

The Generalized Tikhonov Regularization Method for High Order Numerical Derivatives

F. Yang¹, C.L. Fu² and X.X. Li¹

Abstract: Numerical differentiation is a classical ill-posed problem. The generalized Tikhonov regularization method is proposed to solve this problem. The error estimates are obtained for *a priori* and *a posteriori* parameter choice rules, respectively. Numerical examples are presented to illustrate the validity and effectiveness of this method.

Keywords: Numerical differentiation, ill-posed problem, generalized Tikhonov regularization, *A posteriori* parameter choice, Error estimate.

1 Introduction

Numerical differentiation is a classical ill-posed problem in the sense that arbitrarily "small" differences in the input data can result in arbitrarily "large" errors in the approximate derivatives [Kirsch (1996)], and it is very important in science research and practical application. For example, the problems in image process [Deans (1983)], solving Volterra integral equation [Cheng, Hon and Wang (2004); Gorenflo and Vessella (1991)] and identification [Hanke and Scherzer (1999):] had been focused on gaining the numerical differentiation. A number of effective methods have been appeared in the past years: difference methods [Groetsch (1991); Qu (1996); Ramm and Smirnova (2001); Anderssen and Hegland (1999)], Tikhonov regularization methods [Cullum (1971); Hanke and Scherzer (2001); Wang, Jia and Cheng (2002)], mollification methods [Háo (1994); Murio, MejÁa and Zhan (1998)], Wavelets method [Fu, Feng, and Qian (2010); Dou, Fu and Ma (2010)], Fourier method [Qian, Fu, Xiong and Wei (2006):] and quasi-reversibility method [Qian, Fu and Feng (2006)].

The major innovation of the present paper is to give *a posteriori* parameter choice rule. In [Dou, Fu and Ma (2010); Fu, Feng and Qian (2010); Qian, Fu, Xiong and Wei (2006); Qian, Fu and Feng (2006)], the regularization parameters which

¹ School of Science, Lanzhou University of Technology, Lanzhou, Gansu, China.

² School of Mathematics and Statistics, Lanzhou University, Lanzhou, Gansu, China

depend on the noise level and the *a priori* bound are selected by the *a priori* rule. But there is a defect for any *a priori* method, i.e., the *a priori* choice of the regularization parameter depends seriously on the *a priori* bound E of the unknown solution. However, the *a priori* bound E cannot be known exactly in practice, and working with a wrong constant E may lead to the bad regularized solution. We will consider not just the *a priori* choice of the regularization parameter for the generalized Tikhonov regularization method, but also the *a posteriori* choice of the regularization parameter will be given. Under these parameter choice rules, we obtain the Hölder type error estimates which are order optimal.

The outline of the paper is as follows. In Section 2, we give the method to construct approximation solution. The convergence results are given in Section 3. In Section 4, some numerical examples are proposed to show the effectiveness for this method.

2 Ill-posedness of problem and regularization

In this section, we analyze the ill-posedness of numerical differentiation and discuss how to stabilize the numerical derivatives. We consider the function $f(x) \in L^2(\mathbb{R})$. Let \widehat{f} be the Fourier transform of f , i.e.,

$$\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx. \quad (1)$$

Now we consider the k th order derivative of function f , where $k = 1, 2, \dots$. We denote $f^{(k)}(x) = \frac{d^k f(x)}{dx^k}$. Taking the Fourier transform, we have

$$\widehat{f^{(k)}}(\xi) = (i\xi)^k \widehat{f}(\xi), \quad (2)$$

or, equivalently,

$$f^{(k)}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (i\xi)^k \widehat{f}(\xi) e^{i\xi x} d\xi. \quad (3)$$

Note that $f(x) \in H^k(\mathbb{R})$, the above expression is significant. But in practice, in general, we do not know the exact data $f(x)$, instead of it, only a measured data $f_\delta(x)$ which merely belongs to $L^2(\mathbb{R})$ satisfies

$$\|f_\delta - f\| \leq \delta, \quad (4)$$

where $\|\cdot\|$ denotes L^2 -norm, the constant $\delta > 0$ represents a noise level. So, generally, the expression (3) for $f_\delta(x)$ has no sense and some regularization methods are necessary for the numerical derivative. However, before doing that, we impose an

a priori bound on the input data (this is necessary in solving ill-posed problems), i.e.,

$$\|f\|_{H^p} \leq E, \quad p \geq 0, \quad (5)$$

where $E > 0$ is a constant, $\|\cdot\|_{H^p}$ denotes the norm in Sobolev space $H^p(\mathbb{R})$ defined by

$$\|f(\cdot)\|_{H^p} := \left(\int_{-\infty}^{\infty} (1 + \xi^2)^p |\widehat{f}(\xi)|^2 d\xi \right)^{1/2}. \quad (6)$$

From (3), we can obtain

$$f^{(k)}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (i\xi)^k \widehat{f}(\xi) e^{i\xi x} d\xi := Lf. \quad (7)$$

Note that, $|(i\xi)^k| = |\xi|^k$. Therefore, when we consider our problem in $L^2(\mathbb{R})$, the exact data function $\widehat{f}(\xi)$, must decay rapidly as $\xi \rightarrow \infty$. Such a decay is not likely to occur in the Fourier transform of measured noisy temperature history $f_\delta(x)$. In the following, we apply the generalized Tikhonov regularization method to reconstruct a new function $f_{\mu,\delta}(x)$ from the perturbed data $f_\delta(x)$ which minimizes the quantity

$$\Phi(f) = \|f - f_\delta\|^2 + \mu \|Lf\|_{H^p}^2, \quad (8)$$

where μ is the regularization parameter. It can be verified that $\widehat{f}_{\mu,\delta}(\xi)$ is the solution of the following equation [Kirsch (1996)]:

$$(I + \mu(1 + \xi^2)^p \widehat{L}^* \widehat{L}) \widehat{f}_{\mu,\delta} = \widehat{f}_\delta. \quad (9)$$

From (7), we obtain,

$$\widehat{L} = (i\xi)^k, \quad \widehat{L}^* = \overline{(i\xi)^k}. \quad (10)$$

Due to (9) and (10), we obtain

$$\widehat{f}_{\mu,\delta}(\xi) = \frac{1}{1 + \mu(1 + \xi^2)^p |\xi|^{2k}} \widehat{f}_\delta(\xi). \quad (11)$$

Then the approximation solution can be given as

$$\widehat{f}_{\mu,\delta}^{(k)}(\xi) = (i\xi)^k \frac{1}{1 + \mu(1 + \xi^2)^p |\xi|^{2k}} \widehat{f}_\delta(\xi). \quad (12)$$

So

$$f_{\mu,\delta}^{(k)}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\xi x} (i\xi)^k \frac{1}{1 + \mu(1 + \xi^2)^p |\xi|^{2k}} \widehat{f}_\delta(\xi) d\xi = Lf_{\mu,\delta}. \quad (13)$$

3 Choice of regularization parameter and convergence results

In this section, we consider an *a priori* strategy and *a posteriori* choice rule to find the regularization parameter. Under each choice of the regularization parameter, convergence estimates can be obtained. We first give some useful Lemmas.

Lemma 3.1. *If $\xi \in \mathbb{R}$, $p > k$, the following inequality holds:*

$$\frac{|\xi|^{2p}}{(1 + \xi^2)^p |\xi|^{2k}} \leq 1. \quad (14)$$

Proof. As $|\xi| \geq 1$, we obtain

$$\frac{|\xi|^{2p}}{(1 + \xi^2)^p |\xi|^{2k}} \leq \frac{1}{|\xi|^{2k}} \leq 1. \quad (15)$$

As $|\xi| \leq 1$, we obtain

$$\frac{|\xi|^{2p}}{(1 + \xi^2)^p |\xi|^{2k}} \leq |\xi|^{2(p-k)} \leq 1. \quad (16)$$

Combining (15) with (16), we obtain

$$\frac{|\xi|^{2p}}{(1 + \xi^2)^p |\xi|^{2k}} \leq 1. \quad (17)$$

Lemma 3.2. *If $\|Lf\|_{H^p} \leq E$, then we have*

$$\|Lf\| \leq \|f\|^{\frac{p-k}{p}} \|Lf\|_{H^p}^{\frac{k}{p}}. \quad (18)$$

Proof. Using the Hölder inequality and (14), we obtain

$$\begin{aligned} \|Lf\|^2 &= \|\widehat{Lf}\|^2 = \int_{-\infty}^{\infty} |(i\xi)^k \widehat{f}(\xi)|^2 d\xi = \int_{-\infty}^{\infty} |\xi|^{2k} |\widehat{f}(\xi)|^{\frac{2k}{p}} |\widehat{f}(\xi)|^{\frac{2(p-k)}{p}} d\xi \\ &\leq \left[\int_{-\infty}^{\infty} (|\xi|^{2k} |\widehat{f}(\xi)|^{\frac{2k}{p}})^{\frac{p}{k}} d\xi \right]^{\frac{k}{p}} \left[\int_{-\infty}^{\infty} (|\widehat{f}(\xi)|^{\frac{2(p-k)}{p}})^{\frac{p}{p-k}} d\xi \right]^{\frac{p-k}{p}} \\ &= \left[\int_{-\infty}^{\infty} |\xi|^{2p} |\widehat{f}(\xi)|^2 d\xi \right]^{\frac{k}{p}} \left[\int_{-\infty}^{\infty} |\widehat{f}(\xi)|^2 d\xi \right]^{\frac{p-k}{p}} \\ &= \|\widehat{f}\|^{\frac{2(p-k)}{p}} \left[\int_{-\infty}^{\infty} |\xi|^{2p} (1 + \xi^2)^{-p} (1 + \xi^2)^p |\xi|^{-2k} |\xi|^{2k} |\widehat{f}(\xi)|^2 d\xi \right]^{\frac{k}{p}} \\ &= \|f\|^{\frac{2(p-k)}{p}} \left[\int_{-\infty}^{\infty} |\xi|^{2p} (1 + \xi^2)^{-p} |\xi|^{-2k} (1 + \xi^2)^p |\widehat{Lf}|^2 d\xi \right]^{\frac{k}{p}} \\ &\leq \|f\|^{\frac{2(p-k)}{p}} \sup_{\xi \in \mathbb{R}} \frac{|\xi|^{2p}}{(1 + \xi^2)^p |\xi|^{2k}})^{\frac{k}{p}} \|Lf\|_{H^p}^{\frac{2k}{p}} \\ &\leq \|f\|^{\frac{2(p-k)}{p}} \|Lf\|_{H^p}^{\frac{2k}{p}}. \end{aligned}$$

So

$$\|Lf\| \leq \|f\|^{\frac{p-k}{p}} \|Lf\|_{HP}^{\frac{k}{p}}. \quad (19)$$

3.1 The a priori choice rule

We assume that we have obtained an E in (5), then we have Theorem 3.3.

Theorem 3.3. *Let $f^k(x)$ given by (7) be the exact solution and $f_{\mu,\delta}^k(x)$ given by (13) be the general Tikhonov regularized approximation of $f^k(x)$. Let $f_\delta(x)$ be measured data satisfying (4) and the a priori condition (5) holds. If we select*

$$\mu = \left(\frac{\delta}{E}\right)^2, \quad (20)$$

then there holds the following error estimate:

$$\|f^{(k)}(\cdot) - f_{\mu,\delta}^{(k)}(\cdot)\| \leq (\sqrt{2} + 1)\delta^{\frac{p-k}{p}} E^{\frac{k}{p}}. \quad (21)$$

Proof: For $f_{\mu,\delta}(x)$ is the minimizer of (8), we obtain

$$\|f_{\mu,\delta} - f_\delta\|^2 \leq \Phi(f_{\mu,\delta}) \leq \Phi(f) = \|f - f_\delta\|^2 + \mu \|Lf\|_{HP}^2 \leq 2\delta^2 \quad (22)$$

and

$$\|Lf_{\mu,\delta}\|_{HP}^2 \leq \frac{1}{\mu} \Phi(f_{\mu,\delta}) \leq \frac{1}{\mu} \Phi(f) = \frac{1}{\mu} \|f - f_\delta\|^2 + \|Lf\|_{HP}^2 \leq 2E^2. \quad (23)$$

Due to (4) and (22), we obtain

$$\|f - f_{\mu,\delta}\| \leq \|f - f_\delta\| + \|f_\delta - f_{\mu,\delta}\| \leq (\sqrt{2} + 1)\delta. \quad (24)$$

Due to (5) and (23), we obtain

$$\|L(f - f_{\mu,\delta})\|_{HP} \leq \|Lf\|_{HP} + \|Lf_{\mu,\delta}\|_{HP} \leq E + \sqrt{2}E = (\sqrt{2} + 1)E. \quad (25)$$

Due to Lemma 3.2, we obtain,

$$\begin{aligned} \|f^{(k)} - f_{\mu,\delta}^{(k)}\| &\leq \|Lf - Lf_{\mu,\delta}\| \leq \|f - f_{\mu,\delta}\|^{\frac{p-k}{p}} \|L(f - f_{\mu,\delta})\|_{HP}^{\frac{k}{p}} \\ &\leq (\sqrt{2} + 1)\delta^{\frac{p-k}{p}} E^{\frac{k}{p}}. \end{aligned}$$

3.2 The a posteriori choice rule

Morozov's discrepancy principle is used as a *a posteriori* rule, i.e. choosing $\mu = \mu_2$ as the solution of the following equation:

$$\|f_{\mu,\delta} - f_\delta\| = \delta. \quad (26)$$

To establish existence and uniqueness of solution for equation (26), we need the following lemma:

Lemma 3.4. Let $\rho(\mu) := \|f_{\mu,\delta} - f_\delta\|$, then for $\delta < \|f_\delta\|$, there hold

(a) $\rho(\mu)$ is a continuous function;

(b) $\lim_{\mu \rightarrow 0^+} \rho(\mu) = 0$;

(c) $\lim_{\mu \rightarrow +\infty} \rho(\mu) = \|f_\delta\|$;

(d) $\rho(\mu)$ is a strictly increasing function.

The proof is very easy and we omit it here.

Theorem 3.5. Assume the conditions (4), (5) hold and take the solution μ of Eq. (26) as the regularization parameter, then there holds the following error estimate:

$$\|f^{(k)} - f_{\mu,\delta}^{(k)}\| \leq 2\delta^{\frac{p-k}{p}} E^{\frac{k}{p}}. \quad (27)$$

Proof: Since $f_{\mu_2,\delta}(x)$ is the minimizer of (8), we obtain

$$\|f_{\mu_2,\delta} - f_\delta\|^2 + \mu_2 \|Lf_{\mu_2,\delta}\|_{H^p}^2 = \Phi(f_{\mu_2,\delta}) \leq \Phi(f) = \|f - f_\delta\|^2 + \mu_2 \|Lf\|_{H^p}^2. \quad (28)$$

Due to (26) and (28), we obtain

$$\|Lf_{\mu_2,\delta}\|_{H^p}^2 \leq \|Lf\|_{H^p}^2 + \frac{1}{\mu_2} (\|f - f_\delta\|^2 - \delta^2) \leq \|Lf\|_{H^p}^2 \leq E^2. \quad (29)$$

So

$$\|Lf_{\mu_2,\delta} - Lf\|_{H^p} \leq \|Lf_{\mu_2,\delta}\|_{H^p} + \|Lf\|_{H^p} \leq 2E. \quad (30)$$

Due to (26) and (4), we obtain

$$\|f_{\mu_2,\delta} - f\| \leq \|f_{\mu_2,\delta} - f_\delta\| + \|f_\delta - f\| \leq 2\delta. \quad (31)$$

Due to Lemma 3.2, (30) and (31), we obtain

$$\begin{aligned} \|f^{(k)} - f_{\mu_2,\delta}^{(k)}\| &= \|Lf - Lf_{\mu_2,\delta}\| \leq \|f - f_{\mu_2,\delta}\|^{\frac{p-k}{p}} \|L(f - f_{\mu_2,\delta})\|_{H^p}^{\frac{k}{p}} \\ &\leq 2\delta^{\frac{p-k}{p}} E^{\frac{k}{p}}. \end{aligned}$$

The proof of Theorem 3.5. is complete.

4 Several numerical examples

In this section, we will test three kinds of functions to verify the effect of the proposed algorithm. Moreover, we would like to compare numerical results of the *a posteriori* parameter choice (26) with one of the *a priori* parameter choice rule (20). The bisection method is used to solve the Eq. (26). In the following experiments, we consider the numerical derivatives only in the finite interval $x \in [0, 1]$.

Suppose that the sequence $\{f_j\}_{j=0}^n$ represents samples from the function $f(x)$ on an equidistant grid, then we add a random uniform perturbation to each data, which forms the vector f_δ , i.e.,

$$f_\delta = f + \varepsilon \text{randn}(\text{size}(f)), \tag{32}$$

where the function “*randn*(\cdot)” generates arrays of random whose elements are normally distributed with mean 0, variance $\sigma^2 = 1$. “*Randn*(*size*(*g*))” returns an array of random entries that is of the same size as *g*. The total noise level δ can be measured in the sense of Root Mean Square Error(RMSE) according to

$$\delta = \|f_\delta - f\|_{l^2} = \left(\frac{1}{n} \sum_{i=1}^n (f_i - f_{i,\delta})^2\right)^{\frac{1}{2}}. \tag{33}$$

Example 1. First we consider the function

$$f(x) = \exp\left(2 - \frac{1}{x(1-x)}\right). \tag{34}$$

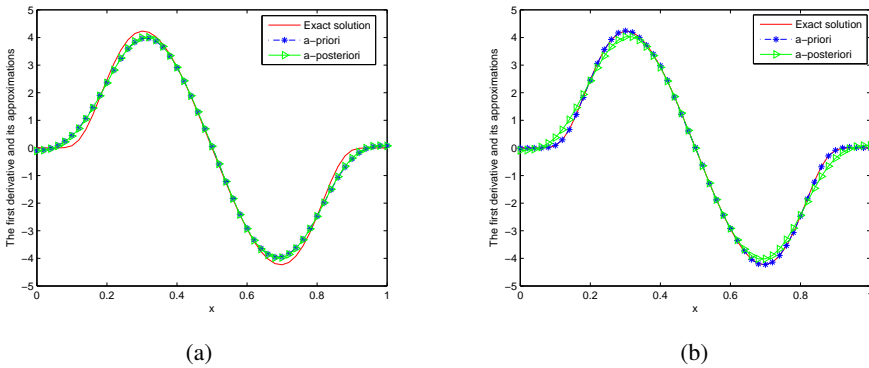
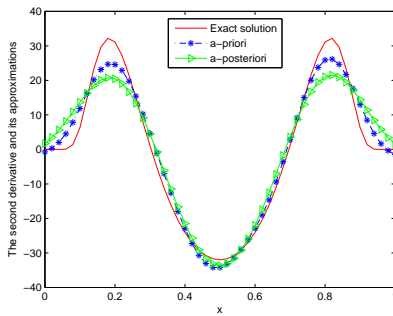
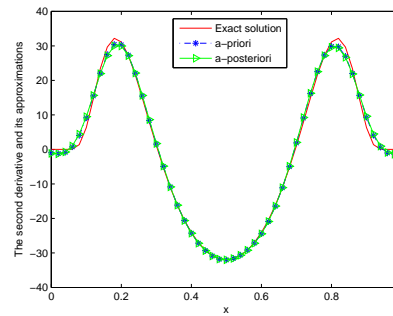


Figure 1: The comparison of the numerical effects between the first derivative solution and its computed approximations with Example 1: (a) $\varepsilon = 0.01$, (b) $\varepsilon = 0.001$.

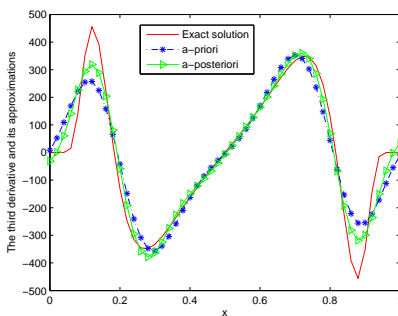


(a)

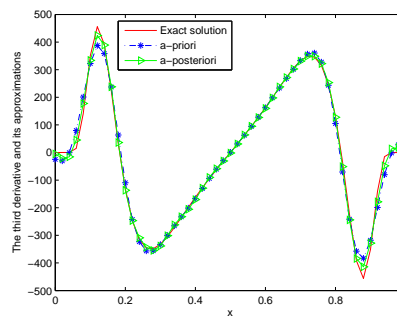


(b)

Figure 2: The comparison of the numerical effects between the second derivative solution and its computed approximations with Example 1: (a) $\varepsilon = 0.01$, (b) $\varepsilon = 0.001$.



(a)



(b)

Figure 3: The comparison of the numerical effects between the third derivative solution and its computed approximations with Example 1: (a) $\varepsilon = 0.01$, (b) $\varepsilon = 0.001$.

It is clear that the above function $f \in H^p(\mathbb{R}), p \geq 0$, and this fact is responsible for good numerical results.

Example 2. Consider the function

$$f(x) = \sin(10\pi x). \tag{35}$$

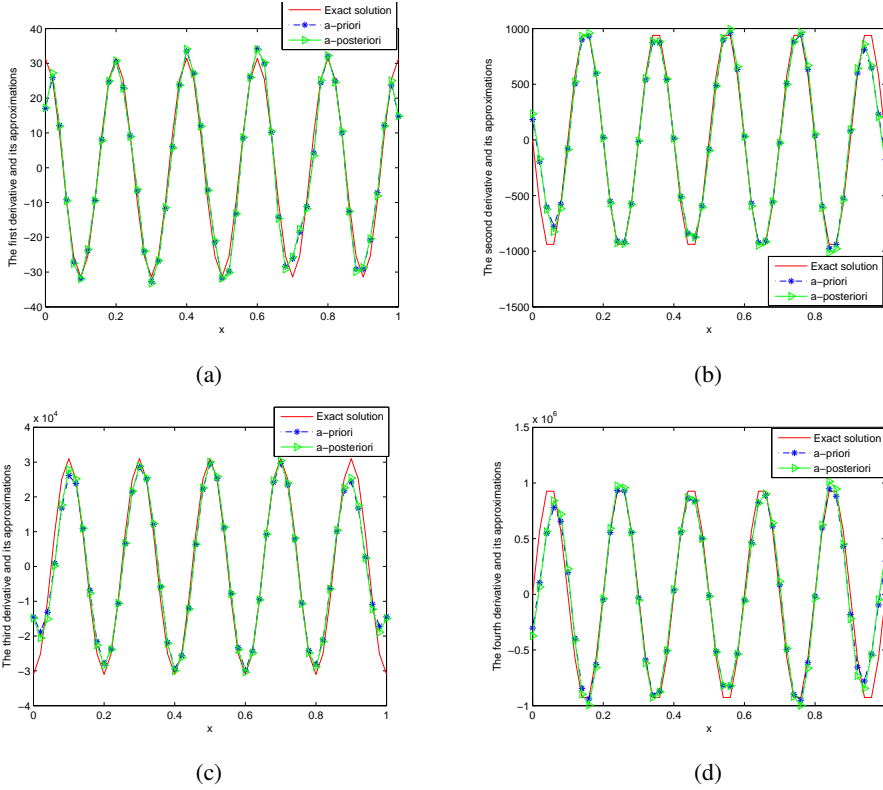


Figure 4: The comparison of the numerical effects between the first, second, third and fourth derivatives and their computed approximations with Example 2: $\varepsilon = 0.1$: (a) the first derivative and its approximations. (b) the second derivative and its approximations. (c) the third derivative and its approximations. (d) the fourth derivative and its approximations.

Example 3. Consider a non-smooth function

$$f(x) = \begin{cases} 0, & 0 \leq x \leq \frac{1}{5}, \\ 2x^2 - \frac{4}{5}x + \frac{2}{25}, & \frac{1}{5} < x \leq \frac{1}{2}, \\ -2x^2 + \frac{16}{5}x - \frac{23}{25}, & \frac{1}{2} < x \leq \frac{4}{5}, \\ \frac{9}{25}, & \frac{4}{5} < x \leq 1. \end{cases} \quad (36)$$

From Figs. 1-5, we firstly find that the smaller ε , the better the computed approximation is, and the smaller k is, the better the computed approximation is. There are consistent with our theoretical analysis. Moreover, we can also easily find that the *a posteriori* parameter choice rule works better than the *a priori* parameter choice

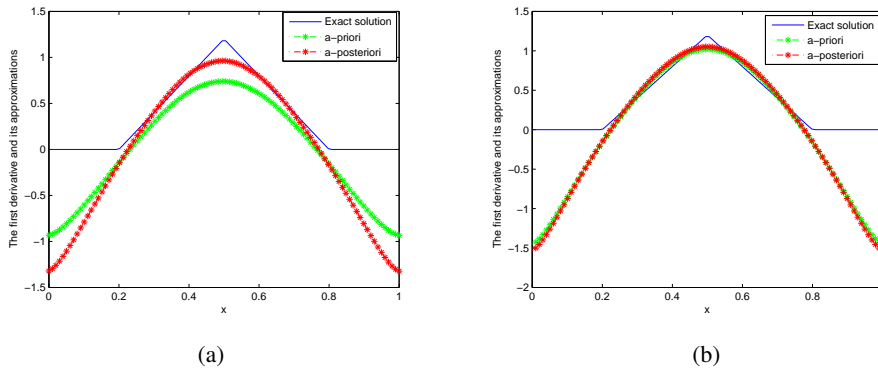


Figure 5: The comparison of the numerical effects between the first derivative solution and its computed approximations with Example 3: (a) $\varepsilon = 0.01$, (b) $\varepsilon = 0.001$.

rule. Finally, From Figs. 1-5, it can be seen that the numerical solutions of Example 3 are less ideal than these of Examples 1 and 2. It is not difficult to see that the well-known Gibbs phenomenon and the recovered data near the discontinuities points are not accurate.

Acknowledgement: The project is supported by the National Natural Science Foundation of China (Nos. 11171136 and 11261032), the Distinguished Young Scholars Fund of Lan Zhou University of Technology (Q201015), the basic scientific research business expenses of Gansu Province College and the Natural Science Foundation of Gansu province(1310RJYA021).

References

- Anderssen, R. S.; Hegland, M.**(1999): For numerical differentiation, dimensionality can be a blessing. *Math. Comput.*, vol.68, no.227, pp.1121-1141.
- Cheng, J.; Hon, Y. C.; Wang, Y. B.**(2004): A numerical method for the discontinuous solutions of Abel integral equations. In inverse problems and spectral theory. vol. 348 of *Contemp. Math.*, pp. 233-243. *Amer. Math. Soc., Providence, RI*.
- Cullum, J.** (1971): Numerical differentiation and regularization. *SIAM J. Numer. Anal.*, vol.8, pp.254-265.
- Deans, S.R.**(1983): *The Radon transform and some of its applications*. A Wiley-Interscience Publication. John Wiley & Sons Inc., New York.
- Dou, F. F.; Fu, C. L.; Ma, Y. J.** (2010): A Wavelet-Galerkin method for high order numerical differentiation, *Appl. Math. Comput.*, vol.215, pp.3702-3712.

- Fu, C.L.; Feng, X.L.; Qian, Z.** (2010) Wavelets and high order numerical differentiation. *Appl. Math. Model.*, vol.34, pp.3008-3021.
- Gorenflo, R.; Vessella, S.**(1991): *Abel integral equations, volume 1461 of Lecture Notes in Mathematics.* Springer-Verlag, Berlin.
- Groetsch, C. W.** (1991): Differentiation of approximately specified functions. *Amer. Math. Monthly*, vol.98, no.9, pp.847-850.
- Hanke, M.; Scherzer, O.**(1999): Error analysis of an equation error method for the identification of the diffusion coefficient in a quasi-linear parabolic differential equation, *SIAM J. Appl. Math.*, vol.59, pp.1012-1027. (electronic)
- Hanke, M.; Scherzer, O.**(2001): Inverse problems light: numerical differentiation. *Amer. Math. Monthly*, vol.108, no.6, pp.512-521.
- Hào, D. N.** (1994): A mollification method for ill-posed problems. *Numer. Math.*, vol.68, pp.469-506.
- Kirsch, A.**(1996): *An introduction to the mathematical theory of inverse problems.* Springer-Verlag, Berlin.
- Murio, D. A.; Mejía, C. E.; Zhan, S.** (1998): Discrete mollification and automatic numerical differentiation. *Computers Math. Appl.*, vol.35, pp.1-16.
- Qian, Z.; Fu, C. L.; Xiong, X. T.; Wei, T.** (2006): Fourier truncation method for high order numerical derivatives, *Appl. Math. Comput.*, vol.181, pp.940-948.
- Qian, Z.; Fu, C. L.; Feng, X. L.**(2006): A modified method for high order numerical derivatives. *Appl. Math. Comput.*, vol.182, pp.1191-1200.
- Qu, R.**(1996): A new approach to numerical differentiation and integration, *Math. Comput. Model.*, vol.24, no.10, pp.55-68.
- Ramm, A. G.; Smirnova, A. B.** (2001): On stable numerical differentiation. *Math. Comput.*, vol.70, pp.1131-1153. (electronic)
- Wang, Y. B.; Jia, X. Z.; Cheng, J.**(2002), A numerical differentiation method and its application to reconstruction of discontinuity. *Inverse Problems*, vol.18, pp.1461-1476.

