INSTITUTE OF PHYSICS PUBLISHING INVERSE PROBLEMS

Inverse Problems **21** (2005) 1773–1790 doi:10.1088/0266-5611/21/5/016

On local regularization methods for linear Volterra equations and nonlinear equations of Hammerstein type

Patricia K Lamm and Zhewei Dai

Department of Mathematics, Michigan State University, E Lansing, MI 48824-1027, USA

E-mail: lamm@math.msu.edu and daizhewe@msu.edu

Received 13 June 2005, in final form 8 August 2005
Published 16 September 2005
Online at stacks.iop.org/IP/21/1773

Abstract

Local regularization methods allow for the application of sequential solution techniques for the solution of Volterra problems, retaining the causal structure of the original Volterra problem and leading to fast solution techniques. Stability and convergence of these methods was shown to hold on a large class of linear Volterra problems, i.e., the class of ν-smoothing problems for $\nu = 1, 2, \ldots$ in Lamm (2005 *Inverse Problems* **21** 785–803). In this paper, we enlarge the family of convergent local regularization methods to include sequential versions of classical regularization methods such as sequential Tikhonov regularization. In fact, sequential Tikhonov regularization was considered earlier by Lamm and Eldén (1997 *SIAM J. Numer. Anal.* **34** 1432–50) but there the theory was limited to the class of discretized one-smoothing Volterra problems. An interesting feature of sequential classical regularization methods is that they involve two regularization parameters: the usual local regularization parameter r controls the size of the local problem while a second parameter α controls the amount of regularization to be applied in each subproblem. This approach suggests a wavelet type of regularization method with the parameter r controlling spatial resolution and α controlling frequency resolution. In this paper, we also show how the 'future polynomial regularization' method of Cinzori (2004 *Inverse Problems* **20** 1791–806) can be viewed as a special case of the general framework of Lamm (2005) in the 1-smoothing case. In addition, we extend the results of Lamm (2005) to nonlinear Volterra problems of Hammerstein type and give numerical results to illustrate the effectiveness of the method in this case.

1. Introduction

In this paper, we consider Volterra equations of the form

$$\mathcal{A}u = f, \tag{1}$$

where for suitable $u : [0, T] \mapsto \mathbb{R}$, the convolution operator $\mathcal{A}$ will be given by

$$\mathcal{A}u(t) = \int_0^t k(t-s)u(s)\,\mathrm{d}s, \qquad t \in [0, T], \tag{2}$$

in the linear case (sections 2 and 3) and in the case of the nonlinear Hammerstein problem (section 4),

$$\mathcal{A}u(t) = \int_0^t k(t-s)S(s, u(s))\,\mathrm{d}s, \qquad t \in [0, T], \tag{3}$$

with $S : [0, T] \times \mathbb{R} \mapsto \mathbb{R}$. Sections 2 and 3 are due to the first author while the work in section 4 is the result of the combined effort of both authors.

A discussion of the existence and uniqueness of solutions of (1) may be found in [7] in the linear case and in [5, 6] in the nonlinear case. Throughout we will assume that $\mathcal{A}$ satisfies a ν-smoothing condition for some $\nu = 1, 2, \ldots$, that is, the kernel k satisfies

$$k \in C^{\nu}[0, T], \qquad k^{(j)}(0) = 0, \qquad j = 0, 1, \ldots, \nu - 2, \qquad k^{(\nu-1)}(0) \neq 0, \tag{4}$$

where without loss of generality we will take $k^{(\nu-1)}(0) = 1$. For any $\nu > 0$ the solution $\bar{u}$ of (1) generally lacks continuous dependence on data f (using either the $L_2(0, T)$ or $L_\infty(0, T)$ topologies on u and f), and it is relevant to note that the degree of the ill-posedness of the problem increases with increasing ν. Since in practice, we will only have access to a perturbed version f^δ of f where for some $\delta > 0$

$$\|f - f^\delta\|_\infty < \delta, \tag{5}$$

regularized solution methods will thus be essential in recovering a reasonable approximation to the 'true' or 'ideal' solution $\bar{u}$ of (1).

We will also assume that the desired $\bar{u}$ of (1) satisfies the Hölder condition

$$|\bar{u}(t) - \bar{u}(s)| \leqslant L_{\bar{u}}|t - s|^{\mu}, \tag{6}$$

for $0 < \mu \leqslant 1$, $L_{\bar{u}} > 0$, and t, s in the interval of interest (although this hypothesis may be relaxed to $\bar{u}$ only continuous, as is discussed in [12]).

A generalized theory for local regularization of linear ν-smoothing Volterra problems was developed in [12]. Local regularization methods allow for the application of sequential solution techniques for the solution of Volterra problems, retaining the causal structure of the original Volterra problem instead of replacing it with a non-causal/non-Volterra problem (such as typically happens when a classical regularization method such as Tikhonov regularization is used) [11]. As a result one is able to apply fast solution techniques to discretizations of Volterra problems.

Until recently it was thought that sequential local regularization methods faced a severe limitation in that stability and convergence could not be guaranteed unless restricted to ν-smoothing Volterra problems with $\nu \leqslant 4$ [16]. However in [12] the first author was able to successfully enlarge the definition of local regularization in order to develop a new method found to be both stable and convergent for *all* ν-smoothing Volterra problems, $\nu = 1, 2, \ldots$. With this finding, there is renewed motivation for exploring other possibilities for the local regularization of general ν-smoothing Volterra problems. The purpose of this paper is to extend some of the ideas in [12] to show how the new generalized local regularization method may be applied in completely new contexts.

In section 2, we show how sequential Tikhonov regularization (the theory of which has been limited to discretized 1-smoothing problems until now [13]) may be viewed as an extension of the theory of [12]. The importance of this extension is that two regularization parameters result: the usual local regularization parameter r controls the size of the local problem while a local Tikhonov parameter α controls the amount of regularization to be applied in each subproblem. This approach suggests a wavelet type of regularization method with r controlling spatial resolution and α controlling frequency resolution. In addition, variable local control over the amount of regularization is possible if α is allowed to vary over the domain of the problem. This feature has been illustrated in numerical examples in [10, 14, 15] for sequential Tikhonov and other sequential local methods in the case of $\mathcal{A}$ 1-smoothing; in this paper we extend these results to general ν-smoothing Volterra problems in the case of constant α. In addition to local/sequential Tikhonov regularization, the theory in section 2 also applies to any classical regularization method (Landweber, truncated singular value expansion, etc) applied in a sequential manner.

In section 3, we show how another local regularization method, 'future polynomial regularization' [1], can be viewed as a special case of the general framework of [12] in the 1-smoothing case. The link between the two methods is directly due to the fact that the theory in [12] allows for the used of signed measures.

Finally, in section 4 we extend the results of [12] and of section 2 to nonlinear Volterra problems of Hammerstein type.

1.1. Generalized local regularization of linear Volterra problems

Because this paper serves to extend the ideas in [12], we will briefly recall here the ideas of the local regularization method introduced in that reference for linear Volterra operators.

We will let $r \in (0, \bar{r}]$ denote a small parameter and assume that equation (1) holds on an extended interval $[0, T + \bar{r}]$ for sufficiently small $\bar{r} > 0$ fixed. If data are not available past the original interval then this can always be accomplished by simply decreasing the size of T slightly. Then the 'true' solution $\bar{u}$ of (1) satisfies

$$\int_0^{t+\rho} k(t + \rho - s)u(s)\,\mathrm{d}s = f(t + \rho), \qquad t \in [0, T], \quad \rho \in [0, r].$$

If we split the integral at t and make a change of integration variable we then have

$$\int_0^t k(t + \rho - s)u(s)\,\mathrm{d}s + \int_0^\rho k(\rho - s)u(t + s)\,\mathrm{d}s = f(t + \rho), \qquad t \in [0, T], \ \ \rho \in [0, r]. \tag{7}$$

For each $t \in [0, T]$, the ρ variable serves to advance the equation slightly into the future. We introduce a Borel measure $\eta_r = \eta_r(\rho)$ in order to 'consolidate' this future information; that is, we integrate both sides of the above equation with respect to $\eta_r(\rho)$ and obtain

$$\int_0^t \int_0^r k(t + \rho - s)\,\mathrm{d}\eta_r(\rho)\,u(s)\,\mathrm{d}s + \int_0^r \int_0^\rho k(\rho - s)u(t + s)\,\mathrm{d}s\,\mathrm{d}\eta_r(\rho)$$
$$= \int_0^r f(t + \rho)\,\mathrm{d}\eta_r(\rho), \qquad t \in [0, T]. \tag{8}$$

Equation (8) is still satisfied by $\bar{u}$ exactly; it is when the 'ideal' data f are replaced by a perturbation f^δ that a regularized form of the equation is needed in order to find a suitable approximation of $\bar{u}$. In [12] the regularized form of equation (8) is obtained by replacing $u(t + s)$ by $u(t)$ in the second term in equation (8), as if the goal is to hold u constant

(temporarily) on the small local interval of length $[t, t+r]$. The length r of this local interval becomes the regularization parameter for this method.

In the case of perturbed data f^δ, the regularized equation is then given by

$$a_r u(t) + \int_0^t \tilde{k}_r(t-s)u(s)\,\mathrm{d}s = \tilde{f}_r^\delta(t), \qquad t \in [0,T], \tag{9}$$

where

$$\tilde{k}_r(t) = \int_0^r k(t+\rho)\,\mathrm{d}\eta_r(\rho), \qquad \tilde{f}_r^\delta(t) = \int_0^r f^\delta(t+\rho)\,\mathrm{d}\eta_r(\rho), \tag{10}$$

$$a_r = \int_0^r \int_0^\rho k(\rho - s)\,\mathrm{d}s\,\mathrm{d}\eta_r(\rho). \tag{11}$$

Sufficient conditions for stability and convergence of solutions u_r^δ to $\bar{u}$ include the hypotheses on the measures η_r given below:

(H1) For $j = 0, 1, \ldots, \nu$, there is some $\sigma \in \mathbb{R}$ and $c_j = c_j(\nu) \in \mathbb{R}$ independent of r, $c_\nu > 0$, such that

$$\int_0^r \rho^j\,\mathrm{d}\eta_r(\rho) = r^{j+\sigma}(c_j + \mathcal{O}(r)), \qquad \text{as} \quad r \to 0. \tag{12}$$

(H2) The parameters c_j, $j = 0, 1, \ldots, \nu$ in (H1) satisfy the condition that all roots of the polynomial $p_\nu(\lambda) = \frac{c_\nu}{\nu!}\lambda^\nu + \frac{c_{\nu-1}}{(\nu-1)!}\lambda^{\nu-1} + \cdots + \frac{c_1}{1!}\lambda + \frac{c_0}{0!}$ have negative real part.

(H3) There exists a $\tilde{C} > 0$ independent of r such that $\left|\int_0^r g(\rho)\,\mathrm{d}\eta_r(\rho)\right| \leqslant \tilde{C}\|g\|_\infty r^\sigma$, for all $g \in C[0,r]$ and all $r > 0$ sufficiently small.

As was shown in [12], one may find an infinite number of continuous and discrete families $\{\eta_r\}_{r>0}$ of measures which are easily constructed and which satisfy (H1)–(H3). Further, under this construction we have from [12] that

$$\left|u_r^\delta(t) - \bar{u}(t)\right| \leqslant C_1\frac{\delta}{r^\nu} + C_2 r^\mu, \qquad t \in [0,T],$$

(for some $C_1, C_2 \geqslant 0$ and for μ the Hölder exponent on $\bar{u}$ in (6)) so that the choice

$$r = r(\delta) \sim \delta^{\frac{1}{\mu+\nu}}$$

gives

$$\left\|u_r^\delta - \bar{u}\right\|_\infty = \mathcal{O}\left(\delta^{\frac{\mu}{\mu+\nu}}\right) \to 0 \quad \text{as} \quad \delta \to 0.$$

2. Sequential classical regularization methods for linear problems

Here we will extend the theory outlined in section 1.1 to sequential versions of classical regularization methods such as Tikhonov regularization. Although more general nonlinear problems will be considered in section 4 our work in this section will be for the case of the ν-smoothing linear operator $\mathcal{A}$ as given by (2) and (4).

2.1. Sequential Tikhonov regularization applied to a discrete form of (1)

In [13] the idea of sequential Tikhonov regularization was introduced for the regularized solution of discretized first-kind Volterra problems. The idea in that paper was that the original problem (1) when discretized typically generates a matrix system of the form

$$\mathbf{A}\mathbf{u} = \mathbf{f}, \tag{13}$$

with $\mathbf{f} = (f_i) \in \mathbb{R}^N$ and $\mathbf{A} \in \mathbb{R}^{N\times N}$ nonsingular, lower-triangular, and Toeplitz, i.e.,

$$\mathbf{A} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ a_2 & a_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_N & a_{N-1} & \cdots & a_1 \end{pmatrix}.$$

Standard Tikhonov regularization applied to (13) requires the solution of the minimization problem

$$\min_{\mathbf{u}\in\mathbb{R}^N} \left\{ \|\mathbf{A}\mathbf{u} - \mathbf{f}\|^2_{\mathbb{R}^N} + \alpha \|\mathbf{u}\|^2_{\mathbb{R}^N} \right\},$$

where $\alpha > 0$ is a given parameter, or equivalently of the system

$$(\mathbf{A}^\top \mathbf{A} + \alpha \mathbf{I})\mathbf{u} = \mathbf{A}^\top \mathbf{f}, \tag{14}$$

where the leading matrix in (14) no longer retains the lower-triangular structure of $\mathbf{A}$.

In contrast, sequential (local) Tikhonov regularization preserves the structure of the Volterra problem and thus leads to a more efficient approach [13]. We motivate this method in what follows. Assuming that $u_1, u_2, \ldots, u_{i-1}$ have already been found, we note that the vector $\mathbf{u}_R^{(i)}$

$$\mathbf{u}_R^{(i)} = (u_i, u_{i+1}, \ldots, u_{i+R-1})^\top,$$

containing the next R components of $\mathbf{u}$ (for $1 \leqslant R \ll N$, $i \leqslant N - R + 1$), is the solution of a smaller subproblem

$$\mathbf{A}_R \mathbf{u}^{(i)} = \mathbf{f}_R^{(i)}, \tag{15}$$

where $\mathbf{A}_R$ is the leading $R \times R$ submatrix of $\mathbf{A}$, $\mathbf{u}^{(i)} \in \mathbb{R}^R$, and the vector $\mathbf{f}_R^{(i)} = \left(f_1^{(i)}, \ldots, f_R^{(i)}\right)^\top \in \mathbb{R}^R$ is given for $i \geqslant 2$ by

$$f_k^{(i)} = f_{i+k-1} - \sum_{j=1}^{i-1} a_{i+k-j} u_j, \qquad k = 1, \ldots, R,$$

with the components of $\mathbf{f}_R^{(1)}$ given by $f_k^{(1)} = f_k, k = 1, \ldots, R$. We then apply Tikhonov regularization to the smaller subproblem (15) and find the vector $\mathbf{u}_{R,\alpha}^{(i)} \in \mathbb{R}^R$ which solves the reduced-dimension regularization problem

$$\mathbf{u}_{R,\alpha}^{(i)} = \left(\mathbf{A}_R^\top \mathbf{A}_R + \alpha \mathbf{I}\right)^{-1} \mathbf{A}_R^\top \mathbf{f}_R^{(i)}, \tag{16}$$

where now $\mathbf{I}$ denotes the $\mathbb{R}^{R\times R}$ identity. For R small, only the first few components of $\mathbf{u}_{R,\alpha}^{(i)}$ are reliable estimates of $\mathbf{u}_R^{(i)}$ since it is well known that numerical solutions of Volterra problems are less accurate towards the end of the reconstruction interval. We therefore approximate

$$u_i = \left\langle \mathbf{u}_{R,\alpha}^{(i)}, \mathbf{x}_R \right\rangle_{\mathbb{R}^R}$$

for some $\mathbf{x}_R \in \mathbb{R}^R$ satisfying

$$\langle \mathbf{1}, \mathbf{x}_R \rangle_{\mathbb{R}^R} = 1, \qquad \mathbf{1} = (1, 1, \ldots, 1)^\top,$$

so that the approximate solution u_i at the end of the ith step is constructed using a weighted sum of the components of $\mathbf{u}_{R,\alpha}^{(i)}$ (although the 'weights' may actually be negative). The vector $\mathbf{u}_{R,\alpha}^{(i)}$ is no longer needed after this point and is discarded before solving the next local Tikhonov problem (i.e., (16) with i now replaced by $i+1$). Although more general choices of $\mathbf{x}_R$ are possible, we note that $\mathbf{x}_R = (1, 0, \ldots, 0)^\top$ was used in [13] so that only the first component of the solution of (16) was retained at the end of the ith step.

It is worth clarifying that sequential Tikhonov regularization is not merely a matter of a decomposition of the original matrix system into smaller subproblems, with standard Tikhonov regularization applied independently to each small problem. Rather the method performs Tikhonov regularization on small overlapping problems updating the definition of each subproblem as new information about the solution is determined [13].

The stability and convergence of this method was established in [13] for the discrete version of the problem in the 1-smoothing case. Numerical examples for 1-smoothing and 2-smoothing problems may be found in [10] and for an infinitely smoothing problem (the inverse heat conduction problem) in [13]. Below we formulate a continuous version of the sequential Tikhonov regularization method and prove stability and convergence of the method for all ν-smoothing operators $\mathcal{A}$. In addition the theory is formulated so that other sequential regularization methods (such as a sequential Landweber method) are possibilities as well.

2.2. *Sequential regularization in the setting of the continuous problem*

The idea of a sequential classical regularization method for the continuous problem (1) can be motivated starting from equation (7) in section 1.1. If we momentarily fix $t \in [0, T]$ and assume that u is already known on $[0, t)$, then equation (7) can be viewed as an equation of Volterra type restricted to the interval $[0, r]$ only, that is,

$$\mathcal{A}_r u_t(\rho) = F_{t,r}^{\delta}(\rho), \qquad \text{a.a.} \quad \rho \in [0, r]. \tag{17}$$

where $\mathcal{A}_r$ is the bounded linear operator on $L_2(0, r)$ given by

$$\mathcal{A}_r v(\rho) = \int_0^\rho k(\rho - s) v(s)\, \mathrm{d}s, \qquad v \in L_2(0, r), \tag{18}$$

and u_t, $F_{t,r}^{\delta} \in L_2(0, r)$ are defined for a.a. $\rho \in [0, r]$ via

$$u_t(\rho) = u(t + \rho), \tag{19}$$

$$F_{t,r}^{\delta}(\rho) = f^{\delta}(t + \rho) - \int_0^t k(t + \rho - s) u(s)\, \mathrm{d}s \tag{20}$$

(where we have already replaced f by f^{δ} in (7)). But the local problem (17) is still ill-posed; making use of classical Tikhonov regularization for this problem results in an approximation $u_t^{\delta}(\cdot; r, \alpha) \in L_2(0, r)$,

$$u_t^{\delta}(\cdot; r, \alpha) = (\mathcal{A}_r^\star \mathcal{A}_r + \alpha I)^{-1} \mathcal{A}_r^\star F_{t,r}^{\delta}$$

for I the identity on $L_2(0, r)$. Or, for a more general classical regularization method we would instead define

$$u_t^{\delta}(\cdot; r, \alpha) = R_\alpha(\mathcal{A}_r^\star \mathcal{A}_r) \mathcal{A}_r^\star F_{t,r}^{\delta}, \tag{21}$$

where the compact self-adjoint operator $R_\alpha(\mathcal{A}_r^\star \mathcal{A}_r)$ on $L_2(0, r)$ is defined in the usual way using a family $\{R_\alpha\}$ of continuous functions $R_\alpha : [0, \Lambda] \mapsto \mathbb{R}$, for some $\Lambda \in (0, 1]$, satisfying

(R1) $R_\alpha(t) \to 1/t$ as $\alpha \to 0$ for each $t > 0$;
(R2) $|tR_\alpha(t)| \leqslant C, t \in [0, \Lambda], \alpha > 0$;
(R3) $t|1 - tR_\alpha(t)| \leqslant \omega(\alpha), t \in [0, \Lambda]$, where $\omega(\alpha) \to 0$ as $\alpha \to 0$.

(See [6, 8].) We note that typically R_α is defined on the interval $[0, \|\mathcal{A}_r\|^2]$, however for $\mathcal{A}_r$ ν-smoothing we will have from (28) below that $\|A_r\| = \mathcal{O}(r^\nu)$ so that we may pick $\Lambda \in (0, 1]$ such that the domain $[0, \Lambda]$ is valid for sufficiently small $r > 0$. It is also worth nothing that for Tikhonov regularization, $R_\alpha(t) = 1/(t + \alpha)$, $C = 1$ and $\omega(\alpha) = \alpha$.

For each fixed t we will not need to make use of all of $u_t(\cdot; r, \alpha)$ obtained in (21) but instead retain only the information obtained in the scalar quantity

$$\left\langle u_t^\delta(\cdot; r, \alpha), \chi_r \right\rangle_r \tag{22}$$

to approximate the value of $\bar{u}$ at t. Here $\langle \cdot, \cdot \rangle_r$ denotes the usual inner product on $L_2(0, r)$ and χ_r is a suitably selected $L_2(0, r)$ function with the property that $\langle \mathbf{1}, \chi_\mathbf{r} \rangle_r = 1$ for $\mathbf{1}(\rho) = 1, \rho \in [0, r]$. This discussion motivates the following as a regularized approximating equation for $\bar{u}$ on all of $[0, T]$ (not just at fixed t); that is, we seek $u = u_{r,\alpha}^\delta : [0, T] \mapsto \mathbb{R}$ satisfying

$$u(t) = \left\langle u_t^\delta(\cdot; r, \alpha), \chi_r \right\rangle_r, \qquad \text{a.a.} \quad t \in [0, T],$$

or

$$\begin{aligned} u(t) &= \left\langle R_\alpha(\mathcal{A}_r^\star \mathcal{A}_r)\mathcal{A}_r^\star F_{t,r}^\delta, \chi_r \right\rangle_r \\ &= \left\langle F_{t,r}^\delta, \mathcal{A}_r R_\alpha(\mathcal{A}_r^\star \mathcal{A}_r)\chi_r \right\rangle_r \\ &= \left\langle F_{t,r}^\delta, R_\alpha(\mathcal{A}_r \mathcal{A}_r^\star)\mathcal{A}_r \chi_r \right\rangle_r \end{aligned}$$

(see for example section 2.1 of [8]). Then using the definition of $F_{t,r}^\delta$ in (20), it is not difficult to see that $u = u_{r,\alpha}^\delta$ is the (unique) solution of the well-posed second-kind Volterra equation

$$u(t) + \int_0^t \tilde{k}_{r,\alpha}(t - s)u(s)\,\mathrm{d}s = \tilde{f}_{r,\alpha}^\delta(t), \qquad \text{a.a.} \quad t \in [0, T], \tag{23}$$

where

$$\tilde{k}_{r,\alpha}(t) = \int_0^r k(t + \rho)R_\alpha(\mathcal{A}_r \mathcal{A}_r^\star)\mathcal{A}_r \chi_r(\rho)\,\mathrm{d}\rho, \tag{24}$$

$$\tilde{f}_{r,\alpha}^\delta(t) = \int_0^r f^\delta(t + \rho)R_\alpha(\mathcal{A}_r \mathcal{A}_r^\star)\mathcal{A}_r \chi_r(\rho)\,\mathrm{d}\rho, \tag{25}$$

for $t \in [0, T]$. The remaining issue is how to select χ_r and the regularization parameters $r = r(\delta), \alpha = \alpha(\delta)$ so that $u_{r(\delta),\alpha(\delta)}^\delta \to \bar{u}$ as the noise level δ in the data goes to zero.

2.3. Preliminary definitions and some technical results

Before turning to the statement of the main convergence result in theorem 2.1, we will first make some assumptions which will hold for the remainder of the section. We will henceforth let $T = 1$ so that (1) is defined on $[0, 1]$. In addition we will let $\mathcal{B}$ denote the compact operator on $L_2(0, 1)$ defined for a.a. $t \in [0, 1]$ by

$$\mathcal{B}v(t) = \int_0^t \frac{(t - s)^{\nu-1}}{(\nu - 1)!} v(s)\,\mathrm{d}s, \qquad v \in L_2(0, 1),$$

and note that $\mathcal{B}$ is a perturbation of the ν-smoothing operator $\mathcal{A}$.

We will assume that ψ is an $L_2(0,1)$ function which satisfies hypothesis (H) given below:

(H) For $j = 0, 1, \ldots, \nu$,

$$\int_0^1 \mathcal{B}\mathcal{B}^\star \rho^j \psi(\rho)\,\mathrm{d}\rho = c_j,$$

where $c_\nu > 0$ and where the polynomial

$$p(\lambda) = \sum_{j=0}^{\nu} \frac{c_j}{j!}\lambda^j$$

has ν roots λ_j, $j = 1, 2, \ldots, \nu$, all with negative real part.

Our first lemma shows that we may always find ψ satisfying hypothesis (H) and in fact we have a sufficiently large selection of ψ to allow for the roots of p to be placed anywhere we wish on the negative real axis.

Lemma 2.1. *Let $\bar{c} > 0$ and $m_i > 0$ for $i = 1, 2, \ldots, \nu$. Then we may find ψ satisfying (H) with $c_\nu = \bar{c}$ and such that the roots λ_i of p are given by $\lambda_i = -m_i$, $i = 1, 2, \ldots, \nu$.*

Proof. Let d_i denote the coefficients of the polynomial $P(\lambda) = \prod_{i=1}^{\nu}(\lambda + m_i)$, i.e.,

$$\lambda^\nu + d_{\nu-1}\lambda^{\nu-1} + \cdots + d_1\lambda + d_0 = \prod_{i=1}^{\nu}(\lambda + m_i).$$

Then we seek the vector $\mathbf{a} = (a_0, a_1, \ldots, a_\nu)^\top$ such that the polynomial ψ defined by $\psi(\rho) = \sum_{\ell=0}^{\nu} a_\ell \rho^\ell$ satisfies $c_j \equiv \int_0^1 \mathcal{B}\mathcal{B}^\star \rho^j \psi(\rho)\,\mathrm{d}\rho = j!\bar{c}d_j/\nu!$, for $j = 0, 1, \ldots, \nu$ ($d_\nu = 1$). Defining $\mathbf{d} = \left(0!\frac{\bar{c}}{\nu!}d_0, 1!\frac{\bar{c}}{\nu!}d_1, \ldots, (\nu-1)!\frac{\bar{c}}{\nu!}d_{\nu-1}, \bar{c}\right)^\top$, the vector $\mathbf{a}$ satisfies $\mathbf{H}\mathbf{a} = \mathbf{d}$ where $\mathbf{H}$ is the $(\nu+1)$-square matrix with entries

$$\mathbf{H}_{i,j} = \int_0^1 \mathcal{B}\mathcal{B}^\star \rho^i \rho^j \mathrm{d}\rho = \int_0^1 \mathcal{B}^\star \rho^i \mathcal{B}^\star \rho^j \mathrm{d}\rho \equiv (\rho^i, \rho^j)_{\mathcal{B}^\star},$$

and $(\cdot,\cdot)_{\mathcal{B}^\star}$ is a valid inner product on $L_2(0,1)$ from the injectivity of $\mathcal{B}^\star$. It follows that $\mathbf{H}$ is a Gram matrix for the independent functions ρ^i, $i = 0, 1, \ldots, \nu$ and as such is nonsingular. Thus ψ satisfies (H) with $c_\nu = \bar{c}$ and $p(\lambda) = \frac{\bar{c}}{\nu!}P(\lambda)$. □

We will use any ψ satisfying hypothesis (H) to construct the χ_r needed in our local regularization method. The following lemma is needed in this construction.

Lemma 2.2. *For any $\psi \in L_2(0,1)$ satisfying (H) and $\psi_r \in L_2(0,r)$ given by*

$$\psi_r(\rho) = \psi(\rho/r), \qquad \text{a.a.} \quad \rho \in [0, r], \tag{26}$$

the quantity $b_r \in \mathbb{R}$ defined by

$$b_r \equiv \langle 1, \mathcal{A}_r^\star(\mathcal{A}_r\mathcal{A}_r^\star)\psi_r\rangle_r \tag{27}$$

is positive for all $r > 0$ sufficiently small. Here $\mathcal{A}_r$ is given by (18).

Proof. Since $b_r = \langle(\mathcal{A}_r\mathcal{A}_r^\star)\mathcal{A}_r 1, \psi_r\rangle_r$ we have

$$\begin{aligned} b_r &= \int_0^r \left(\int_0^\rho k(\rho - s)\int_s^r k(\tau - s)\int_0^\tau k(\tau - w)\,\mathrm{d}w\,\mathrm{d}\tau\,\mathrm{d}s\right)\psi_r(\rho)\,\mathrm{d}\rho \\ &= r^4\int_0^1 \left(\int_0^\rho k(r(\rho - s))\int_s^1 k(r(\tau - s))\int_0^\tau k(r(\tau - w))\,\mathrm{d}w\,\mathrm{d}\tau\,\mathrm{d}s\right)\psi(\rho)\,\mathrm{d}\rho. \end{aligned}$$

But k is ν-smoothing so for $t \in [0, 1]$ we have [12]

$$k(t) = \frac{t^{\nu-1}}{(\nu-1)!} + h(t), \qquad |h(t)| \leqslant \frac{\|k^{(\nu)}\|_\infty t^\nu}{\nu!},$$

or

$$k(rt) = r^{\nu-1}\left(\frac{t^{\nu-1}}{(\nu-1)!} + \mathcal{O}(r)\right) \tag{28}$$

for all $t \in [0, 1]$. Thus for $v \in L_2(0, 1)$,

$$\begin{aligned}\int_0^\rho k(r(\rho-s))v(s)\,\mathrm{d}s &= r^{\nu-1}\left[\int_0^\rho \frac{(\rho-s)^{\nu-1}}{(\nu-1)!}v(s)\,\mathrm{d}s + \mathcal{O}(r)\right] \\ &= r^{\nu-1}\left[\mathcal{B}v(\rho) + \mathcal{O}(r)\right], \qquad \rho \in [0, 1],\end{aligned} \tag{29}$$

from which it follows that

$$\begin{aligned}b_r &= r^4 r^{3(\nu-1)}\left[\int_0^1 \mathcal{B}\mathcal{B}^\star(\mathcal{B}1)(\rho)\psi(\rho)\,\mathrm{d}\rho + \mathcal{O}(r)\right] \\ &= \frac{1}{\nu!}r^{3\nu+1}\left[\int_0^1 \mathcal{B}\mathcal{B}^\star\rho^\nu\psi(\rho)\,\mathrm{d}\rho + \mathcal{O}(r)\right].\end{aligned}$$

We conclude from assumption (H) that

$$b_r = \frac{c_\nu}{\nu!}r^{3\nu+1}(1 + \mathcal{O}(r)) > 0$$

for all $r > 0$ sufficiently small. □

2.4. Definition of χ_r and convergence results

As indicated earlier, we will prove that the sequential (local) regularization method converges for the right choice of χ_r, provided that the regularization parameters r and α are also selected appropriately. The idea behind the theory is that the choice of χ_r should lead to a signed measure η_r which fits into the context of section 1.1.

Theorem 2.1. *Let $\psi_r \in L_2(0, r)$ be given by (26) where $\psi \in L_2(0, 1)$ satisfies hypothesis (H), and for $\alpha > 0$ let $\{R_\alpha\}$ denote a family of continuous functions $R_\alpha : [0, \Lambda] \mapsto \mathbb{R}$ satisfying (R1)–(R3) for some $\Lambda \in (0, 1]$. We will let $u^\delta_{r,\alpha}$ denote the solution of (23) constructed using $\chi_r \in L_2(0, r)$ defined by*

$$\chi_r(\rho) = \frac{1}{\langle 1, \mathcal{A}_r^\star(\mathcal{A}_r\mathcal{A}_r^\star)\psi_r\rangle_r}\mathcal{A}_r^\star(\mathcal{A}_r\mathcal{A}_r^\star)\psi_r, \tag{30}$$

and using data f^δ satisfying (5). Let $\bar{u}$ denote the true solution of (1) which we assume satisfies the Hölder condition (6). Then there is a constant $\hat{C} > 0$ (depending on the c_j defined in (H) and independent of r and α) such that if $\|k^{(\nu)}\|_\infty < \hat{C}$ and if $r = r(\delta)$ is selected satisfying

$$r(\delta) \sim \delta^{\frac{1}{\mu+\nu}}$$

as $\delta \to 0$ and $\alpha = \alpha(\delta)$ is selected to ensure $\alpha(\delta) \to 0$ and

$$\omega(\alpha(\delta)) = \mathcal{O}\big(\delta^{\frac{2\nu+1}{\mu+\nu}}\big)$$

as $\delta \to 0$, it follows that

$$\big\|u^\delta_{r(\delta),\alpha(\delta)} - \bar{u}\big\|_\infty = \mathcal{O}\big(\delta^{\frac{\mu}{\mu+\nu}}\big)$$

as $\delta \to 0$.

Proof. We will assume that $\alpha = \alpha(r)$ has been selected so that $\alpha(r) \to 0$ and

$$\omega(\alpha(r)) = \mathcal{O}(r^{2\nu+1}),$$

as $r \to 0$, so that now r remains as the only regularization parameter. We define a signed measure η_r on $[0, r]$ via

$$\int_0^r g(\rho)\,\mathrm{d}\eta_r(\rho) = \int_0^r g(\rho) R_r(\mathcal{A}_r\mathcal{A}_r^\star)(\mathcal{A}_r\mathcal{A}_r^\star)^2\psi_r(\rho)\,\mathrm{d}\rho,$$

for $g \in L_\infty(0, r)$, where here and henceforth we simplify notation by writing

$$R_r \equiv R_{\alpha(r)},$$

and we will show that η_r satisfies conditions (H1)–(H3). First, for $j = 0, 1, \ldots, \nu$,

$$\begin{aligned}\int_0^r \rho^j\,\mathrm{d}\eta_r(\rho) &= \int_0^r \rho^j R_r(\mathcal{A}_r\mathcal{A}_r^\star)(\mathcal{A}_r\mathcal{A}_r^\star)^2\psi_r(\rho)\,\mathrm{d}\rho \\ &= \int_0^r \rho^j [R_r(\mathcal{A}_r\mathcal{A}_r^\star)(\mathcal{A}_r\mathcal{A}_r^\star)^2\psi_r(\rho) - \mathcal{A}_r\mathcal{A}_r^\star\psi_r(\rho)]\,\mathrm{d}\rho \\ &\quad + \int_0^r \rho^j \mathcal{A}_r\mathcal{A}_r^\star\psi_r(\rho)\,\mathrm{d}\rho \qquad (31)\end{aligned}$$

where in the second term in (31) we have

$$\begin{aligned}\int_0^r \rho^j \mathcal{A}_r\mathcal{A}_r^\star\psi_r(\rho) &= \int_0^r (\mathcal{A}_r\mathcal{A}_r^\star\rho^j)\psi_r(\rho) \\ &= \int_0^r \left(\int_0^\rho k(\rho - s)\int_s^r k(\tau - s)\tau^j\,\mathrm{d}\tau\,\mathrm{d}s\right)\psi_r(\rho)\,\mathrm{d}\rho \\ &= r^{j+3}\int_0^1 \left(\int_0^\rho k(r(\rho - s))\int_s^1 k(r(\tau - s))\tau^j\mathrm{d}\tau\,\mathrm{d}s\right)\psi(\rho)\,\mathrm{d}\rho \\ &= r^{j+3+2(\nu-1)}\left[\int_0^1 (\mathcal{B}\mathcal{B}^\star\rho^j)\psi(\rho)\,\mathrm{d}\rho + \mathcal{O}(r)\right] \\ &= r^{j+(2\nu+1)}(c_j + \mathcal{O}(r)) \qquad (32)\end{aligned}$$

from (29) and the construction of ψ. The first term in (31) satisfies

$$\begin{aligned}&\left|\int_0^r \rho^j [R_r(\mathcal{A}_r\mathcal{A}_r^\star)(\mathcal{A}_r\mathcal{A}_r^\star)^2\psi_r(\rho) - \mathcal{A}_r\mathcal{A}_r^\star\psi_r(\rho)]\,\mathrm{d}\rho\right| \\ &\quad\leqslant \left(\int_0^r \rho^{2j}\,\mathrm{d}\rho\right)^{1/2} \|R_r(\mathcal{A}_r\mathcal{A}_r^\star)(\mathcal{A}_r\mathcal{A}_r^\star)^2\psi_r - \mathcal{A}_r\mathcal{A}_r^\star\psi_r\|_r \\ &\quad= \left(\frac{r^{2j+1}}{2j+1}\right)^{1/2} \|(\mathcal{A}_r\mathcal{A}_r^\star)[I - (\mathcal{A}_r\mathcal{A}_r^\star)R_r(\mathcal{A}_r\mathcal{A}_r^\star)]\psi_r\|_r \\ &\quad\leqslant \frac{r^{j+\frac{1}{2}}}{\sqrt{2j+1}}\omega(\alpha(r))\|\psi_r\|_r \\ &\quad\leqslant \frac{r^{j+\frac{1}{2}}}{\sqrt{2j+1}}\bar{C}r^{2\nu+1}r^{\frac{1}{2}}\|\psi\| \\ &\quad= r^{j+(2\nu+1)}\frac{\bar{C}r}{\sqrt{2j+1}}\|\psi\|, \qquad (33)\end{aligned}$$

for some constant $\bar{C} > 0$ and where $\|\cdot\|_r$ and $\|\cdot\|$ denote the norms on $L_2(0, r)$ and $L_2(0, 1)$ respectively. It follows from (32) and (33) that

$$\int_0^r \rho^j\,\mathrm{d}\eta_r(\rho) = r^{j+\sigma}(c_j + \mathcal{O}(r)), \qquad j = 0, 1, \ldots, \nu,$$

for $\sigma = 2\nu+1$ where the assumptions on the coefficients c_j are given in (H). Thus the measure η_r satisfies conditions (H1)–(H2) from section 1.1. Further, for $g \in C[0, r]$,

$$\begin{aligned}\left|\int_0^r g(\rho)\,\mathrm{d}\eta_r(\rho)\right| &= \left|\int_0^r g(\rho)R_r(\mathcal{A}_r\mathcal{A}_r^\star)(\mathcal{A}_r\mathcal{A}_r^\star)^2\psi_r(\rho)\,\mathrm{d}\rho\right| \\ &\leqslant \|g\|_r\|R_r(\mathcal{A}_r\mathcal{A}_r^\star)(\mathcal{A}_r\mathcal{A}_r^\star)^2\psi_r\|_r \\ &\leqslant r^{\frac{1}{2}}\|g\|_\infty C\|\mathcal{A}_r\mathcal{A}_r^\star\psi_r\|_r\end{aligned}$$

where we have used hypothesis (R2). Again applying (29) we have

$$\begin{aligned}\|\mathcal{A}_r\mathcal{A}_r^\star\psi_r\|_r^2 &= r^5\int_0^1\left(\int_0^\rho k(r(\rho-s))\int_s^1 k(r(\tau-s))\psi(\tau)\,\mathrm{d}\tau\,\mathrm{d}s\right)^2\mathrm{d}\rho \\ &= r^5 r^{4(\nu-1)}\left[\int_0^1(\mathcal{B}\mathcal{B}^\star\psi(\rho))^2\,\mathrm{d}\rho + \mathcal{O}(r)\right] \\ &\leqslant r^5 r^{4(\nu-1)}[\|\mathcal{B}\mathcal{B}^\star\|^2\|\psi\|^2 + \mathcal{O}(r)] \\ &= \|\psi\|^2\mathcal{O}(r^{4\nu+1})\end{aligned}$$

as $r \to 0$, so that

$$\left|\int_0^r g(\rho)\,\mathrm{d}\eta_r(\rho)\right| = r^{\frac{1}{2}}\|g\|_\infty\|\psi\|\mathcal{O}\big(r^{2\nu+\frac{1}{2}}\big) \leqslant \tilde{C}\|g\|_\infty r^\sigma,$$

for some $\tilde{C} > 0$. Therefore the measure η_r satisfies (H3) from section 1.1 as well.

Using the definition of χ_r in (30) we know that for $g \in C[0, r]$

$$\begin{aligned}\int_0^r g(\rho)R_r(\mathcal{A}_r\mathcal{A}_r^\star)\mathcal{A}_r\chi_r(\rho)\,\mathrm{d}\rho &= \frac{1}{b_r}\int_0^r g(\rho)R_r(\mathcal{A}_r\mathcal{A}_r^\star)(\mathcal{A}_r\mathcal{A}_r^\star)^2\psi_r(\rho)\,\mathrm{d}\rho \\ &= \frac{1}{b_r}\int_0^r g(\rho)\,\mathrm{d}\eta_r(\rho)\end{aligned}$$

where b_r defined by (27) is positive. We may therefore rewrite equation (23) in $u = u_{r,\alpha(r)}$ as follows:

$$b_r u(t) + \int_0^t\int_0^r k(t+\rho-s)\,\mathrm{d}\eta_r(\rho)u_r(s)\,\mathrm{d}s = \int_0^r f^\delta(t+\rho)\,\mathrm{d}\eta_r(\rho),$$

for $t \in [0, 1]$, or

$$b_r u(t) + \int_0^t \tilde{k}_r(t-s)u(s)\,\mathrm{d}s = \tilde{f}_r^\delta(t), \qquad t \in [0, 1], \tag{34}$$

where $\tilde{k}_r$ and $\tilde{f}_r^\delta$ are given by (10) in section 1.1.

We note that equation (34) is the same as equation (9) in section 1.1 except for the fact that here b_r replaces a_r as the factor in front of $\bar{u}(t)$. In fact, a_r in (11) satisfies

$$a_r = \langle\mathcal{A}_r 1, R_r(\mathcal{A}_r\mathcal{A}_r^\star)(\mathcal{A}_r\mathcal{A}_r^\star)^2\psi_r\rangle_r,$$

so that again using (R3)

$$\begin{aligned}|b_r - a_r| &= |\langle\mathcal{A}_r 1, (\mathcal{A}_r\mathcal{A}_r^\star)\psi_r - R_r(\mathcal{A}_r\mathcal{A}_r^\star)(\mathcal{A}_r\mathcal{A}_r^\star)^2\psi_r\rangle_r| \\ &\leqslant \|\mathcal{A}_r 1\|_r\bar{C}r^{2\nu+1}r^{1/2}\|\psi\|\end{aligned}$$

where

$$\begin{aligned}\|\mathcal{A}_r 1\|_r^2 &= \int_0^r\left(\int_0^\rho k(\rho-s)\,\mathrm{d}s\right)^2\mathrm{d}\rho \\ &= r^3\int_0^1\left(\int_0^\rho k(r(\rho-s))\,\mathrm{d}s\right)^2\mathrm{d}\rho \\ &= r^3 r^{2(\nu-1)}\left[\int_0^1(\mathcal{B}1(\rho))^2\,\mathrm{d}\rho + \mathcal{O}(r)\right]\end{aligned}$$

so that $\|\mathcal{A}_r 1\|_r = \mathcal{O}(r^{\nu+1/2})$ and $|b_r - a_r| = \mathcal{O}(r^{3\nu+2})$. From lemma 2.1 of [12] (where α_r in that paper corresponds to a_r here) we further have that

$$a_r = \frac{c_\nu}{\nu!} r^{\nu+(2\nu+1)}(1 + \mathcal{O}(r)) > 0,$$

for all r sufficiently small, because η_r satisfies (H1)–(H3). It therefore follows that

$$\frac{|b_r - a_r|}{|a_r|} = \mathcal{O}(r) \tag{35}$$

as $r \to 0$, so that the leading terms in equation (34) above and equation (9) from section 1.1 are quite close.

Now the true solution $\bar{u}$ of (1) satisfies equation (7) so integrating this equation with respect to the measure η_r we have

$$b_r \bar{u}(t) + \int_0^r \int_0^t k(t+\rho-s)\bar{u}(s)\,\mathrm{d}s\,\mathrm{d}\eta_r(\rho) + \int_0^r \int_0^\rho k(\rho-s)\bar{u}(t+s)\,\mathrm{d}s\,\mathrm{d}\eta_r(\rho)$$
$$= \int_0^r f(t+\rho)\,\mathrm{d}\eta_r(\rho) + b_r \bar{u}(t), \qquad t \in [0,1],$$

or

$$b_r \bar{u}(t) + \int_0^t \tilde{k}_r(t-s)\bar{u}(s)\,\mathrm{d}s = \int_0^r f(t+\rho)\,\mathrm{d}\eta_r(\rho)$$
$$+ (b_r - a_r)\bar{u}(t) + \int_0^r \int_0^\rho k(\rho-s)[\bar{u}(t) - \bar{u}(t+s)]\,\mathrm{d}s\,\mathrm{d}\eta_r(\rho).$$

Letting $y_r(t) = u_r(t) - \bar{u}(t)$, the error equation in y_r becomes

$$b_r y(t) + \int_0^t \tilde{k}_r(t-s)y(s)\,\mathrm{d}s = \int_0^r [f^\delta(t+\rho) - f(t+\rho)]\,\mathrm{d}\eta_r(\rho) + (a_r - b_r)\bar{u}(t)$$
$$+ \int_0^r \int_0^\rho k(\rho-s)[\bar{u}(t+s) - \bar{u}(t)]\,\mathrm{d}s\,\mathrm{d}\eta_r(\rho)$$

or

$$K_r \star y_r(t) + y_r(t) = \bar{F}_r^\delta(t),$$

where for $t \in [0,1]$,

$$K_r(t) \equiv \frac{1}{b_r}\tilde{k}_r(t) = \frac{1}{b_r}\int_0^\rho k(t+\rho)\,\mathrm{d}\eta_r(\rho),$$
$$\bar{F}_r^\delta(t) \equiv \frac{1}{b_r}\left(\int_0^r [f^\delta(t+\rho) - f(t+\rho)]\,\mathrm{d}\eta_r(\rho) + (a_r - b_r)\bar{u}(t)\right.$$
$$\left. + \int_0^r \int_0^\rho k(\rho-s)[\bar{u}(t+s) - \bar{u}(t)]\,\mathrm{d}s\,\mathrm{d}\eta_r(\rho)\right).$$

Then using arguments like those in lemma 3.1 from [12], we are able to show

$$K_r^{(p)}(t) = r^{-(\sigma+\nu)}\frac{\nu!}{c_\nu}\int_0^r k^{(p)}(t+\rho)\,\mathrm{d}\eta_r(\rho)(1+\mathcal{O}(r)), \qquad p = 0,\ldots,\nu, \tag{36}$$

$$K_r^{(p)}(0) = \frac{\nu!}{(\nu-(p+1))!}\frac{r^{-(p+1)}}{c_\nu}[c_{\nu-(p+1)} + \mathcal{O}(r)], \qquad p = 0,\ldots,\nu-1, \tag{37}$$

$$\left|F_r^\delta(t)\right| \leqslant C_1\frac{\delta}{r^\nu} + C_2 r^\mu, \qquad t \in [0,T], \tag{38}$$

for nonnegative constants C_1 and C_2 independent of $r, \sigma = 2\nu + 1, \mu \in (0, 1]$ the Hölder exponent for $\bar{u}$, and δ given in (5). Along with the arguments in lemma 3.1 of [12] we have used (35) and the fact that

$$\frac{1}{b_r} = \frac{1}{a_r}(1 + \mathcal{O}(r))$$

as $r \to 0$. We then have enough conditions to guarantee that theorem 3.1 of [12] holds, or that the results of this theorem are true. □

Remark 2.1. We note that the conditions on $\bar{u}$ and f^δ may be relaxed somewhat and refer the reader to [12] for more details.

3. Future polynomial regularization methods

In the 'future polynomial regularization method' [1, 2], the idea is to regularize at each sequential step by finding a (low degree) polynomial which solves the local problem in a least-squares sense; this approach generalizes some of the ideas of [9] where regularization was handled through the use of a zero-degree polynomial. The theory underlying the future polynomial method was established for 1-smoothing operators in [2] for the discrete case and later in [1] for the continuous case. The work in [2] predates the work in [12] and was the first time that a sequential local regularization method made use of a signed measure. Although the theory in [12] was developed in order to address difficulties raised in [16], it is interesting to note that in fact the signed measures defined in [12] for the continuous Volterra problem include the measure used in [1] as a special case, at least when $\mathcal{A}$ is 1-smoothing.

Assuming that u is already known on the interval $[0, t)$ the idea in [1] is to find the best d-degree polynomial p_d which solves (17) in least-squares sense on $[t, t + r]$; after a change of coordinates to the interval $[0, r]$, the value of $p_d(0)$ is then used to approximate the value of u at t. If we write $p_d(\rho) = \sum_{\ell=0}^{d} b_\ell \rho^\ell$ then $p_d(0) = b_0$ where $\mathbf{b} = (b_0, b_1, \ldots, b_d)^\top$ solves the matrix equation

$$\mathbf{C}\mathbf{b} = \mathbf{e}\big(F_{t,r}^\delta\big) \tag{39}$$

where

$$\begin{aligned}
&\mathbf{C} = (\mathbf{c}_0 \mid \mathbf{c}_1 \mid \cdots \mid \mathbf{c}_d),\\
&\mathbf{c}_j = (c_{ij})_{i=0}^d \in \mathbb{R}^{d+1}, && c_{ij} = \langle \mathcal{A}_r \rho^i, \mathcal{A}_r \rho^j \rangle_r,\\
&\mathbf{e}(g) = (e_i(g))_{i=0}^d \in \mathbb{R}^{d+1}, && e_i(g) = \langle g, \mathcal{A}_r \rho^i \rangle_r,
\end{aligned}$$

for $g \in C[0, r]$ and $j = 0, \ldots, d$ (here $\mathcal{A}_r$ and $F_{t,r}^\delta$ are defined in (18) and (20), respectively). If we approximate $u(t)$ via $p_d(0) = b_0$, then this idea can be used to motivate the use of an approximating second-kind equation of the form (9) where the measure η_r arises from the use of Cramer's Rule to find b_0 in the solution of (39) [1]. The measure can be written as follows:

$$\int_0^r g(\rho)\, \mathrm{d}\eta_r(\rho) = \det \mathbf{D}(g) \tag{40}$$

$g \in C[0, r]$, where $\mathbf{D}(g)$ is $(d + 1)$-square,

$$\mathbf{D}(g) = (d_{ij}(g)) = (\mathbf{e}(g) \mid \mathbf{c}_1 \mid \cdots \mid \mathbf{c}_d).$$

Theorem 3.1. *The signed Borel measure η_r defined by (40) satisfies hypotheses (H1)–(H3) in the case of $\mathcal{A}$ 1-smoothing, so that the convergence/stability results of [12] also apply to the future polynomial regularization method in this case.*

Proof. We note that for $\mathcal{A}$ 1-smoothing

$$\begin{aligned} c_{ij} &= \int_0^r \left(\int_0^\rho s^i \, \mathrm{d}s \right) \left(\int_0^\rho s^j \, \mathrm{d}s \right) \mathrm{d}\rho (1 + \mathcal{O}(r)) \\ &= r^{i+j+3} k_{i,j} (1 + \mathcal{O}(r)), \end{aligned}$$

for $k_{i,j}$ independent of r, and similarly

$$e_i(g) = r^{i+2} \hat{k}_i(g; r)(1 + \mathcal{O}(r))$$

where $\hat{k}_i(g; r) = \int_0^1 g(r\rho)\rho^{i+1} \, \mathrm{d}\rho/(i+1)$. Thus using the explicit formula for the determinant (where p denotes a permutation of $\{0, 1, \ldots, d\}$) we have

$$\begin{aligned} \int_0^r g(\rho) \, \mathrm{d}\eta_r(\rho) &= \sum_p \operatorname{sign}(p) \left(\prod_{j=0}^d d_{p(j)j}(g) \right) \\ &= \sum_p \operatorname{sign}(p) e_{p(0)}(g) \left(\prod_{j=1}^d c_{p(j)j} \right) \\ &= \sum_p \operatorname{sign}(p) r^{p(0)+2} \hat{k}_{p(0)}(g; r) \left(\prod_{j=1}^d r^{p(j)+j+3} k_{p(j),j} \right) (1 + \mathcal{O}(r)) \\ &= r^\sigma \sum_p \operatorname{sign}(p) \hat{k}_{p(0)}(g; r) \left(\prod_{j=1}^d k_{p(j),j} \right) (1 + \mathcal{O}(r)), \end{aligned}$$

where $\sigma = 2 + \sum_{j=0}^d j + \sum_{j=1}^d (j+3) = d^2 + 4d + 2$. But $|\hat{k}_i(g; r)| \leqslant \|g\|_\infty \frac{1}{(i+2)(i+1)}$ so it follows that

$$\left| \int_0^r g(\rho) \, \mathrm{d}\eta_r(\rho) \right| \leqslant \hat{C} r^\sigma \|g\|_\infty$$

and (H3) is satisfied. For (H1) and (H2) we let $g(\rho) = \rho^\ell$ for $\ell = 0, 1$ and compute

$$\hat{k}_i(\rho^\ell, r) = \frac{1}{i+1} \int_0^1 (r\rho)^\ell \rho^{i+1} \, \mathrm{d}\rho = \frac{r^\ell}{(i+1)(i+\ell+2)}$$

so that for $i = 0, 1$,

$$\int_0^r \rho^i \, \mathrm{d}\eta_r(\rho) = r^{\sigma+i} c_i (1 + \mathcal{O}(r)),$$

where c_0 and c_1 are independent of r. Note that $c_1 > 0$ since $r^{\sigma+1} c_1 (1 + \mathcal{O}(r)) = \det \mathbf{D}(\rho) = \det \mathbf{C}(1 + \mathcal{O}(r)) > 0$ for r sufficiently small; here we have used the fact that coordinates in the first column of $\mathbf{C}$ satisfy $c_{0j} = \langle \mathcal{A}_r 1, \mathcal{A}_r \rho^j \rangle_r = \langle \rho, \mathcal{A}_r \rho^j \rangle_r (1 + \mathcal{O}(r)) = e_j(\rho)(1 + \mathcal{O}(r)) = d_{0,j}(\rho)(1 + \mathcal{O}(r))$. In addition, $c_0 > 0$ follows from the proof of lemma 4 of [1] where it is shown that $x_1 \equiv \left(\int_0^r 1 \, \mathrm{d}\eta_r(\rho) / \det \mathbf{C} \right) (1 + \mathcal{O}(r))$ satisfies $x_1 > 0$ for all r sufficiently small. Thus (H1) holds and since the root λ of the polynomial p_1 in (H2) is given by $\lambda = -c_0/c_1 < 0$, hypothesis (H2) holds as well. □

4. Application to nonlinear Hammerstein equations

We now turn to the nonlinear Hammerstein equation (1) where $\mathcal{A}$ is defined by (3) with the kernel k of $\mathcal{A}$ satisfying the usual ν-smoothing condition. We assume further that

$S : [0, T] \times \mathbb{R} \mapsto \mathbb{R}$ is continuous with

(S1)

$$\lim_{x \to +\infty} S(t, x) = +\infty, \quad \lim_{x \to -\infty} S(t, x) = -\infty, \qquad t \in [0, T],$$

(S2)

$$(S(t, x) - S(t, y))(x - y) > 0, \qquad \text{for all} \quad t \in [0, T] \quad \text{and} \quad x, y \in \mathbb{R} \text{ with } x \neq y.$$

Then for f sufficiently smooth there exists a unique solution $\bar{u} \in C[0, T]$ of the nonlinear Hammerstein equation. This fact follows from an application of theorem 3 of [5] to the equation which results after (1) has been differentiated $\nu - 1$ times.

Using the same approach as in section 1.1, we assume that $\bar{u}$ satisfies the Hammerstein equation on the extended interval $[0, T + R]$ so that we may define the following nonlinear regularized approximating equation

$$a_r S(t, u(t)) + \int_0^t \tilde{k}_r(t - s) S(s, u(s))\, \mathrm{d}s = \tilde{f}_r^{\delta}(t), \qquad t \in [0, T], \tag{41}$$

where a_r, $\tilde{k}_r$ and $\tilde{f}_r^{\delta}$ are given by (10)–(11) using a signed measure η_r satisfying (H1)–(H3). But if we let $\Sigma(t) = S(t, u(t))$, equation (41) is nothing more than equation (9) in the new variable $\Sigma(t)$, i.e.,

$$a_r \Sigma(t) + \int_0^t \tilde{k}_r(t - s) \Sigma(s)\, \mathrm{d}s = \tilde{f}_r^{\delta}(t), \qquad t \in [0, T]. \tag{42}$$

Then for $f^{\delta} \in C[0, T + R]$, there exists a unique solution $\Sigma_r^{\delta} \in C[0, T]$ of (42); further using (S1) and (S2) it is not difficult to show that there exists a unique $u_r^{\delta} \in C[0, T]$ such that $S\big(t, u_r^{\delta}(t)\big) = \Sigma_r^{\delta}(t)$, $t \in [0, T]$ [5]. Thus there is a unique solution u_r^{δ} of the approximating equation (41). The convergence of u_r^{δ} to $\bar{u}$ follows from the results in [12] and is made precise in the following theorem:

Theorem 4.1. *Consider the nonlinear Hammerstein problem (1) with ν-smoothing $\mathcal{A}$ given by (3) where S, S_t, $S_x : [0, T + R] \times \mathbb{R} \mapsto \mathbb{R}$ are continuous with S_t, S_x bounded on bounded subsets of $[0, T + R] \times \mathbb{R}$, S satisfies (S1) and (S2) for $t \in [0, T + R]$, and the unique solution $\bar{u}$ of (1) satisfies the usual Hölder condition (6) on $[0, T + R]$. We assume further that S satisfies (S3):*

(S3) There exists $\gamma > 0$ such that $(S(t, x) - S(t, y))(x - y) \geqslant \gamma |x - y|^2$, for all $t \in [0, T + R]$ and $x, y \in B$,

where the convex open bounded set $B \in \mathbb{R}$ satisfies $\bar{u}(t) \in B$ for $t \in [0, T + R]$.

Let $\{\eta_r\}$ be a family of signed Borel measures satisfying hypotheses (H1)–(H3) for all $r \in (0, R]$. Then there is a constant $\hat{C} > 0$ (depending only on the c_i defined in (H1) and independent of r) such that if $\|k^{(\nu)}\|_{\infty} < \hat{C}$, then for all $t \in [0, T]$ and in the case of exact data f we have for $t \in [0, T]$,

$$|u_r(t) - \bar{u}(t)| = \mathcal{O}(r^{\mu}) \to 0 \quad \text{as} \quad r \to 0. \tag{43}$$

If in addition $f^{\delta} \in C[0, T + R]$ satisfies (5) then the choice

$$r = r(\delta) \sim \delta^{\frac{1}{\mu + \nu}} \tag{44}$$

gives

$$\big|u_r^{\delta}(t) - \bar{u}(t)\big| = \mathcal{O}\big(\delta^{\frac{\mu}{\mu + \nu}}\big) \quad \text{as} \quad \delta \to 0, \tag{45}$$

for $t \in [0, T]$.

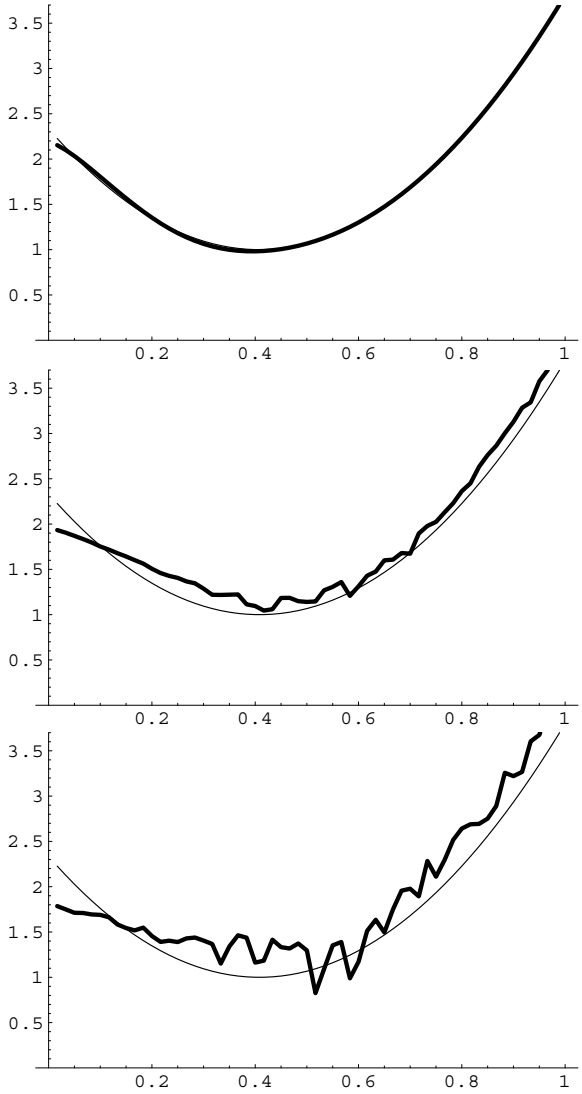

Figure 1. Nonlinear Hammerstein equation with 0%, 0.1% and 1% relative error in the data, with $r = 1/10, 4/10$ and $2/3$, respectively.

Proof. Since $\bar{u}$ satisfies (6) on $[0, T+R]$ the same is true of $\bar{\Sigma}(t) \equiv S(t, \bar{u}(t))$ since

$$|\bar{\Sigma}(t) - \bar{\Sigma}(s)| \leqslant |S(t, \bar{u}(t)) - S(s, \bar{u}(t))| + |S(s, \bar{u}(t)) - S(s, \bar{u}(s))|$$
$$\leqslant K(|t - s| + |\bar{u}(t) - \bar{u}(s)|) \leqslant M|t - s|^{\mu}$$

for suitable constants K, M and $\mu \in (0, 1]$ given in (6). Then from [12] we have

$$\|\bar{\Sigma} - \Sigma_r\|_{\infty} = \mathcal{O}(r^{\mu}) \to 0 \qquad \text{as} \quad r \to 0, \tag{46}$$

where Σ_r is the solution of (42) in the case of exact data f and Σ_r^{δ} the solution in the case of noisy data f^{δ}, with r selected according to (44),

$$\left\|\bar{\Sigma} - \Sigma_r^{\delta}\right\|_{\infty} = \mathcal{O}\left(\delta^{\frac{\mu}{\mu+\nu}}\right) \qquad \text{as} \quad \delta \to 0.$$

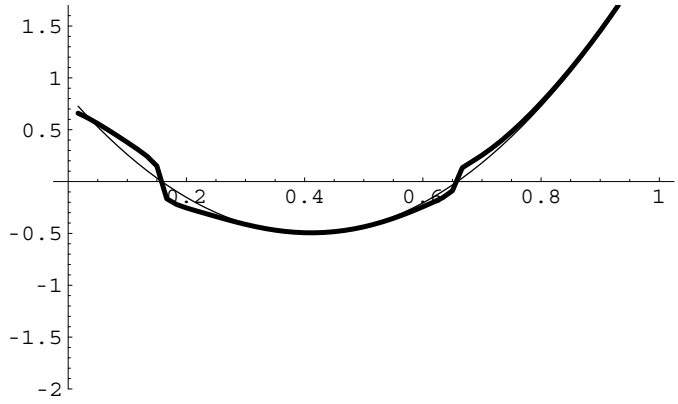

Figure 2. Hammerstein problem with nonlinearity violating hypothesis (S3). Approximate solution using regularization parameter $r = 1/10$ and 0% relative error in the data (compare to first graph in figure 1).

Fix $t \in [0, T]$ and let $r_n \to 0$ as $n \to \infty$ (in the noise-free case). Then (46) implies $\{\Sigma_{r_n}(t)\}_{n=1}^{\infty}$ is bounded and from (S1) it follows that the sequence $\{u_{r_n}(t)\}_{n=1}^{\infty}$ uniquely given by $S(t, u_{r_n}(t)) = \Sigma_{r_n}(t)$ is also bounded. Let $\{u_{r_{n_k}}(t)\}$ be any subsequence of $\{u_{r_n}(t)\}$; this subsequence in turn has a subsequence (which we relabel $u_{r_{n_k}}$) such that $u_{r_{n_k}}(t)$ converges to $\bar{u}(t)$, a consequence of the continuity of S and uniqueness of the solution $\bar{u}(t)$ of $S(t, \bar{u}(t)) = \bar{\Sigma}(t)$. Therefore $u_{r_n}(t) \to \bar{u}(t)$ as $n \to 0$ and for sufficiently large n we have $u_{r_n}(t) \in B$. From (S3) we then have

$$\left|\bar{u}(t) - u_{r_n}(t)\right| \leqslant \frac{1}{\gamma}\left|\bar{\Sigma}(t) - \Sigma_{r_n}(t)\right| = \mathcal{O}\left(r_n^{\mu}\right) \to 0 \qquad \text{as} \quad n \to \infty,$$

from which (43) follows. Similar arguments give (45). □

Remark 4.1. We note that under hypotheses similar to those made in theorem 4.1, the results of theorem 2.1 from section 2 also apply to the nonlinear Hammerstein problem. This extension and the result in theorem 4.1 are consequences of the fact that the Hammerstein problem can be viewed as a composition of the linear Volterra problem (1)–(2) with the nonlinear problem of solving $\Sigma(t) = S(t, u(t))$ for $u(t)$. The local regularization of more general nonlinear Volterra problems obviously requires a different approach as can be seen in the work on (sequential) local regularization methods for the autoconvolution problem

$$\int_0^t u(t-s)u(s) = f(t), \qquad t \in [0, T],$$

in [3, 4].

Example 4.1. To illustrate the local regularization method for a nonlinear Hammerstein equation, we present a numerical example using $k(t) = 0.5t^2$ (a 3-smoothing problem) and $S(t, x) = x^3$. That is, we wish to solve

$$\int_0^t 0.5(t-s)^2 u^3(s)\,\mathrm{d}s = f(t),$$

for u. We illustrate in figure 1 the true solution $\bar{u}(t) = 8(t - 0.4)^2 + 1$ (dashed curve) and the regularized approximate solution u_r^{δ} (solid curve), the latter obtained using an f^{δ} with 0%, 0.1% and 1% relative error in f. The collocation-based discretization is based on a subdivision

of the interval [0, 1] into $N = 60$ subintervals, while the local regularization interval is given by $[0, r]$ where $r = 1/10, 4/10$ and $2/3$, respectively. The measure η_r is a density constructed using discrete measure as described in lemma 2.3 of [12] where the construction has been made in such a way that the polynomial p_3 in (H3) is given by $p_3(\lambda) = (\lambda + 1.35)^3$.

Example 4.2. We repeat the last example but now shift the true solution down so that it crosses the t-axis in two places; in this case we use as our true solution $\bar{u}(t) = 8(t - 0.4)^2 - 0.5$. For this $\bar{u}$, hypothesis (S3) on the nonlinearity $S(\cdot, \cdot)$ is violated. We illustrate in figure 2 the difference that results (noting in particular the instability that begins to appear when the curve crosses the t-axis, as compared to the first graph in figure 1) when the same algorithm as that used in the last example is applied to the noise-free data generated using the new $\bar{u}$ as the true solution.

Acknowledgments

This work was supported in part by the National Science Foundation under contract numbers NSF DMS-0104003 and DMS-0405978. Parts of the research were completed while both authors were at the Institute for Pure and Applied Mathematics (IPAM) at UCLA. We wish to thank IPAM for providing the financial support and hospitable environment conducive to productive research.

References

[1] Cinzori A C 2004 Continuous future polynomial regularization of 1-smoothing Volterra problems *Inverse Problems* **20** 1791–806
[2] Cinzori A C and Lamm P K 2000 Future polynomial regularization of ill-posed Volterra equations *SIAM J. Numer. Anal.* **37** 949–79
[3] Dai Z 2005 Local regularization for the autoconvolution problem *PhD Thesis* Michigan State University
[4] Dai Z and Lamm P K Local regularization of the autoconvolution problem *Preprint*
[5] Deimling K 1995 Nonlinear Volterra integral equations of the first kind *Nonlinear Anal. Theory Methods Appl.* **25** 951–7
[6] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)
[7] Gripenberg G, Londen S O and Staffens O 1990 *Volterra Integral and Functional Equations* (Cambridge: Cambridge University Press)
[8] Groetsch C W 1984 *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind* (Boston: Pitman)
[9] Lamm P K 1995 Future-sequential regularization methods for ill-posed Volterra equations: applications to the inverse heat conduction problem *J. Math. Anal. Appl.* **195** 469–94
[10] Lamm P K 1999 Variable-smoothing regularization methods for inverse problems *Theory and Practice of Control and Systems* ed A Tornambè, G Conte and A M Perdon (Singapore: World Scientific)
[11] Lamm P K 2000 A survey of regularization methods for first-kind Volterra equation *Surveys on Solution Methods for Inverse Problems* ed D Colton, H Engl, A Louis, J Mclaughlin and W Rundell (Vienna: Springer) pp 53–82
[12] Lamm P K 2005 Full convergence of sequential local regularization methods for Volterra inverse problems *Inverse Problems* **21** 785–803
[13] Lamm P K and Eldén L 1997 Numerical solution of first-kind Volterra equations by sequential Tikhonov regularization *SIAM J. Numer. Anal.* **34** 1432–50
[14] Lamm P K and Scofield T L 2000 Sequential predictor–corrector methods for the variable regularization of Volterra inverse problems *Inverse Problems* **16** 373–99
[15] Lamm P K and Scofield T L 2001 Local regularization methods for the stabilization of ill-posed Volterra problems *Numer. Funct. Anal. Opt.* **22** 913–40
[16] Ring W and Prix J 2000 Sequential predictor–corrector regularization methods and their limitations *Inverse Problems* **16** 619–34