

# Integral equation models for image restoration: high accuracy methods and fast algorithms

Yao Lu<sup>1</sup>, Lixin Shen<sup>2</sup> and Yuesheng Xu<sup>2,3,4</sup>

<sup>1</sup> Department of Radiology, University of Michigan, Ann Arbor, MI 48109, USA

<sup>2</sup> Department of Mathematics, Syracuse University, Syracuse, NY 13244, USA

<sup>3</sup> School of Mathematics and Computational Sciences, Sun Yat-sen University, Guangzhou 510275, People's Republic of China

E-mail: [yaol@med.umich.edu](mailto:yaol@med.umich.edu), [lshen03@syr.edu](mailto:lshen03@syr.edu) and [yxu06@syr.edu](mailto:yxu06@syr.edu)

Received 11 August 2009, in final form 17 February 2010

Published 18 March 2010

Online at [stacks.iop.org/IP/26/045006](http://stacks.iop.org/IP/26/045006)

## Abstract

Discrete models are consistently used as practical models for image restoration. They are piecewise constant approximations of true physical (continuous) models, and hence, inevitably impose bottleneck model errors. We propose to work directly with continuous models for image restoration aiming at suppressing the model errors caused by the discrete models. A systematic study is conducted in this paper for the continuous out-of-focus image models which can be formulated as an integral equation of the first kind. The resulting integral equation is regularized by the Lavrentiev method and the Tikhonov method. We develop fast multiscale algorithms having high accuracy to solve the regularized integral equations of the second kind. Numerical experiments show that the methods based on the continuous model perform much better than those based on discrete models, in terms of PSNR values and visual quality of the reconstructed images.

## 1. Introduction

Integral equation models are an important mathematical tool for image processing. However, these models are usually not used directly in the image processing community due to difficulties in dealing with the integral equations. Instead, discrete (matrix) models which are certain piecewise constant approximations of the integral equation models are usually used due to their convenience in implementation and consistence with usual sampling methods. The use of the piecewise constant approximation imposes a bottleneck model error which cannot be compensated by any image processing method. For example, for high resolution image reconstruction, the model of Bose and Boo [4] was obtained by using the piecewise constant

<sup>4</sup> Author to whom any correspondence should be addressed.

approximation of the integrand of the integral that defines the continuous model. The error contributed by the piecewise constant approximation in this model will never be overcome no matter what methods we use to recover the original image. Further, the discrete models make the realization of adaptive approximation very difficult if it is not impossible. In comparison to the ideal integral equation models, the matrix models give much less theoretical insights.

Most of the discrete models have their continuous counterparts. In such a case, we propose to use directly the continuous model instead of an existing discrete model. Here, we list several advantages of using continuous integral equation models. First of all, continuous models are derived directly from physical laws; therefore, they are physically meaningful. Second, continuous models allow us to discretize them with higher order accuracy, and therefore the resulting model error in comparison to the model error associated with piecewise constant discretization can be significantly reduced. Finally, continuous models are necessary, particularly for medical imaging processing, in the development of adaptive mesh approaches for solving the related equations [5, 29].

Direct use of an integral equation model brings new challenges. The integral equation model requires availability of a continuous function which represents the observed image. However, the only available data for the observed image are discrete. How do we obtain the continuous function from the observed discrete image? This is the first challenge which we will face when we use the integral equation model. The study of this problem itself is an active and important research area related to sampling [28] and reconstruction of signals [21]. The second challenge comes from the ill-posedness of the integral equation related to the continuous model. Since in image processing kernels of integral operators are normally smooth, the integral operators are compact in appropriate Banach spaces. Thus, the solutions of the related first kind integral equations are not continuously dependent on the given data, that is, a small perturbation in the data may result in a large perturbation in the solutions. How we can find an appropriate regularization method to turn the ill-posed integral equation to a well-posed equation becomes a challenging issue. The third challenging issue is how we efficiently solve the resulting regularization equation. The new discrete equation normally has a dense coefficient matrix of a large size. We postpone the discussion on the first issue to our future work. In this paper, we mainly focus on the third challenge and at the same time address the second issue to a certain extent. In particular, for the out-of-focus model [13] we use the Lavrentiev regularization and the Tikhonov regularization principles to convert its ill-posed integral equation of the first kind into a well-posed integral equation of the second kind and develop fast solution methods with high accuracy for solving the resulting integral equation of the second kind.

This paper is organized in nine sections with an appendix. In section 2, we describe the integral equation model for out-of-focus image restoration in terms of the Gaussian kernel. The model is expressed in the form of an integral equation of the first kind, which is an ill-posed problem due to the compactness of the corresponding integral operator. To overcome the ill-posedness, we apply the Lavrentiev regularization and introduce a modified Tikhonov regularization for the ill-posed integral equation. Section 3 is devoted to a description of multiscale Galerkin methods for solving the regularized integral equations. We use the piecewise polynomial spaces and the Galerkin principle to discretize the regularized integral equation of the second kind. This allows us to develop high accuracy numerical methods by choosing piecewise polynomials of high orders for solving the image restoration problem. The discretization of the integral equation leads to a linear system with a dense coefficient matrix. A truncation strategy is derived in section 4 (based on properties of the Gaussian kernel) to approximate the dense coefficient matrix by a sparse matrix. We provide in section 5 the convergence analysis for the proposed method. The nonzero entries of the truncated matrix are

integrals involved with the Gaussian kernel. Because high-order derivatives of the Gaussian kernel grow rapidly, numerical integration of such integrals requires careful treatment. We develop in section 6 a numerical quadrature strategy for computing the nonzero entries of the compressed coefficient matrix. In section 7, we show that the quadrature strategy requires only linear (up to a logarithmical factor) number of functional evaluations for computing all nonzero entries of the matrix and preserves the optimal order of convergence for the resulting approximate solution. Numerical examples are presented in section 8 to demonstrate the performance of the proposed algorithm. The numerical performance of the proposed algorithm and the visual quality of the recovered images obtained with it are compared with those of the discrete methods including the discrete Lavrentiev regularization, the discrete Tikhonov regularization and the discrete TV regularization methods. Both numerical and image results show that the proposed method outperforms the well-known discrete methods mentioned above. We draw conclusive remarks in section 9. In the appendix, we include piecewise linear, quadratic and cubic multiscale functions for convenient reference.

## 2. Integral equations for image restoration

We describe in this section integral equation models for image restoration. In particular, we discuss the regularization of the integral equation which governs out-of-focus image models.

Many image processing problems are modeled via integral equations. They can be formulated as follows:

$$\mathbb{K}v = f, \quad (2.1)$$

where  $f$  is an observed image with noise and  $v : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  represents the original image. In image processing,  $\Omega$  is usually chosen as a rectangle domain. In this paper, we will assume that  $\Omega = E \times E$  with  $E := [0, 1]$ . In the context of image analysis,  $\mathbb{K}$  is called a *blurring* operator. Typically, it is a Fredholm integral operator defined in terms of a kernel  $k$  by

$$(\mathbb{K}v)(\mathbf{x}) := \int_{\Omega} k(\mathbf{x}, \mathbf{x}')v(\mathbf{x}') d\mathbf{x}', \quad \mathbf{x} \in \Omega. \quad (2.2)$$

The choice of the kernel depends on the specific application context. We list below a few examples of equation (2.1) with specific kernels. In remote sensing and astronomical imaging, the atmospheric turbulence blur (see [15]) is modeled by (2.1) with the kernel

$$k(\mathbf{x}, \mathbf{x}') = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{1}{2}\left(\frac{x-x'}{\sigma_1}\right)^2 - \frac{1}{2}\left(\frac{y-y'}{\sigma_2}\right)^2\right),$$

where  $\sigma_1, \sigma_2$  are model parameters characterizing the degree of blurring in the  $x$  and  $y$  directions, respectively, and where  $\mathbf{x} = (x, y)$  and  $\mathbf{x}' = (x', y')$ . In medical image processing, confocal microscopy [2] is modeled by (2.1) with the kernel

$$k(\mathbf{x}, \mathbf{x}') = \frac{\sin(\pi w(x-x'))}{\pi w(x-x')} \cdot \frac{\sin(\pi w(y-y'))}{\pi w(y-y')},$$

where the model parameter  $w$  represents the width of an aperture of the collector lens in confocal microscopy. An out-of-focus image is usually modeled (cf [13]) by (2.1) with the kernel

$$k(\mathbf{x}, \mathbf{x}') = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x-x')^2 + (y-y')^2}{2\sigma^2}\right), \quad (2.3)$$

where  $\sigma$  is the model parameter characterizing the degree of accuracy and clearness of the image in the system. In this paper, we will focus on the development of basic concepts of

using integral equation models in the numerical process of image restoration. For this reason, we will use the out-of-focus image model as an example. Image restoration based on other models can be similarly developed.

Out-of-focus image restoration is an important problem in image processing. When a detector in the imaging plane of an optical imaging system is out of focus, we will obtain an out-of-focus image. This physical phenomenon is modeled by equation (2.1) with a blurring operator. By considering practical physical effects, especially aberration and diffraction, the corresponding blurring kernel is approximated by a Gaussian kernel having the form (2.3) (cf [6, 24]). In this case, the blurring operator can be written as a product of two univariate integral operators. The related integral operator is defined for  $u \in L^2(E)$  by

$$(\mathcal{K}_\sigma u)(x) := \int_E K_\sigma(x, y)u(y) dy, \quad x \in E, \quad (2.4)$$

where

$$K_\sigma(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-y)^2}{2\sigma^2}}, \quad x, y \in E.$$

Clearly, the operator  $\mathcal{K}_\sigma$  depends on the standard deviation  $\sigma$ . However, when there is no need to highlight the dependence, we will drop the subscript and simply write  $\mathcal{K}$ . In terms of operator  $\mathcal{K}$ , we write  $\mathbb{K}v = \mathcal{K}_1\mathcal{K}_2v$ , where  $\mathcal{A}_j$  denotes the operator  $\mathcal{A}$  applied to the  $j$ th argument of the function. Consequently, equation (2.1) becomes

$$\mathcal{K}_1\mathcal{K}_2v = f. \quad (2.5)$$

For a given observed image  $f$  with noise, we may obtain the true image by solving two one-dimensional integral equations. Mainly, we first solve

$$(\mathcal{K}v_1(\cdot, y))(x) = f(x, y), \quad x, y \in E,$$

and then solve

$$(\mathcal{K}v(x, \cdot))(y) = v_1(x, y), \quad x, y \in E.$$

Hence, we will focus on the study of the one-dimensional integral equation of the first kind

$$\mathcal{K}u = h, \quad (2.6)$$

where  $h \in L^2(E)$  is a given function. For  $h \in R(\mathcal{K})$ , the range of  $\mathcal{K}$ , we let  $u_* \in L^2(E)$  denote the unique minimum norm solution of equation (2.6) in the sense that

$$\mathcal{K}u_* = h \quad \text{with} \quad \|u_*\| = \inf\{\|v\| : v \in L^2(E), \mathcal{K}v = h\}.$$

Since the integral operator  $\mathcal{K}$  is compact on  $L^2(E)$ , the integral equation (2.6) is ill-posed in the sense that its solution does not continuously depend on the given data  $h$ . Thus, a small perturbation in the given data may cause a large perturbation in the solution. Therefore, the integral equation (2.6) needs to be regularized in order to find its approximate solution. Ill-posed problems are normally treated by regularization methods such as the Tikhonov regularization [31] and the Lavrentiev regularization [14, 25, 26].

Being a convolution-type operator,  $\mathcal{K}$  is self-adjoint. Hence, the Lavrentiev regularization may be applied to (2.6). It leads to the regularization equation

$$(\lambda\mathcal{I} + \mathcal{K})u_\lambda = h, \quad (2.7)$$

where  $\mathcal{I}$  is the identity operator on  $L^2(E)$  and  $\lambda$  is a regularization parameter. We consider the unique solution  $u_\lambda$  of (2.7) as an approximation of  $u_*$ .

By using the Plancherel theorem and the properties of the Fourier transform of the Gaussian function [30], it can be shown that the operator  $\lambda\mathcal{I} + \mathcal{K}$  is strictly coercive and hence,

it has a bounded inverse for any positive number  $\lambda$ . This leads to the following regularization for equation (2.5) for image restoration:

$$(\lambda\mathcal{I} + \mathcal{K})_1(\lambda\mathcal{I} + \mathcal{K})_2 v_\lambda = f. \quad (2.8)$$

The Tikhonov regularization solution  $v_\lambda$  of (2.6) is obtained by solving the following equation:

$$(\lambda\mathcal{I} + \mathcal{K}^* \mathcal{K}) v_\lambda = \mathcal{K}^* h, \quad (2.9)$$

where  $\mathcal{K}^*$  is the adjoint operator of  $\mathcal{K}$ . The regularization solution  $v_\lambda$  can be interpreted in terms of minimization (cf [14]). Since  $\mathcal{K}$  is self-adjoint, that is  $\mathcal{K}^* = \mathcal{K}$ , equation (2.9) becomes

$$(\lambda\mathcal{I} + \mathcal{K}^2) v_\lambda = \mathcal{K} h. \quad (2.10)$$

The operator  $\lambda\mathcal{I} + \mathcal{K}^2$  is strictly coercive and has a bounded inverse for any positive number  $\lambda$ .

Note that equation (2.10) involves the operator  $\mathcal{K}^2$  which is defined by a double integral. Calculation of such an operator is computationally costly. For this reason, we propose a modification. We define a linear operator  $\tilde{\mathcal{K}}_\sigma^*$  for  $u \in L^2(\mathbb{R})$  by

$$\tilde{\mathcal{K}}_\sigma^* u(x) = \int_{\mathbb{R}} K_\sigma(x, y) u(y) dy, \quad x \in E.$$

The notation  $\tilde{\mathcal{K}}_\sigma^* u$  for  $u \in L^2(E)$  should be understood as  $\tilde{\mathcal{K}}_\sigma^*$  applying to the extension  $\tilde{u}$  of  $u$ , where  $\tilde{u}$  equals  $u$  on  $E$  and zero outside of  $E$ .

Using the operator  $\tilde{\mathcal{K}}_\sigma^*$  to replace  $\mathcal{K}_\sigma^*$  in the above Tikhonov regularization method yields a modification of the Tikhonov method. The following proposition shows that  $\tilde{\mathcal{K}}_\sigma^* \mathcal{K}_\sigma$  can be computed by an integral operator with a Gaussian kernel with standard deviation  $\sqrt{2}\sigma$ .

**Proposition 2.1.** *There holds the relation  $\tilde{\mathcal{K}}_\sigma^* \mathcal{K}_\sigma = \mathcal{K}_{\sqrt{2}\sigma}$ .*

**Proof.** For  $u \in L^2(E)$ , by the definition of  $\tilde{\mathcal{K}}_\sigma^*$  and  $\mathcal{K}_\sigma$ , and by the Fubini theorem, we have that

$$\tilde{\mathcal{K}}_\sigma^* \mathcal{K}_\sigma u(x') = \int_E \frac{1}{2\pi\sigma^2} \int_{\mathbb{R}} e^{-\frac{(x-x')^2}{2\sigma^2} - \frac{(x-y)^2}{2\sigma^2}} dx u(y) dy. \quad (2.11)$$

The desired formula is obtained by using the above equation and the identity

$$\int_{\mathbb{R}} \exp\left(-\frac{(x - (x' + y)/2)^2}{\sigma^2}\right) dx = \sqrt{\pi}\sigma. \quad \square$$

Making use of proposition 2.1, the modified Tikhonov regularization method becomes

$$((\lambda\mathcal{I} + \mathcal{K}_{\sqrt{2}\sigma}) v_{1,\lambda}(\cdot, y))(x) = (\tilde{\mathcal{K}}_\sigma^* f(\cdot, y))(x), \quad x, y \in E \quad (2.12)$$

and

$$((\lambda\mathcal{I} + \mathcal{K}_{\sqrt{2}\sigma}) v_\lambda(x, \cdot))(y) = (\tilde{\mathcal{K}}_\sigma^* v_{1,\lambda}(x, \cdot))(y), \quad x, y \in E. \quad (2.13)$$

Both the Lavrentiev and the modified Tikhonov regularization methods led to consideration of solving the following equation:

$$(\lambda\mathcal{I} + \mathcal{K}) u = h, \quad (2.14)$$

where  $h \in L^2(E)$  is a given function and  $u \in L^2(E)$  is the solution to be determined. For any positive parameter  $\lambda$ , equation (2.14) has a unique solution  $u \in L^2(E)$ .

### 3. Multiscale Galerkin methods

The focus of this section is to describe the Galerkin method for solving equation (2.14) by using multiscale piecewise polynomial bases constructed in [19, 20]. We will adopt the multiscale Galerkin method developed in [22] for Fredholm integral equations of the second kind. General information of Galerkin methods for solving Fredholm integral equations of the second kind is found in [1]. Multiscale methods for solving ill-posed problems were considered in [9, 10].

We begin a review of the Galerkin method for solving equation (2.14). Let  $\mathbb{N}$  denote the set of all natural numbers and  $\mathbb{Z}_+$  the set of all nonnegative integers. For  $n \in \mathbb{N}$ , let  $\mathbb{Z}_n := \{0, 1, \dots, n-1\}$ . Suppose that  $\mathbb{X}_n$ ,  $n \in \mathbb{Z}_+$ , is a sequence of finite-dimensional subspaces of  $L^2(E)$  such that

$$L^2(E) = \overline{\bigcup_{n \in \mathbb{Z}_+} \mathbb{X}_n}.$$

Let  $\mathcal{P}_n$  be the orthogonal projection from  $L^2(E)$  onto  $\mathbb{X}_n$ . The Galerkin method for solving equation (2.14) is to find  $u_n \in \mathbb{X}_n$  such that

$$(\lambda \mathcal{I} + \mathcal{K}_n)u_n = \mathcal{P}_n h, \quad (3.1)$$

where  $\mathcal{K}_n := \mathcal{P}_n \mathcal{K}$ . When a basis of the space  $\mathbb{X}_n$  is chosen, the finite-dimensional operator equation (3.1) is equivalent to a system of linear equations. We will choose  $\mathbb{X}_n$  as a piecewise polynomial space and use its multiscale orthonormal basis for equation (3.1).

We now review the construction of a multiscale basis for  $\mathbb{X}_n$ ,  $n \in \mathbb{Z}_+$ , of  $L^2(E)$ , originally introduced in [19]. There are two important ingredients in the construction of the multiscale analysis for a given integer  $\mu \geq 2$ . The first one is a set of contractive mappings on  $E$  defined by  $\phi_\epsilon(t) := (t + \epsilon)/\mu$ ,  $t \in E$ ,  $\epsilon \in \mathbb{Z}_\mu$ . Note that  $E$  is the invariant set of these mappings, that is  $E = \bigcup_{\epsilon \in \mathbb{Z}_\mu} \phi_\epsilon(E)$ , where  $\phi_\epsilon(E)$  is the image of the set  $E$  under the mapping  $\phi_\epsilon$ . Associated with the mappings  $\phi_\epsilon$ , we define linear operators  $\mathcal{T}_\epsilon : L^2(E) \rightarrow L^2(E)$  for  $v \in L^2(E)$  by

$$(\mathcal{T}_\epsilon v)(\cdot) = \sqrt{\mu}(v \circ \phi_\epsilon^{-1})(\cdot) \chi_{\phi_\epsilon(E)}(\cdot), \quad \epsilon \in \mathbb{Z}_\mu,$$

where  $\chi_V$  is the characteristic function on a set  $V \subset E$ . The second one is a subspace, denoted by  $\mathbb{X}_0$ , of  $L^2(E)$ . We choose  $\mathbb{X}_0 = \Pi_p$  which represents the space of all polynomials of degree  $\leq p-1$ . With the given subspace  $\mathbb{X}_0$  and linear operators  $\mathcal{T}_\epsilon$ , according to [19], we form a sequence of subspaces of  $L^2(E)$  by the recursion  $\mathbb{X}_{n+1} := \bigcup_{\epsilon \in \mathbb{Z}_\mu} \mathcal{T}_\epsilon(\mathbb{X}_n)$  and let  $s(n) := \dim \mathbb{X}_n$ . The sequence of subspaces so constructed has the properties

$$\mathbb{X}_n \subset \mathbb{X}_{n+1} \quad \text{and} \quad L^2(E) = \overline{\bigcup_{n \in \mathbb{Z}_+} \mathbb{X}_n}.$$

Furthermore, let  $\mathbb{W}_{n+1}$  be the orthogonal complement of  $\mathbb{X}_n$  in  $\mathbb{X}_{n+1}$ ,  $n \in \mathbb{Z}_+$ . With a convention  $\mathbb{W}_0 := \mathbb{X}_0$ , we have that

$$\mathbb{X}_n = \mathbb{W}_0 \oplus \mathbb{W}_1 \oplus \dots \oplus \mathbb{W}_n. \quad (3.2)$$

Moreover, when  $\mathbb{W}_1$  is available, we have for  $n \in \mathbb{N}$  that

$$\mathbb{W}_{n+1} = \bigcup_{\epsilon \in \mathbb{Z}_\mu} \mathcal{T}_\epsilon(\mathbb{W}_n), \quad (3.3)$$

see [19]. A general construction of  $\mathbb{W}_1$  through solving matrix equations may be found in [20].

We next describe a construction of bases for  $\mathbb{W}_j$ ,  $j \in \mathbb{Z}_{n+1}$ . To this end, for a vector  $\mathbf{s} := (s_0, s_1, \dots, s_{i-1})$  in  $\mathbb{Z}_\mu^i := \mathbb{Z}_\mu \times \dots \times \mathbb{Z}_\mu$  ( $i$  folds), we define a composition

operator  $\mathcal{T}_s := \mathcal{T}_{s_0} \circ \mathcal{T}_{s_1} \circ \dots \circ \mathcal{T}_{s_{i-1}}$ . For any number  $j \in \mathbb{Z}_{\mu^i}$ , there is a unique vector  $\mathbf{s} := (s_0, s_1, \dots, s_{i-1}) \in \mathbb{Z}_{\mu}^i$  such that  $j = \mu^{i-1}s_0 + \dots + \mu s_{i-2} + s_{i-1}$ . We call  $\mathbf{s}$  the  $(\mu, i)$ -form of  $j$  and write  $j = \mu(\mathbf{s})$ . For  $j \in \mathbb{Z}_{n+1}$ , let  $w(j) := \dim \mathbb{W}_j$ . We choose a basis of  $\mathbb{X}_0$  as the (normalized) Legendre polynomials  $e_{0j}$  of degree  $j$ , for  $j \in \mathbb{Z}_p$ . An orthonormal basis  $\{e_{1j} : j \in \mathbb{Z}_{w(1)}\}$  for  $\mathbb{W}_1$  can be constructed from a basis of  $\mathbb{X}_1$  by requiring every element to be orthogonal to the space  $\mathbb{X}_0$ . It is easy to verify that  $w(0) = p$  and  $w(1) = r := (\mu - 1)p$ . For  $j = \mu(\mathbf{s})r + \ell$  with  $i > 1$ ,  $\mathbf{s} \in \mathbb{Z}_{\mu}^{i-1}$  and  $\ell \in \mathbb{Z}_r$ , we define

$$e_{ij} := \mathcal{T}_s e_{1\ell}. \tag{3.4}$$

From (3.3), we know that  $\{e_{ij} : j \in \mathbb{Z}_{w(i)}\}$  forms an orthogonal basis of  $\mathbb{W}_i$  and  $w(i) = \mu^{i-1}r$ . Upon defining the index set  $\mathbb{U}_n := \{(i, j) : i \in \mathbb{Z}_{n+1}, j \in \mathbb{Z}_{w(i)}\}$ , we see from (3.2) that  $\{e_{ij} : (i, j) \in \mathbb{U}_n\}$  forms a multiscale orthonormal basis for  $\mathbb{X}_n$ .

We now show how a partition of  $E$  is generated by the contractive mappings  $\phi_\epsilon, \epsilon \in \mathbb{Z}_{\mu}$ . For  $\mathbf{s} := (s_0, s_1, \dots, s_{i-1}) \in \mathbb{Z}_{\mu}^i$ , we define  $\Phi_s = \phi_{s_0} \circ \phi_{s_1} \circ \dots \circ \phi_{s_{i-1}}$ . For each  $i \in \mathbb{N}$ , we define  $E_{ij} := \Phi_s(E)$ ,  $j \in \mathbb{Z}_{\mu^i}$ , where  $\mathbf{s}$  is the  $(\mu, i)$ -form of  $j$ . Clearly,

$$\bigcup_{j \in \mathbb{Z}_{\mu^i}} E_{ij} = E \quad \text{and} \quad \text{meas}(E_{ij} \cap E_{ij'}) = 0, \quad j \neq j'.$$

In other words, for each fixed  $i \in \mathbb{Z}_+$ ,  $\{E_{ij} : j \in \mathbb{Z}_{\mu^i}\}$  forms a partition of  $E$ .

We list below properties of the partition  $\{E_{ij} : j \in \mathbb{Z}_{\mu^i}\}$ , subspaces  $\mathbb{X}_i$ , and the orthonormal basis  $\{e_{ij} : (i, j) \in \mathbb{U}_n\}$ . In what follows, the notation  $a \sim b$  means that there are two positive constants  $c$  and  $c'$  such that  $ca \leq b \leq c'a$ . We use  $c$  to denote a universal constant which can be distinct at different occurrences.

- If  $d(i) := \max\{\text{meas}(E_{ij}) : j \in \mathbb{Z}_{\mu^i}\}$ , then  $d(i) \sim \mu^{-i}$ ,  $w(i) \sim \mu^i$  and  $s(i) \sim \mu^i$ .
- For  $i \in \mathbb{N}$  and  $j \in \mathbb{Z}_{w(i)}$ , if we write  $j = \mu(\mathbf{s})r + \ell$ , where  $\mathbf{s} \in \mathbb{Z}_{\mu}^{i-1}$  and  $\ell \in \mathbb{Z}_r$ , then  $e_{ij}(t) = 0, t \notin E_{i,\mu(\mathbf{s})}$ . Denoting  $S_{ij} := E_{i,\mu(\mathbf{s})}$ , we see that the support of  $e_{ij}$  is contained in  $S_{ij}$ .
- For  $i \in \mathbb{N}$  and  $j \in \mathbb{Z}_{w(i)}$ ,  $e_{ij}$  has the ‘vanishing moment’ property, that is  $\langle e_{ij}, v \rangle = 0, v \in \Pi_p$ .
- There exists a positive constant  $c$  such that  $\|e_{ij}\| = 1$  and  $\|e_{ij}\|_{\infty} \leq c\mu^{i/2}$ .

Let  $I_n$  denote the  $s(n) \times s(n)$  identity matrix, introduce an  $s(n) \times s(n)$  matrix  $A_n := [\langle \mathcal{K}e_{ij}, e_{i'j'} \rangle : (i, j), (i', j') \in \mathbb{U}_n]$  and define vectors  $\mathbf{h}_n := [\langle h, e_{i'j'} \rangle : (i', j') \in \mathbb{U}_n]$ ,  $\mathbf{u}_n := [u_{ij} : (i, j) \in \mathbb{U}_n]$ . Using these notations, the operator equation (3.1) can be written as an equivalent linear system:

$$(\lambda I_n + A_n)\mathbf{u}_n = \mathbf{h}_n. \tag{3.5}$$

The choice of the parameter  $\lambda$  was discussed in [18] for a general wavelet approach. Upon solving system (3.5) for  $\mathbf{u}_n$ , we obtain the solution  $u_n$  of the integral equation (3.1) by setting

$$u_n = \sum_{(i,j) \in \mathbb{U}_n} u_{ij} e_{ij}.$$

Noting that  $A_n$  is a dense matrix, when its size  $s(n)$  is large, it is expensive to generate it. The use of the multiscale basis allows us to compress the matrix  $A_n$  to a sparse matrix, upon which a fast numerical algorithm can be developed. This will be discussed in the next section.

#### 4. Matrix compression

The main purpose of this section is to develop a matrix compression scheme for the matrix  $A_n$  that appears in system (3.5), based on properties of the Gaussian kernel.

We first estimate the absolute value of the entries  $a_{i'j',ij} := \langle \mathcal{K}e_{ij}, e_{i'j'} \rangle$ ,  $(i, j), (i', j') \in \mathbb{U}_n$ , of the coefficient matrix  $A_n$ . For this purpose, we denote by  $g$  the normalized Gaussian function, that is

$$g(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

In terms of the normalized Gaussian function, the kernel  $K_\sigma$  of the operator  $\mathcal{K}$  has the form  $K_\sigma(x, y) = \sigma^{-1}g((x - y)/\sigma)$ .

The next lemma has the same spirit of lemma 3.1 in [22].

**Lemma 4.1.** *There exists a positive constant  $c$  such that for all  $n \in \mathbb{N}$ , all indices  $(i, j), (i', j') \in \mathbb{U}_n$  and all parameters  $\sigma > 0$ :*

$$|a_{i'j',ij}| \leq c\mu^{(i+i')/2} (d_i d_{i'})^p \sigma^{-(2p+1)} \int_{S_{i'j'}} \int_{S_{ij}} \left| g^{(2p)}\left(\frac{x-y}{\sigma}\right) \right| dy dx, \quad (4.1)$$

where  $p$  is the order of vanishing moments of the multiscale basis used, and  $S_{ij}$  and  $S_{i'j'}$  are supports of  $e_{ij}$  and  $e_{i'j'}$ , respectively.

Note that  $D_x^m D_y^\ell K_\sigma(x, y) = (-1)^\ell \sigma^{-(m+\ell+1)} g^{(m+\ell)}\left(\frac{x-y}{\sigma}\right)$  for any positive integers  $m$  and  $\ell$ . By using this relation of the high-order derivative of the normalized Gaussian function, the proof of lemma 4.1 is essentially the same as that of lemma 3.1 in [22]. Hence we omit its proof here.

Lemma 4.1 shows that when the distance of the supports of two basis functions  $e_{ij}$  and  $e_{i'j'}$  is large enough, the absolute value of the entry  $a_{i'j',ij}$  of  $A_n$  is very small. Hence, the distance of the supports is a suitable indicator to predict the magnitude of  $a_{i'j',ij}$ . This observation leads us to proposing a truncation strategy for the matrix  $A_n$  by using the distances of the supports of the two basis functions which define the entry. For the given matrix  $A_n$ , a matrix truncation refers to keeping some of its entries unchanged and making others zero according to a pre-given rule. The effect of the matrix truncation is that the resulting matrix, denoted by  $\tilde{A}_n$ , is sparse in the sense that most of its entries are zero. By doing so, we are able to develop a fast algorithm for solving the linear system (3.5) with the coefficient matrix  $A_n$  being replaced by  $\tilde{A}_n$ . Clearly, the more sparse the matrix  $\tilde{A}_n$  is, the more efficient the algorithm would be in terms of computational cost. However, the matrix truncation may result in an error in the solution. Therefore, a truncation strategy should be designed to balance two requirements, namely the sparseness of  $\tilde{A}_n$  and the closeness of the solution of the system with the truncated matrix to the solution  $\mathbf{u}_n$  of the original equation (3.5). In the rest of this section, we will propose a matrix truncation scheme and discuss the first requirement. We postpone the discussion of the second requirement to the next section.

We now introduce a truncation strategy for the matrix  $A_n$ . Corresponding to the decomposition (3.2) of the subspace  $\mathbb{X}_n$ ,  $A_n$  is partitioned as a block matrix:

$$A_n = [A_{i'i} : i', i \in \mathbb{Z}_n], \quad \text{where} \quad A_{i'i} = [a_{i'j',ij} : j' \in \mathbb{Z}_{w(i')}, j \in \mathbb{Z}_{w(i)}].$$

We use  $\text{dist}(U, V) := \min\{|x - y| : x \in U, y \in V\}$  to denote the distance between two compact sets  $U$  and  $V$  of  $\mathbb{R}$ . For  $(i', j'), (i, j) \in \mathbb{U}_n$  and a given parameters  $\delta_{i'i}$ , we define

$$\tilde{a}_{i'j',ij} := \begin{cases} a_{i'j',ij}, & \text{if } \text{dist}(S_{i'j'}, S_{ij}) \leq \delta_{i'i}\sigma, \\ 0, & \text{otherwise.} \end{cases} \quad (4.2)$$

For each block  $A_{i'i}$ , we truncate it into a new block  $\tilde{A}_{i'i} = [\tilde{a}_{i'j',ij} : j' \in \mathbb{Z}_{w(i')}, j \in \mathbb{Z}_{w(i)}]$  and define the truncated matrix  $\tilde{A}_n$  by letting  $\tilde{A}_n := [\tilde{A}_{i'i} : i', i \in \mathbb{Z}_n]$ . We then replace the matrix  $A_n$  in (3.5) by  $\tilde{A}_n$  and solve the compressed linear system

$$(\lambda I_n + \tilde{A}_n)\tilde{\mathbf{u}}_n = \mathbf{h}_n. \quad (4.3)$$



Due to the sparsity of the matrix  $\tilde{A}_n$ , system (4.3) leads to a fast algorithm for solving equation (3.1).

We now discuss the choice of the parameters  $\delta_{i'i}$ . In the context of image analysis, we normalize an image so that its domain is  $[0, 1] \times [0, 1]$ . For this reason, the standard deviation  $\sigma$  is assumed to be a positive number smaller than 1. The choice of the parameters  $\delta_{i'i}$  in (4.2) will depend on the standard deviation  $\sigma$ . To introduce the parameters  $\delta_{i'i}$ , for a positive integer  $p$ , we define a function

$$G_{2p}(t) := \int_t^\infty |g^{(2p)}(\xi)| d\xi, \quad t \in [0, \infty).$$

Clearly,  $G_{2p}$  is monotonically decreasing and  $G_{2p}(t) \rightarrow 0$  as  $t \rightarrow \infty$ . For the given positive numbers  $a, \alpha$  and  $\alpha'$ , we let

$$\epsilon_{i'i} := a^{-2p} \mu^{-2p(-n+\alpha(n-i)+\alpha'(n-i'))} \sigma^{2p}. \tag{4.4}$$

If

$$\epsilon_{i'i} < G_{2p}(0), \tag{4.5}$$

we choose  $\delta_{i'i}$  such that

$$G_{2p}(\delta_{i'i}) = \epsilon_{i'i} \tag{4.6}$$

and otherwise, we choose

$$\delta_{i'i} := 0. \tag{4.7}$$

With the proposed truncation strategy and the suggested parameters  $\delta_{i'i}$ , an estimate of the number  $N(\tilde{A}_n)$  of nonzero entries of the  $\tilde{A}_n$  is presented in the next theorem.

**Theorem 4.2.** *Let  $\alpha, \alpha' \leq 1$ . If  $\delta_{i'i}$  is chosen as in (4.6) and (4.7), then there exists a positive constant  $c$  such that for all  $\sigma > 0$  and  $n > 0$*

$$N(\tilde{A}_n) \leq cs(n) \log^{\vartheta(\alpha, \alpha')} s(n),$$

where  $\vartheta(\alpha, \alpha') = 2$  for  $(\alpha, \alpha') = (1, 1)$  and 1 otherwise.

**Proof.** Let  $N(\tilde{A}_{i'i})$  denote the number of nonzero entries of  $\tilde{A}_{i'i}$ . Following [22], we have that

$$N(\tilde{A}_n) = \sum_{i=0}^n \sum_{i'=0}^n N(\tilde{A}_{i'i}) \leq c \sum_{i=0}^n \sum_{i'=0}^n \mu^{i+i'} (d_i + d_{i'} + \delta_{i'i} \sigma). \tag{4.8}$$

Since  $d_i \sim \mu^{-i}$  and  $d_{i'} \sim \mu^{-i'}$ , we obtain that

$$\sum_{i=0}^n \sum_{i'=0}^n \mu^{i+i'} (d_i + d_{i'}) \leq c \left( \sum_{i=0}^n n \mu^i + \sum_{i'=0}^n n \mu^{i'} \right) \leq cn \mu^n. \tag{4.9}$$

It remains to estimate the third term of the sum on the right-hand side of (4.8). Noting that  $\lim_{\xi \rightarrow \infty} g^{(2p)}(\xi) \xi^{2p+1} = 0$ , there exists a constant  $c > 0$  (independent of  $\sigma$ ) such that  $|g^{(2p)}(\xi) \xi^{2p+1}| \leq c$ , which implies that

$$\frac{c}{|\xi|^{2p+1}} \geq |g^{(2p)}(\xi)|.$$

Integrating both sides of the inequality above over  $[\delta_{i'i}, \infty)$  yields

$$\int_{\delta_{i'i}}^\infty \frac{c}{\xi^{2p+1}} d\xi \geq \int_{\delta_{i'i}}^\infty |g^{(2p)}(\xi)| d\xi.$$

If  $\delta_{i'i} \neq 0$ , computing the integral on the left-hand side and using (4.6), we have that

$$\frac{c}{2p} \delta_{i'i}^{-2p} \geq a^{-2p} \mu^{-2p(-n+\alpha(n-i)+\alpha'(n-i'))} \sigma^{2p}.$$

This ensures that there exists a positive constant  $c$  such that for all  $\sigma > 0$  and  $n > 0$

$$\delta_{i'i} \sigma \leq ca \mu^{-n+\alpha(n-i)+\alpha'(n-i')}.$$

Hence,

$$\sum_{i=0}^n \sum_{i'=0}^n \mu^{i+i'} (\delta_{i'i} \sigma) \leq ca \mu^n \sum_{i=0}^n \mu^{(\alpha-1)(n-i)} \sum_{i'=0}^n \mu^{(\alpha'-1)(n-i')} = c \mu^n n^{\vartheta(\alpha, \alpha')},$$

which together with (4.9) proves the desired result of this theorem.  $\square$

Next, we discuss the choice of the truncation parameter  $\delta_{i'i}$ . Choosing the truncation parameters  $\delta_{i'i}$  requires solving the nonlinear equation

$$G_{2p}(t) = \epsilon_{i'i}, \quad (4.10)$$

for a positive solution when  $0 < \epsilon_{i'i} < G_{2p}(0)$ . Since  $G_{2p}$  is a positive strictly decreasing function on  $\mathbb{R}^+$ , when  $0 < \epsilon_{i'i} < G_{2p}(0)$ , equation (4.10) has a unique solution. However, the main difficulty of finding its solution lies in dealing with the absolute value of derivatives of the Gaussian function which defines the function  $G_{2p}$ . This difficulty can be overcome by using root properties of  $g^{(2p)}$ .

We now recall useful root properties of  $g^{(m)}$  for a nonnegative integer  $m$ . The normalized Gaussian function  $g$  is related to the Hermite polynomial. Specifically, the  $m$ th derivative of the Gaussian function  $g$  is related to the Hermite polynomial of degree  $m$  through the Rodrigues formula:

$$H_m(x) = (-1)^m e^{\frac{x^2}{2}} g^{(m)}(x), \quad x \in \mathbb{R}. \quad (4.11)$$

It is well known (see, e.g., [3]) that the Hermite polynomial has the following properties: (i)  $H_m^{(1)} = mH_{m-1}$ ; (ii)  $H_{m+2}(x) - xH_{m+1}(x) + (m+1)H_m(x) = 0$  (recursive relations); (iii)  $H_{2m-1}(0) = 0$  and  $H_{2m}(0) = (-1)^m (2m)!/m!2^m \sqrt{2\pi}$ ; and (iv) the Hermite polynomial  $H_m$  has exactly  $m$  simple real roots. The roots of  $H_m$  and  $H_{m+1}$  strictly interlace. Moreover,  $H_m$  has exactly  $\lfloor \frac{m}{2} \rfloor$  negative roots and exactly  $\lfloor \frac{m}{2} \rfloor$  positive roots. In particular, when  $m$  is odd, the origin is an additional root of  $H_m$ . These properties together with the Rodrigues formula yield the corresponding properties of the Gaussian function  $g$ :

- (1)  $g^{(m+2)}(x) + xg^{(m+1)}(x) + (m+1)g^{(m)}(x) = 0$  (recursive relations);
- (2)  $g^{(2m-1)}(0) = 0$  and  $g^{(2m)}(0) = (-1)^m (2m)!/m!2^m \sqrt{2\pi}$ ;
- (3) the set of roots of the Hermite polynomial  $H_m$  is identical with the set of roots of the  $m$ th derivative of  $g$ .

We now return to solving equation (4.10). Note that  $g^{(2p)}$  does not change its sign between its any two consecutive roots. We split the integral interval  $[t, +\infty)$  for  $t > 0$  according to the roots of  $g^{(2p)}$ . Let  $r_i, i \in \mathbb{Z}_{p+1} \setminus \{0\}$ , be the positive roots of  $g^{(2p)}$ . For notational convenience, we let  $r_0 := 0$ . We consider three cases.

*Case I.* If  $\epsilon_{i'i} = G_{2p}(r_p)$ , we simply choose  $\delta_{i'i} = r_p$ .

*Case II.* If  $\epsilon_{i'i} < G_{2p}(r_p)$ , since  $g^{(2p)}$  does not change sign on the interval  $(r_p, +\infty)$ , we have that

$$\epsilon_{i'i} = \int_t^{+\infty} |g^{(2p)}(x)| dx = \left| \int_t^{+\infty} g^{(2p)}(x) dx \right| = |g^{(2p-1)}(t)|.$$

Equation (4.10) becomes

$$|g^{(2p-1)}(t)| = \epsilon_{i'}^i, \quad t > r_p. \tag{4.12}$$

Note that the largest zero  $r_p$  of  $g^{(2p)}(x)$  is a local extreme point of  $g^{(2p-1)}(x)$ , and  $g^{(2p-1)}(x)$  converges to zero when  $x$  goes to infinity. Hence  $g^{(2p-1)}(x)$  does not change sign and is strictly monotonic on  $(r_p, +\infty)$ , and equation (4.12) can be easily solved by using the bisection method in the following fashion. Define  $x_n := (n + 1)r_p$ . Let  $n_0$  be the smallest integer such that  $|g^{(2p-1)}(x_{n_0+1})| < \epsilon_{i'}^i$ . We find the root of (4.12) on the interval  $[x_{n_0}, x_{n_0+1}]$  by employing the bisection method.

Case III. If  $\epsilon_{i'}^i > G_{2p}(r_p)$ , we then find the largest positive zero  $r_j$  of  $H_{2p}$  such that  $G_{2p}(r_j) < \epsilon_{i'}^i \leq G_{2p}(r_{j-1})$ . Since  $G_{2p}$  is strictly decreasing, the solution  $t_0$  of (4.10) lies in  $[r_{j-1}, r_j]$ . We rewrite (4.10) as

$$\int_t^{r_j} |g^{(2p)}(x)| dx = \epsilon_{i'}^i - G_{2p}(r_j).$$

Noting that  $g^{(2p)}$  does not change sign on the interval  $[r_{j-1}, r_j]$ , we have that

$$|g^{(2p-1)}(r_j) - g^{(2p-1)}(t)| = \epsilon_{i'}^i - G_{2p}(r_j). \tag{4.13}$$

Because  $r_j$  and  $r_{j-1}$  are two consecutive extreme points of  $g^{(2p-1)}$ ,  $g^{(2p-1)}(r_j) - g^{(2p-1)}(t)$  is strictly monotonic and does not change sign for  $t \in [r_{j-1}, r_j]$ . Therefore, we can use the bisection method to solve (4.13) on the interval  $[r_{j-1}, r_j]$ .

For a given tolerance  $10^{-5}$  in the bisection method, the number of iterations used by the bisection method for finding  $\delta_{i'}^i$  in  $A_8$  for  $\sigma = 0.01$  and  $\sigma = 0.02$  are reported respectively, in the following two arrays:

$$\begin{pmatrix} 3 & 3 & 3 & 3 & 3 & 3 & 7 & 9 \\ 3 & 3 & 3 & 3 & 3 & 7 & 9 & 11 \\ 3 & 3 & 3 & 3 & 7 & 9 & 11 & 10 \\ 3 & 3 & 3 & 7 & 9 & 11 & 10 & 14 \\ 3 & 3 & 7 & 9 & 11 & 10 & 14 & 13 \\ 3 & 7 & 9 & 11 & 10 & 14 & 13 & 16 \\ 7 & 9 & 11 & 10 & 14 & 13 & 16 & 17 \\ 9 & 11 & 10 & 14 & 13 & 16 & 17 & 9 \end{pmatrix} \text{ and } \begin{pmatrix} 3 & 3 & 3 & 3 & 3 & 9 & 11 & 12 \\ 3 & 3 & 3 & 3 & 9 & 11 & 12 & 13 \\ 3 & 3 & 3 & 9 & 11 & 12 & 13 & 12 \\ 3 & 3 & 9 & 11 & 12 & 13 & 12 & 14 \\ 3 & 9 & 11 & 12 & 13 & 12 & 14 & 13 \\ 9 & 11 & 12 & 13 & 12 & 14 & 13 & 18 \\ 11 & 12 & 13 & 12 & 14 & 13 & 18 & 17 \\ 12 & 13 & 12 & 14 & 13 & 18 & 17 & 20 \end{pmatrix}.$$

$\sigma = 0.01$   $\sigma = 0.02$

In either case, we choose the parameter  $\delta_{i'}^i$  to be the solution of equation (4.10). The method described above requires evaluating the function  $G_{2p}$  at the roots of  $g^{(2p)}$ . The following lemma gives a convenient way of evaluation of  $G_{2p}$  at a root of  $g^{(2p)}$ .

**Lemma 4.3.** *If the set of nonnegative zeros of  $H_{2m}$  is given by  $\{x_i : i \in \mathbb{Z}_{m+1} \setminus \{0\}\}$  with  $x_1 < x_2 < \dots < x_m$ , then for  $i \in \mathbb{Z}_m \setminus \{0\}$ ,*

$$G_{2m}(x_i) = |g^{(2m-1)}(x_i)| + 2 \sum_{j=i+1}^m |g^{(2m-1)}(x_j)|.$$

**Proof.** Since the Hermite polynomial  $H_{2m}$  has simple real roots,  $H_{2m}$  does not change sign between its any two consecutive roots. Furthermore, for any two consecutive positive roots  $x_q < x_{q+1}$  of  $H_{2m}$ ,  $H_{2m-1}$  has exactly one root in the interval  $(x_q, x_{q+1})$ . Hence,  $H_{2m-1}(x_q)H_{2m-1}(x_{q+1}) < 0$ , that is  $g^{(2m-1)}(x_q)$  and  $g^{(2m-1)}(x_{q+1})$  have different signs.

Combining the above facts together, we have that

$$\begin{aligned}
 G_{2m}(x_i) &= \sum_{j=i}^{m-1} \left| \int_{x_j}^{x_{j+1}} g^{(2m)}(x) dx \right| + \left| \int_{x_m}^{+\infty} g^{(2m)}(x) dx \right| \\
 &= \sum_{j=i}^{m-1} (|g^{(2m-1)}(x_{j+1})| + |g^{(2m-1)}(x_j)|) + |g^{(2m-1)}(x_m)|,
 \end{aligned}$$

proving the desired formula. □

Using lemma 4.3, we can easily compute the values  $G_{2p}(r_i)$ . In fact, according to lemma 4.3 and (4.11), we have that  $G_{2p}(r_i) = |g(r_i)H_{2p-1}(r_i)| + 2 \sum_{j=i+1}^p |g(r_j)H_{2p-1}(r_j)|$ . That is, computing  $G_{2p}(r_i)$  can be performed by evaluating the Gaussian function and the Hermite polynomial  $H_{2p-1}$  at the roots  $r_j, j \geq i$ , of the Hermite polynomial  $H_{2p}$ . This can be efficiently implemented.

### 5. Convergence analysis

We establish in this section the convergence order of the compressed Galerkin method. The solution  $\tilde{\mathbf{u}}_n := [\tilde{u}_{ij} : (i, j) \in \mathbb{U}_n]$  of equation (4.3) defines a function in space  $\mathbb{X}_n$  by

$$\tilde{u}_n := \sum_{(i,j) \in \mathbb{U}_n} \tilde{u}_{ij} e_{ij}.$$

It is an approximation to the exact solution  $u$  of (2.14). The main purpose of this section is to estimate the convergence order in which  $\tilde{u}_n$  approximates  $u$ .

The function  $\tilde{u}_n$  satisfies an operator equation. We let  $\tilde{\mathcal{K}}_n$  be the operator from  $\mathbb{X}_n$  to  $\mathbb{X}_n$  such that its matrix representation under the basis  $e_{ij}, (i, j) \in \mathbb{U}_n$  is  $\tilde{A}_n$ . Then,  $\tilde{u}_n$  is a solution of the operator equation

$$(\lambda \mathcal{I} + \tilde{\mathcal{K}}_n) \tilde{u}_n = \mathcal{P}_n h. \tag{5.1}$$

We will estimate the error between the solution  $u$  of (2.14) and the solution  $\tilde{u}_n$  of (5.1) by studying the error between two operators  $\mathcal{K}_n$  and  $\tilde{\mathcal{K}}_n$ . To this end, we consider the difference between the matrix blocks  $A_{i'i}$  and  $\tilde{A}_{i'i}$  in matrix norms.

**Lemma 5.1.** *Suppose that  $\epsilon_{i'i}$  and  $\delta_{i'i} > 0$  are chosen according to (4.4) and (4.6), respectively. Then, there exists a positive constant  $c$  such that for all  $\sigma > 0$*

$$\|A_{i'i} - \tilde{A}_{i'i}\|_\infty \leq c(d_i d_{i'})^p \mu^{(i-i')/2} \sigma^{-2p} \epsilon_{i'i}$$

and

$$\|A_{i'i} - \tilde{A}_{i'i}\|_1 \leq c(d_i d_{i'})^p \mu^{(i'-i)/2} \sigma^{-2p} \epsilon_{i'i}.$$

**Proof.** We present the proof of the first estimate only, since the other one may be similarly obtained. By (4.2), the set

$$\mathbb{Z}(i'i'j') := \{j : j \in \mathbb{Z}_{w(i)} \text{ and } \text{dist}(S_{ij}, S_{i'j'}) > \delta_{i'i} \sigma\}$$

consists of the indices of nonzero entries in the  $(i', j')$ th row of  $A_{i'i} - \tilde{A}_{i'i}$ . By lemma 4.1, we have that

$$\|A_{i'i} - \tilde{A}_{i'i}\|_\infty = \max_{j' \in \mathbb{Z}_{w(i')}} \sum_{j \in \mathbb{Z}(i'i'j')} |a_{i'j',ij}| \leq c(d_i d_{i'})^p \mu^{(i+i')/2} \sigma^{-(2p+1)} \Delta_{i'i}, \tag{5.2}$$

where

$$\Delta_{i'i} := \max_{j' \in \mathbb{Z}_{w(i')}} \sum_{j \in \mathbb{Z}(i'ij')} \int_{S_{i'j'}} \int_{S_{ij}} \left| g^{(2p)} \left( \frac{x-y}{\sigma} \right) \right| dy dx.$$

By the definition of  $\mathbb{Z}(i'ij')$ , we observe that

$$\Delta_{i'i} \leq c \max_{j' \in \mathbb{Z}_{w(i')}} \int_{S_{i'j'}} \int_{E \setminus \cup \{S_{ij} : \text{dist}(S_{ij}, S_{i'j'}) \leq \delta_{i'i} \sigma\}} \left| g^{(2p)} \left( \frac{x-y}{\sigma} \right) \right| dy dx.$$

Using a change of variables, we obtain that  $\Delta_{i'i} \leq c\sigma \max_{j' \in \mathbb{Z}_{w(i')}} \text{meas}(S_{i'j'}) G_{2p}(\delta_{i'i})$ . Noting that  $\text{meas}(S_{i'j'}) = O(\mu^{-i'})$  and employing the definition of  $\delta_{i'i}$ , we conclude that  $\Delta_{i'i} \leq c\sigma \mu^{i'} \epsilon_{i'i}$ . Upon substituting the above inequality into (5.2), we have the first estimate of this lemma.  $\square$

For an integer  $\gamma \geq 0$ , we denote by  $\mathbb{H}^\gamma$  the Sobolev space of functions  $v$  for which  $\|v^{(\gamma)}\| < \infty$ , where  $v^{(\ell)}$  is the  $\ell$ th weak derivative of  $v$ . The space  $\mathbb{H}^\gamma$  is equipped with the norm

$$\|v\|_{\mathbb{H}^\gamma} = \left( \sum_{\ell=0}^{\gamma} \|v^{(\ell)}\|^2 \right)^{1/2}.$$

The following lemma gives an estimate of the residue of  $\mathcal{K}_n - \tilde{\mathcal{K}}_n$ . The proof of this lemma is adopted from that of theorem 3.3.4 in [7].

**Lemma 5.2.** *Let  $\alpha$  and  $\alpha'$  be real numbers satisfying the condition*

$$\alpha \geq \frac{p+\gamma}{2p}, \quad \alpha' \geq \frac{p+\gamma'}{2p}, \tag{5.3}$$

where  $\gamma, \gamma' \in \mathbb{Z}_{p+1}$ . Let  $u \in \mathbb{H}^\gamma(E)$  and  $v \in \mathbb{H}^{\gamma'}(E)$ . If  $\delta_{i'i}$  is chosen such that (4.6) holds, then there exists a positive number  $\theta$  such that for all positive integer  $n$  and for all  $\sigma > 0$

$$| \langle (\mathcal{K}_n - \tilde{\mathcal{K}}_n) \mathcal{P}_n u, \mathcal{P}_n v \rangle | \leq \theta a^{-2p} \mu^{-n(\gamma+\gamma')} n^\nu \|u\|_{\mathbb{H}^\gamma} \|v\|_{\mathbb{H}^{\gamma'}},$$

where  $\nu$  counts the number of equalities held in inequalities (5.3).

**Proof.** Since the subspace  $\mathbb{X}_n$  is a piecewise polynomial space, there exists a positive number  $c$  such that for all  $n \in \mathbb{N}$

$$\|(\mathcal{I} - \mathcal{P}_n)u\| \leq cd_n^\gamma \|u\|_{\mathbb{H}^\gamma}, \tag{5.4}$$

where  $\mathcal{P}_n$  is the orthogonal projection from  $\mathbb{X}$  onto  $\mathbb{X}_n$ . It follows from (5.4) for  $i \in \mathbb{N}_0$  that

$$\|(\mathcal{P}_i - \mathcal{P}_{i-1})u\| \leq \|(\mathcal{I} - \mathcal{P}_i)u\| + \|(\mathcal{I} - \mathcal{P}_{i-1})u\| \leq cd_{i-1}^\gamma \|u\|_{\mathbb{H}^\gamma}, \tag{5.5}$$

where  $\mathcal{P}_{-1} = 0$  and  $d_{-1} = 1$ .

Let  $\varepsilon_n := \langle (\mathcal{K}_n - \tilde{\mathcal{K}}_n) \mathcal{P}_n u, \mathcal{P}_n u' \rangle$ . Since

$$\mathcal{P}_n u = \sum_{i \in \mathbb{Z}_{n+1}} (\mathcal{P}_i - \mathcal{P}_{i-1})u, \tag{5.6}$$

we have that

$$\varepsilon_n = \sum_{i, i' \in \mathbb{Z}_{n+1}} \langle (\mathcal{K}_n - \tilde{\mathcal{K}}_n) (\mathcal{P}_i - \mathcal{P}_{i-1})u, (\mathcal{P}_{i'} - \mathcal{P}_{i'-1})v \rangle.$$

Since both  $\mathcal{P}_i - \mathcal{P}_{i-1}$  and  $\mathcal{P}_{i'} - \mathcal{P}_{i'-1}$  are projections and  $\mathcal{P}_{i'} - \mathcal{P}_{i'-1}$  is self-adjoint, we observe that

$$\varepsilon_n = \sum_{i, i' \in \mathbb{Z}_{n+1}} \langle [(\mathcal{P}_{i'} - \mathcal{P}_{i'-1})(\mathcal{K}_n - \tilde{\mathcal{K}}_n)(\mathcal{P}_i - \mathcal{P}_{i-1})](\mathcal{P}_i - \mathcal{P}_{i-1})u, (\mathcal{P}_{i'} - \mathcal{P}_{i'-1})v \rangle.$$

Using the Cauchy–Schwarz inequality, we conclude that

$$|\varepsilon_n| \leq \sum_{i,i' \in \mathbb{Z}_{n+1}} \|(\mathcal{P}_{i'} - \mathcal{P}_{i'-1})(\mathcal{K}_n - \tilde{\mathcal{K}}_n)(\mathcal{P}_i - \mathcal{P}_{i-1})\| \|(P_i - P_{i-1})u\| \|(P_{i'} - P_{i'-1})v\|.$$

It can be verified that

$$\|(\mathcal{P}_{i'} - \mathcal{P}_{i'-1})(\mathcal{K}_n - \tilde{\mathcal{K}}_n)(\mathcal{P}_i - \mathcal{P}_{i-1})\| = \|A_{i'i} - \tilde{A}_{i'i}\|_2.$$

From lemma 5.1, we have

$$\|A_{i'i} - \tilde{A}_{i'i}\|_2 \leq (\|A_{i'i} - \tilde{A}_{i'i}\|_1 \|A_{i'i} - \tilde{A}_{i'i}\|_\infty)^{1/2} \leq c(d_i d_{i'})^p \sigma^{-2p} \epsilon_{i'i}. \tag{5.7}$$

It follows from (5.5) and (5.7) that

$$|\varepsilon_n| \leq c(d_i d_{i'})^p \sigma^{-2p} \epsilon_{i'i} d_{i-1}^{\gamma'} d_{i'-1}^{\gamma'} \|u\|_{\mathbb{H}^{\gamma'}} \|v\|_{\mathbb{H}^{\gamma'}}.$$

Because of  $d_i \sim \mu^{-i}$  and the choice of  $\epsilon_{i'i}$ , we have that

$$|\varepsilon_n| \leq ca^{-2p} \sum_{i' \in \mathbb{Z}_{n+1}} \mu^{(p-2p\alpha'+\gamma')(n-i')} \sum_{i \in \mathbb{Z}_{n+1}} \mu^{(p-2p\alpha+\gamma)(n-i)} \mu^{-(\gamma+\gamma')n} \|u\|_{\mathbb{H}^{\gamma'}} \|v\|_{\mathbb{H}^{\gamma'}}.$$

By using the assumption (5.3) for the right-hand side of the last inequality, we obtain the desired estimate of this lemma. We rename the constant  $c$  to  $\theta$  since it will be specifically used later.  $\square$

Lemma 5.2 leads to the following result regarding the stability of the compressed Galerkin method. Since its proof is similar to that for [22, p 267], we skip the proof.

**Lemma 5.3.** *Let  $\alpha$  and  $\alpha'$  be real numbers satisfying the strict inequalities in (5.3). Let  $\theta$  be the constant in lemma 5.2. For a fixed  $\lambda$ , if  $a$  is chosen such that  $a \geq (2\theta/\lambda)^{1/2p}$  and  $\delta_{i'i}$  is chosen such that (4.6) holds, then there exists a positive constant  $q$  such that for all  $n \geq q$ ,  $v \in \mathbb{X}_n$  and  $\sigma > 0$*

$$\|(\lambda\mathcal{I} + \tilde{\mathcal{K}}_n)v\| \geq \frac{1}{2}\lambda\|v\|. \tag{5.8}$$

Lemma 5.3 implies that for all  $n \geq q$  and  $a \geq (2\theta/\lambda)^{1/2p}$ ,  $(\lambda\mathcal{I} + \tilde{\mathcal{K}}_n)^{-1}$  exist and are uniformly bounded. Thus, for sufficiently large  $n$ , the compressed Galerkin method has a unique solution.

Next, we estimate the condition number of the matrix  $\lambda I_n + \tilde{A}_n$ . For a nonsingular square matrix  $A$ , we denote by  $\text{cond}(A)$  its condition number.

**Theorem 5.4.** *Let  $\alpha$  and  $\alpha'$  be the real numbers satisfying the strict inequalities in (5.3). Let  $a$  be chosen such that  $a \geq (2\theta/\lambda)^{1/2p}$ . If  $\delta_{i'i}$  is chosen such that (4.6) holds, then there exist positive constants  $c$  and  $q$  such that for all  $n \geq q$  and  $\sigma > 0$*

$$\text{cond}(\lambda I_n + \tilde{A}_n) \leq (1 + \lambda^{-1})c,$$

where  $c$  is independent of the parameter  $\lambda$ .

The proof of this theorem based on lemmas 5.2 and 5.3 is similar to that for [22, theorem 4.5]. Thus, we omit it. It can be observed from theorem 5.4 that when the regularization parameter  $\lambda$  is small, the condition number of the matrix  $\lambda I_n + \tilde{A}_n$  can be large. In such a case, preconditioning the matrix  $\lambda I_n + \tilde{A}_n$  may be needed (see e.g. [11] for more information).

By using lemma 5.2 and following the proof of [22, theorem 4.6], we establish the convergence order of  $\tilde{u}_n$ , which is stated below.

**Theorem 5.5.** Let  $u \in \mathbb{H}^\gamma(E)$ ,  $\gamma \leq p$ , be the exact solution of equation (2.14). Let  $a$  be chosen such that  $a \geq (2\theta/\lambda)^{1/2p}$ . Suppose that  $\alpha$  and  $\alpha'$  satisfy the conditions

$$\alpha \geq \frac{p+\gamma}{2p} \quad \text{and} \quad \alpha' > \frac{1}{2}. \quad (5.9)$$

If  $\delta_{i_i}$  is chosen such that (4.6) holds, then there exist positive constants  $c$  and  $q$  such that for all  $n \geq q$  and  $\sigma > 0$

$$\|u - \tilde{u}_n\| \leq cs(n)^{-\gamma} \log^\nu s(n) \|u\|_{\mathbb{H}^\nu},$$

where  $\tilde{u}_n$  is the unique solution of (5.1) and  $\nu$  counts the number of the equalities in inequalities (5.9).

We close this section with comments on the regularity hypothesis on  $u$ . In our setting,  $u$  is a restriction of the true image in the  $x$  or  $y$  direction. The Sobolev regularity requirement imposed in the last theorem seems too restricted, since there is a large class of images whose restrictions to the  $x$  or  $y$  direction are not in a Sobolev space. Hence, there is a need to extend the above method and theory to a Besov setting because a Besov regularity hypothesis seems more realistic for images. This will be our future project.

## 6. A numerical quadrature scheme

The nonzero entries of the matrix  $\tilde{A}_n$  are defined in terms of double integrals whose integrands involve products of Gaussian functions and piecewise polynomials. Although Gaussian functions are sufficiently smooth, they behave like singular functions when the variance is small. Hence, integration of the double integrals requires a careful numerical treatment. The main purpose of this section is to develop numerical quadrature schemes for computing the double integrals.

Recall that the nonzero entries  $a_{i'j',ij}$  of  $\tilde{A}_n$  have the form

$$a_{i'j',ij} = \int_{S_{i'j'}} \int_{S_{ij}} g_\sigma(x-y) e_{ij}(y) dy e_{i'j'}(x) dx. \quad (6.1)$$

Since  $g_\sigma(x) := \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$  is not an elementary function, one has to resort to numerical methods for evaluating  $a_{i'j',ij}$ . We first partition  $S_{ij}$  according to the knots of the piecewise polynomial  $e_{ij}$ , that is

$$S_{ij} = \bigcup_{s \in \zeta(i,j)} \Lambda_s^i, \quad (6.2)$$

where  $\Lambda_s^i$  is a maximum interval on which  $e_{ij}$  is a polynomial and  $\text{meas}(\Lambda_s^i \cap \Lambda_t^i) = 0$ , if  $s \neq t$ . By the construction of the basis functions  $e_{ij}$ , the cardinality of  $\zeta(i, j)$  is bounded by a positive constant. We then rewrite  $a_{i'j',ij}$  according to the partitions of  $S_{ij}$  and  $S_{i'j'}$  as

$$a_{i'j',ij} = \sum_{t \in \zeta(i',j')} \sum_{s \in \zeta(i,j)} \int_{\Lambda_t^{i'}} \int_{\Lambda_s^i} g_\sigma(x-y) e_{ij}(y) dy e_{i'j'}(x) dx. \quad (6.3)$$

By using the mapping  $(x, \xi) = (x, x-y)$ , we change the  $xy$ -plane to the  $x\xi$ -plane. A similar mapping was used in [12] in dealing with the double integral appeared in the Galerkin method. We write  $\Lambda_s^i := [y_1, y_2]$  and  $\Lambda_t^{i'} := [x_1, x_2]$ . The rectangle domain  $[x_1, x_2] \times [y_1, y_2]$  is mapped by the above mapping onto the parallelogram domain with vertices  $(x_1, x_1 - y_1)$ ,  $(x_2, x_2 - y_1)$ ,  $(x_1, x_1 - y_2)$  and  $(x_2, x_2 - y_2)$ . The image of  $[x_1, x_2] \times [y_1, y_2]$  under the

mapping can be described by the set  $\{(x, \xi) : \xi \in I_{st} = [x_1 - y_2, x_2 - y_1], x \in D_{st\xi}\}$ , where  $D_{st\xi}$  is the interval varying with  $\xi \in I_{st}$ . Hence,

$$\int_{\Lambda_i^{j'}} \int_{\Lambda_s^i} g_\sigma(x-y)e_{ij}(y) dy e_{i'j'}(x) dx = \int_{I_{st}} g_\sigma(\xi) \int_{D_{st\xi}} e_{i'j'}(x)e_{ij}(x-\xi) dx d\xi. \quad (6.4)$$

For  $\xi \in I_{st}$  we define a function by the integral

$$h_{i'j',ij}(\xi) := \int_{D_{st\xi}} e_{i'j'}(x)e_{ij}(x-\xi) dx, \quad \xi \in I_{st}. \quad (6.5)$$

Since both  $e_{i'j'}(\cdot)$  and  $e_{ij}(\cdot - \xi)$  are the polynomials on  $D_{st\xi}$ , for each  $\xi \in I_{st}$ ,  $h_{i'j',ij}(\xi)$  can be evaluated exactly. Defining a function  $f_{i'j',ij}(\xi) := g_\sigma(\xi)h_{i'j',ij}(\xi)$ ,  $\xi \in I_{st}$ , we note that computing the integral in (6.4) reduces to evaluating the integral

$$\int_{I_{st}} f_{i'j',ij}(\xi) d\xi. \quad (6.6)$$

Understanding the interval  $D_{st\xi}$  is crucial for computing the above integral. For each  $\xi \in I_{st}$ , we denote by  $c(\xi)$  and  $d(\xi)$  the left and right endpoints of the interval  $D_{st\xi}$ , respectively, that is

$$D_{st\xi} := [c(\xi), d(\xi)].$$

It is ready to verify that both  $c(\xi)$  and  $d(\xi)$  are continuous piecewise linear functions of  $\xi$ . Specifically, we have that

$$c(\xi) = \begin{cases} x_1, & \xi \in [x_1 - y_2, x_1 - y_1] \\ \xi + y_1, & \xi \in [x_1 - y_1, x_2 - y_1] \end{cases} \quad (6.7)$$

and

$$d(\xi) = \begin{cases} \xi + y_2, & \xi \in [x_1 - y_2, x_2 - y_2] \\ x_2, & \xi \in [x_2 - y_2, x_2 - y_1]. \end{cases} \quad (6.8)$$

Due to the representations (6.7) and (6.8) of  $c(\xi)$  and  $d(\xi)$ , the function  $h_{i'j',ij}$  is a continuous piecewise polynomial.

Next, we describe a numerical quadrature for computing (6.6) according to the standard deviation  $\sigma$  of the Gaussian function involved in the integrand. To this end, we let  $\Lambda := [-1, 1]$  and note that when  $x$  and  $y$  lie on the interval  $E$ , we have that  $\xi = x - y \in \Lambda$ . We observe that the values of high order derivatives of  $g_\sigma(\xi)$  decay rapidly when  $\xi$  goes away from the origin. This motivates us to subdivide the interval  $\Lambda$  into nonuniform subintervals according to the change of the derivatives of  $g_\sigma$  and apply numerical quadratures with precision of different degrees to different subintervals. Similar ideas for integration were used in [8, 16, 27].

Let  $\epsilon > 0$  be a given fixed number and  $m$  be a positive integer. For the standard deviation  $\sigma$ , we choose  $\beta$  such that

$$1 - \beta < \sigma\beta^{1+1/\epsilon}. \quad (6.9)$$

Clearly, there exists a  $\beta \in (0, 1)$  satisfying inequality (6.9). We then let

$$t_0 := 0, \quad t_\iota := \beta^{m-\iota}, \quad \iota \in \mathbb{Z}_{m+1} \setminus \{0\} \quad (6.10)$$

and choose a sequence of nonnegative integers

$$k_\iota := \lfloor \epsilon \iota \rfloor + 1, \quad \iota \in \mathbb{Z}_m, \quad (6.11)$$

where  $\lfloor x \rfloor$  is the largest integer less than or equal to  $x$ . We define the set of nodes by

$$\Pi_{\beta,m} := \{\pm t_\iota : \iota \in \mathbb{Z}_{m+1}\} \quad (6.12)$$



and we let

$$T_{st} := (\Pi_{\beta,m} \cap I_{st}) \cup \{x_1 - y_1, x_1 - y_2, x_2 - y_1, x_2 - y_2\}. \quad (6.13)$$

Clearly, the set  $T_{st}$  forms a partition for the interval  $I_{st}$ . For every two consecutive points  $\tau_1 < \tau_2$  in the set  $T_{st}$ , there exists  $t \in \mathbb{Z}_m$  such that  $[\tau_1, \tau_2] \subseteq [t, t+1]$  or  $[\tau_1, \tau_2] \subseteq [-t+1, -t]$ . In either case, we compute an approximation of the integral

$$\int_{\tau_1}^{\tau_2} f_{i'j',ij}(\xi) d\xi \quad (6.14)$$

by using a numerical quadrature with a precision of order  $k_t$ . For each subinterval  $[\tau_1, \tau_2]$  of  $I_{st}$ , we compute an approximate value of the integral (6.14) and sum all such values over all possible subintervals defined by the partition  $T_{st}$ . This leads to an approximate value of  $a_{i'j',ij}$ .

We describe below the algorithm for evaluating entries of  $\tilde{A}_n$  by summarizing the above discussion.

**Algorithm 1** (Adaptive numerical quadrature scheme). *The following steps are used to evaluate the entry  $a_{i'j',ij}$  of  $\tilde{A}_n$  in (6.1).*

*Step 1. For a fixed parameter  $\epsilon$  and a positive integer  $m$ , select  $\beta$  according to (6.9), choose a sequence of nonnegative integers  $k_t$  using (6.11) and define the set  $\Pi_{\beta,m}$  of nodes according to (6.12).*

*Step 2. Find the partitions  $\{\Lambda_s^i\}$  of the supports  $S_{ij}$  of  $e_{ij}$  and the partitions  $\{\Lambda_t^{i'}\}$  of the supports  $S_{i'j'}$  of  $e_{i'j'}$  using (6.2). For  $\Lambda_s^i = [y_1, y_2]$  and  $\Lambda_t^{i'} = [x_1, x_2]$ , define  $I_{st} = [x_1 - y_2, x_2 - y_1]$  and generate the set  $T_{st}$  using (6.13).*

*Step 3. Evaluate  $a_{i'j',ij}$  using (6.4), (6.3) and the approximation of (6.14).*

We denote by  $E_{m,k}(f_{i'j',ij})$  the error of the approximation of  $a_{i'j',ij}$ . In the next theorem, we present an estimate of the error.

**Theorem 6.1.** *Let  $\epsilon > 0$  be a given fixed number and  $m$  be a positive integer. Suppose that  $\sigma < 1/2$ . Let  $\beta$  be a number satisfying (6.9). If  $p$  is the order of the polynomials in  $\mathbb{X}_0$ , then there exists a positive constant  $c$  such that for all  $m, \sigma$  and  $\beta$ ,*

$$|E_{m,k}(f_{i'j',ij})| \leq c\mu^{(2p-1)(i+i'-1)}\beta^m.$$

The proof of this theorem requires four technical lemmas. The first one presents an estimate of high order derivatives of  $h_{i'j',ij}$ .

**Lemma 6.2.** *If  $q \geq 0$  and  $\xi \in I_{s,t} \setminus \{x_1 - y_1, x_2 - y_2\}$ , then there exists a positive constant  $c$  such that for all  $i$  and  $i'$  with  $i \leq i'$ ,*

$$|h_{i'j',ij}^{(q)}(\xi)| \leq c\mu^{q(i+i'-1)}.$$

**Proof.** For simplicity, we write  $h := h_{i'j',ij}$ . For  $\xi \in I_{s,t} \setminus \{x_1 - y_1, x_2 - y_2\}$ , we have from (6.5) that

$$\begin{aligned} h^{(1)}(\xi) &= e_{i'j'}(d(\xi))e_{ij}(d(\xi) - \xi)d^{(1)}(\xi) - e_{i'j'}(c(\xi))e_{ij}(c(\xi) - \xi)c^{(1)}(\xi) \\ &\quad - \int_{c(\xi)}^{d(\xi)} e_{i'j'}(x)e_{ij}^{(1)}(x - \xi) dx. \end{aligned} \quad (6.15)$$

From (6.7) and (6.8), we know

$$c^{(1)}(\xi) = \begin{cases} 0, & \xi \in [x_1 - y_2, x_1 - y_1] \\ 1, & \xi \in (x_1 - y_1, x_2 - y_1], \end{cases} \quad d^{(1)}(\xi) = \begin{cases} 1, & \xi \in [x_1 - y_2, x_2 - y_2] \\ 0, & \xi \in (x_2 - y_2, x_2 - y_1], \end{cases}$$

and clearly

$$c^{(1)}(\xi)(c^{(1)}(\xi) - 1) = d^{(1)}(\xi)(d^{(1)}(\xi) - 1) = 0.$$

By using these facts, we differentiate (6.15) recursively and obtain that

$$h^{(q)}(\xi) = \sum_{r=0}^{q-1} (-1)^{q-1-r} [e_{i'j'}^{(r)}(d(\xi))e_{ij}^{(q-1-r)}(d(\xi) - \xi)(d^{(1)}(\xi))^{r+1} - e_{i',j'}^{(r)}(c(\xi))e_{i,j}^{(q-1-r)}(c(\xi) - \xi)(c^{(1)}(\xi))^{r+1}] + (-1)^q \int_{c(\xi)}^{d(\xi)} e_{i'j'}(x)e_{ij}^{(q)}(x - \xi) dx. \tag{6.16}$$

Recalling that  $p$  is the order of the piecewise polynomials in  $\mathbb{W}_1 = \text{span}\{e_{1,j} : j \in \mathbb{Z}_{(p-1)\mu}\}$ , we see that the quantity

$$\Delta := \max \{ \|e_{1j}^{(q)}\|_{\infty} : q \leq p, j \in \mathbb{Z}_{(p-1)\mu} \}$$

is finite. Furthermore, from (3.4), we have that

$$|e_{ij}^{(q)}(t)| = |\mu^{(i-1)/2} \mu^{q(i-1)} e_{1j}^{(q)}(\phi_s^{(-1)}(t))| \leq \Delta \mu^{q(i-1)+i/2}, \quad \mathbf{s} \in \mathbb{Z}_{\mu}^{i-1}.$$

Since the length of the interval  $[c(\xi), d(\xi)]$  is bounded by  $|S_{i'j'}| \leq c\mu^{-i'}$  with a constant  $c$  independent of  $i'$  and  $j'$ , we have from (6.16) and the condition  $i \leq i'$  that

$$|h^{(q)}(\xi)| \leq 2\Delta^2 \sum_{t=0}^{q-1} \mu^{(q-1-t)(i-1)+i/2} \mu^{t(i'-1)+i'/2} + c\Delta^2 \mu^{q(i-1)+i/2} \mu^{i'/2} \mu^{-i'} \leq c\mu^{q(i+i'-1)},$$

proving the lemma. □

In the next three lemmas, we present some useful properties of the Gaussian function.

**Lemma 6.3.** *Let  $q$  be a positive integer.*

(1) *If the nonnegative roots of  $H_{2q+1}$  are denoted by  $\{x_i : i \in \mathbb{Z}_{q+1}\}$  with  $0 = x_0 < x_1 < \dots < x_q$ , then*

$$|g^{(2q)}(x_0)| > |g^{(2q)}(x_1)| > \dots > |g^{(2q)}(x_q)| \quad \text{and} \quad |g^{(2q)}(x_q)| > |g^{(2q)}(x)| \quad \text{for } x > x_q.$$

*Furthermore,  $|g^{(2q)}|$  assumes its global maximum at  $x_0$ .*

(2) *If the positive roots of  $H_{2q}$  are denoted by  $\{x_i : i \in \mathbb{Z}_{q+1} \setminus \{0\}\}$  with  $x_1 < \dots < x_q$ , then*

$$|g^{(2q-1)}(x_1)| > |g^{(2q-1)}(x_2)| > \dots > |g^{(2q-1)}(x_q)| \quad \text{and} \\ |g^{(2q-1)}(x_q)| > |g^{(2q-1)}(x)| \quad \text{for } x > x_q.$$

*Furthermore,  $|g^{(2q-1)}|$  assumes its global maximum at  $x_1$ .*

**Proof.** We present the proof for the first part of the lemma only.

Using the recursive relation of derivatives of the Gaussian function  $g$ , for every nonnegative integer  $j$  we have that

$$\frac{d}{dx} (g^{(j)}(x))^2 = -\frac{1}{j+1} \frac{d}{dx} (g^{(j+1)}(x))^2 - \frac{2x}{j+1} (g^{(j+1)}(x))^2. \tag{6.17}$$

Integrating (6.17) from  $x_i$  to  $x_{i+1}$  with  $j := 2q$  yields

$$(g^{(2q)}(x_{i+1}))^2 - (g^{(2q)}(x_i))^2 = -\int_{x_i}^{x_{i+1}} \frac{2x}{2q+1} (g^{(2q+1)}(x))^2 dx < 0, \quad i \in \mathbb{Z}_q.$$

This leads to  $|g^{(2q)}(x_i)| > |g^{(2q)}(x_{i+1})|$ . By (6.17), for  $x > x_q$ , we have that

$$(g^{(2q)}(x))^2 - (g^{(2q)}(x_q))^2 = -\frac{1}{2q+1}(g^{(2q+1)}(x))^2 - \int_{x_q}^x \frac{2s}{2q+1}(g^{(2q+1)}(s))^2 ds < 0,$$

which implies that  $|g^{(2q)}(x_q)| > |g^{(2q)}(x)|$  for  $x > x_q$ . Since  $g^{(2q)}$  is an even function and has  $x_0, x_1, \dots, x_q$  as its critical points in the positive  $x$ -axis, we can further conclude that  $\|g^{(2q)}\|_\infty = |g^{(2q)}(0)| = \frac{(2q)!}{q!2^q\sqrt{2\pi}}$ .  $\square$

**Lemma 6.4.** *There exists a positive constant  $c$  such that for all  $\ell \in \mathbb{Z}_+$ ,*

$$\|g^{(\ell)}\|_\infty \leq c \|g^{(\ell+1)}\|_\infty.$$

**Proof.** We first consider the case  $\ell = 2q - 1$  with  $q$  being a positive integer. From lemma 6.3, it suffices to show  $|g^{(2q-1)}(x_1)| \leq c |g^{(2q)}(0)|$ , where  $x_1$  is the smallest positive root of  $H_{2q}$ .

Because  $H_{2q-1}(0) = 0$  and  $H_{2q-1}^{(1)}(x) = (2q - 1)H_{2q-2}(x)$ , one has that

$$H_{2q-1}(x_1) = \int_0^{x_1} H_{2q-1}^{(1)}(x) dx = (2q - 1) \int_0^{x_1} H_{2q-2}(x) dx.$$

This result together with the fact  $\|g^{(2q-2)}\|_\infty = \frac{(2q-2)!}{(q-1)!2^{q-1}\sqrt{2\pi}}$  gives

$$|H_{2q-1}(x_1)| \leq (2q - 1)x_1 e^{x_1^2/2} \|g^{(2q-2)}\|_\infty = x_1 e^{x_1^2/2} |H_{2q}(0)| = x_1 e^{x_1^2/2} |g^{(2q)}(0)|.$$

Since the smallest positive zero of  $H_q$  is bounded for all  $q$  [23], we have that

$$|g^{(2q-1)}(x_1)| \leq x_1 |g^{(2q)}(0)| \leq c |g^{(2q)}(0)|,$$

where  $c$  is a constant independent of  $q$ .

Likewise, we can handle the case  $\ell = 2q$ .  $\square$

Lemmas 6.3 and 6.4 yield immediately the following result.

**Lemma 6.5.** *If  $\ell$  and  $q$  are nonnegative integers and  $\ell - q \geq 1$ , then there exists a positive constant  $c$  such that for all  $\ell, q$  and for all  $x \in \mathbb{R}$*

$$\left| \frac{1}{\ell!} \binom{\ell}{q} (g^{(\ell-q+1)}(x) + (\ell - q)g^{(\ell-q-1)}(x)) \right| \leq c.$$

We are now in the position to prove theorem 6.1.

**Proof of theorem 6.1.** Since  $a_{i'j',ij} = a_{ij,i'j'}$ , we consider only the case  $i \leq i'$  in this proof. We know that  $f_{i'j',ij} = g_\sigma h_{i'j',ij}$ . For simplicity, we write  $h := h_{i'j',ij}$ . The error  $E_{m,k_i,\tau_1,\tau_2}(f_{i'j',ij})$  of the numerical quadrature for the integral (6.14) has the bound

$$E_{m,k_i,\tau_1,\tau_2}(f_{i'j',ij}) \leq \frac{c}{k_i!} |(g_\sigma h)^{(k_i)}(\eta)| (\tau_2 - \tau_1)^{k_i+1}, \tag{6.18}$$

where  $\eta$  is a number in  $(\tau_1, \tau_2)$ .

By the Leibnitz formula for the  $k_i$ th derivative of a product of two functions, we obtain that

$$\frac{1}{k_i!} |(g_\sigma h)^{(k_i)}(\eta)| \leq \frac{1}{k_i!} \sum_{q=0}^{\min\{k_i, 2p-1\}} \binom{k_i}{q} |g_\sigma^{(k_i-q)}(\eta)| |h^{(q)}(\eta)|. \tag{6.19}$$

Note that  $g_\sigma^{(k_i-q)}(\eta) = \sigma^{-(k_i-q+1)} g^{(k_i-q)}\left(\frac{\eta}{\sigma}\right)$ . For  $k_i - q = 0$ , we have that

$$\left|g_\sigma^{(k_i-q)}(\eta)\right| = \left|\eta^{-1}\left(\frac{\eta}{\sigma}\right)g\left(\frac{\eta}{\sigma}\right)\right| \leq c|\eta|^{-1}.$$

For  $k_i - q \geq 1$ , the recursive relation of derivatives of the Gaussian function gives

$$\frac{\eta}{\sigma}g^{(k_i-q)}\left(\frac{\eta}{\sigma}\right) = -g^{(k_i-q+1)}\left(\frac{\eta}{\sigma}\right) - (k_i - q)g^{(k_i-q-1)}\left(\frac{\eta}{\sigma}\right).$$

Hence, we conclude from lemma 6.5 that

$$\frac{1}{k_i!}\binom{k_i}{q}\left|g_\sigma^{(k_i-q)}(\eta)\right| \leq c|\eta|^{-1}\sigma^{-(k_i-q)}, \tag{6.20}$$

where  $c$  is a positive constant independent of  $k_i$  or  $q$ . Therefore, from lemma 6.2, equations (6.18), (6.19), (6.20) and the fact of

$$\sum_{q=0}^{\min\{k_i, 2p-1\}} \sigma^{-(k_i-q)} \leq \frac{\sigma^{-k_i}}{1-\sigma},$$

we have the estimate

$$E_{m,k_i,\tau_1,\tau_2}(f_{i'j',ij}) \leq \frac{c\sigma^{-k_i}}{1-\sigma}(t_{i+1} - t_i)^{k_i}(\tau_2 - \tau_1)|\eta|^{-1}\mu^{(2p-1)(i+i'-1)}. \tag{6.21}$$

For  $\iota \neq 0$ , from the definition of  $t_i$  in (6.10) and the fact of  $|\eta|^{-1}t_i < 1$ , we get

$$(t_{i+1} - t_i)^{k_i}|\eta|^{-1} \leq \left(\frac{1}{\beta} - 1\right)^{k_i} t_i^{k_i-1}.$$

Since  $\iota/k_i \leq 1/\epsilon$  and condition (6.9), we obtain

$$\frac{(1-\beta)^{k_i}}{\beta^{k_i+\iota}}\sigma^{-k_i} \leq \left(\frac{1-\beta}{\sigma\beta^{1+1/\epsilon}}\right)^{k_i} \leq 1.$$

We denote by  $E_{m,k_i}(f_{i'j',ij})$  the total error of the numerical quadrature for the integrals in (6.3) whose integration domains lie in  $[t_i, t_{i+1}]$  or  $[-t_{i+1}, -t_i]$ . From (6.21) and the above two inequalities, we have

$$E_{m,k_i}(f_{i'j',ij}) \leq \frac{c}{1-\sigma}\left(\frac{1}{\beta} - 1\right)t_i^{k_i-1}\beta^m\mu^{(2p-1)(i+i'-1)}. \tag{6.22}$$

Moreover,

$$\sum_{i \in \mathbb{Z}_m} t_i^{k_i-1} = \sum_{i=0}^{\lfloor 1/\epsilon \rfloor} t_i^{k_i-1} + \sum_{i=\lfloor 1/\epsilon \rfloor+1}^{m-1} t_i^{k_i-1} \leq \lfloor 1/\epsilon \rfloor + \sum_{i=0}^{m-1} t_i \leq \lfloor 1/\epsilon \rfloor + \frac{\beta}{1-\beta}.$$

Since  $\sigma < 1/2$  and  $\beta$  satisfies (6.9), we derive that  $1/(1-\sigma) < 2$  and  $\beta > 1/2$ . From the above inequality and (6.22), we have that

$$E_{m,k}(f_{i'j',ij}) = \sum_{i=0}^{m-1} E_{m,k_i}(f_{i'j',ij}) \leq c(\lfloor 1/\epsilon \rfloor + 1)\mu^{(2p-1)(i+i'-1)}\beta^m,$$

completing the proof. □

## 7. Analysis of the fully discrete method

This section is devoted to a complete analysis of the fully discrete method developed in the previous sections. Specifically, we will prove that the method has the nearly optimal convergence order and it requires only linear (up to a logarithmic factor) number of functional evaluations for generating the coefficient matrix for the linear system.

Recall that  $\tilde{A}_n = [\tilde{A}_{i'i}]$  is obtained by applying our proposed truncation strategy (4.2) for the matrix  $A_n = [A_{i'i}]$ . The entries  $a_{i'j',ij}$  of  $\tilde{A}_{i'i}$  are numerically evaluated by the numerical quadrature scheme described in the last section. The resulting matrix is denoted by  $\tilde{\tilde{A}}_n = [\tilde{\tilde{A}}_{i'i}]$ . The operator  $\tilde{\mathcal{K}}_n$  associated with  $\tilde{\tilde{A}}_n$  was defined in section 5 and the function  $\tilde{u}_n$  satisfies an operator equation (5.1). We define  $\tilde{\tilde{\mathcal{K}}}_n$  to be the operator from  $\mathbb{X}_n$  to  $\mathbb{X}_n$  such that its matrix representation under the basis  $e_{ij}$ ,  $(i, j) \in \mathbb{U}_n$  is  $\tilde{\tilde{A}}_n$ . Associated with the operator  $\tilde{\tilde{\mathcal{K}}}_n$ , we consider the following operator equation:

$$(\lambda \mathcal{I} + \tilde{\tilde{\mathcal{K}}}_n) \tilde{u}_n = \mathcal{P}_n f, \quad (7.1)$$

where

$$\tilde{u}_n := \sum_{(i,j) \in \mathbb{U}_n} \tilde{u}_{ij} e_{ij}$$

and  $\tilde{u}_{ij}$ ,  $(i, j) \in \mathbb{U}_n$ , denote the components of the solution of the fully discrete system. We will estimate the error between the solution  $u$  of (2.14) and the solution  $\tilde{u}_n$  of (7.1) by studying the error between two operators  $\tilde{\mathcal{K}}_n$  and  $\tilde{\tilde{\mathcal{K}}}_n$ . To establish our main result, we estimate the 1- and  $\infty$ -norms of the error  $\tilde{\tilde{A}}_{i'i} - \tilde{A}_{i'i}$ . Theorem 6.1 immediately leads to the following result.

**Lemma 7.1.** *Let  $\epsilon > 0$  be a given fixed number and  $m$  be the positive integer used in the numerical quadrature. Suppose that  $\sigma < 1/2$ . If  $\beta$  satisfies (6.9) and  $p$  is the order of the polynomials in  $\mathbb{X}_0$ , then there exists a positive constant  $c$  such that for all  $m, \sigma, \beta$  and for all  $i, i' \in \mathbb{Z}_{n+1}$  and  $n \in \mathbb{N}$ ,*

$$\|\tilde{\tilde{A}}_{i'i} - \tilde{A}_{i'i}\|_{\infty} \leq c \mu^i \mu^{(2p-1)(i+i'-1)} \beta^m$$

and

$$\|\tilde{\tilde{A}}_{i'i} - \tilde{A}_{i'i}\|_1 \leq c \mu^{i'} \mu^{(2p-1)(i+i'-1)} \beta^m.$$

To ensure that the approximation solution  $\tilde{u}_n$  preserves the nearly optimal convergence order, the integer  $m$  used in the numerical quadrature (also in lemma 7.1) should vary for different blocks  $\tilde{\tilde{A}}_{i'i}$ . The integer  $m$  for the block  $\tilde{\tilde{A}}_{i'i}$  is denoted by  $m_{i'i}$ .

**Theorem 7.2.** *Let  $\epsilon > 0$  be a given fixed number. Choose  $m_{i'i}$  to satisfy*

$$m_{i'i} \geq \frac{\log \mu}{\log \beta} (-(2p-1)(i+i'-1) - (i+i')/2 + (p-2p\alpha)(n-i) + (p-2p\alpha')(n-i')). \quad (7.2)$$

*Let  $p$  be the order of the polynomials in  $\mathbb{X}_0$ . Suppose that the exact solution  $u$  of equation (2.14) is in  $\mathbb{H}^\gamma(E)$  with  $\gamma \leq p$ . If  $\alpha$  and  $\alpha'$  satisfy condition (5.9), then there exists a positive integer  $q$ , a positive constant  $c$  such that for all  $n \geq q$ ,  $0 < \sigma < 1/2$  and  $\beta$  satisfying (6.9):*

$$\|u - \tilde{u}_n\| \leq cs(n)^{-\gamma} \log^v(s(n)) \|u\|_{\mathbb{H}^\gamma}, \quad (7.3)$$

where  $v$  counts the number of the equalities in inequalities (5.9).

**Proof.** Note that

$$\|u - \tilde{u}_n\| \leq \|u - \tilde{u}_n\| + \|\tilde{u}_n - \tilde{u}_n\|,$$

where  $u$ ,  $\tilde{u}_n$  and  $\tilde{u}_n$  are solutions of (2.14), (5.1) and (7.1), respectively. The estimate of  $\|u - \tilde{u}_n\|$  has been provided in theorem 5.5. By following the proof of [22, theorem 4.6], we only need that there exists a positive constant  $c$  such that for all  $i$  and  $i'$ ,

$$\|\tilde{A}_{i'i} - \tilde{A}_{i'i}\|_2 \leq c\mu^{-p(i+i')}\mu^{-2p(-n+\alpha(n-i)+\alpha'(n-i'))}. \tag{7.4}$$

This is obtained from lemma 7.1 and the choice (7.2) of  $m_{i'i}$ . □

Now we turn to estimating the computational complexity for generating the matrix  $\tilde{A}_n$  in terms of the total number  $\mathcal{M}_n$  of functional evaluations used in computing all nonzero entries of the matrix. We denote by  $\mathcal{M}_{i'i}$  the number of functional evaluations for computing the nonzero entries of  $\tilde{A}_{i'i}$ . Thus,  $\mathcal{M}_n = \sum_{i',i \in \mathbb{Z}_{n+1}} \mathcal{M}_{i'i}$ . The following theorem gives an estimate of  $\mathcal{M}_n$ .

**Theorem 7.3.** *If  $\delta_{i'i}$  is chosen according to (4.6) with  $\alpha \leq 1$  and  $\alpha' \leq 1$ , then there exists a positive constant  $c$  such that for all  $\sigma$  and  $\delta_{i'i}$ ,  $i, i' \in \mathbb{Z}_{n+1}$ , and for all  $n \in \mathbb{N}$*

$$\mathcal{M}_n \leq cs(n) \log^{\vartheta(\alpha, \alpha') + 1} s(n),$$

where  $\vartheta(\alpha, \alpha') = 2$  for  $(\alpha, \alpha') = (1, 1)$  and 1 otherwise.

**Proof.** From (6.1), (6.3) and (6.4), the entry  $a_{i'j',ij}$  of  $A_{i'i}$  is

$$a_{i'j',ij} = \sum_{t \in \zeta(i',j')} \sum_{s \in \zeta(i,j)} \int_{I_{st}} g_\sigma(\xi) h_{i'j',ij}(\xi) d\xi,$$

where  $\zeta(i', j')$ ,  $\zeta(i, j)$ ,  $I_{st}$ ,  $D_{st\xi}$  and  $h_{i'j',ij}$  are defined in the preceding section. Since both  $e_{i'j'}$  and  $e_{ij}$  are polynomials of degree less than  $p$  on the set  $D_{st\xi}$ , for a fixed  $\xi$ , the total number of functional evaluations for computing the integral

$$h_{i'j',ij}(\xi) = \int_{D_{st\xi}} e_{i'j'}(x) e_{ij}(x + \xi) dx$$

is bounded by a positive constant dependent only on  $p$ .

We denote by  $\mathcal{M}(h)$  and  $\mathcal{M}_{i'j',i}$ , respectively, the number of functional evaluations used in the numerical quadrature of the integral  $h$  and the number of functional evaluations in computing the components of the  $i'j'$ th row of  $\tilde{A}_{i'i}$ . That is,

$$\mathcal{M}_{i'j',i} \leq c \sum_{j \in \mathbb{Z}_{i'j',i}} \sum_{t \in \zeta(i',j')} \sum_{s \in \zeta(i,j)} \mathcal{M} \left( \int_{I_{st}} g_\sigma(\xi) h_{i'j',ij}(\xi) d\xi \right), \tag{7.5}$$

where  $\mathbb{Z}_{i'j',i} := \{j \in \mathbb{Z}_{w(i)} : \text{dist}(S_{i'j'}, S_{ij}) \leq \delta_{i'i}\sigma\}$  and  $c$  is a constant dependent only on  $p$ .

For a fixed  $t \in \zeta(i', j')$ , we have that

$$\sum_{j \in \mathbb{Z}_{i'j',i}} \sum_{s \in \zeta(i,j)} \mathcal{M} \left( \int_{I_{st}} g_\sigma(\xi) h_{i'j',ij}(\xi) d\xi \right) \leq c \sum_{j \in \mathbb{Z}_{i'j',i}} \sum_{s \in \zeta(i,j)} \sum_{[\tau_1, \tau_2] \subset I_{st}} k_t,$$

where  $I_{stt} = I_{st} \cap ([t, t_{i+1}] \cup [-t_{i+1}, t])$  and  $[\tau_1, \tau_2]$  is the interval that appeared in (6.14).

Since the number of integrals over the interval  $[\tau_1, \tau_2]$  is bounded by a constant independent of  $j \in \mathbb{Z}_{i'j',i}$  and  $s \in \zeta(i, j)$  for fixed  $(i', j')$  and  $t$ , we get

$$\sum_{j \in \mathbb{Z}_{i'j',i}} \sum_{s \in \zeta(i,j)} \mathcal{M} \left( \int_{I_{st}} g_\sigma(\xi) h_{i'j',ij}(\xi) d\xi \right) \leq c \sum_{[\tau_1, \tau_2] \subseteq \bigcup_{j \in \mathbb{Z}_{i'j',i}} \bigcup_{s \in \zeta(i,j)} I_{st}} k_t.$$

The compression strategy and supports of functions  $e_{i'j'}$  and  $e_{ij}$  ensure that

$$\bigcup_{j \in \mathbb{Z}_{i'j',i}} \bigcup_{s \in \zeta(i,j)} I_{st} \subseteq [-(\delta_{i'i}\sigma + d_i + d_{i'}), \delta_{i'i}\sigma + d_i + d_{i'}] \cap [-1, 1].$$

Associated with the geometric partition in (6.10), there is a unique  $L \in \mathbb{Z}_m$  such that

$$t_L < \min\{\delta_{i'i}\sigma + d_i + d_{i'}, 1\} \quad \text{and} \quad t_{L+1} \geq \min\{\delta_{i'i}\sigma + d_i + d_{i'}, 1\}.$$

Therefore,

$$\sum_{j \in \mathbb{Z}_{i'j',i}} \sum_{s \in \zeta(i,j)} \mathcal{M} \left( \int_{I_{st}} g_\sigma(\xi) h_{i'j',ij}(\xi) d\xi \right) \leq c \sum_{i=0}^L k_i \mu^i (\beta^{m_{i'i} - t - 1} - \beta^{m_{i'i} - t}).$$

Using the facts that  $\text{card}(\zeta(i', j')) \leq c$  for all  $j' \in \mathbb{Z}_{w(i')}$  and  $k_i < \epsilon_l + 1$ , we have from (7.5) that

$$\mathcal{M}_{i',j',i} \leq c \left( \frac{1}{\beta} - 1 \right) \mu^i \sum_{i=0}^L (\epsilon_l + 1) \beta^{m_{i'i} - t} \leq c \mu^i \sum_{i=0}^L (\epsilon_l + 1) \beta^{m_{i'i} - t}.$$

By theorem 7.2, we can set  $m_{i'i} = cn$  for a constant  $c$  which is independent of  $i$  and  $i'$ . From the above inequality, we obtain that  $\mathcal{M}_{i',j',i} \leq c \mu^i n \beta^{m_{i'i} - L}$ , which implies that  $\mathcal{M}_{i'i} \leq cn \mu^{i'+i} \beta^{m_{i'i} - L}$ . Since  $\beta^{m_{i'i} - L} = t_L < \delta_{i'i}\sigma + d_i + d_{i'}$ , we observe that  $\mathcal{M}_{i'i} \leq cn \mu^{i'+i} (\delta_{i'i}\sigma + d_i + d_{i'})$ . By using (4), we further have that  $\mathcal{M}_{i'i} \leq cn \mu^{i'+i} (\mu^{-n+\alpha(n-i)+\alpha'(n-\alpha')} + \mu^{-i} + \mu^{-i'})$ . Following the proof of theorem 4.2, we conclude that

$$\mathcal{M}_n \leq cn \sum_{i' \in \mathbb{Z}_{n+1}} \sum_{i \in \mathbb{Z}_{n+1}} \mu^{i'+i} (\mu^{-n+\alpha(n-i)+\alpha'(n-\alpha')} + \mu^{-i} + \mu^{-i'}) \leq c \mu^n n^{\vartheta(\alpha,\alpha')+1},$$

completing the proof. □

To end this section, we summarize the procedures for solving (2.14) into the following algorithm.

**Algorithm 2.** *Input  $\sigma$  and  $\lambda$ .*

- Step 1. Compute  $\epsilon_{i'i}$  by (4.4) and solve (4.6) for  $\delta_{i'i}$ .*
- Step 2. Use the compression strategy (4.2) to obtain the compressed matrix  $\tilde{A}_n$ .*
- Step 3. Compute the entries of matrix  $\tilde{A}_n$  by algorithm 1 with parameters  $m_{i'i}$ .*
- Step 4. Solve equation (4.3) with  $\tilde{A}_n$  being replaced by  $\tilde{\tilde{A}}_n$ .*

### 8. Numerical experiments

In this section, we present numerical examples to compare the approximation accuracy of the integral equation model with that of the discrete models for deblurring out-of-focus images. The approximation accuracy is measured by both the peak signal-to-noise ratio (PSNR) and the relative  $L^2$  error. Specifically, we will compare the numerical performance and image quality of the continuous Lavrentiev regularization (CLR) and the continuous Tikhonov regularization (CTR) with those of the discrete Lavrentiev regularization (DLR), the discrete Tikhonov regularization (DTR) and the discrete TV regularization (DTV). We also illustrate through a numerical example that the proposed continuous models are suitable and convenient for enlarging the restored images.

We first briefly describe DLR, DTR and DTV methods. We generate a symmetric discrete Gaussian filter

$$\left[ K_\sigma \left( -\frac{1}{256} \frac{N-1}{2} \right), K_\sigma \left( -\frac{1}{256} \frac{N-3}{2} \right), \dots, K_\sigma \left( \frac{1}{256} \frac{N-3}{2} \right), K_\sigma \left( \frac{1}{256} \frac{N-1}{2} \right) \right],$$

with an odd positive integer  $N$ , where  $\sigma$  will be specified later. We then generate a  $257 \times 257$  symmetric Toeplitz matrix  $B$  whose first row is a vector formed by the last  $(N + 1)/2$  elements of the discrete Gaussian filter, followed by  $257 - (N + 1)/2$  zeros. With the matrix  $B$ , we define  $A := B \otimes B$ . The discrete model is formulated as

$$A\bar{v} = \bar{f}, \quad (8.1)$$

where  $\bar{v}$  is the original image and  $\bar{f}$  is the observation in the discrete form. Note that in this section the image will be either considered as a matrix or a vector formed by vectorizing the matrix in the lexicographical order. The DLR and DTR are to solve the following systems:

$$(\lambda I + A)\bar{v} = \bar{f} \quad (8.2)$$

and

$$(\lambda I + A^T A)\bar{v} = A^T \bar{f}, \quad (8.3)$$

respectively. The DTV is to find the minimizer of the following functional:

$$\|A\bar{v} - \bar{f}\|^2 + \lambda \sum_{i,j=1}^{256} \sqrt{(\bar{v}_{i+1,j} - \bar{v}_{i,j})^2 + (\bar{v}_{i,j+1} - \bar{v}_{i,j})^2}, \quad (8.4)$$

where the homogeneous Neumann boundary condition is applied to obtain the pixel values  $\bar{v}_{i,j}$  outside the range  $1 \leq i, j \leq 257$ . We adopt the numerical implementation of the total variation regularization proposed in [32] due to the availability of its source code.

Next, we explain the necessary details of the CLR and CTR methods. The CLR method refers to solving (2.8) while the CTR method refers to solving (2.12) and (2.13). To solve these equations, we need to discuss how the observed continuous functions (images) are generated from the observed discrete images. To this end, we assume that the pixel values of the observed image are the samples, on the uniform grid  $\{(i/256, j/256) : i, j = 0, 1, \dots, 256\}$ , of a continuous intensity function

$$v(x, y) = \sum_{i,j=0}^{256} \bar{v}_{ij} \psi_i(x) \psi_j(y), \quad (8.5)$$

where for  $\ell = 0, 1, \dots, 256$ :

$$\psi_\ell(t) = \begin{cases} (256t - (\ell - 1)), & (\ell - 1)/256 < t \leq \ell/256, \\ ((\ell + 1) - 256t), & \ell/256 < t \leq (\ell + 1)/256, \\ 0, & \text{otherwise.} \end{cases}$$

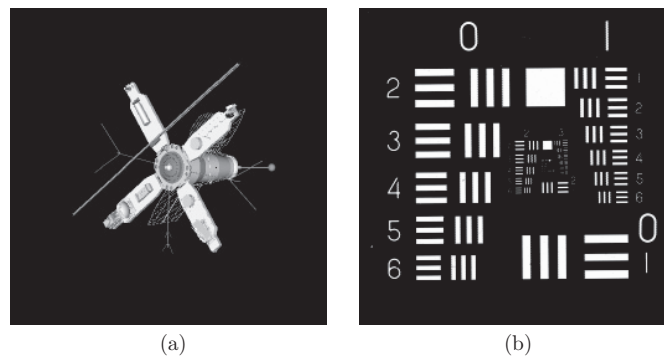
Then the observed continuous function is generated by (2.2) with the kernel  $k$  given by (2.3). Specifically, by using (8.5), we have that

$$f(x', y') = \sum_{i,j=0}^{256} \bar{v}_{ij} \int_0^1 \int_0^1 \frac{1}{2\pi\sigma^2} e^{-\frac{(x-x')^2+(y-y')^2}{2\sigma^2}} \psi_i(x) \psi_j(y) dy dx.$$

The CLR and CTR methods will be combined with the multiscale piecewise linear (L), quadratic (Q) and cubic (C) bases. The corresponding methods are abbreviated as CLRL, CLRQ, CLRC, CTRL, CTRQ and CTRC, respectively. The multiscale piecewise linear, quadratic and cubic bases are explicitly given in the appendix.

Two measures are used in evaluation of the quality of the recovered images by various methods. The first measure is the peak signal-to-noise ratio (PSNR). The larger the PSNR, the better the objective quality of the reconstructed image. The discrete form of the reconstructed images from CLR(L,Q,C) and CTR(L,Q,C) methods is samplings of the reconstructed images





**Figure 1.** (a) Original 'Satellite' image. (b) Original 'Resolution Chart' image.

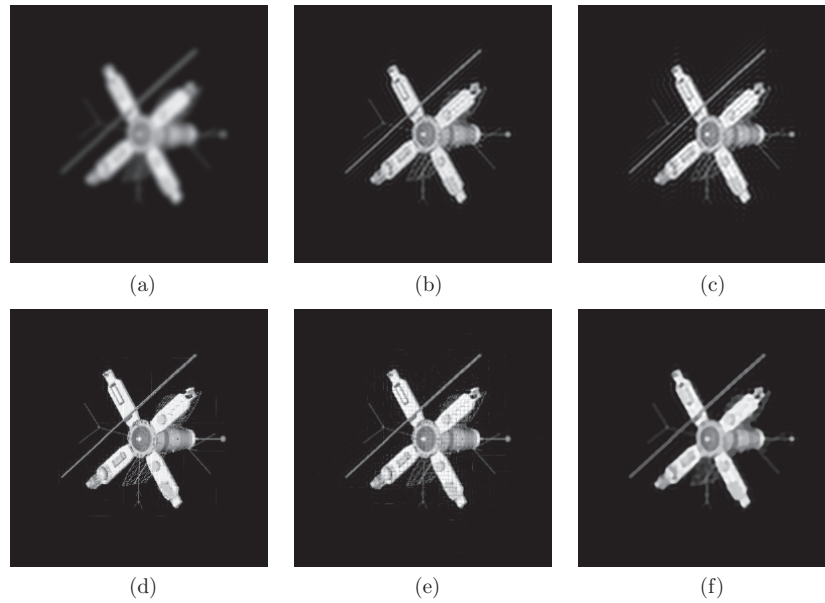
**Table 1.** The values of PSNR and the relative error for  $\sigma = 0.01$  with various methods and noise levels for the 'Satellite' image.

Noise	DLR	CLRL	CLRQ	CLRC	DTR	DTV	CTRL	CTRQ	CTRC
PSNR (dB)									
Free	27.44	29.26	35.20	36.68	27.21	27.45	27.54	29.64	29.65
Noise Std1	19.16	21.89	20.90	20.28	25.67	26.31	26.84	26.90	26.92
Noise Std2	18.10	20.99	19.38	19.02	25.20	26.05	26.60	26.66	26.66
Noise Std3	17.51	20.53	19.07	18.10	24.88	25.83	26.41	26.46	26.47
Relative error ( $\times 10^{-2}$ )									
Free	4.15	0.63	0.05	0.07	4.55	4.08	0.92	0.30	0.47
Noise Std1	24.65	9.31	12.85	14.30	4.97	4.50	2.32	2.29	2.29
Noise Std2	31.86	14.34	15.24	16.85	5.43	4.64	2.52	2.50	2.50
Noise Std3	36.69	15.44	17.48	21.07	5.86	4.82	2.72	2.68	2.67

on the uniform grid  $\{(i/256, j/256) : i, j = 0, 1, \dots, 256\}$ . The second measure is the relative error defined as  $RE := \|u - v\|_2^2 / \|v\|_2^2$ , where  $v$  is the original continuous function and  $u$  is the reconstructed continuous function. When this measure is used, the reconstructed images  $\bar{u}$  by DLR, DTR and DTV will be considered as piecewise constant functions to compute the relative error. The smaller the RE, the better the objective quality of the reconstructed image. Numerical experiments are given for the 'Satellite' image (figure 1(a)) and the 'Resolution Chart' image (figure 1(b)), respectively.

In the experiment for the 'Satellite' image, we consider restoration of the 'Satellite' image from the Gaussian blurring kernels having the standard deviations  $\sigma = 0.01$  and  $\sigma = 0.02$ , separately. One of the reasons for choosing the 'Satellite' image in our experiment is that its pixel values in the vicinity of the image boundary are zeros or close to zeros. This meets the boundary requirement of the proposed algorithms. The image originally consisting of  $256 \times 256$  pixels is extended by simply repeating its last row and column to form an image of size  $257 \times 257$ , in order to fit the numerical requirement of our algorithms.

In our simulation, all regularization parameters are chosen within the range from  $10^{-8}$  to 1 such that the corresponding regularization method produces the best PSNR values. All numerical results are obtained by using a 1.7 GHz Intel Pentium M Processor machine with 1 GB of memory running Windows XP. The summary of the results are tabulated in tables 1–2

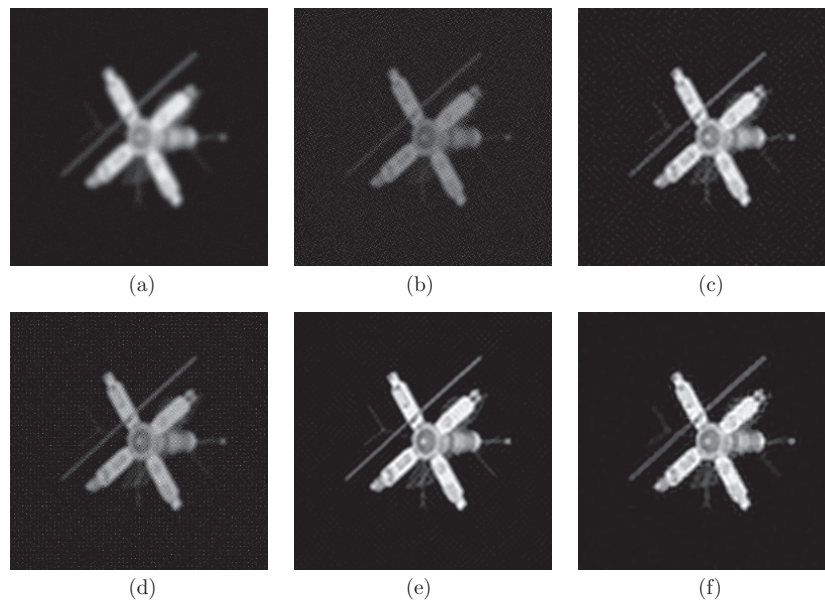


**Figure 2.** Reconstructed results in the noise-free case (with  $\sigma = 0.01$  in the blurring kernel). (a) Samplings of the continuous observed ‘Satellite’ image at a uniform grid; reconstructed ‘Satellite’ images by using (b) DLR ( $\lambda = 2e - 2$ ); (c) DTR ( $\lambda = 2e - 4$ ); (d) CLRC ( $\lambda = 1e - 7$ ); (e) CTRC ( $\lambda = 5e - 7$ ) and (f) DTV ( $\lambda = 3e - 3$ ), respectively.

**Table 2.** The values of PSNR and the relative error for  $\sigma = 0.02$  with various methods and noise levels for the ‘Satellite’ image.

Noise	DLR	CLRL	CLRQ	CLRC	DTR	DTV	CTRL	CTRQ	CTRC
PSNR (dB)									
Free	24.22	27.38	28.56	27.77	23.45	23.85	25.60	25.14	25.13
Noise Std1	18.09	19.93	19.60	19.12	22.67	23.74	24.06	24.06	24.06
Noise Std2	17.15	19.60	18.43	18.13	22.47	23.51	23.85	23.86	23.85
Noise Std3	16.66	19.15	18.16	17.55	22.23	23.43	23.63	23.62	23.62
Relative error ( $\times 10^{-2}$ )									
Free	6.83	1.90	1.38	1.65	8.05	7.22	3.63	4.17	4.16
Noise Std1	31.64	16.18	19.79	20.23	9.63	7.48	6.00	5.99	5.99
Noise Std2	39.74	20.46	21.47	23.23	10.34	7.78	6.39	6.38	6.38
Noise Std3	45.01	22.04	24.28	28.23	10.89	7.94	6.84	6.87	6.85

from the noise-free case to varying levels of noise cases with methods of DLR, DTR, DTV, CLR(L,Q,C) and CTR(L,Q,C). In these tables, ‘Free’ means that the observed image is free of noise. ‘Noise Std1’, ‘Noise Std2’ and ‘Noise Std3’ refer to the cases that the observed images are corrupted by Gaussian white noise with standard deviations 1, 2 and 3, respectively. The numerical results reported in these tables are the averages over five different noise samples. It can be clearly seen that the methods with the continuous model are consistently better than the corresponding methods with the discrete model in terms of PSNR and RE values. For the noise-free case, CLR(L,Q,C) gives the best result. For the noise cases, the methods based on



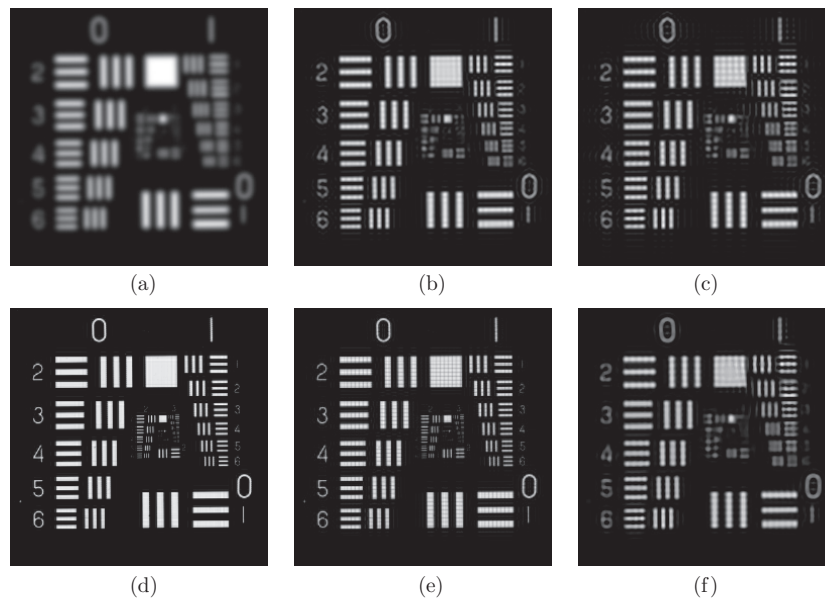
**Figure 3.** Reconstructed results for the noise case 'Noise STD3' (with  $\sigma = 0.01$  in the blurring kernel). (a) Samplings of the continuous observed 'Satellite' image at a uniform grid; reconstructed 'Satellite' images by using (b) DLR ( $\lambda = 0.4$ ); (c) DTR ( $\lambda = 0.02$ ); (d) CLRC ( $\lambda = 0.3$ ); (e) CTRLC ( $\lambda = 0.004$ ) and (f) DTV ( $\lambda = 0.02$ ), respectively.

**Table 3.** The values of PSNR and the relative error for  $\sigma = 0.01$  with various methods and noise levels for the 'Resolution Chart' image.

Noise	DLR	CLRL	CLRQ	CLRC	DTR	DTV	CTRL	CTRQ	CTRC
PSNR (dB)									
Free	19.61	22.52	29.61	31.48	19.14	19.14	20.40	20.54	25.52
Noise Std1	14.65	16.66	15.32	15.14	17.93	18.87	18.77	18.54	18.82
Noise Std2	13.78	15.38	14.98	14.36	17.63	18.51	18.46	18.50	18.48
Noise Std3	13.26	15.25	14.43	13.89	17.41	18.41	18.24	18.34	18.32
Relative error ( $\times 10^{-2}$ )									
Free	11.99	2.11	0.14	0.17	8.56	2.68	3.41	0.63	0.95
Noise Std1	32.38	15.77	17.60	20.56	9.69	6.52	7.66	7.49	7.49
Noise Std2	42.49	20.44	24.64	25.40	9.91	9.66	8.36	8.22	8.22
Noise Std3	42.41	24.78	25.72	27.46	8.96	11.43	8.83	8.71	8.71

the Tikhonov regularization perform better than those based on the Lavrentiev regularization. The visual quality of the reconstructed images are displayed in figure 2 for the noise-free case and figure 3 for the noise case. We can draw the same conclusion as we did from tables 1–2.

It is worthwhile to report computing time for generating the matrix and solving the related linear system by using the proposed methods once the regularization parameter is chosen. Both a direct method and an iterative algorithm are tested and they produce the same results under a given precision for solving equation (4.3) in algorithm 2. The direct method here refers to the Gaussian elimination method. With the direct method, the computing time used to generate the numerical results in table 1 for the noise-free case by using the piecewise

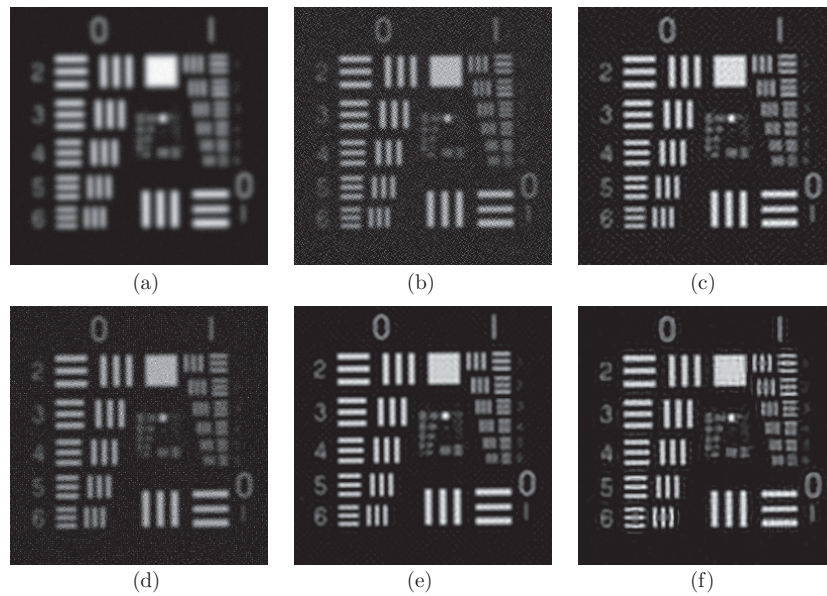


**Figure 4.** Reconstructed results in the noise-free case (with  $\sigma = 0.01$  in the blurring kernel). (a) Samplings of the continuous observed ‘Resolution Chart’ image at a uniform grid; reconstructed ‘Resolution Chart’ images by using (b) DLR ( $\lambda = 2e - 2$ ); (c) DTR ( $\lambda = 2e - 4$ ); (d) CLRC ( $\lambda = 8e - 8$ ); (e) CTRC ( $\lambda = 5e - 7$ ) and (f) DTV ( $\lambda = 4e - 3$ ), respectively.

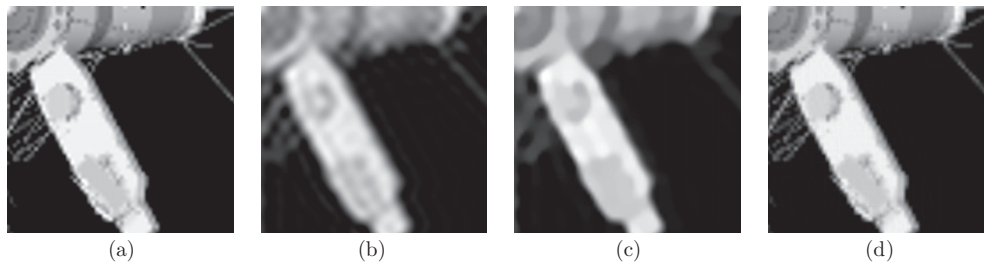
linear basis, the piecewise quadratic basis and the piecewise cubic basis are 16 s, 25 s and 41 s, respectively. The iterative algorithm used is a shadow block iterative algorithm from [17]. When the regularization parameter  $\lambda$  is large, the shadow block iterative algorithm works well. However, the parameter  $\lambda$  is usually small in our experiments, the iterative algorithm may have slow convergence. Preconditioning the matrix  $\lambda I_n + \tilde{A}_n$  may be needed in such a case [11]. It is our future project to develop a fast preconditioning iterative algorithm by employing the sparsity property of the matrix  $\tilde{A}_n$ .

We conduct the numerical experiment for the ‘Resolution Chart’ image under the same setting as we have done for the ‘Satellite’ image. The performance of the proposed algorithms for the ‘Resolution Chart’ image is consistent with that for the ‘Satellite’ image. Therefore, we only list in table 3 the PSNR values and relative errors for  $\sigma = 0.01$ . The visual quality of the reconstructed images are displayed in figure 4 for the noise-free case and in figure 5 for the noise case.

Furthermore, we remark that the continuous integral equation model can be naturally applied for enlarged reconstructed images. By solving the integral equation (2.1), we obtain a continuous function  $v_\lambda$ . To get a discrete image, we can sample the continuous solution  $v_\lambda$  on an arbitrary uniform grid. The advantage for enlarging images with continuous models is illustrated in figure 6 for the noise-free case with  $\sigma = 0.01$ . In figure 6, all images are portions of the enlarged images of size  $2041 \times 2041$ . Figures 6(a)–(c) are enlarged images for the original image, the reconstructed image with DLR and the reconstructed image with DTV, respectively, by using bi-cubic interpolation. Figure 6(d) is the enlarged image obtained by directly sampling the reconstructed images from the continuous model CLRC. The visual quality of the images in figure 6(d) is much better than that of image in figures 6(b), (c).



**Figure 5.** Reconstructed results for the noise case 'Noise STD3' (with  $\sigma = 0.01$  in the blurring kernel). (a) Samplings of the continuous observed 'Resolution Chart' image at a uniform grid; reconstructed 'Resolution Chart' images by using (b) DLR ( $\lambda = 0.3$ ); (c) DTR ( $\lambda = 0.01$ ); (d) CLRC ( $\lambda = 0.3$ ); (e) CTRC ( $\lambda = 0.005$ ) and (f) DTV ( $\lambda = 0.02$ ), respectively.



**Figure 6.** Zoom-in portions of the enlarged 'Satellite' images with size  $2041 \times 2041$ . (a) Enlarged original image by using bi-cubic interpolation; (b) enlarged reconstructed image by using DLR and bi-cubic interpolation; (c) enlarged reconstructed image by using DTV and bi-cubic interpolation; (d) sampling of the continuous reconstructed image by CLRQ.

## 9. Conclusion

Restoration of blurred images can be formulated as an integral equation of the first kind with a Gaussian kernel, which is an ill-posed problem. We can treat it by the Lavrentiev or Tikhonov regularization methods and obtain a well-posed integral equation of the second kind. The regularized integral equation is then discretized by using a multiscale piecewise polynomial basis. The resulting full coefficient matrix is compressed by a matrix compression strategy to obtain a sparse matrix whose remaining nonzero entries are evaluated by a numerical quadrature scheme specially designed for integrals involved with the Gaussian function. A fast numerical method with optimal (high) convergence order for blurred image restoration

is developed. Numerical experiments show that the proposed integral model performs much better than the discrete models with either the Lavrentiev, Tikhonov or TV regularization. Theoretical study and numerical results of this paper suggest that the continuous integral equation is a good alternative model for image restoration.

### Acknowledgments

LS was supported by the US National Science Foundation under grant DMS-0712827. YX was supported in part by the US National Science Foundation under grant DMS-0712827, by US Air Force Office of Scientific Research under grant FA9550-09-1-0511 and by the Natural Science Foundation of China under grants 10371122 and 10631080.

### Appendix

In this appendix, for convenient reference we present multiscale bases used in our numerical experiments. They may be found in [7]. According to the construction of a multiscale analysis  $\mathbb{X}_n$ ,  $n \in \mathbb{Z}_+$ , of  $L^2(E)$  presented in section 3, we only need to describe orthogonal bases for  $\mathbb{X}_0$  and  $\mathbb{W}_1$ . The normalized Legendre polynomials of degree  $j = 1, 2, 3$  are chosen as bases for  $\mathbb{X}_0$ . The corresponding multiscale analysis are called piecewise linear, quadratic and cubic bases of  $L^2(E)$ , respectively.

*Piecewise linear basis.* For  $t \in E$ , we choose

$$e_{00}(t) := 1, \quad e_{01}(t) := 2\sqrt{3}t - \sqrt{3}$$

and

$$e_{10}(t) := \begin{cases} 6t - 1, & t \in [0, \frac{1}{2}], \\ 6t - 5, & t \in (\frac{1}{2}, 1], \end{cases} \quad e_{11}(t) = \begin{cases} -4\sqrt{3}t + \sqrt{3}, & t \in [0, \frac{1}{2}], \\ 4\sqrt{3}t - 3\sqrt{3}, & t \in (\frac{1}{2}, 1]. \end{cases}$$

Then,  $\{e_{00}, e_{01}\}$  forms a basis for  $\mathbb{X}_0$  while  $\{e_{10}, e_{11}\}$  forms a basis for  $\mathbb{W}_1$ .

*Piecewise quadratic basis.* For  $t \in E$ , we choose

$$e_{00}(t) = 1, \quad e_{01}(t) = 2\sqrt{3}t - \sqrt{3}, \quad e_{02}(t) = 6\sqrt{5}t^2 - 6\sqrt{5}t + \sqrt{5}$$

and

$$e_{10}(t) = \begin{cases} -6t + 1, & t \in [0, \frac{1}{2}], \\ -6t + 5, & t \in (\frac{1}{2}, 1], \end{cases} \quad e_{11}(t) = \begin{cases} \frac{\sqrt{93}}{31}(240t^2 - 116t + 9), & t \in [0, \frac{1}{2}], \\ \frac{\sqrt{93}}{31}(-4t + 3), & t \in (\frac{1}{2}, 1], \end{cases}$$

$$e_{12}(t) = \begin{cases} \frac{\sqrt{93}}{31}(4t - 1), & t \in [0, \frac{1}{2}], \\ \frac{\sqrt{93}}{31}(240t^2 - 364t + 133), & t \in (\frac{1}{2}, 1]. \end{cases}$$

Then,  $\{e_{00}, e_{01}, e_{02}\}$  forms a basis for  $\mathbb{X}_0$  while  $\{e_{10}, e_{11}, e_{12}\}$  forms a basis for  $\mathbb{W}_1$ .

*Piecewise cubic basis functions.* For  $t \in E$ , we choose

$$e_{00}(t) = 1, \quad e_{01}(t) = 2\sqrt{3}t - \sqrt{3}, \\ e_{02}(t) = 6\sqrt{5}t^2 - 6\sqrt{5}t + \sqrt{5}, \quad e_{03}(t) = \sqrt{7}(20t^3 - 30t^2 + 12t - 1)$$

and

$$\begin{aligned}
 e_{10}(t) &= \begin{cases} 16\sqrt{5}t^2 - 6\sqrt{5}t + \sqrt{5}/3, & t \in [0, \frac{1}{2}], \\ -16\sqrt{5}t^2 + 26\sqrt{5}t - 155\sqrt{5}/15, & t \in (\frac{1}{2}, 1], \end{cases} \\
 e_{11}(t) &= \begin{cases} \sqrt{3}(30t^2 - 14t + 1), & t \in [0, \frac{1}{2}], \\ \sqrt{3}(30t^2 - 46t + 17), & t \in (\frac{1}{2}, 1], \end{cases} \\
 e_{12}(t) &= \begin{cases} \sqrt{7}(160t^3 - 120t^2 + 24t - 1), & t \in [0, \frac{1}{2}], \\ \sqrt{7}(-160t^3 + 360t^2 - 264t + 63), & t \in (\frac{1}{2}, 1], \end{cases} \\
 e_{13}(t) &= \begin{cases} 420t^3 - 310t^2 + 60t - 7/3, & t \in [0, \frac{1}{2}], \\ 420t^3 - 950t^2 + 700t - 503/3, & t \in (\frac{1}{2}, 1]. \end{cases}
 \end{aligned}$$

Then,  $\{e_{00}, e_{01}, e_{02}, e_{03}\}$  forms a basis for  $\mathbb{X}_0$  while  $\{e_{10}, e_{11}, e_{12}, e_{13}\}$  forms a basis for  $\mathbb{W}_1$ .

## References

- [1] Atkinson K E 1997 *The Numerical Solution of Integral Equations of the Second Kind* (Cambridge: Cambridge University Press)
- [2] Bertero M, Brianzi P and Pike E 1987 Super-resolution in confocal scanning microscopy *Inverse Problems* **3** 195–212
- [3] Borwein P and Erdélyi T 1995 *Polynomials and Polynomial Inequalities (Graduate Texts in Mathematics vol 161)* (New York: Springer)
- [4] Bose N and Boo K 1998 High-resolution image reconstruction with multisensors *Int. J. Imaging Syst. Technol.* **9** 294–304
- [5] Brankov J G, Yang Y and Wernick M N 2004 Tomographic image reconstruction based on a content-adaptive mesh model *IEEE Trans. Med. Imaging* **23** 202–12
- [6] Chaudhuri S and Rajagopalan A 1999 *Depth from Defocus: A Real Aperture Imaging Approach* (New York: Springer)
- [7] Chen Z, Micchelli C and Xu Y 2009 *Fast Multiscale Methods for Fredholm Integral Equations* unpublished monograph
- [8] Chen Z, Wu B and Xu Y 2005 Error control strategies for numerical integrations in fast collocation methods *Northeast. Math. J. Dongbei Shuxue* **21** 233–52
- [9] Chen Z, Xu Y and Yang H 2006 A multilevel augmentation method for solving ill-posed operator equations *Inverse Problems* **22** 155–74
- [10] Chen Z, Xu Y and Yang H 2008 Fast collocation methods for solving ill-posed integral equations of the first kind *Inverse Problems* **24** 065007 (21 pp)
- [11] Egger H 2007 Preconditioning CGNE iteration for inverse problems *Numer. Linear Algebra Appl.* **14** 183–96
- [12] Fang W, Wang Y and Xu Y 2004 An implementation of fast wavelet Galerkin methods for integral equations of the second kind *J. Sci. Comput.* **20** 277–302
- [13] Gonzalez R and Woods R 1993 *Digital Image Processing* (Boston, MA: Addison-Wesley)
- [14] Groetsch C W 1984 *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind (Research Notes in Mathematics vol 105)* (Boston, MA: Pitman)
- [15] Hansen P C 2002 Deconvolution and regularization with Topelitz matrices *Numer. Algor.* **29** 323–78
- [16] Kaneko H and Xu Y 1994 Gauss-type quadratures for weakly singular integrals and their application to Fredholm integral equations of the second kind *Math. Comput.* **62** 739–53
- [17] Lu Y, Shen L and Xu Y 2005 Shadow block iterations for solving linear systems obtained from wavelet transforms *Appl. Comput. Harmon. Anal.* **19** 359–85
- [18] Maass S V, Pereverzev P, Ramlau R and Solodky S G 2001 An adaptive discretization for Tikhonov–Phillips regularization with a posteriori parameter selection *Numer. Math.* **87** 485–502
- [19] Micchelli C and Xu Y 1994 Using the matrix refinement equation for the construction of wavelets on invariant sets *Appl. Comput. Harmon. Anal.* **1** 391–401
- [20] Micchelli C A and Xu Y 1997 Reconstruction and decomposition algorithms for biorthogonal multiwavelets *Multidimens. Syst. Signal Process.* **8** 31–69

- [21] Micchelli C A, Xu Y and Zhang H 2009 Optimal learning of bandlimited functions from localized sampling *J. Complexity* **25** 85–114
- [22] Micchelli C A, Xu Y and Zhao Y 1997 Galerkin methods for second-kind integral equations *J. Comput. Appl. Math.* **86** 251–70
- [23] Moreno-Balcázar J J 2005 Smallest zeros of some types of orthogonal polynomials: asymptotics *J. Comput. Appl. Math.* **179** 289–301
- [24] Pentland A 1987 A new sense for depth of field *IEEE Trans. Pattern Anal. Mach. Intell.* **9** 523–31
- [25] Plato R 1995 Iterative and parametric methods for linear ill-posed equations *Habilitationsschrift* (Berlin: Fachbereich Mathematik, TU Berlin)
- [26] Plato R 1997 The Galerkin scheme for Lavrentiev's  $m$ -times iterated method to solve linear accretive Volterra integral equations of the first kind *BIT* **37** 404–23
- [27] Schwab C 1994 Variable order composite quadrature of singular and nearly singular integrals *Computing* **53** 173–94
- [28] Shannon C E 1949 Communication in the presence of noise *Proc. IRE* **37** 10–21
- [29] Sitek A, Huesman R H and Gullberg G T 2006 Tomographic reconstruction using an adaptive tetrahedral mesh defined by a point cloud *IEEE Trans. Med. Imaging* **25** 1172–9
- [30] Stein E and Weiss G 1971 *Introduction to Fourier Analysis on Euclidean Spaces* (Princeton, NJ: Princeton University Press)
- [31] Tikhonov A and Arsenin V 1977 *Solutions to Ill-Posed Problems* (New York: Wiley)
- [32] Vogel C and Oman M 1998 Fast, robust total variation-based reconstruction of noisy, blurring images *IEEE Trans. Image Process.* **7** 813–24