\ &= -\sum_{m=0}^{n-1}\text{WDCT}_m(y^\delta)c(m)\frac{\sqrt{2}}{\sqrt{n}}\frac{(n-1)m\pi}{n}\sin\left(\left(t\cdot\frac{n-1}{n} + \frac{1}{2n}\right)m\pi\right).\end{aligned} \tag{2.12}$$

We note that in [14] the author considers (2.9) as a robust data smoothing technique with missing sampling data by introducing special weights. For us, the most interesting formula is the continuous reconstructed derivative $\left(\tilde{y}_\alpha^\delta(t)\right)'$ in (2.12). One then need to choose the regularization parameter α carefully to obtain a properly reconstructed derivative.

3. Choices of the regularization parameter

One of the important issues in inverse problems is an appropriate choice of the regularization parameter, which is the key to successfully solving such problems

as well. Among many parameter choice rules there are those with a sound mathematical foundation and others which are heuristic [1,19,20]. In this section, as having been well-studied in the regularization theory, we will reinvestigate several regularization parameter choice rules, respectively. In case of the regularization in general form (i.e. (2.5)) we use the (semi-)norm $\|D_n \cdot\|$ in R^n in the parameter choice rules.

Before we go further, two important functions over the regularization parameter α are presented by using the explicit reconstructed function $\tilde{y}_\alpha^\delta(t)$ in (2.9). As shown in [14], the discrete discrepancy in (2.5) between $\tilde{y}_\alpha^\delta(t)$ and $y^\delta(t)$ can be simplified by following function:

$$\mathrm{Dis}(\alpha) = \sum_{i=0}^{n-1} (y^\delta(t_i) - \tilde{y}_\alpha^\delta(t_i))^2 = \sum_{m=0}^{n-1} \left(\frac{1}{1+\alpha\lambda_m^2} - 1 \right)^2 \mathrm{DCT}_m(y^\delta)^2. \tag{3.1}$$

Another important function is the penalty term in (2.5) which is defined as $\mathrm{Pen}(\alpha)$ such that

$$\mathrm{Pen}(\alpha) = \|D_n \tilde{y}_\alpha^\delta\|^2 = \sum_{m=0}^{n-1} \left(\frac{\lambda_m}{1+\alpha\lambda_m^2} \right)^2 \mathrm{DCT}_m(y^\delta)^2. \tag{3.2}$$

The defined discrepancy and penalty functions over α then formulate some of the parameter choice rules into corresponding explicit forms. The performance comparison among following parameter choice rules is presented in the following section where Gibbs phenomena of the reconstructed derivatives is weakened near the boundary. Here, we mainly revisit these rules and present their explicit forms in our current framework.

3.1. *Generalized cross-validation*

The generalized cross-validation (GCV) parameter choice rule is proposed in [21] and may be the most widely used one in statistics. Ref. [22] shows that the regularization parameter α can be chosen as the minimizer of a function such that

$$\alpha = \arg\min(\mathrm{GCV}(\alpha)) \quad \text{with} \quad \mathrm{GCV}(\alpha) = \frac{\mathrm{Dis}(\alpha)/n}{(1 - \mathrm{Trace}(H)/n)^2}, \tag{3.3}$$

where $H = (I_n + \alpha D_n^T D_n)^{-1}$. As already discussed in [14,23], $\mathrm{GCV}(\alpha)$ of (3.3) can be easily rewritten as follows:

$$\mathrm{GCV}(\alpha) := \frac{n \sum_{m=0}^{n-1} \left(\frac{1}{1+\alpha\lambda_m^2} - 1 \right)^2 \mathrm{DCT}_m^2(y^\delta)}{\left(n - \sum_{m=0}^{n-1} \frac{1}{1+\alpha\lambda_m^2} \right)^2}.$$

3.2. *(Modified) L-curve method*

The L-curve method is based on a graph of the penalty function $\mathrm{Pen}(\alpha)$ versus the discrepancy function $\mathrm{Dis}(\alpha)$ in a log–log scale. The overall shape of the graph often

looks like the letter 'L'. The choice of the regularization parameter is focused on the corner where the vertical line turns to be a horizontal one. Such an idea retrospects back to [24] and was so named much later in [25,26]. The definition of the corner is suggestion in [26] such that the curvature of the graph reaches its maximum where

$$\alpha = \arg\max L(\alpha) \quad \text{with} \quad L(\alpha) = \frac{\rho'(\alpha)\eta''(\alpha) - \rho''(\alpha)\eta''(\alpha)}{((\rho'(\alpha))^2 + (\eta'(\alpha))^2)^{3/2}}$$

by defining $\rho(\alpha) = \log \mathrm{Dis}(\alpha)$ and $\eta(\alpha) = \log \mathrm{Pen}(\alpha)$.

Meanwhile, ref. [27] points out that the L-curve method is equivalent to the following modified L-curve (MLC) minimization approach

$$\alpha = \arg\min MoL(\alpha) \quad \text{with} \quad MoL(\alpha) = \mathrm{Dis}(\alpha) Pen^{\mu}(\alpha). \tag{3.4}$$

$-1/\mu$ is the tangent slope value when the regularization parameter α is chosen by the L-curve method in a log–log scale. In our particular forms of the $\mathrm{Dis}(\alpha)$ and $\mathrm{Pen}(\alpha)$, one has an explicit following function:

$$MoL(\alpha) := \left(\sum_{m=0}^{n-1}\left(\frac{1}{1+\alpha\lambda_m^2} - 1\right)^2 \mathrm{DCT}_m^2(y^{\delta})\right)\left(\sum_{m=0}^{n-1}\left(\frac{\lambda_m}{1+\alpha\lambda_m^2}\right)^2 \mathrm{DCT}_m^2(y^{\delta})\right)^{\mu}.$$

The constant μ is usually *a priori* fixed from the computation experiences, for instance, in this article we fix $\mu = 2$.

3.3. *Discrepancy principle*

The previous two parameter choice rules are heuristic where the noisy level δ is not included. At the same time, when such information is available, some other noise dependent methods become attractive since sound convergence and convergence rate results have been proven (cf. [1]).

For the sake of simplicity, in the context, we mainly reinvestigate the discrepancy principle (DP) [28] which has been frequently used in many applications. The idea behind the DP is to choose a regularization parameter α such that the value of the discrepancy, namely $\mathrm{Dis}(\alpha)$, equals the squared noisy level, i.e.

$$\sum_{i=0}^{n-1}\left(\frac{1}{1+\alpha\lambda_i^2} - 1\right)\mathrm{DCT}_i^2(y^{\delta}) = \delta^2$$

or one can formulate the DP into an equivalent minimization approach

$$\alpha = \arg\min\left|\mathrm{Dis}(\alpha) - \delta^2\right|. \tag{3.5}$$

The explicit discrepancy and penalty functions also works for variants of parameter choice rules, for instance, the Hanke-Raus rule [29], the Gfrerer rule [30] or the monotone error rule [31]. We only mention the names here but encourage those who are interested in for a detailed discussion in these references. One can obtain a convergence analysis near the sampling points for these *a posteriori* parameter choice rules by assuming that the original function $y(t)$ can be expanded in a cosine series. Detailed discussion will be carried out in a forthcoming paper.

Remark 1 All the above mentioned parameter choice rules have in common an explicit function form over the regularization parameter α where standard routined i.e. *fminbnd* function in Matlab, can provide the minimizer. In the following section, we will compare the performance among these rules. One should understand that there is no universal criteria for choosing the regularization parameter yet, and most likely will not exist [20]. The current conclusions among different rules are only valid for this particular problem and the corresponding mathematical setting.

Remark 2 Note that the choice of the regularization parameter in the form of minimization of an explicit function is also valid when the fast Fourier transform is employed to replace the DCT. We mention ref. [16] where a heuristic parameter choice rule is considered.

4. Gibbs phenomena

The original Gibbs phenomena intimately appears in the study of Fourier series when an analytic but non-periodic function is presented in a linear combination of truncated Fourier coefficients. We refer to the review [32] and references therein for those who are interested in this topic.

In our case the situation is slightly different. As one can notice in (2.4) and (2.12) the reconstructed derivative at the boundary of the domain will be small if n is chosen sufficient large. We present Example 4.1 indicating the Gibbs phenomena in current framework.

Example 4.1 This example concerns an unknown exact function $y(t) = (t-0.5)^2$ with $y'(t) = 2(t-0.5)$. Sampling points $\{t_i\}_{i=0}^{99}$ are the same as in Example 2.1. The noisy sampling data y^δ are generated with 1% normally distributed noise. The reconstructed derivative (2.12) with the regularized parameter α chosen by the MLC method (3.4) is obtained from the discrete noisy sampling data $\{y^\delta(t_i)\}_{i=0}^{99}$. Notice $y'(t) \neq 0$ at $t = 0$, 1, Gibbs phenomena, which can be clearly observed from the left graph of Figure 2, appears near both boundaries since each derivative component $g'_m(t)$ in (2.12) is quite small at $t = 0$ and $t = 1$. We then consider a second unknown testing function $y(t) = t^3/3 - t^2/2$ with $y'(t) = t(t-1)$ where $y'(0) = y'(1) = 0$. The sampling points $\{t_i\}_{i=0}^{99}$ and noisy sampling data y^δ are generated in the same way. The reconstructed derivative, shown in the right graph of Figure 2, is obtained in the same routine (2.12) with α chosen by (3.4). One can observe that

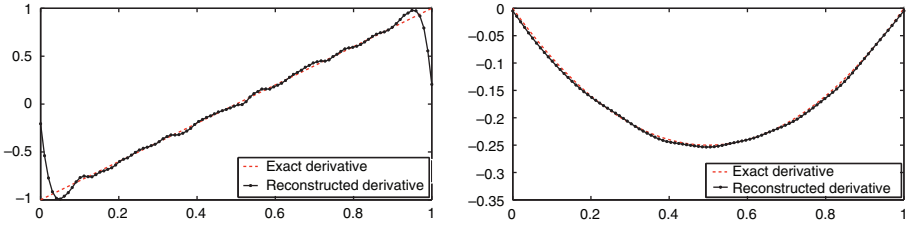

Figure 2. The reconstructed derivative vs. the exact derivative for two testing functions in Example 4.1. The left graph shows a clear Gibbs phenomena on both boundaries while the right graph does not since the exact derivative on the boundaries for the right graph is zero.

the Gibbs phenomena is automatically weakened when the reconstructed derivative $y'(t)$ is quite small near the boundaries.

Since the value of the exact derivative may vary on the boundary, one needs to treat the Gibbs phenomena carefully to obtain properly reconstructed derivatives. The classic approach to weaken the Gibbs phenomena is usually by utilizing the weighted Gegenbauer polynomials based on the knowledge of truncated Fourier coefficients [32]. Such an approach cannot be implemented in current framework since no Fourier transform is considered. One thus needs some other treatments to weaken the Gibbs phenomena appearing in the left graph of Figure 2.

In this section, we propose two simple approaches to weaken the Gibbs phenomena near the boundary when we reconstruct the derivative in the form of (2.12). Both approaches are motivated by the numerical experiments in Example 4.1.

4.1. *Even expansion of the sampling data (even expansion approach)*

First attempt to weaken the Gibbs phenomena is motivated by the left graph of Figure 2. As one may notice, the Gibbs phenomena only influences the reconstructed derivative in a short interval near the boundary since the DCT uses reflective boundary conditions [23]. The other part of the reconstructed derivative, for instance in $(0.2, 0.8)$, looks quite good. A natural idea is then to expand the sampling data y^δ from both boundaries in an even way so that the underlying expanded exact derivative contains the original one and is continuous in the whole expanded interval. When we implement the regularization method on the expanded sampling data, the Gibbs phenomena of the reconstructed expanded derivative will only appear near the expanded boundaries. One then cuts off the expanded interval to obtain the reconstructed derivative for the original sampling data where Gibbs phenomena is weakened. A similar approach has been widely used in the Fourier series method for partial differential equations (see, e.g. [33]).

We use the same function in left graph of Figure 2 as an illustration, i.e. the expanded interval is $[-1, 2]$ and the size of the expanded sampling data is 298. The expanded exact function $y\exp$ and the expanded exact derivative $y\exp'(t)$ are

$$y\exp(t) = \begin{cases} -(t+0.5)^2 + 0.5, & t \in [-1, 0) \\ (t-0.5)^2, & t \in [0, 1] \\ -(t-1.5)^2 + 0.5, & t \in (1, 2] \end{cases}, \quad y\exp'(t) = \begin{cases} -2(t+0.5), & t \in [-1, 0) \\ 2(t-0.5), & t \in [0, 1] \\ -2(t-1.5), & t \in (1, 2] \end{cases}.$$

The expanded noisy sampling data and the original one are shown in the left graph of Figure 3. The overall interval of the expanded noisy sample data is three times of the original one while the whole original sampling data are contained in the expanded one away from the expanded boundaries. We implement the same RDT (2.12) and the same parameter choice rule (3.4) on the expanded sampling data. The reconstructed expanded derivative is shown in the middle graph of Figure 3. As one can observe, though the Gibbs phenomena appears near the expanded boundaries the derivative in the original interval approximates the exact derivative quite well. We then remove the expanded interval to obtain the reconstructed derivative which is presented in the right graph of Figure 3. Compared with that in

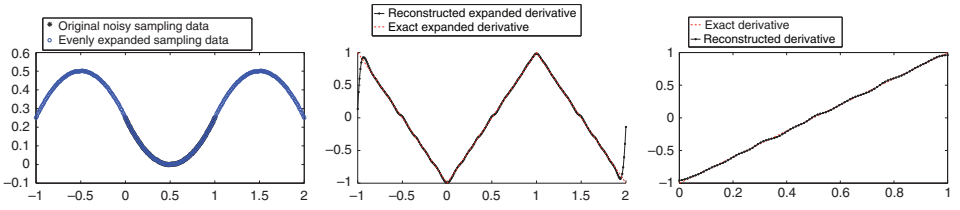

Figure 3. In the left graph, the solid ○ are the expanded noisy sampling data after implementing the even expansion. The original noisy sampling data is presented with * in the same graph. The reconstructed derivative of the expanded noisy sampling data vs. the exact expanded derivative is shown in the middle graph where Gibbs phenomena appears on small intervals near the expanded boundaries −1 and 2. We show the reconstructed derivative on the original interval vs. the exact original derivative in the right graph of Figure 3. The Gibbs phenomena is essentially weakened on both boundaries in the right graph.

the left panel in Figure 2, the Gibbs phenomena is essentially weakened. The relative error between the reconstructed derivative and the exact one is around 2%.

4.2. *The zero derivative of the sampling data (zero derivative approach)*

This approach is motivated by the right graph of Figure 2 where almost no Gibbs phenomena can be observed when the exact derivative equals zero on both boundaries.

Assuming the exact derivative $y'(t)$ has the value $y'(0)$ and $y'(1)$ on both boundaries, we would like to construct a substituted derivative $y\,\mathrm{sub}'(t)$, with $y\,\mathrm{sub}(t)$ denoting the substituted exact function, such that both boundary value $y\,\mathrm{sub}'(0) = y\,\mathrm{sub}'(1) = 0$. This can be easily realized by adding a linear function $(y'(0) - y'(1))t - y'(0)$ to the exact derivative $y'(t)$ to obtain such a substituted one

$$y\,\mathrm{sub}'(t) = y'(t) + (y'(0) - y'(1))t - y'(0). \tag{4.1}$$

The substituted exact function $y\,\mathrm{sub}(t)$ can then be constructed as follows:

$$y\,\mathrm{sub}(t) = y(t) + \frac{(y'(0) - y'(1))}{2}t^2 - y'(0)t,$$

where a constant is ignored in the construction. When dealing with the discrete sampling data $\{y(t_i)\}_{i=0}^{n-1}$, one only need to add a new sampling data set $\left\{\frac{(y'(0)-y'(1))}{2}t_i^2 - y'(0)t_i\right\}_{i=0}^{n-1}$ to construct the substituted one $\{y\,\mathrm{sub}(t_i)\}_{i=0}^{n-1}$ where the corresponding derivative $y\,\mathrm{sub}'(t)$ equals zero on both boundaries. The Gibbs phenomena then can also be weakened concerning the numerical illustration in the right graph of Figure 2 when we reconstruct the derivative $y\,\mathrm{sub}'(t)$. One then obtain, the reconstructed derivative $y'(t)$ by reducing the added term $(y'(0) - y'(1))t - y'(0)$ from $y\,\mathrm{sub}'(t)$ according to (4.1). Final issue in the current approach is to obtain an approximated boundary derivative within the knowledge of discrete sampling data; one actually can use forward (backward) finite difference formulae (or multiple-point finite difference formulae [34]) to obtain the needed information.

Such an idea can also be implemented on the noisy sampling data $\{y^\delta(t_i)\}_{i=0}^{n-1}$. The substitute sampling data is constructed as follows:

$$y\,\mathrm{sub}^\delta(t_i) = y^\delta(t_i) + \frac{(y0' - y1')}{2} t_i^2 - y0' t_i, \quad \text{for } i = 0, 1, \ldots, n-1,$$

where $y0' = (y^\delta(t_2) - y^\delta(t_1))(n-1)$ approximates the boundary derivative $y'(0)$ and $y1' = (y^\delta(t_{n-1}) - y^\delta(t_{n-2}))(n-1)$ approximates the boundary derivative $y'(1)$, respectively. One then follows the approach in (2.12) to obtain the reconstructed derivative $(\widetilde{y\,\mathrm{sub}}_\alpha^\delta(t))'$ of the noisy substitute sampling data $\{y\,\mathrm{sub}^\delta(t_i)\}_{i=0}^{n-1}$ where the reconstructed derivative $\left(\tilde{y}_\alpha^\delta(t)\right)'$ of the noisy sampling data $\{y^\delta(t_i)\}_{i=0}^{n-1}$ can be obtained by the following formula:

$$\left(\tilde{y}_\alpha^\delta(t)\right)' = \left(\widetilde{y\,\mathrm{sub}}_\alpha^\delta(t)\right)' - (y0' - y1')t + y0'.$$

Hereby we use a numerical example to show the performance of the proposed approach. We note in the right graph in Figure 1 that the reconstructed derivative of the exact sampling data is also obtained by using the zero derivative approach.

Example 4.2 We again consider the exact function $y(t) = (t-0.5)^2$ and exact derivative $y'(t) = 2(t-0.5)$. The discrete sampling data of the exact function $y(t)$ and substituted function $y\,\mathrm{sub}(t)$ are shown in the upper left graph of Figure 4. The corresponding reconstructed derivative obtained by (2.12) and (3.4) versus the exact derivative is shown in the upper right graph of Figure 4. The relative error between them are up to 10^{-4} which indicates that the reconstructed derivative is quite accurate.

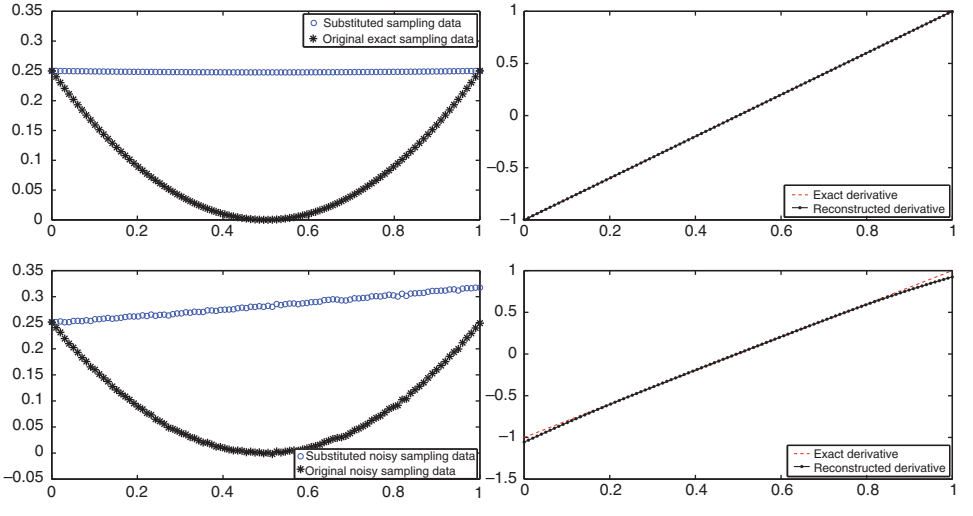

Figure 4. The upper left graph presents the substituted sampling data ○ vs. the exact sampling data *. The reconstructed derivative vs. the exact derivative is shown in the upper right graph. The lower left graph presents the substituted sampling data ○ vs. the noisy sampling data * with 1% normally distributed noise. Corresponding reconstructed derivative vs. the exact derivative is shown in the lower right graph.

The noisy sampling data are obtained with 1% normally distributed noise over the exact sampling data. The graph of both noisy and substituted sampling data is shown in the lower left graph of Figure 4. The corresponding reconstructed derivative of the noisy sampling data obtained by (2.12) and (3.4) versus the exact derivative is shown in the lower right graph of Figure 4. In this graph, the relative error is around 2.5%. One can observe that in both noise-free and noisy cases the Gibbs phenomena is essentially weakened.

As shown in the numerical examples, both proposed approaches can essentially weaken the Gibbs phenomena near the boundary. Between these two approaches, the first one consumes more computation time since the overall sampling data are increased approximately three times. Meanwhile, the second approach highly depends on an accurate knowledge of the sampling data on both boundaries. We would recommend the audience to use the first approach since the computational complexity for n-points DCT is of $O(n \log n)$ according to [35]. Then the corresponding computational complexity for the expanded sampling data is still of the same order compared with the original one. Nevertheless, the second approach is highly recommended if some of the boundary sampling data are accurately known.

Algorithm 1 Numerical differentiation by the Tikhonov regularization

Input: $\{y^\delta(t_i)\}_{i=0}^{n-1}$, $\{t_i\}_{i=0}^{n-1}$

1: Calculate the DCT coefficients $\{\mathrm{DCT}_m(y^\delta)\}$ over $\{y^\delta(t_i)\}$ when an approach to weaken the Gibbs phenomena is included.
2: Call the minimization procedure to find a proper regularization parameter α.
3: Calculate the weighted DCT coefficients $\{\mathrm{WDCT}_m(y^\delta)\}$, for instance, following the formula (2.10).
4: Return the reconstructed derivative in the form of (2.12).

We present Algorithm 1 summarizing the contents in previous sections. As one can observe, the overall computational complexity of Algorithm 1 is $O(n \log n)$ when we choose either approach to weaken the Gibbs phenomena. Moreover, we avoid the matrices generation which may consume additional computational costs.

Finally, we present the comparison between the performance of different parameter choice rules in the following example where the Gibbs phenomena is weakened by the proposed approaches in this section.

Example 4.3 The exact function $y(t) = (t - 0.5)^2$ and exact derivative $y'(t) = 2(t - 0.5)$ are considered in this example. The compared parameter choice rules are the GCV criteria (3.3), the MLC method (3.4) with $\mu = 2$ and the DP (3.5). We note here that the GCV criteria usually provides a small regularization parameter so that in this example all the regularization parameter chosen by the GCV criteria are multiplied by a constant 100 which is also proposed in [14].

We simulate 1% normally distributed noise 50 times in order to see the robustness of different approaches. In Figure 5, we show the relative error between the reconstructed derivative and the exact one with a circle. The smaller the relative error, the better the reconstructed derivative. In the left graph of Figure 5, we compare the even expansion approach and the zero derivative approach under the

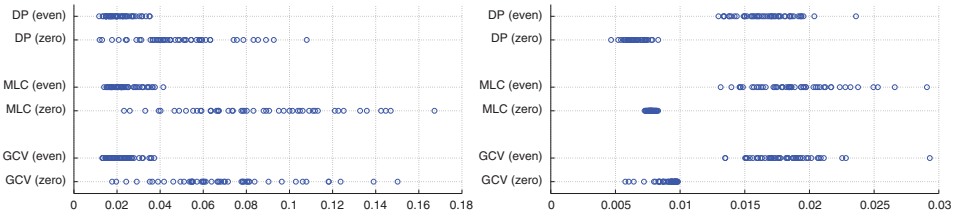

Figure 5. Comparison between different approaches in Example 4.3. Both graphs are obtained with 1% normally distributed noise simulated 50 times. Each circle presents the relative error. The left graph is obtained without any additional tricks. The right graph is obtained by setting the first two and last two sampling data equal the exact ones.

GCV criteria, the MLC method and the DP. One can see that the even expansion approach performs more accurate and more stable than the zero derivative approach. Meanwhile, the performance among the proposed three parameter choice rules are similar to each other where the DP is slightly better since more a priori information is known.

The right graph of Figure 5 is obtained when two boundary sampling data are accurately known, for instance, the boundary derivative obtained by the forward (backward) finite difference formula is approximated quite accurately. One then obtains great improvement by implementing the zero derivative approach. Once again, the DP outperforms the other two heuristic parameter choice rules.

5. Multidimensional extension

In the final section, we extend our approach to the multidimensional numerical differentiation. Some recent results on two-dimensional numerical differentiation can be found in [9,36] based on the Tikhonov regularization (1.2) discretized by radial basis functions and thin plate splines respectively.

5.1. *Multidimensional numerical differentiation formulation*

The multidimensional extension of (2.9) has been well-described in [14,23]. For the sake of completeness, we will adopt some formulae there to explain ours.

As shown in [15], a multidimensional DCTs is a composition of multiple one-dimensional DCTs along each dimension. Assume that the exact function $y(t^1, t^2, \ldots, t^N)$ is an N-variable function in the domain $[0,1]^N$ where N-dimensional evenly distributed noisy sampling data $y^\delta = y^\delta(t^1_{i_1}, t^2_{i_2}, \ldots, t^N_{i_N})$ is produced. Here we use n_j denoting the size of sampling data along j-th dimension such that $i_j = 0, \ldots, n_j - 1$. We then rewrite (2.9) as follows:

$$\tilde{y}^\delta_\alpha(t^1_{i_1}, t^2_{i_2}, \ldots, t^N_{i_N}) = \mathrm{IDCTN}(\Pi^N(\alpha) \circ \mathrm{DCTN}(y^\delta)) \tag{5.1}$$

where DCTN and IDCTN are the N-dimensional DCT and IDCT respectively, and $\circ$ is the Schur (elementwise) product. $\Pi^N(\alpha)$ is a tensor of rank N in the form of

$$\Pi^N = 1^N \div (1^N + \alpha \Gamma^N \circ \Gamma^N),$$

where $\div$ is the elements-by-elements division and 1^N is an N-rank tensor of ones. The N-rank eigenvalue tensor Γ^N stands for a diagonal tensor and the diagonal element corresponding to the point $(i_1, i_2, \ldots, i_N)$ is [23]

$$\Gamma^N_{i_1,i_2,..,i_N} = \sum_{k=1}^{N}\left(-2 + 2\cos\frac{i_k\pi}{n_k}\right).$$

Following (5.1), consequently, the multidimensional numerical differentiation by a Tikhonov regularization based on the DCT yields corresponding multiple one-dimensional numerical differentiation

$$\frac{\partial\tilde{y}^{\delta}_{\alpha}(t^1_{i_1}, t^2_{i_2}, \ldots t^j, \ldots, t^N_{i_N})}{\partial t^j} = \mathrm{RDT}N_j(\Pi^N(\alpha) \circ \mathrm{DCT}N(y^{\delta})) \tag{5.2}$$

for $j = 1, \ldots, N$ where $\mathrm{RDT}N_j$ is a linear operator in the modified form of N-dimensional IDCT where the IDCT-2 component in the j-th dimension

$$g_m(t^j) = c(m)\frac{\sqrt{2}}{\sqrt{n_j}}\cos\left(\left(\frac{n_j - 1}{n_j}t^j + \frac{1}{2n_j}\right)m\pi\right), \quad \text{for } m = 0, \ldots, n_j - 1$$

is replaced by the j-th derivative component

$$g'_m(t^j) = -c(m)\frac{\sqrt{2}}{\sqrt{n_j}}\frac{(n_j - 1)m\pi}{n_j}\sin\left(\left(\frac{n_j - 1}{n_j}t^j + \frac{1}{2n_j}\right)m\pi\right), \quad \text{for } m = 0, \ldots, n_j - 1.$$

Remark 1 In the multidimensional numerical differentiation (5.2), one only obtains a continuous function of the j-th component t^j following the proposed approach. The grid of the sampling points of other dimensions should be fine enough to obtain a good resolution of the reconstructed derivative.

5.2. *Parameter choice rules extension*

The parameter choice rules presented in Section 3 can be extended in the case of N-dimensional numerical differentiation. We first present the multidimensional GCV criteria which can also be found in [14,23] as follows:

$$\alpha = \arg\min(\mathrm{GCV}N(\alpha))$$

with

$$\mathrm{GCV}N(\alpha) = \frac{n_1 n_2 \cdots n_N \|(\Pi^N - 1^N) \circ \mathrm{DCT}N(y^{\delta})\|_F^2}{(n_1 n_2 \cdots n_N - \|\Pi^N\|_1)^2},$$

where $\|\cdot\|_F$ is the Frobenius norm and $\|\cdot\|_1$ is the 1-norm. The MLC method in the N-dimensional case can be rewritten as

$$\alpha = \arg\min(MoLN(\alpha))$$

with

$$MoLN(\alpha) = \|(\Pi^N - 1^N) \circ \mathrm{DCT}N(y^{\delta})\|_F^2 \|\Gamma^N \div (1^N + \alpha\Gamma^N \circ \Gamma^N) \circ \mathrm{DCT}N(y^{\delta})\|_F^{2\mu}.$$

Finally, the DP in the N-dimensional numerical differentiation is

$$\alpha = \arg\min |\|(\Pi^N - 1^N) \circ \mathrm{DCT}N(y^\delta)\|_F^2 - \delta^2|, \tag{5.3}$$

where δ is the Frobenius norm of the N-dimensional noise tensor.

5.3. *Gibbs phenomena in multidimensional numerical differentiation*

The Gibbs phenomena in the multidimensional numerical differentiation can be weakened similarly following the approaches in Section 4.

Concerning the j-th partial derivative of $\tilde{y}_\alpha^\delta(t^1, t^2, \ldots, t^N)$, the even expansion approach produces an expanded noisy sampling data along the j-th dimension with the expanded size $(n_1, n_2, \ldots, 3n_j - 2, \ldots, n_N)$. One then calls the regularized approach (5.2) and the parameter choice rule to obtain an expanded partial derivative where Gibbs phenomena appears on the expanded boundaries. The reconstructed j-th partial derivative of the original noisy sampling data can be obtained after cutting off the expanded intervals along the j-th dimension.

The zero derivative approach, similar to that in Section 4, produces firstly an N-dimensional tensor

$$\mathrm{sub}((t_{i_1}^1, t_{i_2}^2, \ldots, t_{i_N}^N)) = (y0' - y1')(t_{i_j}^j)^2/2 + y0' t_{i_j}^j.$$

where

$$y0' = (y^\delta(t_{i_1}^1, t_{i_2}^2, \ldots, t_2^j, \ldots t_{i_N}^N) - y^\delta(t_{i_1}^1, t_{i_2}^2, \ldots, t_1^j, \ldots t_{i_N}^N))(n_j - 1)$$

is a discrete approximated boundary derivative of $\frac{\partial y(t^1, t^2, \ldots, t^N)}{\partial t^j}$ at $t^j = 0$ and

$$y1' = (y^\delta(t_{i_1}^1, t_{i_2}^2, \ldots, t_{n_j-1}^j, \ldots t_{i_N}^N) - y^\delta(t_{i_1}^1, t_{i_1}^2, \ldots, t_{n_j-2}^j, \ldots t_{i_N}^N))(n_j - 1)$$

is another discrete approximated boundary derivative of $\frac{\partial y(t^1, t^2, \ldots, t^N)}{\partial t^j}$ at $t^j = 1$, respectively. N-th dimensional substitute sampling data $\{y\,\mathrm{sub}^\delta((t_{i_1}^1, t_{i_2}^2, \ldots, t_{i_N}^N))\}$ with the j-th partial derivative equalling zero on the boundaries are constructed by the following formula:

$$y\,\mathrm{sub}^\delta((t_{i_1}^1, t_{i_2}^2, \ldots, t_{i_N}^N)) = y^\delta((t_{i_1}^1, t_{i_2}^2, \ldots, t_{i_N}^N)) + \mathrm{sub}((t_{i_1}^1, t_{i_2}^2, \ldots, t_{i_N}^N)).$$

After implementing the regularization method on the substituted sampling data, one can obtain the j-th partial derivative $\frac{\partial \tilde{y}_\alpha^\delta(t^1, t^2, \ldots, t^N)}{\partial t^j}$ of the original noisy sampling data from the j-th partial derivative of $\widetilde{y\,\mathrm{sub}^\delta}(t^1, t^2, \ldots, t^N)$ by reducing $\mathrm{subde}((t^1, t^2, \ldots, t^N))$ with

$$\mathrm{subde}(t^1, t^2, \ldots, t^N) = (y0' - y1')t^j - y0'.$$

We finally use two numerical examples to show the efficiency of our proposed approach in the multidimensional numerical differentiation.

Example 5.1 The first example is adopted from [36] where the exact two-dimensional function $y(t, s) = \sin(\pi t)\sin(\pi s)\mathrm{e}^{-(t^2+s^2)}$ is discretized uniformly along a square domain $[-2, 2] \times [-2, 2]$.

In the first test, we take $n_1 = 101$ and $n_2 = 21$ with $\delta = 0.0469$. The noise is normally distributed with the same level as in [36] (Case 1). The regularization

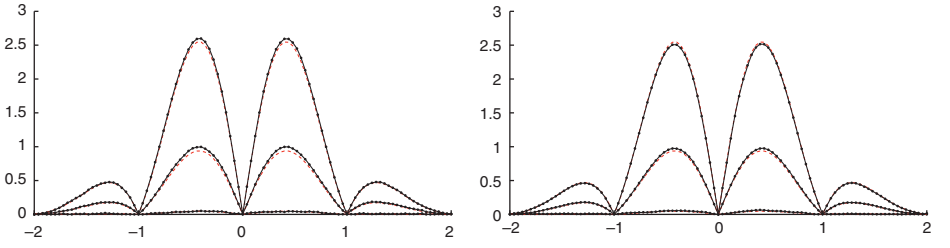

Figure 6. The curves of $|\nabla \tilde{y}_\alpha^\delta(t,s)(t,s)|$ (solid dotted line) vs. $|\nabla y(t,s)|$ (dashed line) at $s=0,1,2$, respectively.

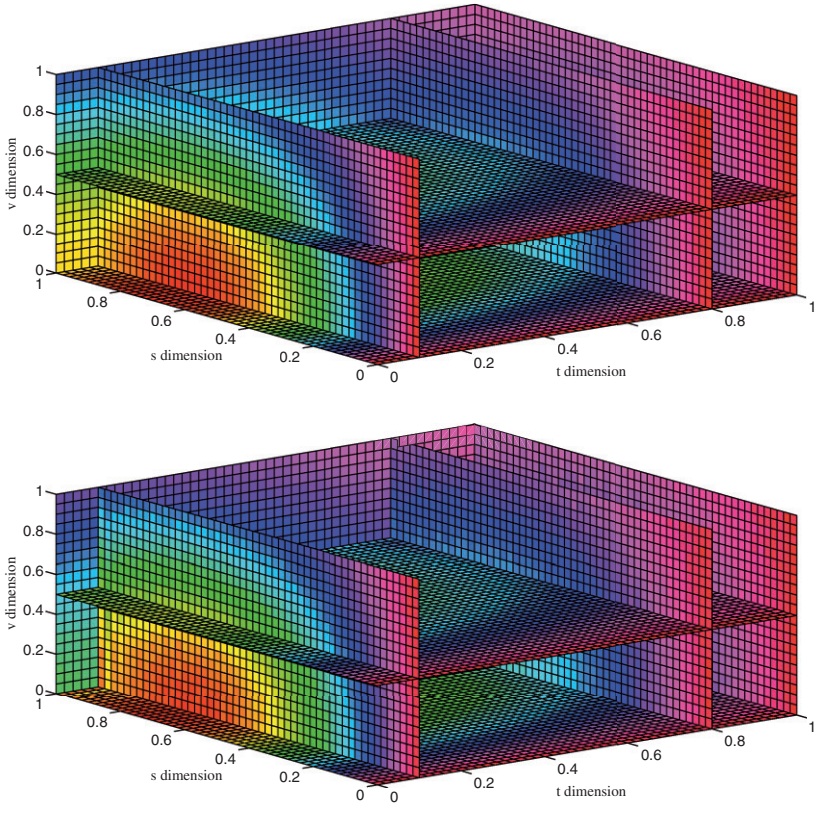

Figure 7. The upper graph is the exact partial derivative $\frac{\partial y(t,s,v)}{\partial s}$ and the lower graph is the reconstructed partial derivative $\frac{\partial \tilde{y}_\alpha^\delta(t,s,v)}{\partial s}$.

parameter is chosen by the DP (5.3). The reconstructed derivative semi-norm $|\nabla \tilde{y}_\alpha^\delta(t,s)|$ versus the exact derivative semi-norm $|\nabla y(t,s)|$ at $s=0,1,2$ is shown in the left graph of Figure 6 which is quite similar to those in [36], for instance the root mean square error is 0.0220 compared with 0.0157 there.

In the second test, we take a finer mesh where $n_1=101$ and $n_2=41$ with the same $\delta=0.0469$. The graph of $|\nabla \tilde{y}_\alpha^\delta(t,s)|$ versus $|\nabla y(t,s)|$ is shown in the right graph of Figure 6 where the root mean square error equals 0.0128. In both graphs, the reconstructed derivative approximates the exact one in a proper manner.

Example 5.2 This final example concerns the exact three-dimensional function $y(t,s,v) = \mathrm{e}^{-(t^2+s^2+v^2)}/2$ which is discretized uniformly along a cube domain $[0,1]\times[0,1]\times[0,1]$. We take each discretization level $n_t = 51$, $n_s = 51$, $n_v = 21$ and normally distributed noise with $\delta = 0.0158$. The regularization parameter α is chosen by the DP (23). The reconstructed second partial derivative $\frac{\partial y_\alpha^\delta(t,s,v)}{\partial s}$ versus the exact second partial derivative $\frac{\partial y(t,s,v)}{\partial s}$ is shown in Figure 7 where the relative error between them is 5% and the Gibbs phenomena is weakened by the even expansion approach.

Acknowledgements

We are grateful to anonymous referees for careful reading and helpful comments, as well as pointing out that ref. [16,19] that have improved the presentation of the manuscript essentially. This work is supported by the Shanghai Science and Technology Commission Grant: 11ZR1402800. The second author is a research fellow of the Alexander von Humboldt Foundation. He thanks Nan Chen (Fudan University) and Sergei V Pereverzev (RICAM) for careful reading of the manuscript and for many useful comments which improved the presentation.

References

[1] H. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
[2] C.W. Groetsch, *Differentiation of approximately specified functions*, Am. Math. Monthly 98 (1991), pp. 847–850.
[3] M. Hanke and O. Scherzer, *Inverse problems light numerical differentiation*, Am. Math. Monthly 108 (2001), pp. 512–521.
[4] L. Ling, *Finding numerical derivatives for unstructured and noisy data by multiscale kernels*, SIAM J. Numer. Anal. 44 (2006), pp. 1780–1800.
[5] S. Lu and S.V. Pereverzev, *Numerical differentiation from a viewpoint of regularization theory*, Math. Comput. 75 (2006), pp. 1853–1870.
[6] A.G. Ramm and A.B. Smirnova, *On stable numerical differentiation*, Math. Comput. 70 (2001), pp. 1131–1153.
[7] G. Wahba, *Spline Models for Observational Data*, SIAM, Philadelphia, 1990.
[8] Y.B. Wang, X.Z. Jia, and J. Cheng, *A numerical differentiation method and its application to reconstruction of discontinuity*, Inverse Probl. 18 (2002), pp. 1461–1476.
[9] T. Wei and Y.C. Hon, *Numerical differentiation by radial basis function approximation*, Adv. Comput. Math. 27 (2007), pp. 247–272.
[10] P. Mathe and S.V. Pereverzev, *The use of higher order finite difference schemes is not dangerous*, J. Complex. 25 (2009), pp. 3–10.
[11] P. Mathe and S.V. Pereverzev, *Geometry of linear ill-posed problems in variable Hilbert scales*, Inverse Probl. 19 (2003), pp. 789–803.
[12] E.T. Whittaker, *On a new method of graduation*, Proc. Edinb. Math. Soc. 41 (1923), pp. 62–75.
[13] H.L. Weinert, *Efficient computation for Whittaker-Henderson smoothing*, Comput. Stat. Data Anal. 52 (2007), pp. 959–974.
[14] D. Garcia, *Robust smoothing of gridded data in one and higher dimensions with missing values*, Comput. Stat. Data Anal. 54 (2010), pp. 1167–1178.
[15] G. Strang, *The discrete cosine transform*, SIAM Rev. 41 (1999), pp. 135–147.

[16] R.S. Anderssen and P. Bloomfield, *Numerical differentiation procedure for non-exact data*, Numer. Math. 22 (1974), pp. 157–182.
[17] P.C. Hanse, *Regularization tools: A Matlab package for analysis and solutions of discrete ill-posed problems*, Numer. Algorithms 6 (1994), pp. 1–35.
[18] W.C. Yueh, *Eigenvalues of several tridiagonal matrices*, Appl. Math. E-Notes 5 (2005), pp. 66–74.
[19] F. Bauer and M.A. Lukas, *Comparing parameter choice methods for regularization of ill-posed problems*, Math. Comput. Simul. 81 (2011), pp. 1795–1841.
[20] M. Hanke and P.C. Hansen, *Regularization methods for large-scale problems*, Surv. Math. Ind. 3 (1993), pp. 253–315.
[21] P. Craven and G. Wahba, *Smoothing noisy data with spline functions. Estimating the correct degree of smoothing by the method of generalized cross validation*, Numer. Math. 31 (1978), pp. 377–403.
[22] G.H. Golub, M. Heath, and H. Wahba, *Generalized cross-validation as a method for choosing a good ridge parameter*, Technometrics 21 (1979), pp. 215–223.
[23] M.J. Buckley, *Fast computation of a discretized thin-plate smoothing spline for image data*, Biometrika 81 (1994), pp. 247–258.
[24] K. Miller, *Least squares methods for ill-posed problems with a prescribed bound*, SIAM J. Math. Anal. 1970 (1970), pp. 52–74.
[25] P.C. Hansen, *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Rev. 34 (1992), pp. 561–580.
[26] P.C. Hansen and D.P. O'Leary, *The use of the L-curve in the regularization of discrete ill-posed problems*, SIAM J. Sci. Comput. 14 (1993), pp. 1487–1503.
[27] T. Regińska, *A regularization parameter in discrete ill-posed problems*, SIAM J. Sci. Comput. 17 (1996), pp. 740–749.
[28] V.A. Morozov, *On the solution of functional equations by the method of regularization*, Sov. Math. Dokl. 7 (1966), pp. 414–417.
[29] M. Hanke and T. Raus, *A general heuristic for choosing the regularization parameter in ill-posed problems*, SIAM J. Sci. Comput. 17 (1996), pp. 956–972.
[30] H. Gfrerer, *An a posteriori choice for ordinary and iterative Tikhonov regularization of ill-posed problems leading to optimal convergence rates*, Math. Comput. 49 (1987), pp. 507–522.
[31] U. Hämarik and U. Tautenhahn, *On the monotone error rule for parameter choice in iterative and continuous regularization methods*, BIT 41 (2001), pp. 1029–1038.
[32] D. Gottlieb and C.W. Shu, *On the Gibbs phenomenon and its resolution*, SIAM Rev. 39 (1997), pp. 644–668.
[33] R.A. Bernatz, *Fourier Series and Numerical Methods for Partial Differential Equations*, John Wiley & Sons, Hoboken, NJ, 2010.
[34] R. Qu, *A new approach to numerical differentiation and integration*, Math. Comput. Model. 24 (1996), pp. 55–68.
[35] M.J. Narasimha and A.M. Peterson, *On the computation of the discrete cosine transform*, IEEE Trans. Commun. 26 (1978), pp. 934–936.
[36] T. Wei, Y.C. Hon, and Y.B. Wang, *Reconstruction of numerical derivatives from scattered noisy data*, Inverse Probl. 21 (2005), pp. 657–672.