Adv Comput Math (2009) 31:421–455
DOI 10.1007/s10444-008-9077-4

Multiparameter regularization for Volterra kernel identification via multiscale collocation methods

Marty Brenner · Ying Jiang · Yuesheng Xu

Received: 24 July 2007 / Accepted: 11 March 2008 /
Published online: 16 May 2008

Abstract Identification of the Volterra system is an ill-posed problem. We propose a regularization method for solving this ill-posed problem via a multiscale collocation method with multiple regularization parameters corresponding to the multiple scales. Many highly nonlinear problems such as flight data analysis demand identifying the system of a *high* order. This task requires huge computational costs due to processing a dense matrix of a large order. To overcome this difficulty a compression strategy is introduced to approximate the full matrix resulted in collocation of the Volterra kernel by an appropriate sparse matrix. A numerical quadrature strategy is designed to efficiently compute the entries of the compressed matrix. Finally, numerical results of three simulation experiments are presented to demonstrate the accuracy and efficiency of the proposed method.

Communicated by Juan Manuel Peña.

Supported in part by the US National Aeronautics and Space Administration under Cooperative Agreement NNX07AC37A, by the US National Science Foundation under grants CCR-0407476 and DMS-0712827, by the Natural Science Foundation of China under grants 10371122 and 10631080, and by the Education Ministry of the People's Republic of China under the Changjiang Scholar Chair Professorship Program through Sun Yat-sen University.

M. Brenner
NASA Dryden Flight Research Center, MS 4830D/RS,
Edwards, CA 93523-0273, USA
e-mail: martin.brenner@dfrc.nasa.gov

Y. Jiang · Y. Xu (✉)
Department of Mathematics, Syracuse University,
Syracuse, NY 13244-1150, USA
e-mail: yxu06@syr.edu

Y. Jiang
e-mail: yjiang02@syr.edu

Keywords Multiparameter regularization · Volterra kernel identification · Multiscale collocation method

Mathematics Subject Classifications (2000) 47A52 · 45D05

1 Introduction

Volterra kernels have been widely employed in identification of nonlinear systems in understanding the nonlinearity of the systems [1, 4, 16, 18, 20, 28, 29, 35, 41, 48]. Specifically, the Volterra model of order two was used in [18] for control design in the chemical/petrochemical industries, and in [28] for predictive control of a simulated multivariable polymerization reactor. In a studying of nerve networks, Volterra kernels were used to simulate the nonlinearity of the somatosensory evoked potentials and to predicate the responses of mechanoreceptor neurons to physiological inputs [20, 29, 35]. In clinical medicine, the Volterra model was utilized to study lung tissue viscoelasticity and to present relationship between renal blood pressure and blood flow (cf., [16, 48]). Volterra models were also frequently used in artificial intelligence and signal processing. For instance, a nonlinear system in pattern recognition was modeled by a Volterra model [1]. Since many signals of interest are generated by nonlinear sources or are processed by nonlinear systems, Volterra modeling was introduced in signal processing for estimating and controlling an active noise arising from dynamical systems [41]. In applications of the Volterra system, an important and difficult issue is identification of Volterra kernels of a given system. Many methods such as the sampled methods, finite element methods, and wavelet methods have been employed to obtain discrete representations of Volterra systems, see [8, 23, 34, 37, 45].

In identifying a Volterra kernel, we face two major computational challenges. The first challenge comes from huge computational costs in identifying the kernel. Discretization of a Volterra system leads to a full matrix. Computation with a full matrix is a very costly task. Many highly nonlinear problems require the use of the Volterra system of a high order to characterize their high order nonlinearity. In other words, we need to use high order Volterra kernels to simulate the nonlinear component of the system. Increase in order of kernels results in huge increase in the computational cost. To overcome this difficulty, we propose to employ the fast multiscale collocation method developed in [11, 13, 14, 46] to compress the matrix obtained from identifying the Volterra kernels and use the augmentation method developed in [12] to solve the resulting linear system.

Identification of nonlinear systems requires detecting their underlying structure by estimating certain unknown parameters. This process is ill-posed in the sense that a small perturbation in the given data will result in a large perturbation in solutions. That is, the solutions do not continuously depend on the given data. This presents us the second computational challenge. Regularization is an approach to treat this challenging issue. The classical regularization method for solving ill-posed problems was proposed independently by Phillips [36] and Tikhonov [42]. Although it has been proved to be an efficient method to tackle the ill-posedness, (cf., [19, 22, 36, 43]), it has also serious limitation. The hypothesis for the *single* parameter regularization is that noise *effect* to an ill-posed problem is uniformly distributed in all frequency bands of the solution. A uniform penalty is introduced to every frequency band of the solution or only the high-frequency band of the solution. The first case may result in solutions that are too smooth to preserve certain useful features of the original data and in the second case, the regularization solutions may be affected by low-frequency noise. In practical applications, we often observe different circumstances where noise distributes differently in different frequency bands. These lead to consideration of *multiparameter* regularization [9]. In particular, in regularization of the ill-possed problem related to identification of the Volterra kernel, we need to distinguish the noise effect to its linear component and to its nonlinear components, and as well as to distinguish the noise effect to different frequency bands of these components.

Multiparameter regularization has been used to treat linear systems in a few different contexts (cf., [3, 5, 9, 17, 26, 27]). A choice of multiple parameters was proposed in [3] by using the generalized L-curve method. A multiparameter regularization algorithm for the solution of over-determined, ill-conditioned linear systems was proposed in [5], where numerical examples were presented to demonstrate that the proposed algorithm is stable and robust. A multiparameter regularization for solving ill-posed operator equations was proposed in [9] and convergence theorems and error estimates were established there. In [17], the authors used a multi-parameter regularization method for atmospheric remote sensing. Multiparameter regularization for certain issues of sinal and image processing was studied in [26, 27].

Because the focus of this paper is on developing an efficient identification method for the Volterra system by multiparameter regularization via multiscale collocation methods, we will not emphasize on the choice of the parameters or convergence of the regularization method. Those who are interested in these important issues are referred to [3] for choices of the parameters and [9] for convergence results of the regularization. The emphasis of this paper is on the fast solution of the well-posed integral equations of high dimensions resulting from the regularization of the ill-posed identification problem.

This paper is organized in eight sections. In Section 2, we formulate regularization methods with different parameters for different orders of the Volterra kernels and for different frequency scales. We introduce in Section 3 a multiscale collocation method for solving the Euler equation resulting from the regularization described in Section 2. Section 4 is devoted to a description of block matrix compression strategy for the collocation method and Sections 5 and 6 are for the complexity and convergence analysis of the compression algorithm, respectively. Since the entries of the coefficient matrix resulting from the compression strategy are integrals, we develop in Section 7 a numerical quadrature rule for computing these entries and a strategy for controlling the error contribution from the quadrature method. Finally in Section 8, we present three numerical experiments that demonstrate the performance of the proposed method.

2 Regularization methods

In this section, we briefly introduce the Volterra system and formulate the Volterra kernel identification problem as an operator equation of the first kind. Since the related integral operator is compact, the Volterra kernel identification problem is ill-posed. We then describe a multiparameter regularization method for solving the ill-posed problem and derive its corresponding Euler equation.

For a positive integer n and $t \in \mathbb{R}^+$, we let

$$\Omega_n(t) := \{(\xi_1, \xi_2, \ldots, \xi_n) : 0 \le \xi_n \le \xi_{n-1} \le \cdots \le \xi_1 \le t\},$$

and for a given fixed $T \in \mathbb{R}^+$, we use Ω_n for $\Omega_n(T)$. Let $H_n := L^2(\Omega_n)$. For a given real-valued function u and for $n \in \mathbb{N}$, we define operator $\mathcal{U}_n : H_n \to H_1$ for $h \in H_n$ by

$$(\mathcal{U}_n h)(t) := \int_0^t \int_0^{\xi_1} \cdots \int_0^{\xi_{n-1}} h(\xi_1, \ldots, \xi_n) \prod_{i=1}^{n} u(t - \xi_i) d\xi_1 \cdots d\xi_n, \qquad t \in [0, T]. \quad (1)$$

Note that operator $\mathcal{U}_n$ maps a multivariate function into a univariate function. Associated with the fixed positive integer ν, we set $\mathbf{H} := H_1 \times H_2 \times \cdots \times H_\nu$. Clearly, $\mathbf{H}$ is a Hilbert space with the inter product $(\mathbf{h}, \mathbf{g}) := \sum_{n \in \mathbb{Z}_\nu} \int_{\Omega_n} h_n(\xi) g_n(\xi) d\xi$, for $\mathbf{h}, \mathbf{g} \in \mathbf{H}$, where $\mathbb{Z}_\nu := \{1, 2, \ldots, \nu\}$. We define operator $\mathcal{U} : \mathbf{H} \to H_1$ by $\mathcal{U}\mathbf{h} := \sum_{n \in \mathbb{Z}_\nu} \mathcal{U}_n h_n$, for $\mathbf{h} := (h_1, h_2, \ldots, h_\nu)^T \in \mathbf{H}$. The operator $\mathcal{U}$ is completely determined by the function u. For a given $v \in H_1$, the Volterra kernel identification problem is to find $\mathbf{h} \in \mathbf{H}$ such that

$$v = \mathcal{U}\mathbf{h}. \quad (2)$$

We will call $\mathcal{U}$ the Volterra operator with the input signal u, $\mathbf{h}$ the Volterra kernel vector, h_n the Volterra kernel of the nth order, and v the output signal.

If $u \in L^2(0, T)$, then the operator $\mathcal{U}_n$ can be written as an integral operator with a Hilbert-Schmidt kernel, and the operator $\mathcal{U}_n$ is compact (see Sections 7.1 and 7.5 of [39]). Specifically, we have the following lemma.

Lemma 1 *If $u \in L^2(0, T)$, then $\mathcal{U}$ is compact as an operator from $\mathbf{H}$ to H_1.*

Hence, Eq. 2 is ill-posed. It means that the solution $\mathbf{h}$ of (2) does not depend continuously on v. To overcome this challenge, single parameter regularization methods were employed in the literature, such as the Tikhonov regularization method. The single parameter regularization methods impose the same penalty parameter to the Volterra kernels of all orders. However, it is observed that the noise effects to Volterra kernels of different orders might be different, and as well as the noise effect to different frequency bands of these components might be different. These motivate us to introduce different parameters for Volterra kernels of different orders and for different frequency bands. To this end, we present a multi-parameter regularization method via a multiscale decomposition of space H_n.

For a fixed $n \in \mathbb{N}$, we introduce an orthogonal decomposition of the Hilbert space H_n and the related orthogonal projections. Let $\mathbb{N}_0 := \{0, 1, \ldots\}$. We assume that $\{\mathbb{H}_{n,i} : i \in \mathbb{N}_0\}$ is a sequence of multiscale finite dimensional spaces of H_n satisfying $\mathbb{H}_{n,i} \subset \mathbb{H}_{n,i+1}$, for $i \in \mathbb{N}_0$ and $\overline{\bigcup_{i \in \mathbb{N}_0} \mathbb{H}_{n,i}} = H_n$. Thus, for each $i \in \mathbb{N}$, there is a subspace $\mathbb{W}_{n,i} \subset \mathbb{H}_{n,i}$ which is the orthogonal complement of $\mathbb{H}_{n,i-1}$ in $\mathbb{H}_{n,i}$. By letting $\mathbb{W}_{n,0} := \mathbb{H}_{n,0}$, we obtain that

$$H_n = \overline{\bigoplus_{i \in \mathbb{N}_0}{}^{\perp} \mathbb{W}_{n,i}}.$$

This decomposition was studied in [31, 32], where an orthogonal wavelet decomposition was constructed on invariant sets. In this paper, we assume that our finite dimensional subspaces $\mathbb{H}_{n,i}$ are spaces of piecewise polynomials of certain orders on domains Ω_n with a scale corresponding to i. For specific constructions of such decompositions, see [11, 14, 31, 33, 46]. For each $i \in \mathbb{N}_0$, we let $P_{n,i}$ be the orthogonal projection from H_n onto $\mathbb{H}_{n,i}$. It follows that $Q_{n,i} := P_{n,i} - P_{n,i-1}$ is the orthogonal projection from H_n onto $\mathbb{W}_{n,i}$ for $i \geq 1$. By letting $Q_{n,0} := P_{n,0}$, we have that $P_{n,i} = \sum_{j=0}^{i} Q_{n,j}$.

Given a set of parameters $\Delta := \{\lambda_{n,i} : \lambda_{n,i} > 0,\ n \in \mathbb{Z}_\nu, i \in \mathbb{N}_0\}$, we find $\mathbf{h}_\Delta \in \mathbf{H}$ such that

$$\mathbf{h}_\Delta \in \arg\min_{\mathbf{h}\in\mathbf{H}} \left\{ \|\mathcal{U}\mathbf{h} - v\|_{H_1}^2 + \sum_{n\in\mathbb{Z}_\nu} \sum_{i\in\mathbb{N}_0} \lambda_{n,i} \|Q_{n,i} h_n\|_{H_n}^2 \right\}. \tag{3}$$

We next show the Euler equation of (3) in Theorem 1. To this end, we define an operator $\Lambda : \mathbf{H} \to \mathbf{H}$ for a given Δ by $\Lambda\mathbf{h} := \left[\sum_{i\in\mathbb{N}_0} \lambda_{n,i} Q_{n,i} h_n :\ n \in \mathbb{Z}_\nu\right]$. We let $\lambda' := \inf\{\lambda_{n,i} : n \in \mathbb{Z}_\nu, i \in \mathbb{N}_0\}$ and $\lambda^* := \sup\{\lambda_{n,i} : n \in \mathbb{Z}_\nu, i \in \mathbb{N}_0\}$. Clearly, when $\lambda' > 0$ and $\lambda^* < +\infty$, Λ is bounded, positive definite and invertible. We also let $\mathcal{U}_n^* : H_1 \to H_n$ be the conjugate operator of $\mathcal{U}_n$. The operator $\mathcal{U}_n^*$ has the form

$$(\mathcal{U}_n^* f)(\xi_1, \ldots, \xi_n) = \int_{\xi_1}^{T} \prod_{i=1}^{n} u(t - \xi_i) f(t) dt, \quad f \in H_1.$$

Thus, the conjugate operator $\mathcal{U}^* : H_1 \to \mathbf{H}$ of $\mathcal{U}$ has the form $\mathcal{U}^* f = \left(\mathcal{U}_1^* f, \mathcal{U}_2^* f, \ldots, \mathcal{U}_\nu^* f\right)^T$, for $f \in H_1$.

Theorem 1 *Suppose that a set Δ of parameters is given and satisfy $\lambda' > 0$. If $\mathcal{U} : \mathbf{H} \to H_1$ is the bounded linear operator defined through (1), then $\mathbf{h}_\Delta \in \mathbf{H}$ is the solution of (3) if and only if $\mathbf{h}_\Delta$ is the solution of the equation*

$$\mathcal{U}^*\mathcal{U}\mathbf{h}_\Delta + \Lambda\mathbf{h}_\Delta = \mathcal{U}^* v. \tag{4}$$

Moreover, $\mathbf{h}_\Delta$ depends continuously on v.

The proof of this theorem follows from a standard argument (cf., [7]). Note that if $\lambda_{n,i} = \lambda_n$ for all $i \in \mathbb{N}_0$, (3) or (4) reduces to the multiparameter Tikhonov regularization method, with one parameter per kernel.

To close this section, we estimate $\left\|\mathcal{U}(\mathcal{U}^*\mathcal{U} + \Lambda_\Delta)^{-1}\right\|_{\mathbf{H}\to H_1}$ that will be used in analysis of the multiscale collocation method to be developed in Section 4.

Lemma 2 *If $\lambda' > 0$, then*

$$\left\|\mathcal{U}(\mathcal{U}^*\mathcal{U} + \Lambda_\Delta)^{-1}\right\|_{\mathbf{H}\to H_1} \leq \frac{1}{\sqrt{\lambda'}}. \tag{5}$$

Proof We first recall a basic inequality. Let $\mathbb{X}$ and $\mathbb{Y}$ be two Hilbert spaces with norms $\|\cdot\|_{\mathbb{X}}$ and $\|\cdot\|_{\mathbb{Y}}$, respectively. For a compact operator $\mathcal{A} : \mathbb{X} \to \mathbb{Y}$, by using the singular value system of $\mathcal{A}$, we find that

$$\left\|\mathcal{A}\left(\mathcal{A}^*\mathcal{A} + \mathcal{I}\right)^{-1}\right\|_{\mathbb{X}\to\mathbb{Y}} \leq 1. \tag{6}$$

Letting $\mathcal{A} := \mathcal{U}\Lambda_{\Delta}^{-1/2}$, we have that

$$\mathcal{U}\left(\mathcal{U}^*\mathcal{U} + \Lambda_{\Delta}\right)^{-1} = \mathcal{A}\left(\mathcal{A}^*\mathcal{A} + \mathcal{I}\right)^{-1}\Lambda_{\Delta}^{-1/2}. \tag{7}$$

We define $\Lambda_{\Delta}^{1/2}$ for all $\mathbf{h} := [h_1, h_2, \ldots, h_\nu] \in \mathbf{H}$ by

$$\Lambda_{\Delta}^{1/2}\mathbf{h} := \left[\sum_{i\in\mathbb{N}_0} \sqrt{\lambda_{n,i}}\, Q_{n,i} h_n : n \in \mathbb{Z}_\nu\right].$$

It follows for all $\mathbf{h} := [h_1, h_2, \ldots, h_\nu] \in \mathbf{H}$ that

$$\Lambda_{\Delta}^{-1/2}\mathbf{h} := \left[\sum_{i\in\mathbb{N}_0} \frac{1}{\sqrt{\lambda_{n,i}}} Q_{n,i} h_n : n \in \mathbb{Z}_\nu\right].$$

Thus, we have that

$$\left\|\Lambda_{\Delta}^{-1/2}\mathbf{h}\right\|_{\mathbf{H}}^2 = \sum_{n\in\mathbb{Z}_\nu} \left\|\sum_{i\in\mathbb{N}_0} \frac{1}{\sqrt{\lambda_{n,i}}} Q_{n,i} h_n\right\|_{H_n}^2. \tag{8}$$

Since $Q_{n,i}$, $i \in \mathbb{N}_0$ are orthogonal projections, form (8), we observe that

$$\left\|\Lambda_{\Delta}^{-1/2}\mathbf{h}\right\|_{\mathbf{H}}^2 = \sum_{n\in\mathbb{Z}_\nu} \sum_{i\in\mathbb{N}_0} \frac{1}{\lambda_{n,i}} \| Q_{n,i} h_n \|_{H_n}^2. \tag{9}$$

Note that for all $n \in \mathbb{Z}_\nu$, $i \in \mathbb{N}_0$, $\lambda' \leq \lambda_{n,i}$. From (9), we conclude that

$$\left\|\Lambda_{\Delta}^{-1/2}\mathbf{h}\right\|_{\mathbf{H}}^2 \leq \frac{1}{\lambda'} \sum_{n\in\mathbb{Z}_\nu} \sum_{i\in\mathbb{N}_0} \| Q_{n,i} h_n \|_{H_n}^2 \leq \frac{1}{\lambda'} \|\mathbf{h}\|_{\mathbf{H}}^2.$$

In other words,

$$\left\|\Lambda_{\Delta}^{-1/2}\right\|_{\mathbf{H}} \leq \frac{1}{\sqrt{\lambda'}}. \tag{10}$$

Substituting (6) and (10) into (7), we obtain inequality (5). □

3 The multiscale collocation method

The multiscale regularization method converts the Volterra kernel identification problem to the second kind integral Eq. 4. A natural way to discretize this equation would be the Galerkin method since it naturally fits into the Hilbert space setting for regularization. However, the Galerkin method will double the dimension of the integrals involved in Eq. 4. In particular, for Eq. 4, it will result in huge computational costs in generating the coefficient matrix of the discrete system, since Eq. 4 involves integrals of high dimensions. For this reason, we propose to use the multiscale collocation method introduced in [11] to discretize the equation.

There are two difficulties in using the multiscale collocation method for solving Eq. 4. For analysis of the collocation method, a natural norm to use is the L^∞-norm, while Eq. 4 was derived by regularization in L^2 spaces. The change from the L^2-norm to the L^∞-norm requires a careful treatment. Note that the solution of Eq. 4 is the minimum norm solution, in the L^2 sense, of the original Volterra kernel identification problem. While the collocation solution of Eq. 4 will approximate its exact solution in the L^∞ sense. By the fact that the L^2 error is bounded by a constant time of the L^∞ error, the collocation solution approximates the exact solution of Eq. 4 in the L^2 sense as well. The second difficulty comes from the fact that the solution of Eq. 4 is a vector-value function with different components being functions of different dimensions. This presents challenges to incorporate the basis functions and collocation functionals having different dimensions.

Let $\mathbb{H}_{n,i}$ be the space of piecewise polynomials of total degree $k_n - 1$ on Ω_n with the partition $\Xi_{n,i} := \{\Xi_{n,i,j} : j \in \mathbb{Z}_{N_n(i)}\}$ described in [14], where $N_n(i)$ is a positive integer. The partition $\Xi_{n,i}$ satisfies that for each $j \in \mathbb{Z}_{N_n(i)}$, there exists an affine mapping $\phi_{n,i,j}$ from $\Omega' := \Omega_n(1)$ to $\Xi_{n,i,j}$. We will use this sequence of finite dimensional subspaces $\mathbb{H}_{n,i}$, $i \in \mathbb{N}_0$ to develop the multiscale collocation method. Although, the subspaces $\mathbb{H}_{n,i}$ are dense in H_n for a fixed $n \in \mathbb{Z}_\nu$, they are spaces of piecewise polynomials in different scales. We use H_n^∞ to denote $L^\infty(\Omega_n)$. For each $n \in \mathbb{Z}_\nu$, we let $\mathbb{W}_n^\infty := \overline{\bigoplus_{i\in\mathbb{N}_0} \mathbb{W}_{n,i}}$, where the closure is in the L^∞ sense. Hence, we conclude that $C(\Omega_n) \subset \mathbb{W}_n^\infty \subset H_n^\infty$. The bases for the multiscale spaces $\mathbb{W}_{n,i}$ will be used to generate the collocation solution.

To apply the collocation method to Eq. 4, we need to consider another important issue if for all $n \in \mathbb{Z}_\nu$ the component $h_{n,\Delta}$ of the solution $\mathbf{h}_\Delta$ of (4) belongs to $\mathbb{W}_n^\infty$. As a preparation for answering the question, we next show that $\|P_{n,i}\|_\infty$ is uniformly bounded by a constant independent of i. The case $n = 1$ of this result was established in [38].

Lemma 3 *For a fixed $n \in \mathbb{Z}_\nu$, there exists a constant $c > 0$ such that $\left\|P_{n,i}\right\|_\infty \le c$, for all $i \in \mathbb{N}_0$.*

Proof Let $\mathbf{S}$ be a family of subsets of $\mathbb{R}^n$ for which there exists a fixed $\widetilde{\Omega} \in \mathbf{S}$ such that for each element $\Omega \in \mathbf{S}$, there is an affine mapping from $\widetilde{\Omega}$ to Ω. For each $\Omega \in \mathbf{S}$, we define the orthogonal projection Q_Ω from $L^2(\Omega)$ onto the space of polynomials of total degree $k-1$ on Ω. To prove this lemma, we first show that there exists a positive constant c such that for all $\Omega \in \mathbf{S}$, $\|Q_\Omega\|_\infty \le c$.

We let Ω be an arbitrary element in $\mathbf{S}$, and ϕ be the affine mapping from $\widetilde{\Omega}$ to Ω. Let w_j, $j \in \mathbb{Z}_{r_k}$ denote the orthogonal polynomials of total degree $\le k-1$ on $\widetilde{\Omega}$, where $r_k := \binom{k+n-1}{n}$. We assume that these polynomials are normalized so that $\|w_j\|_\infty = 1$ and set $a_j := \|w_j\|_2$. It can be verified that $w_j \circ \phi^{-1}$, $j \in \mathbb{Z}_{r_k}$, are the orthogonal polynomials of total degree $\le k-1$ on Ω and $\|w_j \circ \phi^{-1}\|_\infty = 1$. Since ϕ is an affine mapping, the Jacobian $J(\phi)$ is a constant. We use J_ϕ to denote the constant $\sqrt{|J(\phi)|}$. In this notation, we have that $\|w_j \circ \phi^{-1}\|_2 = a_j J_\phi$.

Let $\alpha := \left(\sum_{j\in\mathbb{Z}_{r_k}} a_j^{-2}\right)^{1/2}$. For $f \in L^\infty(\Omega)$, we write

$$Q_\Omega f := \sum_{j\in\mathbb{Z}_{r_k}} c_j w_j \circ \phi^{-1}$$

and obtain

$$\begin{aligned}\|Q_\Omega f\|_\infty &\le \sum_{j\in\mathbb{Z}_{r_k}} |c_j| \le \frac{\alpha}{J_\phi}\left(\sum_{j\in\mathbb{Z}_{r_k}} |c_j|^2\alpha_j^2 J_\phi^2\right)^{1/2} = \frac{\alpha}{J_\phi}\|Q_\Omega f\|_2\\ &\le \frac{\alpha}{J_\phi}\|f\|_2 = \alpha\|f\circ\phi\|_2 \le \alpha\mathrm{meas}(\widetilde{\Omega})\|f\circ\phi\|_\infty = \alpha\mathrm{meas}(\widetilde{\Omega})\|f\|_\infty.\end{aligned} \tag{11}$$

Since for all $i \in \mathbb{N}_0$ and $j \in \mathbb{Z}_{N_n(i)}$,

$$P_{n,i} f|_{\Xi_{n,i,j}} = Q_{\Xi_{n,i,j}}\left(f|_{\Xi_{n,i,j}}\right).$$

and $\Xi_{n,i,j}$ can be affinely mapped from the unit simplex Ω', the result of this lemma is a straightforward consequence of inequality (11). □

Now, we review a multiscale property of $\mathbb{H}_{n,i}$ (cf., [14]). We use $d(A)$ to denote the diameter of a set A, let $\rho_{n,i} := \max\{d(\Xi_{n,i,j}) : j \in \mathbb{Z}_{N_n(i)}\}$ and observe that $\rho_{n,i}$ have the property:

(I) For each $n \in \mathbb{Z}_\nu$, there exist positive constants c_1, c_2 and an integer $\mu_n > 1$ such that for all $i \in \mathbb{N}_0$

$$c_1\mu_n^{-i/n} \le \rho_{n,i} \le c_2\mu_n^{-i/n}.$$

As usual, for a positive integer k, we use $W^{k,\infty}(\Omega_n)$ to denote the space of all functions h on Ω_n such that $D^\alpha h \in H_n^\infty$, for $|\alpha| \le k$, with the norm $\|h\|_{k,\infty} := \max\{\|D^\alpha h\| : |\alpha| \le k\}$. We define Λ_n for $h \in \mathbb{W}_n^\infty$ by

$$\Lambda_n h = \sum_{i\in\mathbb{N}_0} \lambda_{n,i} Q_{n,i} h.$$

Thus, Λ_n^{-1} has the form

$$\Lambda_n^{-1} y = \sum_{i\in\mathbb{N}_0} \frac{1}{\lambda_{n,i}} Q_{n,i} y$$

for $y \in W^{k,\infty}(\Omega_n)$. We next show that the inverse Λ_n^{-1} is bounded as a map from $W^{k,\infty}(\Omega_n)$ to $\mathbb{W}_n^\infty$.

Lemma 4 *If Δ is given and $\lambda' > 0$, then there exists a positive constant c such that for all $y \in W^{1,\infty}(\Omega_n)$,*

$$\sum_{i\in\mathbb{N}_0} \|Q_{n,i} y\|_\infty \le c\|y\|_{1,\infty}. \tag{12}$$

Therefore, $\Lambda_n^{-1} : W^{1,\infty}(\Omega_n) \to \mathbb{W}_n^\infty$ is bounded.

Proof Let $y \in W^{1,\infty}(\Omega_n)$. We first prove inequality (12). Noting that $Q_{n,i+1} := P_{n,i+1} - P_{n,i}$ for $i \in \mathbb{N}_0$, by the triangle inequality, we get that

$$\|Q_{n,i+1} y\|_\infty \le \|P_{n,i+1}(y - y_i)\|_\infty + \|P_{n,i}(y - y_i)\|_\infty, \quad \text{for any } \ y_i \in \mathbb{H}_{n,i}.$$

Now that for each $i \in \mathbb{N}_0$, $\mathbb{H}_{n,i}$ is a piecewise polynomial space, from Lemma 3, there is a positive constant c such that for all $i \in \mathbb{N}_0$, $\|P_{n,i}\|_\infty \le c$. Thus, there exists a positive constant c such that for $i \in \mathbb{N}_0$,

$$\|Q_{n,i+1}y\|_\infty \le c\|y - y_i\|_\infty. \tag{13}$$

On the other hand, since $y \in W^{1,\infty}(\Omega_n)$, from Property (I), we know that for each $i \in \mathbb{N}_0$, there exists $y_i \in \mathbb{H}_{n,i}$ such that

$$\|y - y_i\|_\infty \le c\mu_n^{-i/n}\|y\|_{1,\infty}. \tag{14}$$

Note that $\mu_n > 1$. Substituting (14) into (13) and using the summability of the geometric series, we obtain (12).

From the definition of Λ_n^{-1}, we have that

$$\left\|\Lambda_n^{-1}y\right\|_\infty \le \frac{1}{\lambda'}\sum_{i\in\mathbb{N}_0}\|Q_{n,i}y\|_\infty. \tag{15}$$

Substituting (12) into (15), we conclude that $\Lambda_n^{-1}y \in \mathbb{W}_n^\infty$ and $\Lambda_n^{-1} : W^{1,\infty}(\Omega_n) \to \mathbb{W}_n^\infty$ is bounded. □

In the next lemma, we show that the component $h_{n,\Delta}$ of the solution $\mathbf{h}_\Delta$ of equation (4) belongs to $\mathbb{W}_n^\infty$. To this end, we remark that the operator $\mathcal{U}_{n'}^*\mathcal{U}_n$ is an integral operator with the kernel

$$K_{n',n}(\eta,\xi) := \int_{\max\{\eta_{n-1},\xi_1\}}^{1}\left(\prod_{i=1}^{n}u(t-\xi_i)\right)\left(\prod_{i=1}^{n'}u(t-\eta_i)\right)dt, \tag{16}$$

for $\eta := (\eta_1,\ldots,\eta_n)^T \in \mathbb{R}^n$ and $\xi := (\xi_1,\ldots,\xi_{n'})^T \in \mathbb{R}^{n'}$. When the input signal $u \in C^1[0,T]$, the kernel $K_{n',n}$ has continuous partial derivatives $D_\eta^\alpha D_\xi^\beta K_{n',n}(\eta,\xi)$ for all $(\eta,\xi) \in \Omega_{n'} \times \Omega_n$ and for $|\alpha| \le 1$ and $|\beta| \le 1$. Specifically, there exists a positive constant c such that for $|\alpha| \le 1$ and $|\beta| \le 1$ and for all $(\eta,\xi) \in \Omega_{n'} \times \Omega_n$, $\left|D_\eta^\alpha D_\xi^\beta K_{n',n}(\eta,\xi)\right| \le c$. Thus, the range of $\mathcal{U}_{n'}^*\mathcal{U}_n$ is a subspace of $W^{1,\infty}(\Omega_n)$.

Lemma 5 *If Δ is given, $u \in C^1[0,T]$ and $\lambda' > 0$, then the component $h_{n,\Delta}$ of the solution $\mathbf{h}_\Delta$ of Eq. 4 belongs to $\mathbb{W}_n^\infty$.*

Proof We let $y := \mathcal{U}_n^*v - \mathcal{U}_n^*\mathcal{U}\mathbf{h}_\Delta$. From Eq. 4, we have that $\Lambda_n h_{n,\Delta} = y$. Since Λ_n is invertible on H_n, we obtain that $h_{n,\Delta} = \Lambda_n^{-1}y$. Since $u \in C^1[0,T]$, we get that $y \in W^{1,\infty}(\Omega_n)$. By Lemma 4 with $n = 1$, we conclude that $h_{n,\Delta} \in \mathbb{W}_n^\infty$. □

In practice, we use only finite number of regularization parameters. In the rest of this section, we consider the case where finite regularization parameters are used. We set for $n \in \mathbb{N}$, $\mathbb{Z}_n^0 := \{0, 1, 2, \ldots, n-1\}$, and $\Delta := \left\{\lambda_{n,i} : \lambda_{n,i} > 0, n \in \mathbb{Z}_\nu, i \in \mathbb{Z}_{\varsigma_n+1}^0\right\}$, where ς_n is a positive integer. In this situation, Λ takes the form

$$\Lambda_\Delta \mathbf{h} := \left[\sum_{i\in\mathbb{Z}_{\varsigma_n}^0}\lambda_{n,i}Q_{n,i}h_n + \lambda_{n,\varsigma_n}\left(\mathcal{I} - P_{n,\varsigma_n-1}\right)h_n : n \in \mathbb{Z}_\nu\right],$$

for all $\mathbf{h} := [h_1, h_2, \ldots, h_\nu] \in \mathbf{H}$.

To close this section, we describe the discrete equation of the collocation method for solving (4). The collocation method requires the availability of multiscale collocation functionals. We let $\mathbb{L}_n$ be the dual space of $C(\Omega_n)$. For $\ell \in \mathbb{L}_n$ and $h \in C(\Omega_n)$, we use $\langle \ell, h \rangle$ to denote the value of the continuous linear functional ℓ at h and use $\|\ell\|$ to denote the norm of ℓ. By using the Hahn-Banach Theorem (see [44], Theorem 3.1), we can obtain a norm preserving extension of ℓ to H_n^∞. We use the same notation for the extensions. A specific extension of the point evaluation functional from the continuous function space to the piecewise continuous function space is found in [2]. Let $\mathbb{F}$ be a finite dimensional subspace of $\mathbb{L}_n$ and $\mathbb{W}$ be a finite dimensional subspace of H_n^∞. We call $\mathbb{F}$ a $\mathbb{W}$-unisolvent if for $w, v \in \mathbb{W}$, $\langle \ell, w \rangle = \langle \ell, v \rangle$, for all $\ell \in \mathbb{F}$ implies that $w = v$. Following [11] (see also [10, 30]), we construct for each n a sequence of finite dimensional spaces $\mathbb{F}_{n,i} \subset \mathbb{L}_n$ such that $\mathbb{F}_{n,i}$ is $\mathbb{W}_{n,i}$-unisolvent and $\mathbb{L}_{n,i} := \mathbb{F}_{n,0} \oplus \mathbb{F}_{n,1} \oplus \cdots \oplus \mathbb{F}_{n,i}$ is $\mathbb{H}_{n,i}$-unisolvent.

To describe the collocation method in the abstract sense, we define the interpolatory projection $\mathcal{P}_{n,i}$ from H_n^∞ onto $\mathbb{H}_{n,i}$, $n \in \mathbb{Z}_\nu$, for $h \in H_n^\infty$ by $\langle \ell, \mathcal{P}_{n,i} h \rangle = \langle \ell, h \rangle$, for all $\ell \in \mathbb{L}_{n,i}$. We set

$$\mathbf{H}^\infty := H_1^\infty \times H_2^\infty \times \cdots \times H_\nu^\infty.$$

For a ν dimensional vector $L := (l_1, l_2, \ldots, l_\nu)^T$, $l_n \in \mathbb{N}_0$, $n \in \mathbb{Z}_\nu$, we define a subspace $\mathbb{H}_L$ of $\mathbf{H}^\infty$ by

$$\mathbb{H}_L := \mathbb{H}_{1,l_1} \times \mathbb{H}_{2,l_2} \times \cdots \times \mathbb{H}_{\nu,l_\nu}.$$

The projection $\mathcal{P}_L$ from $\mathbf{H}^\infty$ onto $\mathbb{H}_L$ is defined by $\mathcal{P}_L \mathbf{h} := [\mathcal{P}_{i,l_i} h_i : i \in \mathbb{Z}_\nu]^T$, for $\mathbf{h} \in \mathbf{H}^\infty$. With this projection, we let $\mho_L$ be the projection $\mathcal{P}_L$ applied to $\mathcal{U}^*\mathcal{U}$ with restriction on $\mathbb{H}_L$, that is $\mho_L := \mathcal{P}_L \mathcal{U}^*\mathcal{U}|_{\mathbb{H}_L}$, and we also let $\Lambda_L := \Lambda|_{\mathbb{H}_L}$. For a given vector L and a parameter set Δ, the collocation method is to find an $\mathbf{h}_{L,\Delta} \in \mathbb{H}_L$ such that

$$(\mho_L + \Lambda_L)\mathbf{h}_{L,\Delta} = \mathcal{P}_L \mathcal{U}^* v. \tag{17}$$

The multiscale collocation method is the collocation method (17) using the multiscale bases for $\mathbb{W}_{n,i}$ and $\mathbb{F}_{n,i}$. We call the multiscale basis for $\mathbb{F}_{n,i}$ the multiscale collocation functionals. The multiscale basis functions and the corresponding multiscale collocation functionals on an n dimensional simplex were constructed in [14]. We will not describe their construction here. Instead, we only define necessary notations sufficient for us to formulate the discrete form of the multiscale collocation method.

We let $w_n(i) := \dim(\mathbb{W}_{n,i})$. Then there exists a positive integer m_n such that $w_n(i+1) = m_n w_n(i)$. In fact, in the construction presented in [14] the integer m_n is the number of affine contractive mappings which define the simplex Ω_n. Suppose that $\mathbb{W}_{n,i} := \operatorname{span}\{w_{n,i,j} : j \in \mathbb{Z}^0_{w_n(i)}\}$, $i \in \mathbb{N}_0$. We let $\Omega_{n,i,j} \subset \Omega_n$ be the support of $w_{n,i,j}$. We remark that the set $\{\Omega_{n,i,j} : j \in \mathbb{Z}^0_{w_n(i)}\}$ is not a partition of Ω_n when $k_n > 1$. However, a fixed number of its elements correspond to the same element in the partition $\Xi_{n,i}$. Hence, their diameters satisfy Property (I). Moreover, for all $j, j' \in \mathbb{Z}^0_{w_n(i)}$, $\operatorname{meas}(\Omega_{n,i,j} \cap \Omega_{n,i,j'}) = 0$ or $\Omega_{n,i,j} = \Omega_{n,i,j'}$. Let $\mathbb{L}_{n,i} := \operatorname{span}\{\ell_{n,i,j} : j \in \mathbb{Z}^0_{w_n(i)}\}$, $i \in \mathbb{N}_0$. We use $\delta_{n,s}$ to denote the linear functional in $\mathbb{L}_n$ defined for $v \in C(\Omega_n)$ by $\langle \delta_{n,s}, v \rangle = v(s)$. The linear functional $\ell_{n,i,j}$ is a finite sum of point evaluations, that is, $\ell_{n,i,j} := \sum_{s \in \widehat{\Omega}_{n,i,j}} c_s \delta_s$, where c_s are constants and $\widehat{\Omega}_{n,i,j}$ is a finite subset of distinct points in $\Omega_{n,i,j}$ with the cardinality bounded independent of $n \in \mathbb{Z}_\nu$ and $(i, j) \in \mathbb{U}_n := \{(i, j) : i \in \mathbb{N}_0, j \in \mathbb{Z}^0_{w_n(i)}\}$. It is notable that the supports of the basis functions $w_{n,i,j}$, $(i, j) \in \mathbb{U}_n$, are shrinking as the level increases and so are the supports of the collocation functionals $\ell_{n,i,j}$.

Set $\mathbb{U}_{n,l_n} := \{(i, j) : i \in \mathbb{Z}^0_{l_n+1}, j \in \mathbb{Z}^0_{w_n(i)}\}$. Let $\mho_{n',n} := \mathcal{U}^*_{n'}\mathcal{U}_n$. We define the matrix block $(\mathbf{U}_{n',n})$ by

$$(\mathbf{U}_{n',n}) := \left[\langle \ell_{n',i',j'}, \mho_{n'n} w_{n,i,j}\rangle : (i', j') \in \mathbb{U}_{n',l'_n}, (i, j) \in \mathbb{U}_{n,l_n}\right].$$

Then the matrix representation of $\mho_L$ is given by

$$\mathbf{U}_L := \left[\mathbf{U}_{n',n} : n', n \in \mathbb{Z}_\nu\right]. \tag{18}$$

We set $\mathbf{E}_{n,L,\Delta} := \left[\lambda_{n,i} \langle \ell_{n,i',j'}, w_{n,i,j}\rangle : (i', j'), (i, j) \in \mathbb{U}_{n,l_n}\right]$ and obtain the matrix form $\mathbf{E}_{L,\Delta}$ of Λ_L, $\mathbf{E}_{L,\Delta} := \mathrm{diag}\left[\mathbf{E}_{n,L,\Delta} : n \in \mathbb{Z}_\nu\right]$. Let $\mathbf{v}_L := \left[\langle \ell_{n,i,j}, \mathcal{U}^* v\rangle : (i, j) \in \mathbb{U}_{n,l_n}, n \in \mathbb{Z}_\nu\right]$. We then obtain the discrete form of (17) given by

$$(\mathbf{U}_L + \mathbf{E}_{L,\Delta})\bar{\mathbf{h}}_{L,\Delta} = \mathbf{v}_L. \tag{19}$$

For notational convenience, we will drop the subscript Δ from Eq. 19 because in the next two sections we will focus on developing fast solutions for this equation after a set of parameters are chosen. Specifically, we let $\mathbf{E}_L := \mathbf{E}_{L,\Delta}$. Hence, Eq. 19 is written as

$$(\mathbf{U}_L + \mathbf{E}_L)\bar{\mathbf{h}}_L = \mathbf{v}_L. \tag{20}$$

This is a Volterra integral equation of the second kind. For traditional collocation methods for solving a Volterra integral equation of the second kind, the readers are referred to [6].

4 Compression strategy

We propose in this section a matrix compression strategy which approximates the coefficient matrix of linear system (20) by a sparse matrix. The compression strategy is based on multiscale properties of the basis and collocation functionals used in the collocation method.

We begin with a review of the multiscale properties of the basis functions and collocation functionals.

(II) For each $n \in \mathbb{Z}_\nu$, there exist positive constants c_1, c_2 such that for all $i \in \mathbb{N}_0$

$$c_1 \mu_n^i \leq \dim \mathbb{W}_{n,i} \leq c_2 \mu_n^i,$$

where μ_n is the constant defined in Property (I).

(III) For any $n \in \mathbb{N}$, $\langle \ell_{n,i',j'}, w_{n,i,j}\rangle = \delta_{i'i}\delta_{j'j}$, $\quad i \leq i'$, $\quad j' \in \mathbb{Z}^0_{w_n(i')}$, $\quad j \in \mathbb{Z}^0_{w_n(i)}$, and there exists a positive constant γ_n such that for all $(i', j'), (i, j) \in \mathbb{U}_n$, $i > i'$, $\sum_{j \in \mathbb{Z}^0_{w_n(i)}} \left|\langle \ell_{n,i',j'}, w_{n,i,j}\rangle\right| \leq \gamma_n$.

(IV) For any polynomial p of total degree $k_n - 1$ on Ω_n, $\langle \ell_{n,i,j}, p \rangle = 0$ and $(w_{n,i,j}, p)_n = 0$, for $(i, j) \in \mathbb{U}_n$, where $(\cdot, \cdot)_n$ denotes the inner product in $L^2(\Omega_n)$, $n \in \mathbb{Z}_\nu$.
Property (IV) is called the vanishing moment property of the collocation functionals and the basis functions. It is crucial for establishing the matrix compression scheme.

(V) There exists a positive constant c_3 such that for all $n \in \mathbb{Z}_\nu$, $(i, j) \in \mathbb{U}_n$, $\|\ell_{n,i,j}\| + \|w_{n,i,j}\|_\infty \le c_3$.

Property (V) means that $\ell_{n,i,j}$ and $w_{n,i,j}$ are uniformly bounded.

As a preparation for compression of the matrix $\mathbf{U}_{n',n}$, we estimate the norm of the matrix block

$$(\mathbf{U}_{n',n})_{i',i} := \left[\langle \ell_{n',i',j'}, \mho_{n',n} w_{n,i,j} \rangle : j' \in \mathbb{Z}^0_{w_n(i')}, j \in \mathbb{Z}^0_{w_n(i)} \right]$$

of matrix $\mathbf{U}_{n',n}$. Let $k' := \max\{k_n : n \in \mathbb{Z}_\nu\}$.

Lemma 6 *If $u \in C^{k'}[0, T]$, then there exists a positive constant c such that for all $i' \in \mathbb{Z}^0_{l_{n'}+1}$, $i \in \mathbb{Z}^0_{l_n+1}$,*

$$\left\| \left(\mathbf{U}_{n',n} \right)_{i',i} \right\|_\infty \le c \mu_n^{-k_n i/n - (\log_{\mu_n} \mu_{n'}) k_{n'} i'/n'}. \tag{21}$$

Proof Note that the support of $w_{n,i,j}$ is $\Omega_{n,i,j}$. As proceeded in [11] by using the Taylor expansion and by Property (IV), and the fact that the kernels are smooth, we have that

$$\left| \left(\mathbf{U}_{n',n} \right)_{i',j';i,j} \right| \le c \|\ell_{n',i',j'}\| \|w_{n,i,j}\|_\infty \rho_{n',i'}^{k_{n'}} \rho_{n,i}^{k_n+n}. \tag{22}$$

By the definition of the infinity norm of a matrix, we obtain that

$$\left\| \left(\mathbf{U}_{n',n} \right)_{i',i} \right\|_\infty = \max_{j' \in \mathbb{Z}^0_{w_{n'}(i')}} \sum_{j \in \mathbb{Z}^0_{w_n(i)}} \left| \left(\mathbf{U}_{n',n} \right)_{i',j';i,j} \right|. \tag{23}$$

By Property (V) and substituting (22) into (23), we conclude that there exits a positive constant c such that

$$\left\| \left(\mathbf{U}_{n',n} \right)_{i',i} \right\|_\infty \le c \rho_{n',i'}^{k_{n'}} \rho_{n,i}^{k_n+n} w_n(i). \tag{24}$$

From Properties (I), (II) and inequality (24), we obtain that

$$\left\| \left(\mathbf{U}_{n',n} \right)_{i',i} \right\|_\infty \le c \mu_{n'}^{-k_{n'} i'/n'} \mu_n^{-k_n i/n} = c \mu_n^{-k_n i/n - (\log_{\mu_n} \mu_{n'}) k_{n'} i'/n'},$$

proving the lemma. □

Lemma 6 leads to a block compression strategy for the matrix $\mathbf{U}_{n',n}$. Specifically, we wish to form a new matrix $\mathbf{V}_{n',n}$ by assigning most of blocks in $\mathbf{U}_{n',n}$ to be zero without ruining the convergence order of the corresponding approximate solution. Hence, we require the matrix $\mathbf{V}_{n',n}$ to have the estimate that there exists a positive constant c such that

$$\sum_{i \in \mathbb{Z}^0_{l_n+1}} \left\| (\mathbf{U}_{n',n})_{i',i} - (\mathbf{V}_{n',n})_{i',i} \right\| \mu^{k_n (l_n - i)/n} \le c.$$

This yields the following block truncation strategy. For $n, n' \in \mathbb{Z}_\nu$, we let $\kappa_{n',n} := \frac{k_n n'}{k_{n'} n(\log_{\mu_n} \mu_{n'})}$ and for a given L, we set

$$\left(\mathbf{V}_{n',n}\right)_{i'i} := \begin{cases} \left(\mathbf{U}_{n',n}\right)_{i'i}, & 2i + i'/\kappa_{n',n} \leq l_n, \\ 0, & \text{otherwise}. \end{cases} \tag{25}$$

Thus, for a given L, we have the compression matrix

$$\mathbf{V}_L := \left[\mathbf{V}_{n',n} : n', n \in \mathbb{Z}_\nu\right], \tag{26}$$

which is an approximation of $\mathbf{U}_L$. In Eq. 20 we replace $\mathbf{U}_L$ by $\mathbf{V}_L$ and have the compression method

$$(\mathbf{V}_L + \mathbf{E}_L)\,\bar{\mathbf{h}}_L = \mathbf{v}_L. \tag{27}$$

The coefficient matrix of this equation is now sparse, which leads to a fast solution.

5 Computational complexity

In this section, we consider the computational complexity of the compression method by showing that the number of nonzero entries of the compression matrix $\mathbf{V}_L$ is bounded by a constant multiple of the size of the matrix. For a matrix $\mathbf{A}$, we use $\mathcal{N}(\mathbf{A})$ to denote the number of the nonzero entries of $\mathbf{A}$. By N_L we denote the size of matrix $\mathbf{V}_L$ and we will prove that $\mathcal{N}(\mathbf{V}_L) = \mathcal{O}(N_L)$. To this end, we recall the definition of $\kappa_{n',n}$ and assume that there is a positive constant ρ such that the components $l_n, n \in \mathbb{Z}_\nu$, of the vector L satisfy

$$|\kappa_{n',n} l_n - l_{n'}| \leq \rho. \tag{28}$$

We remark that condition (28) limits the choice of $\{l_n : n \in \mathbb{Z}_\nu\}$. This condition balances the levels of approximation for different kernels.

Theorem 2 *If $\mathbf{V}_L$ is chosen by compression strategy (25) for L satisfying (28), then there exists a positive constant c such that*

$$\mathcal{N}(\mathbf{V}_L) \leq c N_L. \tag{29}$$

Proof By the definition of $\mathbf{V}_L$, it suffices to estimate $\mathcal{N}(\mathbf{V}_{n',n})$. For fixed $n, n' \in \mathbb{Z}_\nu$, we let $\gamma_i := \lceil \kappa_{n',n}(l_n - 2i) \rceil + 1$, $i \in \mathbb{Z}_\nu$. From the compression strategy, if $i' \geq \gamma_i$ or $i \geq \left\lceil \frac{l_n}{2} \right\rceil + 1$, then block $(\mathbf{V}_{n',n})_{i',i}$ is equal to zero. Hence, we get that

$$\mathcal{N}\left(\mathbf{V}_{n',n}\right) \leq \sum_{i \in \mathbb{Z}^0_{\lceil \frac{l_n}{2} \rceil + 1}} \sum_{i' \in \mathbb{Z}^0_{\gamma_i}} \mathcal{N}\left((\mathbf{V}_{n',n})_{i',i}\right). \tag{30}$$

Noting that, from Property (II), there is a positive constant c such that for all $n', n \in \mathbb{Z}_\nu$ and $i' \in \mathbb{Z}^0_{l_{n'}+1}$, $i \in \mathbb{Z}^0_{l_n+1}$, the number of entries of $(V_{n',n})_{i',i}$ is less than or equal to $c\mu_{n'}^{i'}\mu_n^i$. Substituting it into (30) and using the summability of the geometric series, we get that

$$\mathcal{N}\left(\mathbf{V}_{n',n}\right) \leq c \sum_{i \in \mathbb{Z}^0_{\lceil \frac{l_n}{2} \rceil + 1}} \mu_n^i \mu_{n'}^{\lceil \kappa_{n',n}(l_n - 2i) \rceil + 1} \leq c\mu_{n'}^{\kappa_{n',n} l_n} \sum_{i \in \mathbb{Z}^0_{\lceil \frac{l_n}{2} \rceil + 1}} \left(\frac{\mu_n}{\mu_{n'}^{2\kappa_{n',n}}} \right)^i. \tag{31}$$

We next consider three different cases.

Case 1: $\mu_n = \mu_{n'}^{2\kappa_{n',n}}$. In this case,

$$\mathcal{N}\left(\mathbf{V}_{n',n}\right) \leq c\left(\left\lceil \frac{l_n}{2} \right\rceil + 1\right)\mu_{n'}^{\kappa_{n',n} l_n} \leq c\left(\left\lceil \frac{l_n}{2} \right\rceil + 1\right)\mu_n^{l_n/2}.$$

Case 2: $\mu_n > \mu_{n'}^{2\kappa_{n',n}}$. Substituting this condition into (31), we see that $\mathcal{N}\left(\mathbf{V}_{n',n}\right) \leq c\mu_n^{l_n}$.

Case 3: $\mu_n < \mu_{n'}^{2\kappa_{n',n}}$. In this case, it follows from (31) that $\mathcal{N}\left(\mathbf{V}_{n',n}\right) \leq c\mu_{n'}^{\kappa_{n',n} l_n}$. Form (28), we know that $\kappa_{n',n} l_n \leq l'_n + c$. Hence, we obtain that $\mathcal{N}\left(\mathbf{V}_{n',n}\right) \leq c\mu_{n'}^{l_{n'}}$.

Summing all $\mathcal{N}\left(\mathbf{V}_{n',n}\right)$, for $n, n' \in \mathbb{Z}_\nu$, we conclude that

$$\mathcal{N}(\mathbf{V}_L) \leq c \sum_{n \in \mathbb{Z}_\nu} \mu_n^{l_n} \leq c N_L,$$

where the second inequality follows from Property (II), and obtain the desired estimate (29). □

Theorem 2 reveals that the compression method (27) reduces the number of nonzero entries of the coefficient matrix from $\mathcal{O}\left(N_L^2\right)$ to $\mathcal{O}(N_L)$. This significant reduction in the computational cost serves as a base for a fast algorithm for solving the Volterra identification problem and it makes it possible to use the Volterra system of high order (with order greater than two). In the next section, we will show that this compression strategy will *not* ruin the order of convergence of the original collocation method.

6 Convergence analysis

In this section, we analyze the convergence order of the compression algorithm. We will show that the compression method converges in a nearly optimal order.

To establish the convergence result, we first discuss a somewhat more general problem described below. Let $\mathbf{K}_{n',n}$ be an approximation of $\mathbf{U}_{n',n}$ and $\mathbf{K}_L := \left[\mathbf{K}_{n',n} : n', n \in \mathbb{Z}_\nu\right]$. We assume that the matrices $\mathbf{K}_{n',n}$, $n', n \in \mathbb{Z}_\nu$ satisfy the condition that there exists a positive constant c such that for all $h \in W^{k_n,\infty}$,

$$\left\| \left(\mathbf{U}_{n',n} - \mathbf{K}_{n',n}\right)\bar{h} \right\|_\infty \leq c\mu_n^{-k_n l_n/n} \|h\|_{k_n,\infty} \tag{32}$$

and for all $h \in H_n^{\infty}$,

$$\left\| \left(\mathbf{U}_{n',n} - \mathbf{K}_{n',n}\right) \bar{h} \right\|_{\infty} \leq c\mu_n^{-\frac{1}{2}k_n l_n/n} \|h\|_{\infty}, \tag{33}$$

where $\bar{h} := \left[h_{i,j} : (i,j) \in \mathbb{U}_{n,l_n}\right]$ with

$$\mathcal{P}_{n,l_n} h = \sum_{(i,j)\in\mathbb{U}_{n,l_n}} h_{i,j} w_{n,i,j}. \tag{34}$$

In Eq. 20, we replace $\mathbf{U}_L$ by $\mathbf{K}_L$ and obtain an approximate equation of (20)

$$(\mathbf{K}_L + \mathbf{E}_L)\,\overline{\mathbf{h}}_L = \mathbf{v}_L. \tag{35}$$

We write the solution $\overline{\mathbf{h}}_L$ of (35) as $\overline{\mathbf{h}}_L = \left[\left(\overline{\mathbf{h}}_L\right)_{n,i,j} : n \in \mathbb{Z}_\nu,\ (i,j) \in \mathbb{U}_{n,l_n}\right]$, and let

$$\mathbf{h}_{n,l_n} := \sum_{i=0}^{l_n} \sum_{j=0}^{w_n(i)-1} \left[\overline{\mathbf{h}}_L\right]_{n,i,j} w_{n,i,j},\ \ n \in \mathbb{Z}_\nu \ \text{ and } \ \mathbf{h}_L := \left[\mathbf{h}_{n,l_n}, n \in \mathbb{Z}_\nu\right]^T.$$

We also write the solution of Eq. 4 as $\mathbf{h}_\Delta = [\mathbf{h}_n : n \in \mathbb{Z}_\nu]^T$. Our convergence result concerns the bound for $\|\mathbf{h}_n - \mathbf{h}_{n,l_n}\|_\infty$.

To estimate the bound, we let $\overline{\mathcal{B}}_{n',n} : \mathbb{H}_{n,l_n} \to \mathbb{H}_{n',l_{n'}}$ be the operator whose discrete form is $\mathbf{K}_{n',n}$ in the same basis and collocation functionals, and define the operator $\overline{\mathcal{B}}_L : \mathbb{H}_L \to \mathbb{H}_L$ by $\overline{\mathcal{B}}_L := \left[\overline{\mathcal{B}}_{n',n} : n', n \in \mathbb{Z}_\nu\right]$. From the definitions of $\overline{\mathcal{B}}_L$, $\mathbf{h}_L$ and $\mathbf{v}_L$, we observe that Eq. 35 has the operator equation formulation

$$\left(\overline{\mathcal{B}}_L + \Lambda_L\right)\mathbf{h}_L = \mathcal{P}_L \mathcal{U}^* v. \tag{36}$$

We next present two technical lemmas. For $n, n' \in \mathbb{Z}_\nu$ we let $\mathcal{B}_{n',n} := \mathcal{P}_{n',l_{n'}} \mho_{n'n}|_{\mathbb{H}_{n,l_n}}$. It can be verified that matrix $\mathcal{U}_{n',n}$ is the discrete form of $\mathcal{B}_{n',n}$ in the basis and collocation functionals used in the multiscale collocation method. We define two operators $\overline{\mathcal{B}}_n := \left[\overline{\mathcal{B}}_{n,m} : m \in \mathbb{Z}_\nu\right]$ and $\mho_n := \mathcal{P}_{l_n} \mathcal{U}_n^* \mathcal{U}|_{\mathbb{H}_L}$ which map from $\mathbb{H}_L \to \mathbb{H}_{n,l_n}$. In the next lemma, we estimate the error $\mho_n - \overline{\mathcal{B}}_n$. To this end, for $\mathbf{k} := [k_n : n \in \mathbb{Z}_\nu]$ we define the space

$$\mathbf{W}^{\mathbf{k},\infty} := W^{k_1,\infty}(\Omega_1) \times W^{k_2,\infty}(\Omega_2) \times \cdots \times W^{k_\nu,\infty}(\Omega_\nu).$$

To prepare for the development of this result, we present another property of the basis functions and the collocation functionals:

(VI) There exist positive constants c_1 and c_2 such that for all $n \in \mathbb{Z}_\nu$, $l_n \in \mathbb{N}_0$ and for all $h \in \mathbb{H}_{n,l_n}$ with $h := \sum_{(i,j)\in\mathbb{U}_{n,l_n}} h_{i,j} w_{n,i,j}$ there holds

$$c_1 \|\bar{h}\|_\infty \leq \|h\|_\infty \leq c_2(l_n + 1)\|\mathbf{E}_{n,l_n}\bar{h}\|_\infty,$$

where $\bar{h} := \left[h_{i,j}, (i,j) \in \mathbb{U}_{n,l_n}\right]$, and $\mathbf{E}_{n,l_n} := \left[\left\langle \ell_{n,i',j'}, w_{n,i,j}\right\rangle : (i',j'), (i,j) \in \mathbb{U}_{n,l_n}\right]$.

Lemma 7 *Let $u \in C^{k'}[0,T]$. If (32) and (33) hold, then there exists a positive constant c such that for all $\mathbf{h} \in \mathbf{W}^{\mathbf{k},\infty}$,*

$$\left\|(\mho_n - \overline{\mathcal{B}}_n)\mathcal{P}_L\mathbf{h}\right\|_\infty \le c(l_n+1)\sum_{m\in\mathbb{Z}_\nu} \mu_m^{-k_m l_m/m}\|h_m\|_{k_m,\infty} \tag{37}$$

and for all $\mathbf{h} \in \mathbf{H}^\infty$,

$$\left\|\left(\mho_n - \overline{\mathcal{B}}_n\right)\mathcal{P}_L\mathbf{h}\right\|_\infty \le c(l_n+1)\sum_{m\in\mathbb{Z}_\nu} \mu_m^{-\frac{1}{2}k_m l_m/m}\|h_m\|_\infty . \tag{38}$$

Proof The proof of this lemma demands the following two estimates: There exists a positive constant c such that for all $h \in W^{k_n,\infty}(\Omega_n)$

$$\left\|\left(\mathcal{B}_{n',n} - \overline{\mathcal{B}}_{n',n}\right)\mathcal{P}_{n,l_n}h\right\|_\infty \le c(l_{n'}+1)\mu_n^{-k_n l_n/n}\|h\|_{k_n,\infty}, \tag{39}$$

and for all $h \in \mathbf{H}^\infty$,

$$\left\|\left(\mathcal{B}_{n',n} - \overline{\mathcal{B}}_{n',n}\right)\mathcal{P}_{n,l_n}h\right\|_\infty \le c(l_{n'}+1)\mu_n^{-\frac{1}{2}k_n l_n/n}\|h\|_\infty . \tag{40}$$

We now present a proof for (39) and (40). For $h \in W^{k_n,\infty}$, we write

$$\left(\mathcal{B}_{n',n} - \overline{\mathcal{B}}_{n',n}\right)\mathcal{P}_{n,l_n}h = \sum_{(i,j)\in\mathbb{U}_{n',l_{n'}}} e_{i,j}w_{n',i,j}, \tag{41}$$

where $e_{i,j} \in \mathbb{R}$. Let $\bar{h} := \left[h_{i,j} : (i,j) \in \mathbb{U}_{n,l_n}\right]$ and $e := \left[e_{i,j} : (i,j) \in \mathbb{U}_{n',l_{n'}}\right]$. Using Property (VI) in (41), we see that there exists a positive constant c such that

$$\left\|\left(\mathcal{B}_{n',n} - \overline{\mathcal{B}}_{n',n}\right)\mathcal{P}_{n,l_n}h\right\|_\infty \le c(l_{n'}+1)\left\|\mathbf{E}_{n',l_{n'}}e\right\|_\infty .$$

Applying functionals $\ell_{n,i,j}$, $(i,j) \in \mathbb{U}_{n',l_{n'}}$, to both sides of (41), we obtain the equation

$$\left(\mathbf{U}_{n',n} - \mathbf{K}_{n',n}\right)\bar{h} = \mathbf{E}_{n',l_{n'}}e.$$

Hence, we get that

$$\left\|\left(\mathcal{B}_{n',n} - \overline{\mathcal{B}}_{n',n}\right)\mathcal{P}_{n,l_n}h\right\|_\infty \le c(l_{n'}+1)\left\|\left(\mathbf{U}_{n',n} - \mathbf{K}_{n',n}\right)\bar{h}\right\|_\infty . \tag{42}$$

Substituting (32) into (42), we get estimate (39). By substituting (33) into (42) proves the desired estimate (40).

Estimates (37) and (38) are obtained by using (39) and (40), respectively, in conjunction with the inequality

$$\left\|\left(\mho_n - \overline{\mathcal{B}}_n\right)\mathcal{P}_L\mathbf{h}\right\|_\infty \le \sum_{m\in\mathbb{Z}_\nu}\left\|\left(\mathcal{B}_{n,m} - \overline{\mathcal{B}}_{n,m}\right)\mathcal{P}_{m,l_m}h_m\right\|_\infty ,$$

completing the proof. □

Next, we estimate the error $\left\|\overline{\mathcal{B}}_n - \mathcal{U}_n^*\mathcal{U}\right\|_\infty$.

Lemma 8 *Let* $u \in C^{k'}[0, T]$. *If* (32) *and* (33) *hold, then there exists a positive constant* c *such that for all* $\mathbf{h} \in \mathbf{H}^\infty$,

$$\left\|\left(\overline{\mathcal{B}}_n - \mathcal{U}_n^*\mathcal{U}\right)\mathcal{P}_L\mathbf{h}\right\|_\infty \le c\left[(l_n+1)\sum_{m\in\mathbb{Z}_\nu}\mu_m^{-\frac{1}{2}k_m l_m/m}\,\|h_m\|_{k_m,\infty} + \mu_n{}^{-k_n l_n/n}\|\mathbf{h}\|_\infty\right]. \tag{43}$$

Proof Since for all $\mathbf{h} \in \mathbb{H}_L$, we have that $\mathcal{U}_n^*\mathcal{U}\mathbf{h} \in W^{k_n,\infty}(\Omega_n)$, there exists a positive constant c such that

$$\left\|(\mho_n - \mathcal{U}_n^*\mathcal{U})\mathbf{h}\right\|_\infty \le \|(\mathcal{I} - \mathcal{P}_{n,l_n})\mathcal{U}_n^*\mathcal{U}\mathbf{h}\|_\infty \le c\mu_n{}^{-k_n l_n/n}\|\mathbf{h}\|_\infty. \tag{44}$$

Therefore, estimate (43) follows directly from the second estimate of Lemma 7, inequality (44) and the triangle inequality. □

To give a bound for $\|\mathbf{h}_n - \mathbf{h}_{n,l_n}\|_\infty$, we estimate $\|(\mathcal{U}^*\mathcal{U} + \Lambda_\Delta)^{-1}\|_\infty$. According to the definition of $\mathcal{U}^*$, we know that for all $f \in L^2([0, T])$,

$$\|\mathcal{U}^* f\|_\infty \le W\|f\|_2, \tag{45}$$

where

$$W := \max_{n\in\mathbb{Z}_\nu}\ \sup_{(\xi_1,\xi_2,\ldots,\xi_n)\in\Omega_n}\left(\int_{\xi_1}^T\left|\prod_{i=1}^n u(t-\xi_i)\right|dt\right)^{\frac{1}{2}}.$$

Lemma 9 *For a given* Δ, *there exists a positive constant* c *such that*

$$\left\|\left(\Lambda_\Delta + \mathcal{U}^*\mathcal{U}\right)^{-1}\right\|_\infty \le c. \tag{46}$$

Proof For $w \in \mathbf{H}^\infty$ we set

$$v := \left(\Lambda_\Delta + \mathcal{U}^*\mathcal{U}\right)^{-1} w \ \text{ and } \ z := \mathcal{U}\left(\Lambda_\Delta + \mathcal{U}^*\mathcal{U}\right)^{-1} w.$$

It follows that $v = \Lambda_\Delta^{-1}(w - \mathcal{U}^* z)$. From Lemma 3, we know that for each $n \in \mathbb{Z}_\nu$, $\|Q_{n,i}\|_\infty, i \in \mathbb{N}_0$ are uniform bounded. Recalling that for all $\mathbf{h} := [h_1, h_2, \ldots, h_\nu] \in \mathbf{H}$,

$$\Lambda_\Delta^{-1}\mathbf{h} := \left[\sum_{i=0}^{\varsigma_n-1}\lambda_{n,i}^{-1}Q_{n,i}h_n + \lambda_{n,\varsigma_n}^{-1}\left(\mathcal{I} - P_{n,\varsigma_n-1}\right)h_n : n \in \mathbb{Z}_\nu\right],$$

we have that there exists a positive constant $c_1 > 0$ such that $\left\|\Lambda_\Delta^{-1}\right\|_\infty \le c_1\frac{1}{\lambda'}$. From (5) and (45), there exists a positive constant $c_2 > 0$ such that

$$\|\mathcal{U}^* z\|_\infty \le c_2\frac{W}{\sqrt{\lambda'}}\|w\|_\infty.$$

Hence, we know that there exist a positive constant c such that

$$\left\|\left(\Lambda_\Delta+\mathcal{U}^*\mathcal{U}\right)^{-1}w\right\|_\infty=\|v\|_\infty=\|\Lambda_\Delta^{-1}\left(w-\mathcal{U}^*z\right)\|_\infty$$
$$\leq c\frac{1}{\lambda'}\left(\|w\|_\infty+\|\mathcal{U}^*z\|_\infty\right)$$
$$\leq c\frac{1}{\lambda'}\left(1+\frac{W}{\sqrt{\lambda'}}\right)\|w\|_\infty.$$

This leads to estimate (46). □

Now, we estimate the bound for $\|\mathbf{h}_n-\mathbf{h}_{n,l_n}\|_\infty$.

Theorem 3 *For a given* Δ, *let* $\mathbf{h}_\Delta\in\mathbf{W}^{\mathbf{k},\infty}$ *be the solution of Eq. 4 and* $u\in C^{k'}[0,T]$. *If (32) and (33) hold, then there exist a positive constant c and a positive integer* $\ell\in\mathbb{N}$ *such that for all* $L\in\mathbb{N}^\nu$ *satisfying* $l':=\inf\{l_n:n\in\mathbb{Z}_\nu\}>\ell$ *and for all* $n\in\mathbb{Z}_\nu$,

$$\|\mathbf{h}_n-\mathbf{h}_{n,l_n}\|_\infty\leq c\left(\mu_n^{-k_nl_n/n}\|\mathbf{h}_n\|_{k_n,\infty}+(l_n+1)\sum_{m\in\mathbb{Z}_\nu}\mu_m^{-k_ml_m/m}\|\mathbf{h}_m\|_{k_m,\infty}\right). \tag{47}$$

Proof By Lemmas 8 and 9, there exist a positive constant c and an $\ell\in\mathbb{N}$ such that for all $L\in\mathbb{N}^\nu$ satisfying $l'>\ell$ and for all $\mathbf{h}\in\mathbb{H}_L$, $c\|(\Lambda_L+\overline{\mathcal{B}}_L)\mathbf{h}\|_\infty\geq\|\mathbf{h}\|_\infty$. It follows from the fact $\mathcal{P}_L\mathbf{h}_\Delta-\mathbf{h}_L\in\mathbb{H}_L$ that

$$\|\mathbf{h}_\Delta-\mathbf{h}_L\|_\infty\leq\|\mathbf{h}_\Delta-\mathcal{P}_L\mathbf{h}_\Delta\|_\infty+c\|\left(\Lambda_L+\overline{\mathcal{B}}_L\right)(\mathcal{P}_L\mathbf{h}_\Delta-\mathbf{h}_L)\|_\infty. \tag{48}$$

From Eqs. 4 and 36, we know that $\mathcal{P}_L\left(\Lambda+\mathcal{U}^*\mathcal{U}\right)\mathbf{h}_\Delta=\left(\Lambda_L+\overline{\mathcal{B}}_L\right)\mathbf{h}_L$, which leads to the formula

$$\left(\Lambda_L+\overline{\mathcal{B}}_L\right)(\mathcal{P}_L\mathbf{h}_\Delta-\mathbf{h}_L)=\mathcal{P}_L(\Lambda+\mathcal{U}^*\mathcal{U})(\mathcal{P}_L\mathbf{h}_\Delta-\mathbf{h}_\Delta)+\left(\overline{\mathcal{B}}_L-\mho_L\right)\mathcal{P}_L\mathbf{h}_\Delta. \tag{49}$$

Substituting (49) into (48), we get that

$$\|\mathbf{h}_\Delta-\mathbf{h}_L\|_\infty\leq(1+c\|\Lambda+\mathcal{U}^*\mathcal{U}\|_\infty)\|\mathbf{h}_\Delta-\mathcal{P}_L\mathbf{h}_\Delta\|_\infty+c\|\left(\overline{\mathcal{B}}_L-\mho_L\right)\mathcal{P}_L\mathbf{h}_\Delta\|_\infty.$$

In particular, for $n\in\mathbb{Z}_\nu$ we have that

$$\|\mathbf{h}_n-\mathbf{h}_{n,l_n}\|_\infty\leq(1+c\|\Lambda+\mathcal{U}^*\mathcal{U}\|_\infty)\|\mathbf{h}_n-\mathcal{P}_{n,l_n}\mathbf{h}_n\|_\infty+c\|\left(\overline{\mathcal{B}}_n-\mho_n\right)\mathcal{P}_L\mathbf{h}_\Delta\|_\infty.$$

Finally, since $\mathbf{h}_\Delta\in\mathbf{W}^{\mathbf{k},\infty}$, there exists a positive constant c such that for $n\in\mathbb{Z}_\nu$,

$$\|\mathbf{h}_n-\mathcal{P}_L\mathbf{h}_n\|_\infty\leq c\mu_n^{-k_nl_n/n}\|\mathbf{h}_n\|_{k_n,\infty}. \tag{50}$$

From the first estimate of Lemma 7 and (50), we obtain (47). □

In the rest of this section, we consider the convergence property of the compression method (27). For this purpose, we prove that estimates (32) and (33) hold when $\mathbf{K}_{n',n} := \mathbf{V}_{n',n}$. For fixed n, n', we set $\mathbb{S}_{i'} := \{i : \kappa_{n',n}(l_n - 2i) - i' < 0\}$.

Lemma 10 *If $u \in C^{k'}[0, T]$, then there exists a positive constant c such that for $h \in W^{k_n,\infty}$,*

$$\left\| \left(\mathbf{U}_{n',n} - \mathbf{V}_{n',n}\right) \bar{h} \right\|_{\infty} \le c\mu_n^{-k_n l_n/n} \|h\|_{k_n,\infty} \tag{51}$$

and for $h \in H_n^{\infty}$,

$$\left\| \left(\mathbf{U}_{n',n} - \mathbf{V}_{n',n}\right) \bar{h} \right\|_{\infty} \le c\mu_n^{-\frac{1}{2}k_n l_n/n} \|h\|_{\infty}, \tag{52}$$

where $\bar{h} := [h_{i,j} : (i, j) \in \mathbb{U}_{n,l_n}]$ satisfy (34).

Proof We first prove Eq. 51. Since $h \in W^{k_n,\infty}$, by Proposition 4.1 of [14], there exists a positive constant c such that

$$\max\left\{|h_{i,j}| : j \in \mathbb{Z}^0_{w_n(i)}\right\} \le c\mu_n^{-k_n i/n}\|h\|_{k_n,\infty}, \quad \text{for all } i \in \mathbb{Z}^0_{l_n+1}.$$

It follows that

$$\left\| \left(\mathbf{U}_{n',n} - \mathbf{V}_{n',n}\right) \bar{h} \right\|_{\infty} \le c \max_{i' \in \mathbb{Z}^0_{l_{n'}+1}} \sum_{i \in \mathbb{Z}^0_{l_n+1}} \left\| \left(\mathbf{U}_{n',n}\right)_{i',i} - \left(\mathbf{V}_{n',n}\right)_{i',i} \right\|_{\infty} \mu_n^{-k_n i/n} \|h\|_{k_n,\infty}. \tag{53}$$

Using the compression strategy and substituting (21) into (53), we obtain that

$$\left\| \left(\mathbf{U}_{n',n} - \mathbf{V}_{n',n}\right) \bar{h} \right\|_{\infty} \le c\mu_n^{-k_n l_n/n} \|h\|_{k_n,\infty} \max_{i' \in \mathbb{Z}^0_{l_{n'}+1}} \sum_{i \in \mathbb{S}_{i'}} \mu_n^{k_n/n\left(l_n - 2i - \frac{i'}{\kappa_{n',n}}\right)}. \tag{54}$$

To bound the right-hand side of the above inequality, we set $i_{i',0} := \min\{i : i \in \mathbb{S}_{i'}\}$. Using the summability of geometric series, we have the estimate

$$\left\| \left(\mathbf{U}_{n',n} - \mathbf{V}_{n',n}\right) \bar{h} \right\|_{\infty} \le c\frac{\mu_n^{-k_n l_n/n} \|h\|_{k_n,\infty}}{\left(1 - \mu_n^{-2k_n/n}\right)} \max_{i' \in \mathbb{Z}^0_{l_{n'}+1}} \mu_n^{k_n/n\left(l_n - 2i_{i',0} - \frac{i'}{\kappa_{n',n}}\right)}. \tag{55}$$

Since $i_{i',0} \in \mathbb{S}_{i'}$, by the definition of the set $\mathbb{S}_{i'}$, we obtain that for all i', $\mu_n^{k_n/n\left(l_n - 2i_{i',0} - \frac{i'}{\kappa_{n',n}}\right)} < 1$. From (55), we obtain estimate (51).

To prove estimate (52), we set $\mathbf{B} := \mu_n^{\frac{1}{2}k_n l_n/n}\left(\mathbf{U}_{n',n} - \mathbf{V}_{n',n}\right)$. By Property (VI), there exists a positive constant c such that

$$\left\| \left(\mathbf{U}_{n',n} - \mathbf{V}_{n',n}\right) \bar{h} \right\|_{\infty} \le c\mu_n^{-\frac{1}{2}k_n l_n/n} \|\mathbf{B}\|_{\infty} \|h\|_{\infty}. \tag{56}$$

Since $u \in C^{k'}[0, T]$, by Lemma 6, there exists a positive constant c such that

$$\|\mathbf{B}\|_\infty \leq c \max_{i' \in \mathbb{Z}^0_{l_{n'}}} \sum_{i \in \mathbb{S}_{i'}} \mu_n^{k_n/n\left(\frac{1}{2}l_n - i - \frac{i'}{\kappa_{n',n}}\right)} \leq c. \tag{57}$$

By substituting (57) into (56), we obtain estimate (52). □

We are now ready to present the main result of this section.

Theorem 4 *For a given Δ, let $\mathbf{h}_\Delta \in \mathbf{W}^{\mathbf{k},\infty}$ be the solution of Eq. 4. If $u \in C^{k'}[0, T]$, then there exist a positive constants c and a positive integer $\ell \in \mathbb{N}$ such that for all $L \in \mathbb{N}^\nu$ satisfying $l' > \ell$ and for all $n \in \mathbb{Z}_\nu$,*

$$\|\mathbf{h}_n - \mathbf{h}_{n,l_n}\|_\infty \leq c\left(\mu_n^{-k_n l_n/n}\|\mathbf{h}_n\|_{k_n,\infty} + (l_n + 1)\sum_{m \in \mathbb{Z}_\nu} \mu_m^{-k_m l_m/m}\|\mathbf{h}_m\|_{k_m,\infty}\right). \tag{58}$$

Proof Estimate (58) follows directly from Theorem 3 and Lemma 10. □

7 A numerical quadrature strategy

The numerical solution of the compression method (27) requires generating the coefficient matrix $\mathbf{V}_L$. Hence, we have to compute the entries in the blocks $(\mathbf{V}_{n',n})_{i',i}$ for $i' \in \mathbb{Z}^0_{l_{n'}+1}$, $i \in \mathbb{Z}^0_{l_n+1}$with $2i + i'/\kappa_{n',n} \leq l_n$, for $n', n \in \mathbb{Z}_\nu$. These entries have the integral form

$$I_{n',i',j',n,i,j} := \int_{\Omega_n} \left\langle \ell_{n',i',j'}, K_{n',n}(\cdot, \xi)\right\rangle w_{n,i,j}(\xi)d\xi, \quad \text{for } i', j' \in \mathbb{U}_{n',l_{n'}}, i \in \mathbb{S}_{i'}, j \in \mathbb{Z}^0_{w_n(i)}, \tag{59}$$

where $K_{n',n}(\eta, \xi)$ is defined by (16). In this section, we propose a numerical quadrature strategy for computing (59). Ideally, such a quadrature strategy should preserve the convergence order for the compression method and use only $\mathcal{O}(N_L)$ number of functional evaluations in computing all nonzero entries of the matrix $\mathbf{V}_L$. The quadrature strategy proposed in this section is influenced by the work in [21, 25, 40, 47].

To describe the quadrature formula for $I_{n',i',j',n,i,j}$, we set $f_{i',j'}(\xi) := \left\langle \ell_{n',i',j'}, K_{n',n}(\cdot, \xi)\right\rangle$, for $\xi \in \Omega_n$, and $g_{i',j',i,j}(\xi) := f_{i',j'}(\xi)w_{n,i,j}(\xi)$, for $\xi \in \Omega_n$. Thus, we have that

$$I_{n',i',j',n,i,j} = \int_{\Omega_n} g_{i',j',i,j}(\xi)d\xi.$$

Fix n' and n. For $i' \in \mathbb{Z}^0_{l_{n'}+1}$, we choose a partition $D_{i'} := \left\{D_{i',\iota} : D_{i',\iota} \subset \Omega_n, \iota \in \mathbb{Z}^0_{\tau_{i'}}\right\}$ of Ω_n satisfying that $\Omega_n = \bigcup_{\iota \in \mathbb{Z}^0_{\tau_{i'}}} D_{i',\iota}$ and for $\iota_1, \iota_2 \in \mathbb{Z}^0_{\tau_{i'}}$, $\text{int}(D_{i',\iota_1}) \cap \text{int}(D_{i',\iota_2}) = \varnothing$, $\iota_1 \neq \iota_2$, where $\tau_{i'}$ is a positive integer and $\text{int}(A)$ denotes the interior of A. Appropriate choices of the partition $D_{i'}$ will be given later to ensure that the quadrature strategy has the desired convergence and computational complexity properties.

Let q_n be a positive integer. We choose a set of points $X_{i',\iota} := \left\{\xi^l_{i',\iota} : l \in \mathbb{Z}^0_{q_n}\right\} \subset D_{i',\iota}$ and evaluate $F_{j',\iota,l} := f_{i',j'}\left(\xi^l_{i',\iota}\right)$ for all $j' \in \mathbb{Z}^0_{w_{n'}(i')}$, $l \in \mathbb{Z}^0_{q_n}$ and $\iota \in \mathbb{Z}^0_{\tau_{i'}}$. We also choose the weights $A_{i',\iota} := \left\{a^l_{i',\iota} : l \in \mathbb{Z}^0_{q_n}\right\}$ and introduce two index sets $\mathbb{V}_{i'} := \left\{(i, j) : i \in \mathbb{S}_{i'}, j \in \mathbb{Z}^0_{w_n(i)}\right\}$ and $\mathbb{S}_{i',i,j} := \left\{\iota \in \mathbb{Z}^0_{\tau_{i'}} : D_{i',\iota} \subset \Omega_{n,i,j}\right\}$ for $(i', j') \in \mathbb{U}_{n',l_{n'}}$, $(i, j) \in \mathbb{V}_{i'}$. For $j' \in \mathbb{Z}^0_{w_{n'}(i')}$ and $(i, j) \in \mathbb{V}_{i'}$, we compute the quadrature

$$\tilde{I}_{n',i',j',n,i,j} := \sum_{\iota \in \mathbb{S}_{i',i,j}} \sum_{l \in \mathbb{Z}^0_{qn}} a^l_{i',\iota} F_{j',\iota,l} w_{n,i,j}\left(\xi^l_{i',\iota}\right). \tag{60}$$

Note that the values $F_{j',\iota,l}$ are used repeatedly when computing $\tilde{I}_{n',i',j',n,i,j}$ for $(i, j) \in \mathbb{V}_{i'}$. Repeating using these values reduces significantly the number of functional evaluations.

We let $\left(\mathbf{M}_{n',n}\right)_{i',i}$ denote the block $\left(\mathbf{V}_{n',n}\right)_{i',i}$ with $I_{n',i',j',n,i,j}$ being replaced by $\tilde{I}_{n',i',j',n,i,j}$ for $j' \in \mathbb{Z}^0_{w_{n'}(i')}$, $j \in \mathbb{Z}^0_{w_n(i)}$. Setting $\left(\mathbf{M}_{n',n}\right) := \left[\left(\mathbf{M}_{n',n}\right)_{i',i} : i' \in \mathbb{Z}^0_{l'_n+1},\ i \in \mathbb{Z}^0_{l_n+1}\right]$ and $\mathbf{M}_L := \left[\mathbf{M}_{n',n} : n', n \in \mathbb{Z}_\nu\right]$, we have the completely discrete equation

$$(\mathbf{E}_L + \mathbf{M}_L)\mathbf{h}_L = \mathbf{v}_L, \tag{61}$$

which is an approximate equation for (27).

Next, we turn our attention to the choices of partitions $D_{i'}$. We require that they satisfy the conditions:

(C-1) For $i' \in \mathbb{Z}^0_{l_{n'}+1}$, $(i, j) \in \mathbb{V}_{i'}$ and $\iota \in \mathbb{Z}^0_{\tau_{i'}}$, $\Omega_{n,i,j} \cap \mathrm{int}(D_{i',\iota}) = \varnothing$ or $\Omega_{n,i,j} \cap \mathrm{int}(D_{i',\iota}) = \mathrm{int}(D_{i',\iota})$.

(C-2) For $i' \in \mathbb{Z}^0_{l_{n'}+1}$, $i \in \mathbb{S}_{i'}$, $j \in \mathbb{Z}^0_{w_n(i)}$ and $\iota \in \mathbb{Z}^0_{\tau_{i'}}$, $w_{n,i,j}|_{\mathrm{int}(D_{i',\iota})} \in C^p(\mathrm{int}(D_{i',\iota}))$.

Condition (C-1) ensures that for each $\iota \in \mathbb{Z}^0_{\tau_{i'}}$, $D_{i',\iota}$ is contained in the support of the basis function $w_{n,i,j}$. Condition (C-2) implies that on $D_{i',\iota}$, $w_{n,i,j}$ is a smooth function.

We choose $p := \left\lceil \frac{3k_{n'}n}{2n'} \right\rceil$ and require that the quadrature formula chosen in (60) satisfies the condition that for all $h \in W^{p,\infty}$,

$$\int_{D_{i',\iota}} h(\xi)d\xi = \sum_{l \in \mathbb{Z}^0_{qn}} a^l_{i',\iota} h\left(\xi^l_{i',\iota}\right) + E(D_{i',\iota}, A_{i',\iota}, X_{i',\iota}) \tag{62}$$

with

$$|E(D_{i',\iota}, A_{i',\iota}, X_{i',\iota})| \le c \max_{|\alpha|=p} \left\{|(D^\alpha h)(\xi_{i',\iota,\alpha})|\right\} |d(D_{i',\iota})|^p \mathrm{meas}(D_{i',\iota}) \tag{63}$$

for $\xi_{i',\iota,\alpha} \in D_{i',\iota}$. We also let $d_{D_{i'}} := \max\left\{d(D_{i',\iota}) : \iota \in \mathbb{Z}^0_{\tau_{i'}}\right\}$ and $e_{i',j',i,j} := |I_{n',i',j',n,i,j} - \tilde{I}_{n',i',j',n,i,j}|$, for all $(i', j') \in \mathbb{U}_{n',l_{n'}}$, $(i, j) \in \mathbb{V}_{i'}$.

Lemma 11 *If $u \in C^{k^*}[0, T]$ with $k^* = p + k' + 1$, then there exists a positive constant c such that for $(i', j') \in \mathbb{U}_{n',l_{n'}}$, $(i, j) \in \mathbb{V}_{i'}$,*

$$e_{i',j',i,j} \leq c d_{D_{i'}}^{p} \mu_{n'}^{-i'k_{n'}/n'} \mu_n^{ik_n/n} \text{meas}(\Omega_{n,i,j}), \tag{64}$$

and

$$\|\mathbf{V}_{n',n} - \mathbf{M}_{n',n}\|_\infty \leq c \max_{(i',j') \in \mathbb{U}_{n',l_n'}} \left\{ d_{D_{i'}}^{p} \mu_{n'}^{-3i'k_{n'}/(2n')} \mu_n^{l_n k_n/(2n)} \right\}. \tag{65}$$

Proof Since $u \in C^{k^*}[0, T]$, we know that $f_{i',j'} \in W^{k^*,\infty}(\Omega_n)$. Hence, from condition (C-2), we have that $g_{i',j',i,j} \in W^{k^*,\infty}(D_{i',\iota})$. Note that $k^* > p$. By the definition of $e_{i',j',i,j}$ and (63), we observe that

$$e_{i',j',i,j} \leq c \sum_{\iota \in \mathbb{S}_{i',i,j}} \max_{|\alpha|=p} \left\{ |\left(D^\alpha g_{i',j',i,j}\right)\left(\xi_{i',\iota,\alpha}\right)| \right\} |d\left(D_{i',\iota}\right)|^p \text{meas}\left(D_{i',\iota}\right)$$

with $\xi_{i',\iota,\alpha} \in D_{i',\iota}$. From conditions (C-2), we get that

$$e_{i',j',i,j} \leq c d_{D_{i'}}^{p} \text{meas}(\Omega_{n,i,j}) \max_{|\alpha|=p} \left\{ |(D^\alpha g_{i',j',i,j})(\xi_{i',\iota,\alpha})| \right\}. \tag{66}$$

We next estimate $\max_{|\alpha|=p} \left\{ |(D^\alpha g_{i',j',i,j})(\xi_{i',\iota,\alpha})| \right\}$. We let $K_\alpha(\eta) := D^\alpha K_{n',n}(\eta, \xi_{i',\iota,\alpha})$ for $|\alpha| \leq p$. Since $u \in C^{k^*}[0, T]$, we know that $K_\alpha(\eta)$, $|\alpha| \leq p$, has continuous partial derivatives $D^\beta K_\alpha(\eta)$ for $\eta \in \Omega_{n'}$, when $|\beta| \leq k_{n'}$. Hence, as proceeded in [11] by using the Taylor expansion and by properties (I), (IV) and (V), we obtain that there exists a positive constant c such that

$$\left| D^\alpha f_{i',j'}(\xi) \right| \leq c \mu_{n'}^{-i'k_{n'}/n'}, \tag{67}$$

for $|\alpha| \leq p$. It follows from the proof for Lemma 5.2 of [14] that there is a positive constant c such that for all $|\alpha| \leq p$,

$$\left| D^\alpha w_{n,i,j}(\xi) \right| \leq c \mu_n^{ik_n/n}. \tag{68}$$

Noting that $g_{i',j',i,j}(\xi) = f_{i',j'}(\xi) w_{n,i,j}(\xi)$ and $g_{i',j',i,j} \in W^{k^*,\infty}(D_{i',\iota})$, it follows from (67) and (68) that there exists a positive constant c such that

$$\max_{|\alpha|=p} \left\{ |\left(D^\alpha g_{i',j',i,j}\right)\left(\xi_{i',\iota,\alpha}\right)| \right\} \leq c \mu_{n'}^{-i'k_{n'}/n'} \mu_n^{ik_n/n}. \tag{69}$$

Substituting (69) into (66), we get the estimation (64).

By (64) and the relation $\sum_{j \in \mathbb{Z}_{w_n(i)}^0} \text{meas}(\Omega_{n,i,j}) = \text{meas}(\Omega_n)$, there exists a positive constant c such that

$$\|\mathbf{V}_{n',n} - \mathbf{M}_{n',n}\|_\infty \leq c \max_{(i',j') \in \mathbb{U}_{n',l_n'}} \left\{ d_{D_{i'}}^{p} \mu_{n'}^{-i'k_{n'}/n'} \sum_{i \in \mathbb{Z}_{l_n+1}^0 / \mathbb{S}_{i'}} \mu_n^{ik_n/n} \right\} \tag{70}$$

By the definition of $\mathbb{S}_{i'}$, we conclude that there exists a positive constant c such that

$$\sum_{i\in\mathbb{Z}^0_{l_n+1}/\mathbb{S}_{i'}} \mu_n^{ik_n/n} \le c\mu_n^{l_n k_n/(2n)} \mu_{n'}^{-i' k_{n'}/(2n')}. \tag{71}$$

Substituting (71) into (70) yields estimate (65). □

In order to ensure that the quadrature will not ruin the convergence result, we need the matrix obtained via the quadrature formula to satisfy

$$\left\|\mathbf{V}_{n',n} - \mathbf{M}_{n',n}\right\|_\infty \le c_1\mu_n^{-l_n k_n/n},$$

where c_1 is a positive constant. Hence, it follows from Lemma 11 that we demand that

$$d^p_{D_{i'}} \mu_{n'}^{-3i' k_{n'}/(2n')} \mu_n^{l_n k_n/(2n)} \le c_1\mu_n^{-l_n k_n/n},$$

or equivalently,

$$d_{D_{i'}} \le c_1\mu_{n'}^{-\varrho_{i'}}, \tag{72}$$

where $\varrho_{i'} := \frac{3k_{n'}}{2pd_{n'}}\left(\kappa_{n',n} l_n - i'\right)$. This imposes an additional hypothesis on the partition.

Lemma 12 *Let* $u \in C^{k^*}[0, T]$ *with* $k^* = p + k' + 1$. *If* $D_{i'}$, $i' \in \mathbb{Z}^0_{l_{n'}+1}$ *satisfies* (72)*, then there exists a positive constant c such that*

$$\|\mathbf{V}_{n',n} - \mathbf{M}_{n',n}\|_\infty \le c\mu_n^{-l_n k_n/n}. \tag{73}$$

Proof By substituting (72) into (65), we obtain estimate (73). □

In order to reduce the complexity of quadrature, we also require that $d_{D_{i'}} \ge c_2\mu_{n'}^{-\varrho_{i'}}$, where c_2 is a positive constant. This, combined with (72), leads to the following condition on the partition.

(C-3) For $i' \in \mathbb{Z}^0_{l_{n'}}$ and $\iota \in \mathbb{Z}^0_{\tau_{i'}}$, there exists two positive constants c_1 and c_2 such that

$$c_2\mu_{n'}^{-\varrho_{i'}} \le d(D_{i',\iota}) \le c_1\mu_{n'}^{-\varrho_{i'}}. \tag{74}$$

On the other hand, since $\cup_{\iota\in\mathbb{Z}^0_{\tau_{i'}}} D_{i',\iota} = \Omega_n$, from the left inequality of (74), we get that for all $i' \in \mathbb{Z}^0_{l_{n'}+1}$,

$$\tau_{i'} \le c\mu_{n'}^{\varrho_{i'}} \tag{75}$$

where c is a positive constant.

Table 1 Regularization parameters related to the uncompressed matrix $\mathbf{U}_L$ and the compressed matrix $\mathbf{V}_L$

	Parameters related to $\mathbf{U}_L$	Parameters related to $\mathbf{V}_L$
λ_1	$7.50E-9$	$1.00E-8$
λ_2	$1.50E-8$	$3.00E-8$
λ_3	$2.00E-6$	$5.00E-6$

We next show that the quadrature strategy does not ruin the convergence rate of the compression method (27). Let $\bar{\mathbf{h}}_L$ be the solution of Eq. 61 and let

$$\mathbf{h}_{n,l_n} := \sum_{i=0}^{l_n} \sum_{j=0}^{w_n(i)-1} \left(\bar{\mathbf{h}}_L\right)_{n,i,j} w_{n,i,j}, \quad n \in \mathbb{Z}_\nu \text{ and } \mathbf{h}_L := \left[\mathbf{h}_{n,l_n}, n \in \mathbb{Z}_\nu\right]^T.$$

Theorem 5 *For a given* Δ, *let* $\mathbf{h}_\Delta \in \mathbf{W}^{\mathbf{k},\infty}$ *be the solution of Eq. 4. If* $u \in C^{k'}[0, T]$ *with* $k^* = k' + p + 1$ *and* $\mathbf{h}_L \in \mathbf{W}^{\mathbf{k},\infty}$, *then there exist a positive constants* c *and an* $\ell \in \mathbb{N}$ *such that for all* $L \in \mathbb{N}^\nu$ *satisfying* $l_n > \ell$ *and for all* $n \in \mathbb{Z}_\nu$,

$$\|\mathbf{h}_n - \mathbf{h}_{n,l_n}\|_\infty \le c\left(\mu_n^{-k_n l_n/n}\|\mathbf{h}_n\|_{k_n,\infty} + (l_n+1)\sum_{m\in\mathbb{Z}_\nu}\mu_m^{-k_m l_m/m}\|\mathbf{h}_m\|_{k_m,\infty}\right). \tag{76}$$

Proof By Theorem 3, it suffices to prove that inequalities (32) and (33) hold with $\mathbf{K}_{n,m} := \mathbf{M}_{n,m}$.

Let $h \in W^{k_m,\infty}$ and $\bar{h}$ satisfy (34). By the triangle inequality, we get that

$$\left\|\left(\mathbf{U}_{n,m} - \mathbf{M}_{n,m}\right)\bar{h}\right\|_\infty \le \left\|\left(\mathbf{U}_{n,m} - \mathbf{V}_{n,m}\right)\bar{h}\right\|_\infty + \left\|\left(\mathbf{V}_{n,m} - \mathbf{M}_{n,m}\right)\bar{h}\right\|_\infty. \tag{77}$$

From (53), we have that $\left\|\bar{h}\right\|_\infty \le c\|h\|_{k_m,\infty}$. Combining this inequality with Lemma 12, we obtain that

$$\left\|\left(\mathbf{V}_{n,m} - \mathbf{M}_{n,m}\right)\bar{h}\right\|_\infty \le c\mu_m^{-k_m l_m/m}\|h\|_{k_m,\infty}. \tag{78}$$

By substituting (78) and (51) into (77), we get (32).

Table 2 Output errors of the Volterra systems identified by the uncompressed matrix $\mathbf{U}_L$ and the compressed matrix $\mathbf{V}_L$

	Output errors related to $\mathbf{U}_L$	Output errors related to $\mathbf{V}_L$
e_1	$1.44E-3$	$1.52E-3$
e_2	$4.57E-2$	$4.82E-2$
e_3	$3.65E-2$	$3.73E-2$
e_4	$4.87E-2$	$5.09E-2$

Table 3 Computing time for generating matrices $\mathbf{U}_L$ and $\mathbf{V}_L$, and solution time for solving Eqs. 20 and 27

	CT (s)	ST (s)
UCM	4374.47	180.34
CM	551.63	1.20

CT computing time, *ST* solution time, *UCM* uncompressed coefficient matrix, *CM* coefficient matrix

It remains to show (33). Let $h \in H^{\infty}$. From Property (VI), we know that $\|\bar{h}\|_{\infty} \leq c\|h\|_{\infty}$. Employing the same procedure as in the proof of (32), we obtain (33). □

We next estimate the number $\mathcal{M}_{n',n}$ of the point evaluations of $K_{n',n}$, $n', n \in \mathbb{Z}_{\nu}$ for generating the matrix $\mathbf{M}_L$. The total number of the point evaluations of $K_{n',n}$, $n', n \in \mathbb{Z}_{\nu}$ for computing the entries $\mathbf{M}_L$ is then given by

$$\mathcal{M}_L := \sum_{n' \in \mathbb{Z}_{\nu}} \sum_{n \in \mathbb{Z}_{\nu}} \mathcal{M}_{n',n}.$$

In the next theorem, we give an estimate for $\mathcal{M}_L$.

Theorem 6 *If* $\mathbf{M}_L$ *is generated by the compression strategy* (25) *and the quadrature strategy* (60)*, then there exists a positive constant c such that*

$$\mathcal{M}_L \leq c N_L. \tag{79}$$

Proof We first estimate $\mathcal{M}_{n',n}$. Let $\mathcal{M}_{n',n,i'}$ be the number of the point evaluations of $K_{n',n}$ for computing $[(\mathbf{M}_{n',n})_{i',i} : i \in \mathbb{S}_{i'}]$. Since for a fixed $i' \in \mathbb{Z}^0_{l_{n'}+1}$, the evaluation points of $K_{n',n}$ for computing $(\mathbf{M}_{n',n})_{i',i}$ are the same for all $i \in \mathbb{Z}^0_{l_{n'}+1}$, we only need to evaluate $K_{n',n}$ at these points once. Hence, from Property (II) and (75), we know that there exists a positive constant c such that $\mathcal{M}_{n',n,i'} \leq c\mu_{n'}^{i'+\varrho_{i'}}$. Let $r := \frac{3k_{n'}n}{2pn'}$. By the definition of $\varrho_{i'}$, we obtain that $\mathcal{M}_{n',n,i'} \leq c\mu_{n'}^{r\kappa_{n',n}l_n+(1-\tilde{c})i'}$. Note that $p \geq \frac{3k_{n'}n}{2n'}$ and $r < 1$. By summing all $\mathcal{M}_{n',n,i'}$, $i' \in \mathbb{Z}^0_{l_{n'}+1}$, we get that

$$\mathcal{M}_{n',n} \leq c\mu_{n'}^{r(\kappa_{n',n}l_n - l_{n'})+l_{n'}}. \tag{80}$$

It follows from (28) and (80) that there is a positive constant c such that $\mathcal{M}_{n',n} \leq c\mu_{n'}^{l_{n'}}$. Upon summing all $\mathcal{M}_{n',n}$ for $n', n \in \mathbb{Z}_{\nu}$, we get the estimation (79). □

Table 4 Parameters used to identify the Volterra systems of order two and three

Parameters
Parameters of order two
$\lambda_1 = 1.00E{-}8$,
$\lambda_2 = 4.50E{-}6$,
Parameters of order three
$\lambda_1 = 1.00E{-}8$,
$\lambda_2 = 3.00E{-}8$,
$\lambda_3 = 5.00E{-}6$,

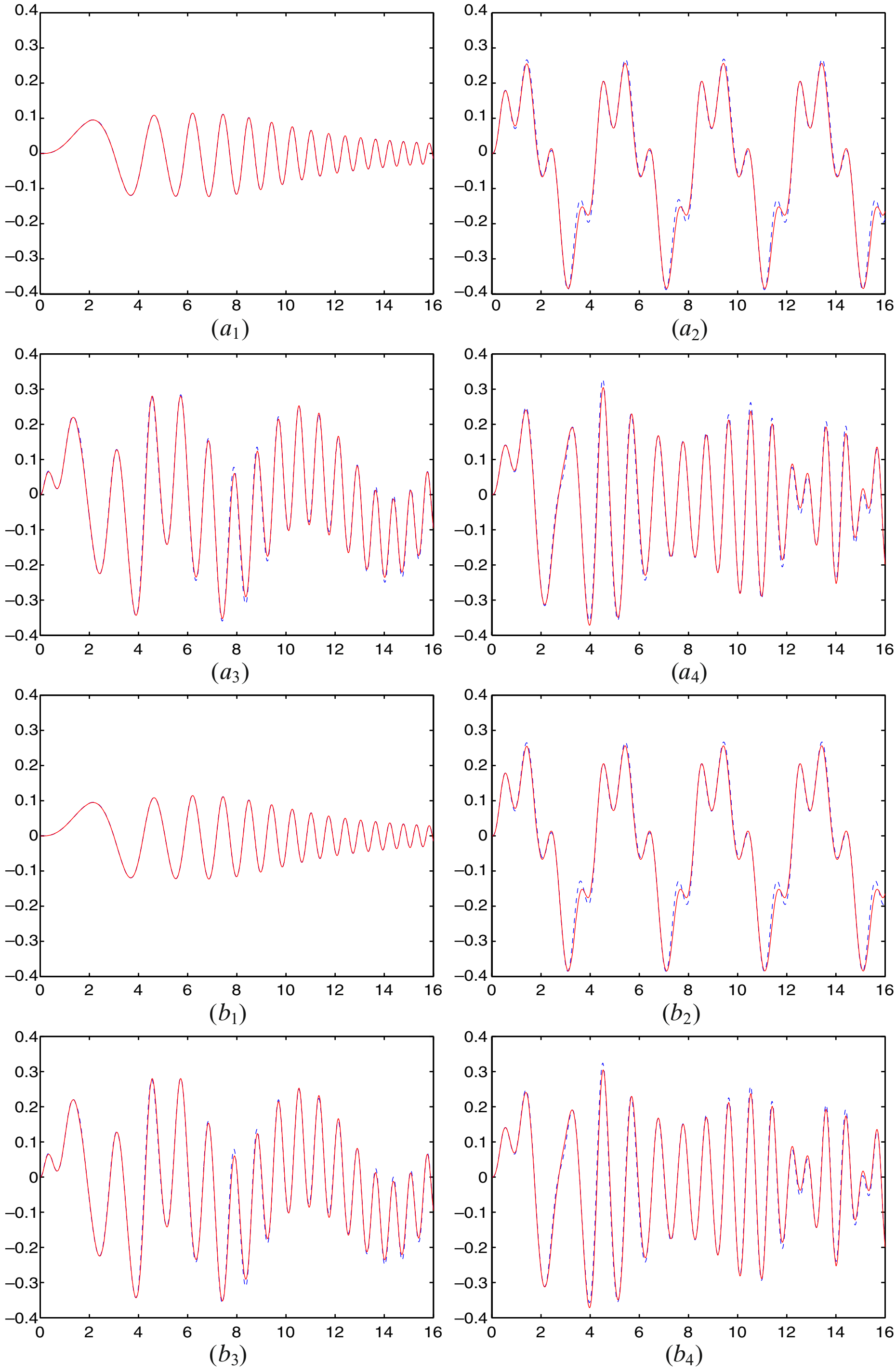

Fig. 1 Outputs of the simulation system related to the uncompressed matrix and the compressed matrix

8 Numerical experiments

We report in this section results of three numerical experiments for the multiparameter regularization via the multiscale collocation method. Experiment 1 is to confirm approximation of the compressed coefficient matrix, Experiment 2 is to compare the third order Volterra system with the second order Volterra system, and Experiment 3 is to test the effect of number of parameters used in regularization.

We suppose that our true system obeys the initial value problem of the nonlinear oscillatory system

$$\begin{cases} v''(t) + 6v'(t) + 4\pi^2 v(t) + 4\pi^2 v^2(t) = u(t),\ 0 < t < T, \\ v(0) = v'(0) = 0, \end{cases} \tag{81}$$

where u is the input and the solution v of (81) is the output. We will use the multiparameter regularization method to construct its Volterra simulation system of order three

$$(\mathcal{V}u)(t) = v(t), \quad t \in [0, T].$$

In our experiments, we first solve Eq. 81 with the right-hand-side being the input $u_0(t) = \sin(0.125\pi t^2)$ by the adaptive Runge-Kutta-Fehlberg method (a Matlab subroutine) and obtain the output v_0. We then use the input u_0 and output v_0 to identify the Volterra system. In this identification process, the wavelet bases on the unit interval (1-D), the unit triangle (2-D) and the unit simplex (3-D) and their corresponding collocation functionals constructed in [14] are used where $k_1 = k_2 = k_3 = 2$. We choose $l_1 = 8$, $l_2 = 5$ and $l_3 = 2$.

To test the resulting Volterra simulation system, we compare the outputs of the true system (81) and the Volterra simulation system. We use

$$u_1(t) = 4\sin\left(0.125\pi t^2\right), \quad u_2(t) = 8\sin\left(\frac{\pi t}{2}\right) + 4\sin(2\pi t),$$

$$u_3(t) = 8\sin\left(\frac{\pi t^{\frac{3}{2}}}{2}\right) + 4\sin\left(2\pi t^{\frac{1}{2}}\right), \quad u_4(t) = 8\sin\left(\frac{\pi t^{\frac{3}{2}}}{2}\right) + 4\sin(2\pi t)$$

as our testing input signals. We denote by v_i and $\tilde{v}_i$ the output of the true system and Volterra simulation system, respectively, corresponding to input u_i, for $i = 1, 2, 3, 4$. We use

$$e_i := \frac{\|v_i - \tilde{v}_i\|_2}{\|v_i\|_2}, \quad i = 1, 2, 3, 4$$

Table 5 Output errors of the Volterra systems of order two and three

	Output errors of order two	Output errors of order three
e_1	$4.86E-3$	$1.52E-3$
e_2	$1.12E-1$	$4.82E-2$
e_3	$7.69E-2$	$3.73E-2$
e_4	$6.97E-2$	$5.09E-2$

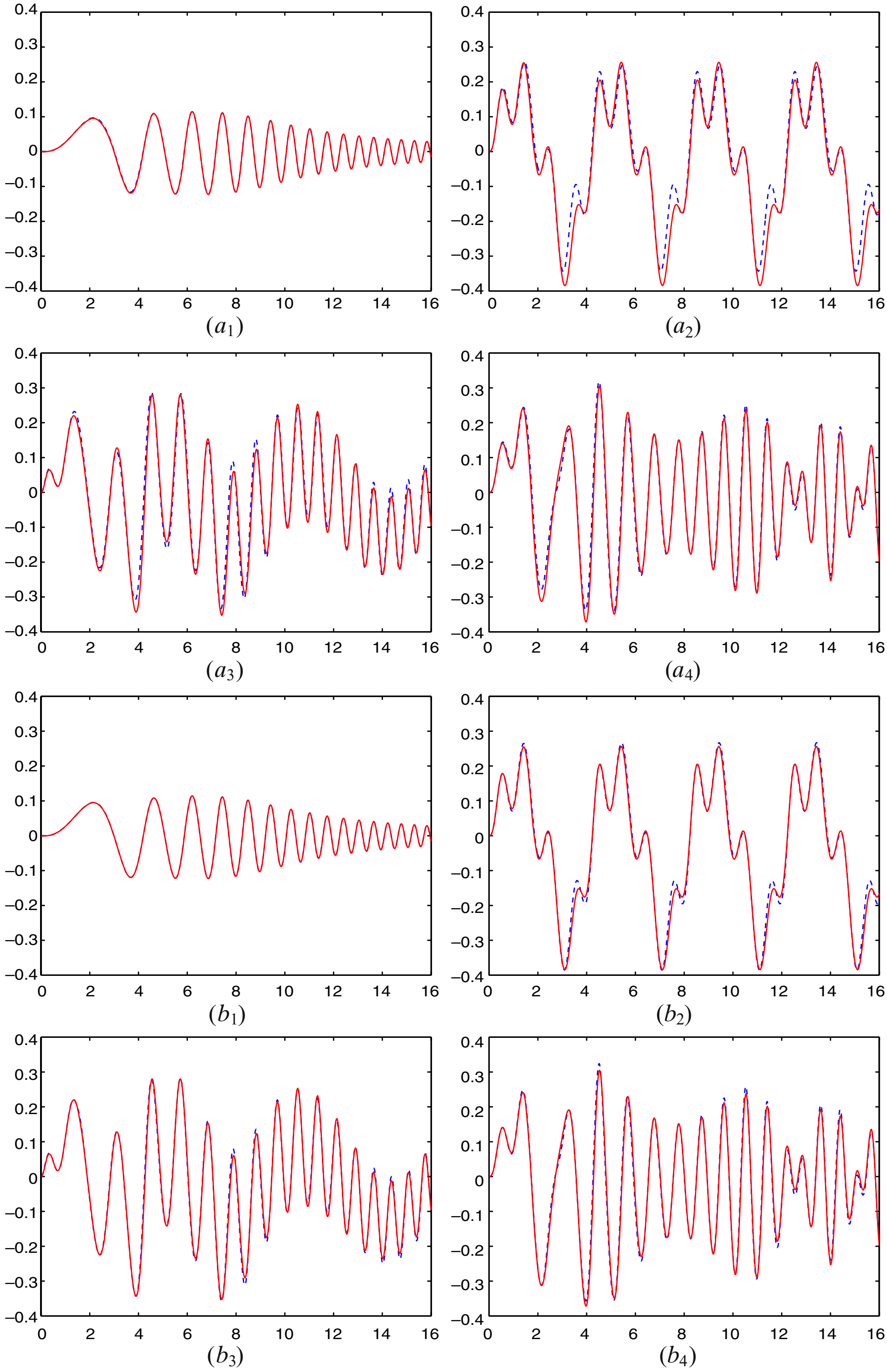

Fig. 2 Outputs $\tilde{v}_i$, $i = 1, 2, 3, 4$ of the Volterra systems of order two and three

to present the error of the output of the true system and that of the simulation system. For each case in all experiments, the set of parameters for the case is chosen to minimize the errors e_i, $i = 1, 2, 3, 4$, corresponding to the case. One may also use the L-curve principle developed in [3] to choose the parameters. The numerical computation is performed in a personal computer with 1.6 GHz CPU, 512M memory and operating system Windows XP.

8.1 Approximation by a compressed matrix

The numerical results presented in this subsection confirm that the compressed matrix gives a satisfactory approximation property. The coefficient matrix $\mathbf{U}_L$ defined by (18) is approximated by a compressed matrix $\mathbf{V}_L$ defined by (26). To illustrate the approximation property of the compressed matrix, we compare the errors of outputs of the Volterra system identified by $\mathbf{U}_L$ and the Volterra system identified by $\mathbf{V}_L$ relative to the true system (81). Since matrix $\mathbf{U}_L$ is full and matrix $\mathbf{V}_L$ is spare, we use Gaussian elimination algorithm to solve Eq. 20 and use the multilevel augmentation method in [12, 15] to solve Eq. 27. To show the efficiency of compression strategy, we also compare the computing time for generating the matrices $\mathbf{U}_L$, $\mathbf{V}_L$ and the time for solving Eqs. 20 and 27.

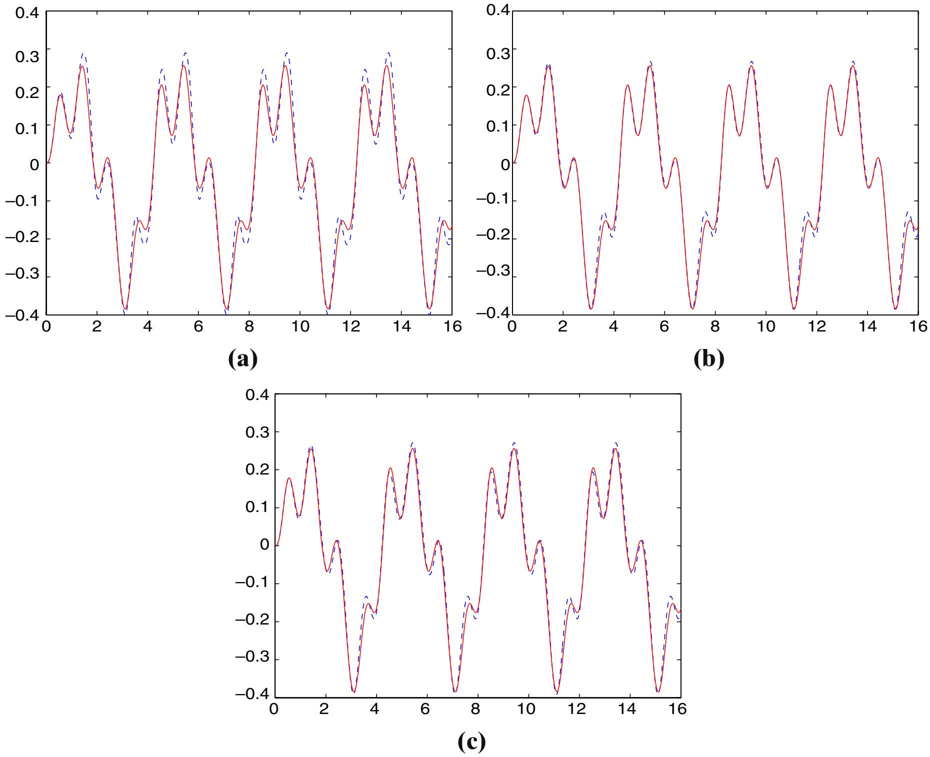

Fig. 3 Outputs $\tilde{v}_2$ of the Volterra systems of order three identified by multiparameter regularization methods using output signal v_0 without noise (**a**–**c**)

In this experiment, we use multiparameter regularization with on parameter for each order term. For each of Eqs. 20 and 27, we choose a set of three regularization parameters λ_i, $i = 1, 2, 3$ and let $\lambda_{i,j} := \lambda_i$, $i = 1, 2, 3$ and $j \in \mathbb{N}_0$, where $\lambda_{i,j}$ is a regularization parameter of (3). We list in Table 1 the parameters and in Table 2 the errors e_i, $i = 1, 2, 3, 4$.

Figure 1 illustrates the outputs of the Volterra systems obtained by using the uncompressed matrix and the compressed matrix, corresponding to inputs u_i, $i = 1, 2, 3, 4$. In this figure, the red solid lines present outputs v_i of the true system, the blue dash lines in row (a_i) present outputs $\tilde{v}_i$ of the Volterra system identified by the uncompressed matrix $\mathbf{U}_L$ and the blue dash lines in row (b_i), present outputs $\tilde{v}_i$ of the Volterra system identified by the compressed matrix $\mathbf{V}_L$. In Table 3, we show the computing time for generating the coefficient matrices and the solution time for solving the linear systems, with both uncompressed coefficient matrix and compressed coefficient matrix, measured in seconds. Figure 1 and Table 2 show that the outputs of the simulation systems related to the uncompressed matrix and to the compressed matrix are similar. Table 3 confirms the compression strategy reduces computing time significantly for both generating the coefficient matrix and solving the corresponding linear system.

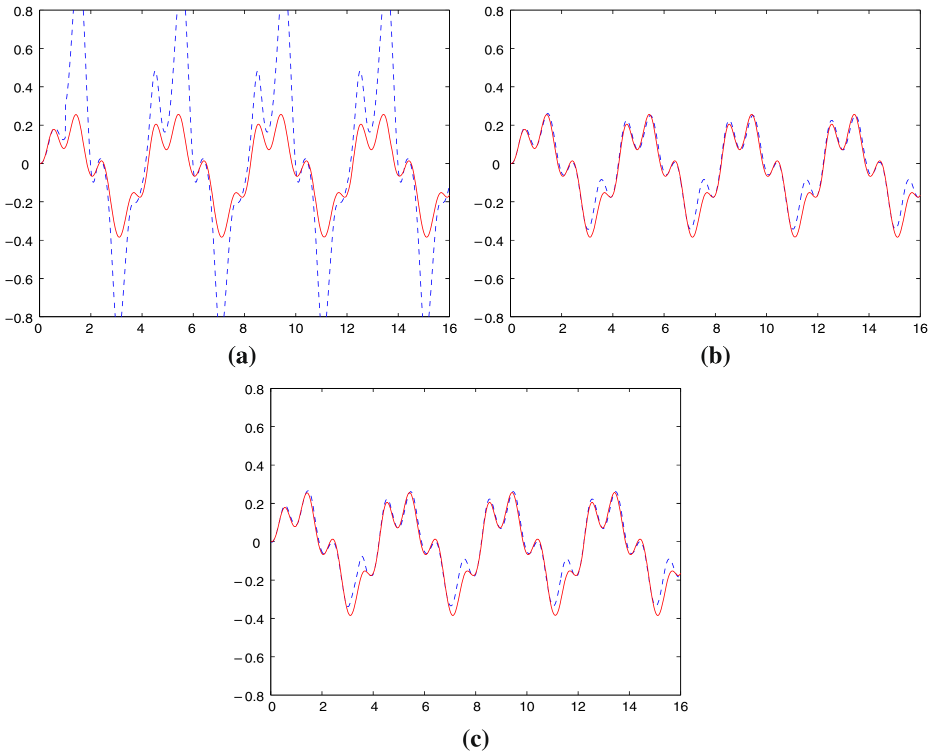

Fig. 4 Outputs $\tilde{v}_2$ of the Volterra systems identified using output signal v_0 with noise $\delta = 10\%$ (**a**–**c**)

8.2 Approximation by Volterra systems of higher orders

In this experiment, we demonstrate that when the nonlinear system (81) is approximated by Volterra systems, the approximate result given by those of order three is better than that given by those of order two. Specifically, we compare the errors e_i corresponding to the Volterra system of order three with the error e_i corresponding to the Volterra system of order two.

Parameters λ_i used in this experiment for both systems are listed in Table 4. Again we choose the regularization parameters in (3) by $\lambda_{i,j} := \lambda_i$, for $i = 1, 2, 3,\ j \in \mathbb{N}_0$. In Table 5, we list the errors e_i, $i = 1, 2, 3, 4$, for both systems. The outputs $\tilde{v}_i$, $i = 1, 2, 3, 4$ of the Volterra system of order two and that of order three are displayed in Fig. 2, where the red solid lines represent outputs v_i, of the true system, the blue dash lines in row (a_i) are outputs $\tilde{v}_i$ of the Volterra system of order two,a nd the blue dash lines in row (b_i) are outputs $\tilde{v}_i$ of the Volterra system of order three, for $i = 1, 2, 3, 4$. This experiment shows that the Volterra system of order three catches the nonlinearity of the true system (81) better than that of order two does.

8.3 Multiple parameters corresponding to different kernels and different scales

In this subsection, we show improvement of using the multiple parameters corresponding to different kernels and different scales in the regularization method for

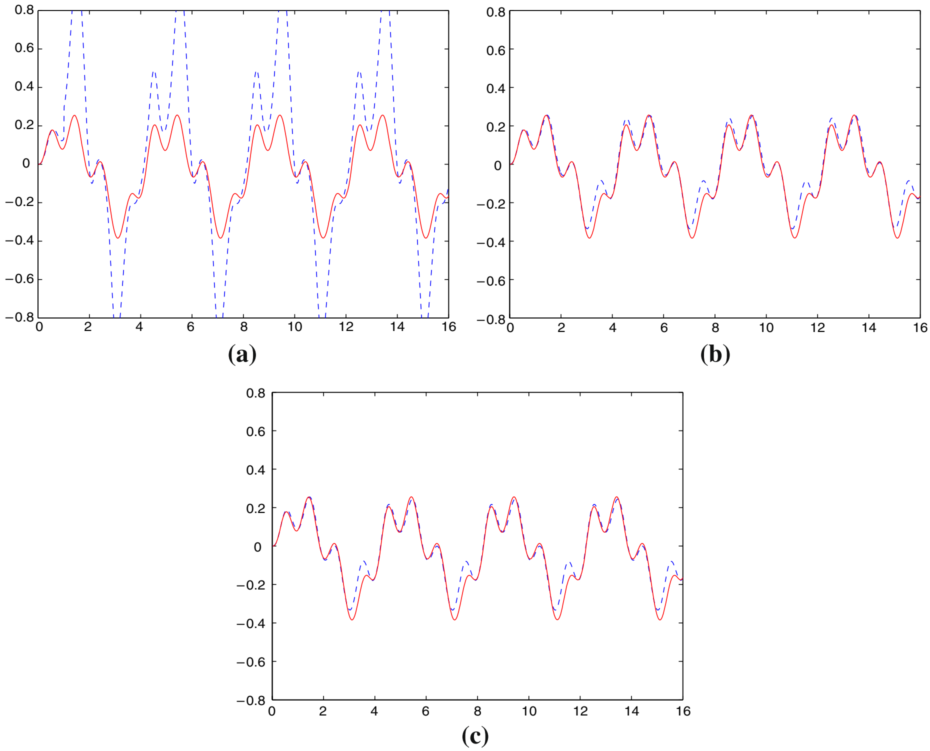

Fig. 5 Outputs $\tilde{v}_2$ of the Volterra systems identified using output signal v_0 with noise $\delta = 20\%$ (**a**–**c**)

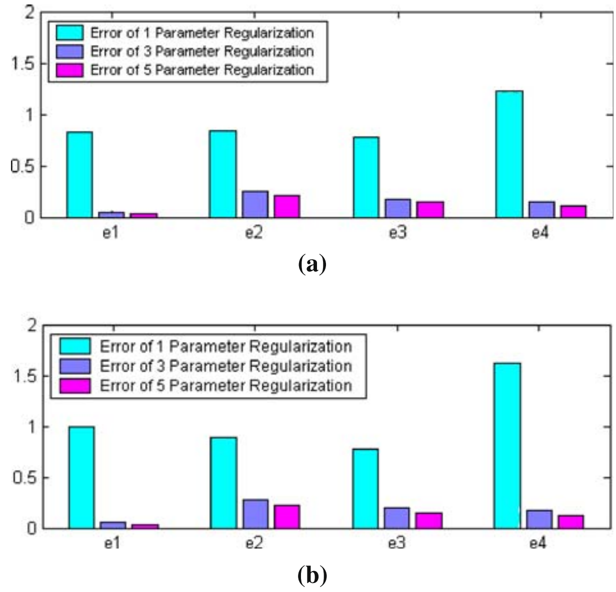

Fig. 6 An illustration for errors of the identification with multiparameters using output signal v_0 with noise. **a** Error of identification with multiparameters for signals with noise $\delta = 10\%$. **b** Error of identification with multiparameters for signals with noise $\delta = 20\%$

identifying the Volterra system. Specifically, we identify the Volterra systems of order three by the regularization method using one parameter, three parameters (one for every kernel) and five parameters (two for the kernel of order one, two for the kernel of order two, and one for the kernel of order three), respectively. In particular, for the regularization method (3) with five parameters, we let $\lambda_{1,i} = \lambda_1$, for $i = 0, 1, \ldots, 5$, $\lambda_{1,i} = \lambda_2$, for $i = 6, 7, 8$, $\lambda_{2,i} = \lambda_3$, for $i = 0, 1, 2, 3, 4$, $\lambda_{2,5} = \lambda_4$, for $i = 5$ and $\lambda_{3,i} = \lambda_5$, for $i = 0, 1, 2$. In this subsection, we use the outputs $\tilde{v}_2$ of the Volterra systems of order three to show the effects of the multiparameter regularization method. We consider two cases: signals without noise and noisy signals. In Figs. 3, 4 and 5, the red solid lines represent the outputs v_2 of the true system (81), the blue dash lines in image (*a*) represent the outputs $\tilde{v}_2$ of the Volterra system identified with one parameter, the blue dash lines in image (*b*) represent the outputs $\tilde{v}_2$ of the

Table 6 Parameters for the Volterra system identification using the output signal v_0 without noise

Parameters
1 parameter
$\lambda = 2.00E - 8$
3 parameters
$\lambda_1 = 1.00E{-}8$
$\lambda_2 = 3.00E{-}8$
$\lambda_3 = 5.00E{-}6$
5 parameters
$\lambda_1 = 1.00E{-}8$
$\lambda_2 = 1.50E{-}7$
$\lambda_3 = 7.50E{-}9$
$\lambda_4 = 5.00E{-}3$
$\lambda_5 = 5.00E{-}6$

Table 7 Output errors of the Volterra system identification using output signal v_0 without noise

	1 parameter	3 parameters	5 parameters
e_1	$4.41E-2$	$1.52E-3$	$1.22E-3$
e_2	$2.67E-1$	$4.82E-2$	$4.91E-2$
e_3	$5.42E-1$	$3.73E-2$	$3.62E-2$
e_4	$9.04E-1$	$5.07E-2$	$4.38E-2$

Volterra system identified with three parameters, and the blue dash lines in image (c) represent the outputs $\tilde{v}_2$, of the Volterra system identified with five parameters.

8.3.1 Signals without noise

In this experiment, we test the performance of the multiparameter regularization method solving the Volterra kernel identification problem when the signal v_0 contains no noise. We list in Table 6 the parameters used in this experiment and in Table 7 the errors $e_i, i = 1, 2, 3, 4$ corresponding to each case. This example confirms that using multiple parameters for different kernels and different scales in Volterra system identification improves significantly the approximation results of the Volterra simulation system.

8.3.2 Signals with noise

In this experiment, we conduct the same experiment described in the last subsection for signals with noise. We test the identification method for the output signal v_0 perturbed by the Gauss noise with mean 0 and two different variances which reflecting two different noise levels $\delta = 10\%$ and $\delta = 20\%$, where $\delta := \frac{\|v_0^\delta - v_0\|_2}{\|v_0\|_2}$. We let v_0^δ denote the noisy output signal. We list the parameters used in this experiment in Table 8. In Table 9 we list the errors $e_i, i = 1, 2, 3, 4$ of the Volterra systems identified by the multiparameter regularization methods with different noise levels. We also illustrate the errors in Fig. 6.

Table 8 Parameters for the Volterra system identification using output signal v_0 with noise $\delta = 10\%$ and $\delta = 20\%$

	1 parameter	3 parameters	5 parameters
$\delta = 10\%$	$\lambda = 3.00E{-}1$	$\lambda_1 = 6.00E{-}3$, $\lambda_2 = 3.00E{-}2$, $\lambda_3 = 3.00E{-}1$	$\lambda_1 = 1.00E{-}6$, $\lambda_2 = 1.00E{-}3$, $\lambda_3 = 1.00E{-}2$, $\lambda_4 = 6.00E{-}1$, $\lambda_5 = 4.00E{-}1$
$\delta = 10\%$	$\lambda = 4.90E{-}1$	$\lambda_1 = 1.00E{-}2$, $\lambda_2 = 1.00E{-}1$, $\lambda_3 = 4.00E{-}1$	$\lambda_1 = 1.00E{-}3$, $\lambda_2 = 1.60E{-}2$, $\lambda_3 = 2.30E{-}1$, $\lambda_4 = 3.40E{-}1$, $\lambda_5 = 5.30E{-}1$

Table 9 Errors of the identification using output signal v_0 with noise

	$\delta = 10\%$			$\delta = 20\%$		
	1 parameter	3 parameters	5 parameters	1 parameter	3 parameters	5 parameters
e_1	$8.21E-1$	$4.80E-2$	$2.89E-2$	$9.92E-1$	$5.93E-2$	$3.59E-2$
e_2	$8.34E-1$	$2.48E-1$	$2.13E-1$	$8.91E-1$	$2.78E-1$	$2.27E-1$
e_3	$7.82E-1$	$1.75E-1$	$1.46E-1$	$7.82E-1$	$2.04E-1$	$1.46E-1$
e_4	$1.22E-0$	$1.45E-1$	$1.14E-1$	$1.62E-0$	$1.76E-1$	$1.20E-1$

The experiments presented in this subsection reveal that when the output signal contains noise of a higher level, the multiparameter regularization method for the Volterra system identification is more efficient.

References

1. Aiordachioaiea, D., Ceangaa, E., Keyserb, R., Naka, Y.: Detection and classification of non-linearities based on Volterra kernels processing. Eng. Appl. Artif. Intell. **14**, 497–503 (2001)
2. Atkinson, K., Graham, I., Sloan, I.: Piecewise continuous collocation for integral equations. SIAM J. Numer. Anal. **20**, 172–186 (1983)
3. Belge, M., Kilmer, M.E., Miller, E.L.: Efficient determination of multiple regularization parameters in a generalized L-curve framework. Inverse Problems **18**, 1161–1183 (2002)
4. Brenner, M., Xu, Y.: A factorization method for identification of Volterra systems. J. Comput. Appl. Math. **144**, 105–117 (2002)
5. Brezinski, C., Redivo-Zaglia, M., Rodriguez, G., Seatzu, S.: Multi-parameter regularization techniques for ill-conditioned linear system. Numer. Math. **94**, 203–228 (2003)
6. Brunner, H.: Collocation Methods for Volterra Integral and Related Functional Differential Equations. Cambridge University Press, Cambridge (2004)
7. Ciarlet, P.: Introduction to Linear Algebra and Optimization. Cambridge University Press, Cambridge (1989)
8. Chen, L., Narendra, K.: Identification and control of a nonlinear discrete-time system based on its linearization: a unified framework. IEEE Trans. Neural Netw. **15**, 663–673 (2004)
9. Chen, Z., Lu, Y., Xu, Y., Yang, H.: Multi-parameter Tikhonov regularization for ill-posed operator equations. J. Comput. Math. **26**, 37–55 (2008)
10. Chen, Z., Micchelli, C., Xu, Y.: A construction of interpolating wavelets on invariant sets. Comput. Math. **68**, 1560–1587 (1999)
11. Chen, Z., Micchelli, C., Xu, Y.: Fast collocation methods for second kind integral equations. SIAM J. Numer. Anal. **40**, 344–375 (2002)
12. Chen, Z., Wu, B., Xu, Y.: Multilevel augmentation methods for solving operator equations. Numer. Math. J. Chinese Univ. **14**, 31–55 (2005)
13. Chen, Z., Wu, B., Xu, Y.: Fast numerical collocation solutions of integral equations. Comm. Pure Appl. Anal. **6**, 643–666 (2007)
14. Chen, Z., Wu, B., Xu, Y.: Fast collocation methods for high dimensional weakly singular integral equations. J. Integral Equations Appl. **20**, 49–92 (2008)
15. Chen, Z., Xu, Y., Yang, H.: A multilevel augmentation method for solving ill-posed operator equations. Inverse Problems **22**, 155–174 (2006)
16. Chon, K., Chen, Y., Holstein-Rathlou, N., Marmarelis, V.: Nonlinear system analysis of renal autoregulation in normotensive and hypertensive rats. IEEE Trans. Biomed. Eng. **45**, 342–353 (1998)
17. Doicu, A., Schreier, F., Hilgers, S., Hess, M.: Multi-parameter regularization method for atmospheric remote sensing. Comput. Phys. Comm. **165**, 1–9 (2005)
18. Doyle, F., Ogunnaike, B., Pearson, R.: Nonlinear model control using second-order volterra models. Automatica **31**, 697–714 (1995)
19. Engl, H.W., Hanke, M., Neubauer, A.: Regularization of Inverse Problems. Kluwer, Dordrecht (1996)

20. French, A., Sekizawa, S., Hoger, U., Torkkeli, P.: Predicting the responses of mechanoreceptor neurons to physiological inputs by nonlinear system identification. Ann. Biomed. Eng. **29**, 187–194 (2001)
21. Ford, W.F., Xu, Y., Zhao, Y.: Derivative correction for quadrature formulas. Adv. Comput. Math. **6**, 139–157 (1996)
22. Groetsch, C.W.: The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind. Research Notes in Mathematics, 105. Pitman, Boston (1984)
23. Glentis, G., Koukoulas, P., Kalouptsidis, N.: Efficient algorithms for volterra system identification. IEEE Trans. Signal Process. **47**, 3042–3057 (1999)
24. Khouaja, A., Kibangou, A., Favier, G.: Third-order Volterra kernels complexity reduction using PARAFAC. In: Proc. 1st IEEE International Symposium on Control, Communications and Signal Processing (ISCCSP 04), pp. 857–860. IEEE, Piscataway (2004)
25. Kaneko, H., Xu, Y.: Gauss-type quadratures for weakly singular integrals and their application to Fredholm integral equations of the second kind. Math. Comp. **62**, 739–753 (1994)
26. Lu, Y., Shen, L., Xu, Y.: Shadow block iteration for solving linear systems obtained from wavelet transforms. Appl. Comput. Harmon. Anal. **19**, 359–385 (2005)
27. Lu, Y., Shen, L., Xu, Y.: Multi-parameter regularization methods for high-resolution image reconstruction with displacement errors. IEEE Trans. Circuits Systems I **54**, 1788–1799 (2007)
28. Maner, B., Doyle, F., Ogunnaike, B., Pearson, R.: Nonlinear model predictvie control of a simulated multivariable polymerization reactor using second-order volterra models. Automatica **32**, 1285–1301 (1996)
29. Marmarelis, V., Juusola, M., French, A.: Principal dynamic mode analysis of nonlinear transduction in a spider mechanoreceptor. Ann. Biomed. Eng. **27**, 391–402 (1999)
30. Micchelli, C.A., Sauer, T., Xu, Y.: A construction of refinable sets for interpolating wavelets. Results Math. **34**, 359–372 (1998)
31. Micchelli, C.A., Xu, Y.: Using the matrix refinement equation for the construction of wavelets on invariant sets. Appl. Comput. Harmon. Anal. **1**, 391–401 (1994)
32. Micchelli, C.A., Xu, Y.: Reconstruction and decomposition algorithms for biorthogonal wavelets. Multidimens. Systems Signal Process. **8**, 31–69 (1997)
33. Micchelli, C.A., Xu, Y., Zhao, Y.: Wavelet Galerkin methods for second-kind integral equations. J. Comput. Appl. Math. **86**, 251–270 (1997)
34. Nemeth, J., Kollar, I., Schoukens, J.: Identification of volterra kernels using interpolation. IEEE Trans. Instrum. Meas. **51**, 770–775 (2002)
35. Parsa, V., Parker, P., Scott, R.: Adaptive stimulus artifact reduction in noncortical somatosensory evoked potential studies. IEEE Trans. Biomed. Eng. **45**, 165–179 (1998)
36. Phillips, D.L.: A technique for the numerical solution of certain integral equations of the first kind. J. Assoc. Comput. Mach. **9**, 84–97 (1962)
37. Prazenica, R., Kurdila, A.: Volterra kernel identification using triangular wavelet. J. Vib. Control **10**, 597–622 (2004)
38. Richter, G.R.: Supercovergence of piecewise polynomial galerkin approximations, for Fredholm integral equation of the second kind. Numer. Math. **31**, 63–70 (1978)
39. Renardy, M., Rogers, R.C.: An Introduction to Partial Differential Equations, 2nd edn. Springer, New York (2004)
40. Schwab, C.: Variable order composite quadrature of singular and nearly singular integrals. Computing **53**, 173–194 (1994)
41. Strauch, P., Mulgrew, B.: Active control of nonlinear noise processes in a linear duct. IEEE Trans. Signal Process. **46**, 2404–2412 (1998)
42. Tikhonov, A.N.: On the solution of ill posed problems and the method of regularization (Russian). Dokl. Akad. Nauk SSSR **151**, 501–504 (1963)
43. Tikhonov, A.N., Arsenin, V.Y.: Solutions of Ill-posed Problems. Wiley, New York (1977)
44. Taylor, A.E., Lay, D.C.: Intruduction to Functional Analysis, 2nd edn. Wiley, New York (1980)
45. Tsoulkas, V., Koukoulas, P., Kalouptsidis, N.: Identification of input-output bilinear systems using cumulants. IEEE Trans. Signal Process. **49**, 2753–2761 (2001)
46. Wang, Y., Xu, Y.: A fast wavelet collocation method for integral equations on polygons. J. Integral Equations Appl. **17**, 227–330 (2005)
47. Xu, Y., Chen, H.L., Zou, Q.: Limit values of derivatives of the Cauchy integrals and computation of the logarithmic potential. Computing **73**, 295–237 (2004)
48. Zhang, Q., Suki, B., Westwick, D., Lutchen, K.: Factors affecting Volterra kernel estimation: emphasis on lung tissue viscoelasticity. Ann. Biomed. Eng. **26**, 103–116 (1998)