JOURNAL OF INTEGRAL EQUATIONS
AND APPLICATIONS
Volume 25, Number 2, Summer 2013

SOLVING VOLTERRA INTEGRAL EQUATIONS OF THE SECOND KIND BY SIGMOIDAL FUNCTIONS APPROXIMATION

DANILO COSTARELLI AND RENATO SPIGLER

Communicated by Hermann Brunner

ABSTRACT. In this paper, a numerical collocation method is developed for solving *linear* and *nonlinear* Volterra integral equations of the second kind. The method is based on the approximation of the (exact) solution by a superposition of *sigmoidal* functions and allows one to solve a large class of integral equations having either continuous or L^p solutions. Special computational advantages are obtained using unit step functions, and analytical approximations of the solution are also at hand. The numerical errors are discussed, and *a priori* as well as *a posteriori* estimates are derived for them. Numerical examples are given for the purpose of illustration.

1. Introduction. A collocation solution to a Volterra integral equation on an interval $[a, b]$, is an element from some finite-dimensional function space (the collocation space), which satisfies the equation on an appropriate finite subset of points in $[a, b]$. The latter is the set of collocation points, whose cardinality matches the dimension of the collocation spaces. In this paper, we introduce a new collocation method for solving linear as well as nonlinear Volterra integral equations of the second kind, of the form

$$\text{(I)} \qquad y(t) = f(t) + \int_a^t K(t, s)\, y(t)\, ds, \quad t \in [a, b],$$

and

$$\text{(II)} \qquad y(t) = f(t) + \int_a^t K(t, s; y(s))\, ds, \quad t \in [a, b],$$

2010 AMS *Mathematics subject classification.* 65R20, 45D05, 45G10.
Keywords and phrases. Collocation methods, sigmoidal functions, unit step functions, linear Volterra integral equations, nonlinear Volterra integral equations.
This work was supported, in part, by the GNAMPA and the GNFM of the Italian INdAM.
Received by the editors on January 24, 2012, and in revised form on June 19, 2012.
DOI:10.1216/JIE-2013-25-2-193

where $f : [a, b] \to \mathbf{R}$ and the kernels K are sufficiently smooth. Our method is based on approximating the exact solutions by a superposition of "sigmoidal functions," i.e., functions $\sigma : \mathbf{R} \to \mathbf{R}$ such that $\lim_{x\to-\infty} \sigma(x) = 0$ and $\lim_{x\to+\infty} \sigma(x) = 1$. In particular, we shall use unit step functions (Heaviside functions).

Every either continuous or L^p real-valued function on $[a, b]$ can be approximated uniformly or in L^p-norm by a finite superposition of bounded sigmoidal functions, like

$$\text{(III)} \qquad \sum_{k=0}^{N} \alpha_k \sigma(w(t - t_k)), \quad t \in \mathbf{R},$$

where σ is a sigmoidal function, $\alpha_k \in \mathbf{R}$, and the t_k's, $k = 0, 1, \ldots, N$, $N \in \mathbf{N}^+$, are suitable real points, and $w > 0$ is a suitable scaling parameter [**9, 10, 12**]. The sums above represent a kind of univariate Neural Networks, largely used in a number of engineering problems and in approximation theory as approximants. Many authors, such as Cybenko, Mhaskar and Micchelli have studied the approximation properties of sums like those in (III), and basic results were proved in [**1, 12, 13, 18, 24**]. In particular, in [**12**], the coefficients α_k, the abscissae t_k and the parameter w in (III) were obtained explicitly, in connection to a given either continuous or L^p function to be approximated.

On the other hand, collocation methods have been widely used to solve integral equations like those in (I) and (II). The most popular of these methods are based on piecewise polynomial collocation spaces (see [**3, 7**], e.g.). Other methods are based on wavelets or on Bernstein polynomial approximation (see, e.g., [**21, 26**]).

Here, the choice of the collocation spaces generated by the unit step functions allows one to solve a large class of integral equations having either continuous or L^p solutions, with some computational advantages. In fact, in the case of equation (I), the method reduces merely to solve a linear *lower triangular* algebraic system. In the nonlinear case (II), the method instead leads to a nonlinear system that can always be solved explicitly by means of a direct formula, without using any iteration (such as, e.g., Newton's methods). In both cases, an analytical form for the approximate solution can be obtained, and the numerical algorithm is very fast. In the linear case (I), with a *convolution* kernel, $K(t, s) \equiv K(t - s)$, we can show that the square matrices, M_N, of the related linear algebraic system, turn out to be lower triangular *Toeplitz* matrices. Moreover, under suitable conditions on K, some estimates for the condition number of M_N in the infinity norm, can also be derived.

The paper is organized as follows. In Section 2, we review some results concerning the sigmoidal functions approximation. In Sections 3 and 4, we introduce the collocation method based on unit step functions, to solve numerically linear and nonlinear Volterra integral equations, respectively. In Section 5, we discuss the numerical errors, both *a priori* and *a posteriori*, affecting our approach. Finally, in Section 6, a few numerical examples are given to illustrate the performance of our method.

2. Approximation results by superposition of sigmoidal functions. Applications of neural networks involve many areas of research. In particular, neural networks play an important role in approximation theory, where they play the role of approximants. There is much interest for networks having a *sigmoidal* function as activation function. Recall that

Definition 2.1. A function $\sigma : \mathbf{R} \to \mathbf{R}$ is named a "sigmoidal function" whenever

$$\lim_{x\to-\infty} \sigma(x) = 0 \quad \text{and} \quad \lim_{x\to+\infty} \sigma(x) = 1.$$

Sometimes, boundedness, continuity and/or monotonicity are prescribed in addition.

Many authors have studied how a given continuous real-valued function, $f : [a,b] \to \mathbf{R}$, could be uniformly approximated by a superposition of finitely many sigmoidal functions, i.e., by neural networks of the form in (III). In [**12**], the following was proved:

Theorem 2.2. *Let σ be a bounded sigmoidal function, and let $f \in C^0[a,b]$ be fixed. For every $\varepsilon > 0$, there exist $N \in \mathbf{N}^+$ and $w > 0$ (depending on N and σ), such that, if*

$$(1)\quad (G_N f)(t) := \sum_{k=1}^{N} [f(t_k) - f(t_{k-1})]\, \sigma(w(t-t_k)) + f(t_0)\, \sigma(w(t-t_{-1}))$$

for $t \in [a,b]$*,* $h := (b-a)/N$*, and* $t_k := a + kh$*,* $k = -1, 0, 1, \dots, N$*, then*

$$\|G_N f - f\|_\infty := \max_{t\in[a,b]} |(G_N f)(t) - f(t)| < \varepsilon.$$

We stress that *continuity* of σ is *not* required. Theorem 2.2 represents a modified form of a result earlier proved in [**9, 10**]. This was, in turn, a constructive version of a non-constructive theorem due to Cybenko [**13**].

Remark 2.3. (a) The form of the coefficients in (1) is independent of the choice of the sigmoidal function σ. Therefore, one can provide various approximations of f using different sigmoidal functions, keeping the same coefficients. The scaling parameter $w > 0$ in (1) depends on N and σ.

(b) Assuming σ to be continuous, Theorem 2.2 provides a density result in $C^0[a,b]$ for the set of functions of the form in (1), with respect to the uniform norm.

(c) The error made approximating a given function by a superposition of sigmoidal functions was studied in [**4, 8, 14, 15, 19, 20**]. In particular, it was shown that the error made approximating functions of bounded variation or Lipschitz continuous functions with $G_N f$ is of order of $\mathcal{O}(1/N)$ (for N sufficiently large), N being the number of the superposed sigmoidal functions.

In [**12**], the following constructive approximation theorem was also established for $L^p[a,b]$ functions, with $1 \le p < \infty$:

Theorem 2.4. *Let* σ *be a bounded sigmoidal function, and let* $f \in L^p[a,b]$*,* $1 \le p < \infty$*, be fixed. Set* $f_n := \rho_n * \widetilde{f}$*, where* $(\rho_n)_{n\in\mathbf{N}^+}$ *is a fixed sequence of mollifiers,* $*$ *is the usual convolution product, and*

$$\widetilde{f}(t) := \begin{cases} f(t) & t \in [a,b], \\ 0 & \textit{otherwise}. \end{cases} \tag{2}$$

Then, for every $\varepsilon > 0$ *there exist* N*,* $n \in \mathbf{N}^+$*, and* $w > 0$ *(depending on* N *and* σ*), such that*

$$\|G_N f_n - f\|_{L^p[a,b]} < \varepsilon,$$

where $G_N f_n$ *is defined in* (1)*.*

By choosing a specific sequence of mollifiers, and a sufficiently large $n \in \mathbf{N}^+$, one can obtain the analytic form of the coefficients of $G_N f_n$ of Theorem 2.4, for every $N \in \mathbf{N}^+$. Moreover, the same observation in Remark 2.3 (a) can be made for the case of $L^p[a,b]$ functions, $1 \le p < \infty$.

A useful example of sigmoidal function is given by the *logistic* function, defined as $\sigma(t) := (1+e^{-t})^{-1}$, $t \in \mathbf{R}$. Logistic functions are largely used in many fields, such as biology, physics, biomathematics, statistics, economics and demography (see, e.g., [**6, 16**]). Clearly, σ is bounded, with $0 < \sigma(t) < 1$, for all $t \in \mathbf{R}$. Using logistic functions and Theorem 2.2, we obtain the following

Corollary 2.5 (see [**12**]). *Let $\sigma(t) := (1+e^{-t})^{-1}$, $t \in \mathbf{R}$. For any given $f \in C^0[a,b]$, and for every $\varepsilon > 0$, there exists an $N \in \mathbf{N}^+$ such that*

$$\|G_N f - f\|_\infty < \varepsilon,$$

for every $G_N f$ defined as in (1) *with $w > [N/(b-a)] \ln(N-1)$.*

Corollary 2.5 provides an estimate for $w > 0$, for every $N \in \mathbf{N}^+$, in the case of uniformly approximations of continuous functions by superposition of logistic functions.

Other interesting (and useful) examples of sigmoidal functions are given by the class of Gompertz functions, defined by

$$\sigma_{\alpha\beta}(t) := e^{-\alpha e^{-\beta t}}, \quad t \in \mathbf{R},$$

where α and $\beta > 0$ represent an effective translation and a scaling, respectively. Gompertz functions find applications, e.g., in modeling tumor growth [**2, 11, 23**]. For every pair α and $\beta > 0$, we have $0 < \sigma_{\alpha\beta}(t) < 1$, $t \in \mathbf{R}$, and, similarly to the case of logistic functions, we have

Corollary 2.6 (see [**12**]). *Let $\sigma_{\alpha\beta}(t) := e^{-\alpha e^{-\beta t}}$, $t \in \mathbf{R}$, α, $\beta > 0$. For every given $f \in C^0[a,b]$, and for every $\varepsilon > 0$, there exists an $N \in \mathbf{N}^+$ such that*

$$\|G_N f - f\|_\infty < \varepsilon,$$

for every $G_N f$ *defined in* (1) *with*

$$w > \frac{N}{(b-a)\beta} \max\left\{ \left| \ln\left(-\frac{1}{\alpha} \ln\left(\frac{N-1}{N} \right)\right)\right|, \ \left| \ln\left(\frac{1}{\alpha} \ln\left(N\right)\right)\right| \right\}.$$

As in Corollary 2.5, Corollary 2.6 provides an estimate for the parameter $w > 0$, when Gompertz sigmoidal functions are used.

Remark 2.7. Note that the estimates for $w > 0$ provided by Corollaries 2.5 and 2.6 for the special cases of logistic and Gompertz functions, also hold when $L^p[a, b]$ functions are being approximated.

Finally, we consider the special case of unit step (or Heaviside) functions, $H(t) := 1$ for $t \geq 0$, and $H(t) := 0$ for $t < 0$. In this case, the results established in Theorem 2.2 and in Theorem 2.4 hold true. In the case of Theorem 2.2, sums of the form (1) reduce to

$$(3) \qquad (G_N f)(t) := \sum_{k=1}^{N} [f(t_k) - f(t_{k-1})] H(t - t_k) + f(t_0) H(t - t_{-1}),$$

$t \in \mathbf{R}$, where $f \in C^0[a, b]$, $h := (b - a)/N$, and $t_k := a + hk$, $k = -1, 0, 1 \dots, N$, [**12**]. Note that in (3) $G_N f$ becomes independent of the scaling parameter, $w > 0$, and the same happens in Theorem 2.4 when applied to the case of Heaviside functions.

Remark 2.8. Set $H_k(t) := H(t - t_k)$, with $H_k : [a, b] \to \mathbf{R}$, $t_k := a + hk$, $h := (b - a)/N$, for $k = -1, 1, \dots, N$, and

$$\Sigma_N := \text{span}\left\{ H_k : k = -1, 1, 2, \dots, N \right\}.$$

Then, the vector function space Σ_N is an $N + 1$ dimensional space, and the set $\{H_k : \ k = -1, 1, 2, \dots, N\}$ is a basis for Σ_N. Indeed, it can be proved that the functions H_k's are linearly independent: if $\sum_{k=1}^{N} \alpha_k H_k + \alpha_0 H_{-1} \equiv 0$, i.e.,

$$\sum_{k=1}^{N} \alpha_k H(t - t_k) + \alpha_0 H(t - t_{-1}) = 0,$$

for every $t \in [a, b]$, we have, in particular,

$$\sum_{k=1}^{N} \alpha_k H(t_i - t_k) + \alpha_0 H(t_i - t_{-1}) = 0,$$

for every $i = 0, 1, \dots, N$. Then, for $i = 0$, we have $\alpha_0 = 0$, and for $i > 0$, we obtain $\sum_{k=i}^{N} \alpha_k + \alpha_0 = 0$; hence, necessarily, $\alpha_0 = \alpha_1 = \cdots = \alpha_N = 0$.

3. Collocation method for linear equations based on unit step functions. In this section, we describe a collocation method for solving linear Volterra integral equations of the second kind, of the form

$$y(t) = f(t) + \int_a^t K(t, s)\, y(s)\, ds, \quad t \in [a, b], \tag{4}$$

$a, b \in \mathbf{R}$, where the function $f : [a, b] \to \mathbf{R}$ and the kernel $K : D \to \mathbf{R}$, $D := \{(t, s) :\ a \le t, s \le b\}$, are sufficiently smooth.

Our method, based on unit step functions, consists of determining approximate solutions to equation (4), of the form $G_N y$ as defined in (3), i.e., $G_N y$ belonging to the collocation space Σ_N, $N \in N^+$. Set

$$(G_N y)(t) = \sum_{k=1}^{N} y_k H(t - t_k) + y_0 H(t - t_{-1}), \quad t \in [a, b], \tag{5}$$

where the coefficients $y_0, \dots, y_N$, $N \in \mathbf{N}^+$, in (5) are unknowns, and $t_k := a + hk$, $k = -1, 0, \dots, N$, $h := (b - a)/N$. Inserting $G_N y$ in place of the exact solution y in (4), we obtain

$$(G_N y)(t) = f(t) + \int_a^t K(t, s)\, (G_N y)(s)\, ds, \quad t \in [a, b],$$

and, rearranging all terms, we have

$$\sum_{k=1}^{N} y_k \left[H(t - t_k) - \int_a^t K(t, s)\, H(s - t_k)\, ds \right] + y_0 \left[H(t - t_{-1}) - \int_a^t K(t, s)\, H(s - t_{-1})\, ds \right] = f(t),$$

for every $t \in [a,b]$. If $\mathcal{C}_N := \{t_0, t_1, \dots, t_N\}$ is the set of the *collocation points*, we can evaluate the equation above at such points. Set

$$m_{i0} := H(t_i - t_{-1}) - \int_a^{t_i} K(t_i, s)\, H(s - t_{-1})\, ds,$$
$$m_{ik} := H(t_i - t_k) - \int_a^{t_i} K(t_i, s)\, H(s - t_k)\, ds,$$

for $t_i \in \mathcal{C}_N$, $i = 0, \dots, N$, and $k = 1, 2, \dots, N$. We obtain the following linear algebraic system of $N+1$ equations,

$$\sum_{k=0}^{N} m_{ik}\, y_k = f(t_i), \tag{6}$$

for $i = 0, 1, \dots, N$. Now, setting $M_N := (m_{ik})_{i,k=0,1,\dots,N}$, $Y_N := (y_0, y_1, \dots, y_N)^t$ and $F_N := (f(x_0), f(x_1), \dots, f(x_N))^t$, the linear system (6) can be written as $M_N Y_N = F_N$, $N \in \mathbf{N}^+$. Solving (6), we can determine $y_0, y_1, \dots, y_N$, the coefficients providing an analytical representation of the solution, $y(t)$, of (4), as a superposition of unit step functions as in (5).

Remark 3.1. By Theorems 2.2 and 2.4, the collocation method based on unit step functions can be applied to *linear* Volterra integral equations with either regular solutions on $[a,b]$, or solutions in $L^p[a,b]$, $1 \leq p < \infty$, such as equations with singular kernels.

We can now prove the following

Theorem 3.2. *The collocation method for solving* (4)*, based on Heaviside functions, admits a unique solution. Moreover, the square matrix M_N of the linear system associated with the method is lower triangular, for every $N \in \mathbf{N}^+$.*

Proof. Let $i = 0, 1, \dots, N$ be fixed. We have $H(t_i - t_k) = 0$ for every $k > i$, and $H(t_i - t_k) = 1$ for $k \leq i$. Besides,

$$\int_a^{t_i} K(t_i, s)\, H(s - t_k)\, ds = 0, \tag{7}$$

for $k > i$, since $H(\cdot - t_k) = 0$ on $[a, t_i]$ (equation (7) also holds for $i = 0$, i.e., for $t_0 = a$). Furthermore,

$$\int_a^{t_i} K(t_i, s)\, H(s - t_{-1})\, ds = \int_a^{t_i} K(t_i, s)\, ds,$$

and

$$\int_a^{t_i} K(t_i, s)\, H(s - t_k)\, ds = \int_{t_k}^{t_i} K(t_i, s)\, ds,$$

for $k \le i$, with $k \ne 0$, since $H(\cdot - t_k) = 0$ on $[a, t_k)$ and $H(\cdot - t_k) = 1$ on $[t_k, t_i]$. Hence, we obtain

$$m_{ik} := \begin{cases} 0 & \text{for } k > i, \\ 1 & \text{for } k = i, \\ 1 - \int_{t_k}^{t_i} K(t_i, s)\, ds & \text{for } k < i, \end{cases} \tag{8}$$

for $i, k = 0, 1, \ldots, N$. Then, the $(N+1) \times (N+1)$ matrix M_N of our method is lower triangular, for every $N \in \mathbf{N}^+$, i.e.,

$$M_N := \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{10} & 1 & 0 & \cdots & 0 \\ m_{20} & m_{21} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ m_{N0} & m_{N1} & \cdots & m_{NN-1} & 1 \end{bmatrix} \tag{9}$$

Hence, $\det(M_N) = 1$, and the linear system $M_N Y_N = F_N$ admits a unique solution, for every $N \in \mathbf{N}^+$. □

Note that, in Theorem 3.2, an integrability assumption on the kernel K of the integral equation in (4) is needed. The entries of M_N of the form $1 - \int_{t_k}^{t_i} K(t_i, s)\, ds$ can be evaluated by exact (analytical) integration, in many instances, or, more generally, upon numerical quadratures.

The method based on unit step functions can be implemented easily, and, in addition, it is definitely characterized by an extremely low computational cost.

In the special (but noteworthy) case of integral equations of the *convolution* type, the linear Volterra integral equations of the second kind,

$$y(t) = f(t) + \int_a^t K(t-s)\, y(s)\, ds, \quad t \in [a,b], \tag{10}$$

we have the following

Corollary 3.3. *The collocation method for solving* (10)*, based on unit step functions, admits a unique solution. Moreover, the real-valued matrix* M_N *is a* lower triangular Toeplitz *matrix, for every* $N \in \mathbf{N}^+$.

Proof. By Theorem 3.2, M_N is lower triangular with $\det(M_N) = 1$, for every $N \in \mathbf{N}^+$; then the method admits a unique solution. Moreover, if $h := (b-a)/N$ is the step-size separating the collocation points t_i, $i = 0, 1, \ldots, N$, we obtain the changing variable, $s = z - h$,

$$\begin{aligned} \int_{t_k}^{t_i} K(t_i - s)\, ds &= \int_{t_k+h}^{t_i+h} K(t_i + h - z)\, dz \\ &= \int_{t_{k+1}}^{t_{i+1}} K(t_{i+1} - z)\, dz. \end{aligned}$$

Then, $m_{i,k} = m_{i+1,k+1}$ for every $k < i$ (see (8)), and thus, M_N is seen to be constant along all its diagonals. This means that M_N is a Toeplitz matrix, and it can be represented as

$$M_N := \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_1 & 1 & 0 & \ddots & 0 \\ m_2 & m_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ m_N & \cdots & m_2 & m_1 & 1 \end{bmatrix}, \tag{11}$$

where

$$m_i := 1 - \int_a^{t_i} K(t_i - s)\, dt,$$

for every $i = 1, 2, \ldots, N$. $\square$

In the case of equations with convolution kernel, the required computational cost is much lower than in the case of general kernels. In fact, by Corollary 3.3, for every $N \in \mathbf{N}^+$, the linear system to be solved is characterized by lower triangular Toeplitz matrices, and thus to compute all entries of M_N, it suffices to evaluate only the N terms $m_1, \dots, m_N$.

Remark 3.4. The approach proposed above can also be useful to determine approximate solutions to initial value problems (IVP) for ordinary differential equations. In fact, generally speaking, every linear IVP, e.g., of the second order, say

$$y'' + A(t)y' + B(t)y = g(t), \quad y(a) = c_1,\ y'(a) = c_2,$$

where A, B and g are sufficiently smooth functions, is equivalent to a linear Volterra integral equation of the second kind (see, e.g., [**17**]), like that in (4), where

$$f(t) := \int_a^t (t-s)\, g(s)\, ds + (t-a)[c_1 A(a) + c_2] + c_1,$$

and

$$K(t,s) := (s-t)[B(s) - A'(s)] - A(s).$$

At this point, we introduce some notation. Given the kernel K, we define

$$\mathcal{K}(t) := \int_a^t K(t,s)\, ds, \quad t \in [a,b]. \tag{12}$$

Under suitable conditions on $\mathcal{K}$, we can obtain some estimates for $\kappa(M_N)$, the *condition number* of M_N in the infinity norm, in case of integral equations of the convolution type, (10). Recall that, for every nonsingular real-valued matrix $A := (a_{i,j})_{i,j=0,1,\dots,N}$,

$$\kappa(A) := \|A\|_\infty \|A^{-1}\|_\infty,$$

where $\|A\|_\infty := \max_{i=0,1,\dots,N} \sum_{j=0}^N |a_{i,j}|$.

We can establish the following

Theorem 3.5. *Let* (10) *be a Volterra integral equation with convolution kernel, K. Let $\mathcal{K}$ be defined in* (12)*, and such that*

(i) $0 \leq \mathcal{K}(t) \leq 1$*, for every* $t \in [a,b]$;

(ii) $\mathcal{K}$ *is non decreasing.*

Then, the square matrix M_N, $N \in \mathbf{N}^+$, obtained applying our collocation method to (10)*, enjoys the property*

$$\kappa(M_N) \leq \begin{cases} \frac{2}{1-\mathcal{K}(b)} \left(1-\mathcal{K}(b)^{[N/2]+1}\right)(N(1-\mathcal{K}(t_1))+1) & \mathcal{K}(b)<1, \\ 2\left(\left[\frac{N}{2}\right]+1\right)(N(1-\mathcal{K}(t_1))+1) & \mathcal{K}(b)=1, \end{cases}$$

where $[\ ]$ *means taking the integer part, for every* $N \in \mathbf{N}^+$*. In particular, if* $\mathcal{K}(b) < 1$*, we have*

$$\kappa(M_N) < \frac{2}{1-\mathcal{K}(b)}\left(N(1-\mathcal{K}(t_1))+1\right).$$

Proof. By Corollary 3.3, M_N is a lower triangular Toeplitz matrix of the form in (11). By (i) and (ii), we have $0 \leq \mathcal{K}(t_0) \leq \mathcal{K}(t_1) \leq \cdots \leq \mathcal{K}(t_N) \leq 1$, where t_i, $i = 0,1,\ldots,N$, are the collocation points, and then, the elements of M_N satisfy the inequalities

$$1 \geq m_1 \geq m_2 \geq \cdots \geq m_N \geq 0.$$

Hence, by well-known results concerning lower triangular Toeplitz matrices with non-increasing monotonic entries ([**5**, Theorem 1.1] and [**25**]), we obtain the bounds

$$\|M_N^{-1}\|_\infty \leq \begin{cases} \frac{2}{1-\mathcal{K}(b)}\left(1-\mathcal{K}(b)^{[N/2]+1}\right) & \mathcal{K}(b)<1, \\ 2\left(\left[\frac{N}{2}\right]+1\right) & \mathcal{K}(b)=1. \end{cases}$$

Now, since $\|M_N\|_\infty \leq N(1-\mathcal{K}(t_1))+1$, and $\kappa(M_N) := \|M_N\|_\infty \|M_N^{-1}\|_\infty$, the proof of the first part of the theorem is complete. Moreover, if we note that in the case of $\mathcal{K}(b) < 1$, there is $(1-\mathcal{K}(b)^{[N/2]+1}) < 1$, we have

$$\kappa(M_N) < \frac{2}{1-\mathcal{K}(b)}\left(N(1-\mathcal{K}(t_1))+1\right). \qquad \square$$

4. Nonlinear Volterra integral equations. In this section, we use the collocation method based on unit step functions to solve general nonlinear Volterra integral equations of the second kind, like

$$y(t) = f(t) + \int_a^t K(t,s;y(s))\,ds, \quad t \in [a,b], \tag{13}$$

$a, b \in \mathbf{R}$, where the function $f : [a,b] \to \mathbf{R}$ and the kernel $K : \Omega \to \mathbf{R}$, $\Omega := D \times \mathbf{R}$, are sufficiently smooth. Proceeding as in Section 3, we can approximate the solution to (13) on the interval $[a,b]$ by means of (5), i.e.,

$$(G_N y)(t) := \sum_{k=1}^{N} y_k H(t-t_k) + y_0 H(t-t_{-1}),$$

where H is the Heaviside function, $t_k := a + hk$ with $h := (b-a)/N$, $k = -1, 0, 1, \ldots, N$, and the coefficients $y_0, y_1, \ldots, y_N$ are unknown. Inserting $G_N y$ for y in (13), we obtain

$$(G_N y)(t) = f(t) + \int_a^t K(t,s;(G_N y)(s))\,ds, \quad t \in [a,b].$$

Again, to obtain an approximate solution to (13) in the form of a superposition of unit step functions, we should determine the unknown coefficients $y_0, y_1, \ldots, y_N$. Given the set $\mathcal{C}_N := \{t_0, t_1, \ldots, t_N\}$ of collocation points, we evaluate the equation at such points as in the linear case. Now we obtain the system of $N+1$ nonlinear equations

$$\begin{aligned} &\sum_{k=1}^{N} y_k H(t_i - t_k) + y_0 H(t_i - t_{-1}) \\ &= f(t_i) + \int_a^{t_i} K\left(t_i, s; \sum_{k=1}^{N} y_k H(s-t_k) + y_0 H(s-t_{-1})\right) ds, \\ &\qquad\qquad i = 0, 1, \ldots, N. \end{aligned} \tag{14}$$

It is easy to check that, for $i = 0$, (14) reduces to $y_0 = f(t_0)$. Now let $i > 0$ be fixed. In this case, $H(t_i - t_k) = 1$ for every $k = 0, \ldots, i$ and

$H(t_i - t_k) = 0$ for $k > i$. Moreover,

$$\int_a^{t_i} K\left(t_i, s; \sum_{k=1}^{N} y_k H(s-t_k) + y_0 H(s - t_{-1})\right) dt$$
$$= \sum_{\nu=1}^{i} \int_{t_{\nu-1}}^{t_\nu} K\left(t_i, s; \sum_{k=1}^{N} y_k H(s-t_k) + y_0\right) ds$$
$$= \sum_{\nu=1}^{i} \int_{t_{\nu-1}}^{t_\nu} K\left(t_i, s; \sum_{k=0}^{\nu-1} y_k\right) ds,$$

since, for every $\nu = 1, \dots, i$, there is an $H(\cdot - t_k) = 1$ on $[t_{\nu-1}, t_\nu]$ for $k = 0, \dots, \nu - 1$ and $H(\cdot - t_k) = 0$ on $[t_{\nu-1}, t_\nu]$ for $k \geq \nu$. Therefore, (14) reduces to the nonlinear system

$$y_0 = f(t_0),$$
$$\sum_{k=0}^{i} y_k = f(t_i) + \left[\sum_{\nu=1}^{i} \int_{t_{\nu-1}}^{t_\nu} K\left(t_i, s; \sum_{k=0}^{\nu-1} y_k\right) ds\right], \quad i = 1, \dots, N,$$

which admits a unique solution that can be given by the formula

(15)
$$\begin{cases} y_0 = f(t_0), \\ y_i = f(t_i) + \sum_{\nu=1}^{i} \left[\int_{t_{\nu-1}}^{t_\nu} K(t_i, s; \sum_{k=0}^{\nu-1} y_k)\, ds\right] \\ \qquad - \sum_{k=0}^{i-1} y_k \qquad\qquad i = 1, \dots, N, \end{cases}$$

for every $N \in \mathbf{N}^+$.

We stress that the nonlinear system in (14) can be solved explicitly, and its solution does not required any iterative method (such as Newton's method, e.g.). Note also that (15) provides an algorithm for solving a large class of nonlinear Volterra integral equations. Only an integrability assumption on the kernel $K(t, \cdot; y)$ on $[a, t]$ for every $y \in \mathbf{R}$ and $t \in [a, b]$ is required.

In the special case of Volterra-Hammerstein integral equations of the second kind, i.e., when the kernel is of the form $K(t, s; y(s)) :=$

$\widetilde{K}(t,s)G(y(s))$, where G and $\widetilde{K}$ are sufficiently smooth functions, (15) reduces to

$$\begin{cases} y_0 = f(t_0), \\ y_i = f(t_i) + \sum_{\nu=1}^{i} \left[G\left(\sum_{k=0}^{\nu-1} y_k \right) \int_{t_{\nu-1}}^{t_\nu} \widetilde{K}(t_i,s)\,ds \right] \\ \qquad - \sum_{k=0}^{i-1} y_k \qquad\qquad\qquad\qquad i = 1, \dots, N. \end{cases}$$

Also, in this case, the integrals $\int_{t_{\nu-1}}^{t_\nu} \widetilde{K}(t_i,s)\,ds$ can be evaluated by an exact (analytical) integration, in many specific instances, or, more generally by numerical quadrature. As in the case of linear integral equations, our collocation method can be applied to nonlinear Volterra integral equations having either regular or $L^p[a,b]$ solutions (with $1 \le p < \infty$), in view of the results on approximation through superposition of sigmoidal functions discussed in Section 2.

5. Error analysis. In this section, we discuss the various sources of errors which affect our method. Our numerical method is based on using for the sought solution y its approximate representation in terms of bounded sigmoidal functions, (see $G_N f$ in (1)). In view of Theorem 2.2 (and also Theorem 2.4), we can write

$$y(t) = (G_N y)(t) + e_N(t), \tag{16}$$

provided that $y(t)$ is continuous (or in $L^p[a,b]$), where the error term, $e_N(t)$, can be estimated uniformly (or in L^p-norm), for every $\varepsilon > 0$, as

$$\|e_N(t)\|_\infty < \varepsilon, \tag{17}$$

for a suitable $N \in \mathbf{N}^+$ (and $w > 0$, depending on N).

Clearly, (17) holds when $G_N y$ is written with the coefficients computed in Section 2. In the following theorem, we establish an estimate for e_N, defined in (16), when $G_N y$ is represented in terms of unit step functions (recall that in this case $G_N y$ is independent of $w > 0$), and with the coefficients y_k determined applying our collocation method to the nonlinear integral equation in (13).

Theorem 5.1. *Let* (13) *be a given nonlinear Volterra integral equation of the second kind. Suppose that the function f is Lipschitz* *continuous, with Lipschitz constant* $L_f > 0$*, and that the kernel* $K \in C^0(\Omega)$ *satisfies the following conditions*:

(i) *there exist* $L_1 > 0$ *and* $L_2 > 0$ *such that*

$$|K(t_1, s, y) - K(t_2, s, y)| \leq L_1 |t_1 - t_2|, \quad \text{for all } (t_1, s, y),\ (t_2, s, y) \in \Omega,$$
$$|K(t, s, y_1) - K(t, s, y_2)| \leq L_2 |y_1 - y_2|, \quad \text{for all } (t, s, y_1),\ (t, s, y_2) \in \Omega;$$

(ii) *for every bounded function* $y : [a, b] \to \mathbf{R}$*, there exists a* $C = C(y) > 0$ *such that*

$$|K(t, s, y(s))| \leq C, \quad \textit{for all } (t, s) \in D.$$

Then, for every $N \in \mathbf{N}^+$*, we have*

$$|e_N(t)| \leq \frac{(b-a)}{N} [L_f + L_1(b-a) + C]\, e^{L_2(t-a)}, \quad t \in [a, b],$$

where e_N *is defined in* (16)*,* $G_N y$ *being represented in terms of unit step functions and with coefficients,* y_k*, determined applying our collocation method to* (13)*.*

Proof. Let $N \in \mathbf{N}^+$ and $t \in [a, b]$ be fixed. Define $j := \max\{i : t_i \leq t, i = 0, 1, \ldots, N\}$, where $t_i := a + ih$, $h := (b-a)/N$, $i = 0, 1, \ldots, N$ are the collocation points. We can write

$$|e_N(t)| = |y(t) - (G_N y)(t)| \leq |y(t) - y(t_j)| + |y(t_j) - (G_N y)(t)|.$$

Now, observing that $(G_N y)(t) = (G_N y)(t_j)$ (since $G_N y$ is written in terms of unit step functions), we obtain

$$\begin{aligned}
|e_N(t)| &\leq |y(t) - y(t_j)| + |y(t_j) - (G_N y)(t_j)| \\
&= \left| f(t) + \int_a^t K(t, s, y(s))\, ds - f(t_j) - \int_a^{t_j} K(t_j, s, y(s))\, ds \right| \\
&\quad + \left| \int_a^{t_j} K(t_j, s, y(s))\, ds - \int_a^{t_j} K(t_j, s, (G_N y)(s))\, ds \right| \\
&\leq |f(t) - f(t_j)| + \int_a^{t_j} |K(t, s, y(s)) - K(t_j, s, y(s))|\, ds \\
&\quad + \int_{t_j}^t |K(t, s, y(s))|\, ds \\
&\quad + \int_a^{t_j} |K(t_j, s, y(s)) - K(t_j, s, (G_N y)(s))|\, ds.
\end{aligned}$$

Using conditions (i), (ii) and the Lipschitz continuity of f, we have

$$\begin{aligned}|e_N(t)| &\le L_f(t-t_j) + L_1(t-t_j)(t_j-a) + C(t-t_j)\\ &\quad + L_2\int_a^{t_j} |y(s)-(G_N y)(s)|\,ds\\ &\le (t-t_j)[L_f + L_1(b-a) + C] + L_2\int_a^{t_j} |e_N(s)|\,ds\\ &\le \frac{(b-a)}{N}[L_f + L_1(b-a) + C] + L_2\int_a^{t} |e_N(s)|\,ds.\end{aligned}$$

The inequality above holds for every $t \in [a,b]$, and then, by Gronwall's lemma, we obtain

$$|e_N(t)| \le \frac{(b-a)}{N}(L_f + L_1(b-a) + C)\Big[1 + L_2\int_a^t e^{L_2(t-s)}\,ds\Big],$$

for every $t \in [a,b]$. Since $L_2\int_a^t e^{L_2(t-s)}\,ds = e^{L_2(t-a)} - 1$, it follows that

$$|e_N(t)| \le \frac{(b-a)}{N}(L_f + L_1(b-a) + C)\,e^{L_2(t-a)}, \quad t \in [a,b]. \qquad \square$$

Theorem 5.1 provides an *a priori* estimate for the approximation errors of our collocation method *applied to* the nonlinear equations in (13). In addition, we can infer from Theorem 5.1 that $\|e_N\|_\infty \to 0$ as $N \to +\infty$, and then that the sequence approximations for the solution to (13), as determined by our method, converges uniformly to the (exact) solution y.

Remark 5.2. In Theorem 5.1, the Lipschitz condition on the kernel K, with respect to y, is global. Hence, Theorem 5.1 *cannot* cover, e.g., the case of nonlinear equations like that in (13) with kernels of the form $K(t,s;y) = \widetilde{K}(t,s)\,y^p$, with $p > 1$.

Clearly, in the special case $K(t,s,y(s)) = \widetilde{K}(t,s)\,y(s)$, equation (13) reduces to the linear equation (4) with kernel $\widetilde{K}$. Therefore, if f is Lipschitz continuous with Lipschitz constant L_f and $\widetilde{K} \in C^0(D)$ is such that

$$|\widetilde{K}(t_1,s) - \widetilde{K}(t_2,s)| \le L_1|t_1 - t_2|,$$

for all (t_1, s), $(t_2, s) \in D$ and some positive constants L_1, we infer from Theorem 5.1 that

$$|e_N(t)| \leq \frac{(b-a)}{N}(L_f + L_1(b-a) + M\|y\|_\infty)\, e^{M(t-a)}, \quad t \in [a,b], \tag{18}$$

where y is the (exact) continuous solution to (4) with kernel $\widetilde{K}$ and $M := \max_{(t,s)\in D} |\widetilde{K}(t,s)|$, for every $N \in \mathbf{N}^+$. Also, in this case, we obtain that $\|e_N\|_\infty \to 0$ as $N \to +\infty$. Now, since $\|y\|_\infty = \|G_N y + e_N\|_\infty$, (18) becomes

$$\|e_N\|_\infty \leq \frac{(b-a)}{N}(L_f + L_1(b-a) + M\|G_N y\|_\infty + M\|e_N\|_\infty)\, e^{M(b-a)},$$

and we have

$$\|e_N\|_\infty \left(1 - M\frac{(b-a)}{N}e^{M(b-a)}\right) \leq \frac{(b-a)}{N}(L_f + L_1(b-a) + M\|G_N y\|_\infty)\, e^{M(b-a)}.$$

Now, for N sufficiently large, we have $M[(b-a)/N]e^{M(b-a)} < 1$, and hence,

$$\|e_N\|_\infty \leq \frac{(b-a)}{N - M(b-a)\, e^{M(b-a)}}(L_f + L_1(b-a) + M\|G_N y\|_\infty)\, e^{M(b-a)}, \tag{19}$$

which represents an *a posteriori* estimate for the approximation error made in case of the linear equations (4), when $G_N y$ is given in terms of unit step functions.

Now we can use the estimate provided by Theorem 5.1 *to derive* another interesting estimate for $e_N(t)$. This can be obtained when the approximate solution is expressed by a superposition of general bounded sigmoidal functions. We denote with $G_N^H y$ the collocation solution to (13) represented in terms of unit step functions, and with $G_N^\sigma y$ the solution obtained by superposing general bounded sigmoidal functions. Note that the collocation solution can also be represented by $G_N^\sigma y$, in view of Remark 2.3 (a). Setting

$$(G_N^\sigma y)(t) = (G_N^H y)(t) + s_N(t), \quad t \in [a,b],$$

we obtain

$$\begin{aligned}|s_N(t)| &\le |(G_N^\sigma y)(t) - (G_N^H y)(t)| \\ &= \left| \sum_{k=1}^{N} y_k[\sigma(w(t-t_k)) - H(t-t_k)] + y_0[\sigma(w(t-t_{-1})) - 1] \right|.\end{aligned}$$

Observing that

$$\sigma(w(t-t_k)) - H(t-t_k) := \begin{cases} \sigma(w(t-t_k)) - 1 & t \ge t_k \\ \sigma(w(t-t_k)) & t < t_k, \end{cases}$$

for every $k = 1, \dots, N$, we have

$$|s_N(t)| \le \sum_{k: t_k \le t} |y_k||\sigma(w(t-t_k)) - 1| + \sum_{k: t_k > t} |y_k||\sigma(w(t-t_k))|.$$

Therefore, under the conditions of Theorem 5.1, we obtain

$$\begin{aligned}|e_N(t)| = |y(t) - (G_N^\sigma y)(t)| &= |y(t) - (G_N^H y)(t) - s_N(t)| \\ &\le |y(t) - (G_N^H y)(t)| + |s_N(t)| \\ &\le \frac{(b-a)}{N}\,[L_f + L_1(b-a) + C]\; e^{L_2(t-a)} \\ &\quad + \sum_{k: t_k \le t} |y_k||\sigma(w(t-t_k)) - 1| \\ &\quad + \sum_{k: t_k > t} |y_k||\sigma(w(t-t_k))|,\end{aligned}$$

for every $t \in [a, b]$. Now, we know by Definition 2.1 that, for $w > 0$ sufficiently large, the terms $|\sigma(w(t-t_k)) - 1|$ and $|\sigma(w(t-t_k))|$ in (20) are small.

In (20), a bound is given for the approximation error, in terms of the y_k's. These have been computed in our collocation method, and thus this can be viewed as an *a posteriori* estimate. Note that, in general, from (20) we cannot infer that $G_N^\sigma y$ converges to y.

Using (18), considerations similar to those above can be made to obtain an a posteriori bound for e_N in the case of linear equations, where the approximate solution is given by $G_N^\sigma y$.

6. Numerical examples. In this section, we apply the method developed above, in this paper, to solve numerically some linear as well as nonlinear Volterra integral equations.

6.1. Linear equations. Here are some examples of linear Volterra equations of the second kind.

Example 6.1.1. Consider the Volterra equation (4) with

$$K(t,s) = e^{ts}, \qquad f(t) = e^{2t} - \frac{1}{t+2}(e^{t(t+2)} - e^{-(t+2)}),$$

on the interval $[a,b] = [-1,1]$. Its solution is $y(t) = e^{2t}$.

We test our collocation method based on step functions for solving this equation. By Remark 2.3 (a), we know that various approximations of $y(t)$ can be obtained using different sigmoidal functions, using the same coefficients. We denoted by

$$\varepsilon_N^i := \frac{\|G_N^i y - y\|_\infty}{\|y\|_\infty}, \quad i = 1,2,3, \tag{21}$$

where y is the exact solution of the integral equation and $G_N^i y$, $i = 1,2,3$, is its approximation obtained as a superposition of sigmoidal functions, of the Heaviside, logistic, and Gompertz (with $\alpha = 0.85$ and $\beta = 0.1$) type, respectively. The $G_N^i y$'s are all obtained evaluating the coefficients $y_0, y_1, \ldots, y_N$, solution of the linear system $M_N Y_N = F_N$. In the cases of $G_N^2 y$ and $G_N^3 y$, the scaling parameter w was chosen accordingly to Corollaries 2.5 and 2.6, respectively, yielding $w = N^2/(b-a)$ for $G_N^2 y$, and $w = N^2/[(b-a)\alpha\beta]$ for $G_N^3 y$. In Table 1, the relative errors ε_N^i are shown. As for the condition number in the infinity norm of the matrices M_N, say $\kappa(M_N)$, we note that they are not very small, being for instance $\kappa(M_{10}) \approx 39.84$, $\kappa(M_{50}) \approx 125.73$, $\kappa(M_{500}) \approx 816,39$, but its value is not very important, since we stress that our method, used to solve the linear equations, is *not* iterative, but rather a pure *direct elimination* method. In Figures 1 and 2, the approximate solutions $G_N^1 y$, $G_N^2 y$ for $N = 20$ and $N = 60$ are plotted, respectively.

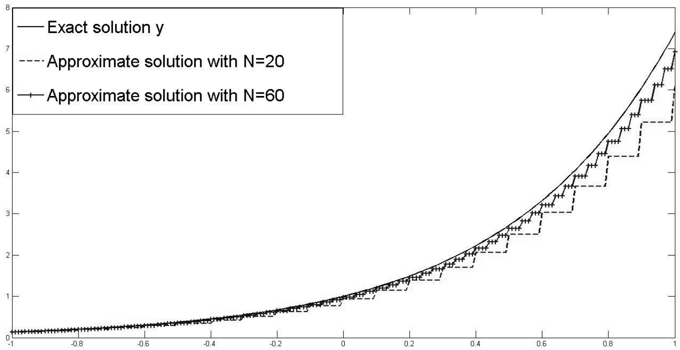

FIGURE 1. Approximate solution $G_N^1 y$ of Example 6.1.1, for $N = 20$ and $N = 60$.

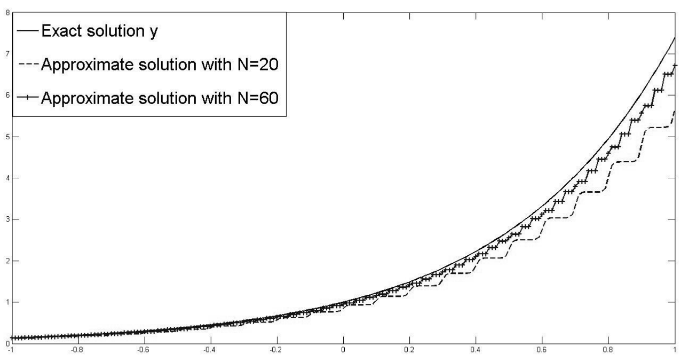

FIGURE 2. Approximate solution $G_N^2 y$ of Example 6.1.1, for $N = 20$ and $N = 60$.

TABLE 1. Numerical results for Example 6.1.1.

N	ε_N^1	ε_N^2	ε_N^3
10	4.45×10^{-1}	3.80×10^{-1}	4.15×10^{-1}
20	2.73×10^{-1}	2.59×10^{-1}	2.73×10^{-1}
30	1.95×10^{-1}	1.94×10^{-1}	1.95×10^{-1}
40	1.50×10^{-1}	1.50×10^{-1}	1.50×10^{-1}
50	1.20×10^{-1}	1.20×10^{-1}	1.20×10^{-1}
60	9.98×10^{-2}	9.98×10^{-2}	9.98×10^{-2}
500	1.16×10^{-2}	1.21×10^{-2}	1.26×10^{-2}
1000	4.10×10^{-3}	6.10×10^{-3}	6.40×10^{-2}

Example 6.1.2. Consider the following initial value problem of the second order,

$$y'' - \frac{t}{2}y' + y = \frac{t}{2}\sin t, \quad y(0) = -1,\ y'(0) = 0.$$

Its solution is $y(t) = t^2 + \cos t - 2$.

To such an IVP, a linear Volterra integral equation like that in (4) can be associated (indeed, it is equivalent to it), with

$$K(t,s) := 2s - \frac{3}{2}t, \qquad f(t) := -\frac{t}{2}\sin t - \cos t,$$

see Remark 3.4. Consider such an integral equation on the interval $[a,b] = [0,1]$. In Table 2, the relative errors ε_N^i, $i = 1,2,3$ are shown, as in Example 6.1.1. The same observation can be made on the condition number of the M_N's. In Figures 3 and 4, the approximate solutions $G_N^1 y$, $G_N^2 y$ for $N = 10$ and $N = 50$ are shown.

Our method can be compared with other classical collocation methods, e.g., those based on piecewise polynomials [**7**]. We have compared the numerical errors made applying both methods to Example 6.1.2, choosing for the latter method quadratic polynomials on the subintervals of $[0,1]$, when the same number of collocation points are used. Taking M subintervals, we need $N = 3M$ collocation points. The

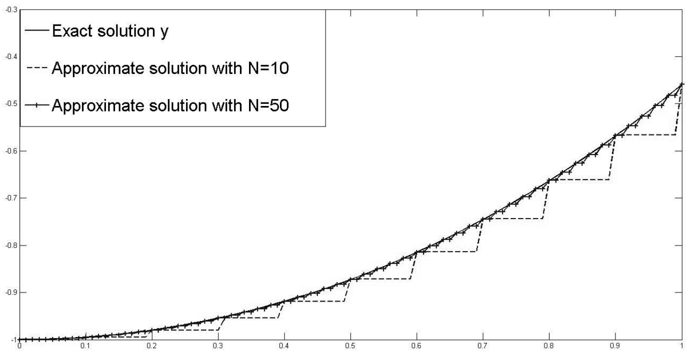

FIGURE 3. Approximate solution $G_N^1 y$ of Example 6.1.2, for $N = 10$ and $N = 50$.

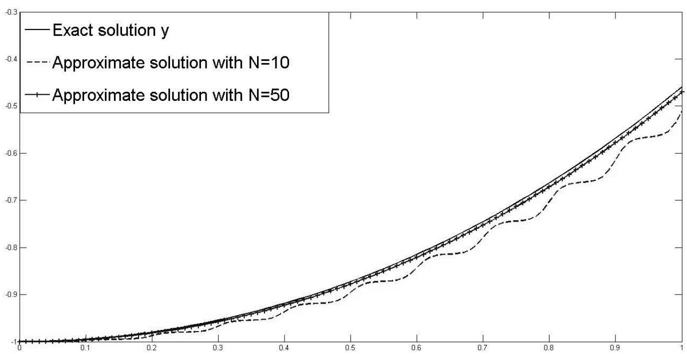

FIGURE 4. Approximate solution $G_N^2 y$ of Example 6.1.2, for $N = 10$ and $N = 50$.

TABLE 2. Numerical results for Example 6.1.2.

N	ε_N^1	ε_N^2	ε_N^3
10	9.52×10^{-2}	7.08×10^{-2}	8.82×10^{-2}
20	4.31×10^{-2}	4.21×10^{-2}	4.31×10^{-2}
30	3.02×10^{-2}	2.86×10^{-2}	3.02×10^{-2}
40	2.12×10^{-2}	2.12×10^{-2}	2.12×10^{-2}
50	1.06×10^{-2}	1.07×10^{-2}	1.24×10^{-2}
500	8.09×10^{-5}	1.10×10^{-3}	1.20×10^{-3}
1000	4.06×10^{-5}	5.38×10^{-4}	6.22×10^{-4}

relative approximation errors, ε_M^p, made with $M = 7$, $M = 10$, and $M = 15$, turn out to be $\varepsilon_7^p = 3.03 \times 10^{-2}$, $\varepsilon_{10}^p = 2.82 \times 10^{-2}$, and $\varepsilon_{15}^p = 2.55 \times 10^{-2}$, respectively. These should be compared with the results shown in Table 2.

Example 6.1.3. Consider the following *singular* Volterra integral equation (of Abel's type), with convolution kernel as in (10), with

$$K(t,s) = -\frac{1}{\sqrt{t-s}}, \qquad f(t) = t^2 + \frac{16}{15}t^{5/2},$$

whose solution is $y(t) = t^2$, see, e.g., [**26**].

We consider this equation on the interval $[a,b] = [0,1]$. The numerical results obtained by our collocation method with unit step functions are described in Table 3. The computed relative errors ε_N^i, $i = 1,2,3$ are those defined in (21).

Example 6.1.4. Finally, consider the singular Volterra integral equation (4) with

$$K(t,s) = -\frac{1}{\sqrt{t-s}}, \qquad f(t) = 1 + 3\sqrt{t} + \frac{\pi}{2}t,$$

on the interval $[a,b] = [0,1]$. Its solution is $y(t) = 1 + \sqrt{t}$, which has an unbounded derivative at $t = 0$.

TABLE 3. Numerical results for Example 6.1.3.

N	ε_N^1	ε_N^2	ε_N^3
10	1.07×10^{-1}	6.47×10^{-2}	9.53×10^{-2}
20	4.61×10^{-2}	4.43×10^{-2}	4.61×10^{-2}
30	3.56×10^{-2}	3.27×10^{-2}	3.56×10^{-2}
40	2.32×10^{-2}	2.32×10^{-2}	2.32×10^{-2}
50	1.28×10^{-2}	7.20×10^{-3}	1.00×10^{-2}
500	1.20×10^{-3}	7.85×10^{-4}	1.10×10^{-3}
1000	6.02×10^{-4}	3.97×10^{-4}	5.42×10^{-4}

TABLE 4. Numerical results for Example 6.1.4.

N	ε_N^1	ε_N^2	ε_N^3
10	1.5×10^{-1}	1.21×10^{-1}	1.41×10^{-1}
20	10^{-1}	9×10^{-2}	10^{-1}
30	8.66×10^{-2}	8.1×10^{-2}	8.66×10^{-2}
40	7.07×10^{-2}	7.07×10^{-2}	7.07×10^{-2}
50	5×10^{-2}	5×10^{-2}	5×10^{-2}
500	8×10^{-4}	1.8×10^{-3}	2.2×10^{-3}
1000	3.78×10^{-4}	9.01×10^{-4}	1.1×10^{-3}

The numerical results for such an example, obtained by our collocation method with unit step functions, are given in Table 4.

From the tables, one can observe that the convergence of the method is rather slow, and its accuracy poor. This is due to the basic approximation result based on sigmoidal functions, see Remark 2.3 (c). One should note, however, that the method can be applied under very weak assumptions on the kernel and the data. A large class of integral equations can then be solved in this way, and analytical representations of the solutions can also be obtained (as a superposition of sigmoidal functions), at a low computational cost. In fact, all coefficients (of $G_N y$) needed to approximate a given solution, are evaluated just solving a lower triangular algebraic system. In the special case of integral equations of the convolution type, such matrices are Toeplitz matrices; hence, only N integrals must be computed. It follows that the methods we propose may actually be very fast, but the so-obtained solution could also be used as a starting point of other methods.

6.2. Nonlinear equations. Here are some examples of nonlinear Volterra equations.

Example 6.2.1. Consider the nonlinear Volterra-Hammerstein equation (13) with

$$\begin{aligned} K(t,s;y(s)) &:= \widetilde{K}(t,s)\,G(y(s)) \\ &:= e^{y(s)}\cos s, \\ f(t) &:= \sin t - e^{\sin t} + 1, \end{aligned}$$

on the interval $[a,b]=[0,1]$. The solution is $y(t)=\sin t$, see, e.g., [**22**].

The numerical results obtained applying our method with unit step functions (introduced in Section 4), are shown in Table 5. As above, we computed the relative errors ε_N^i, $i=1,2,3$, defined in (21). In Figures 5 and 6, the approximate solutions $G_N^1 y$, $G_N^2 y$ are depicted, for $N=30$ and $N=80$.

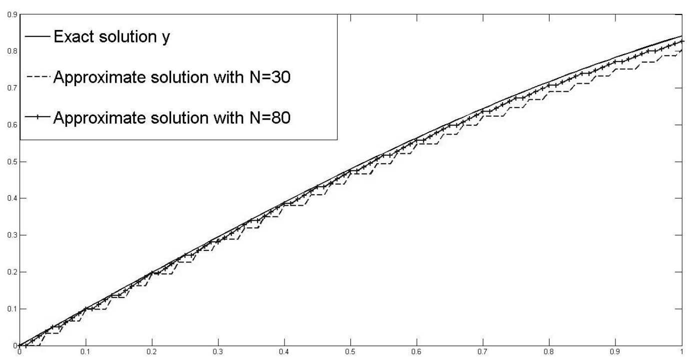

FIGURE 5. Approximate solution $G_N^1 y$ of Example 6.2.1, for $N=30$ and $N=80$.

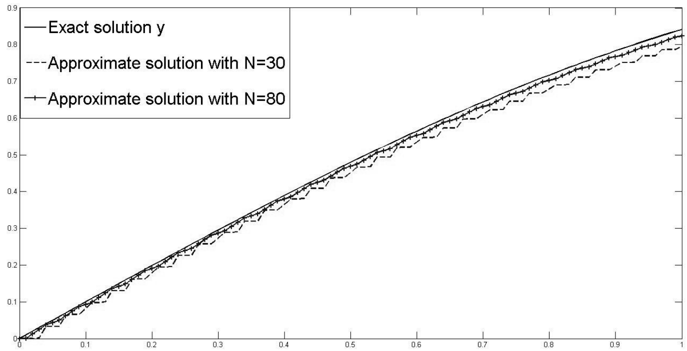

FIGURE 6. Approximate solution $G_N^2 y$ of Example 6.2.1, for $N = 30$ and $N = 80$.

TABLE 5. Numerical results for Example 6.2.1.

N	ε_N^1	ε_N^2	ε_N^3
10	1.66×10^{-1}	1.53×10^{-1}	1.63×10^{-1}
20	8.76×10^{-2}	8.70×10^{-2}	8.76×10^{-2}
30	6.00×10^{-2}	5.90×10^{-2}	6.00×10^{-2}
40	4.53×10^{-2}	4.53×10^{-2}	4.53×10^{-2}
50	3.34×10^{-2}	3.37×10^{-2}	3.46×10^{-2}
60	3.09×10^{-2}	3.09×10^{-2}	3.09×10^{-2}
70	2.74×10^{-2}	2.74×10^{-2}	2.74×10^{-2}
80	2.31×10^{-2}	2.31×10^{-2}	2.31×10^{-2}
500	2.90×10^{-3}	3.50×10^{-3}	3.60×10^{-3}
1000	1.40×10^{-3}	1.80×10^{-3}	1.80×10^{-3}

Example 6.2.2. Consider the nonlinear Volterra-Hammerstein equation (13) with

$$K(t, s; y(s)) := \widetilde{K}(t, s)\, G(y(s)) = e^{s-t}\left(e^{-y(s)} + y(s)\right), \quad f(t) := e^{-t},$$

on the interval $[a, b] = [0, 1]$. The solution is $y(t) = \ln(t + e)$, see, e.g., [**22**].

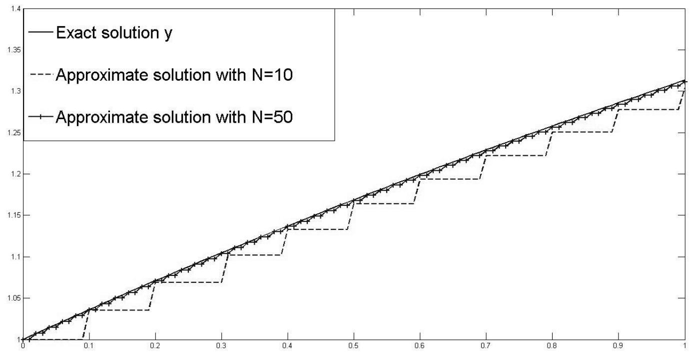

FIGURE 7. Approximate solution $G_N^1 y$ of Example 6.2.2, for $N = 10$ and $N = 50$.

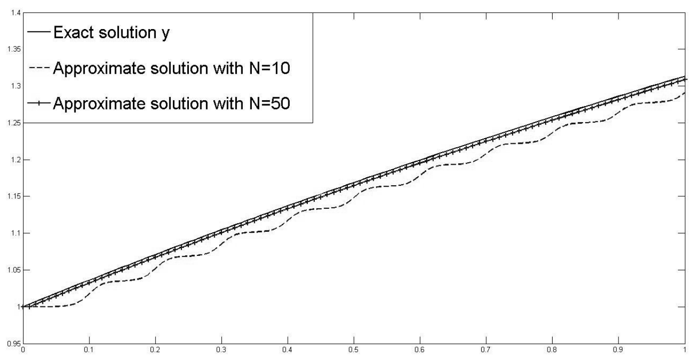

FIGURE 8. Approximate solution $G_N^2 y$ of Example 6.2.2, for $N = 10$ and $N = 50$.

TABLE 6. Numerical results for Example 6.2.2.

N	ε_N^1	ε_N^2	ε_N^3
10	2.73×10^{-2}	2.06×10^{-2}	2.38×10^{-2}
20	1.38×10^{-2}	1.14×10^{-2}	1.16×10^{-2}
30	8.40×10^{-3}	8.10×10^{-3}	8.40×10^{-3}
40	6.90×10^{-3}	5.80×10^{-3}	5.80×10^{-3}
50	3.40×10^{-3}	3.40×10^{-3}	3.70×10^{-3}
500	1.40×10^{-4}	3.45×10^{-4}	3.75×10^{-4}
1000	7.03×10^{-5}	1.72×10^{-4}	1.87×10^{-4}

As above, we show in Table 6 the relative numerical errors ε^i_N, $i = 1, 2, 3$. Again, we plotted in Figures 7 and 8 the approximate solutions $G^1_N y$, $G^2_N y$, for $N = 10$ and $N = 50$.

As in the case of linear equations, our collocation method exhibits slow convergence and poor accuracy. However, the procedure in (15) quickly yields all coefficients that can be used in an analytical approximate representation for the solutions in terms of sigmoidal functions, to a large class of nonlinear equations. The low computational cost of the algorithm allows one to considerably increase the number N of collocation points, and hence the number of the superposed sigmoidal functions, so to obtain a higher accuracy.

REFERENCES

1. A.V. Adrianov, *An analog of the Jackson-Nikol'skij theorem on the approximation by superposition of sigmoidal functions*, Math. Notes **59** (1996), 661–664 (transl. of Matem. Z., Russian Academy of Sciences).

2. G. Albano and V. Giorno, *On the first exit time problem for a Gompertz-type tumor growth*, Lect. Notes Comp. Sci. **5717/2009** (2009), 113–120.

3. K.E. Atkinson, *The numerical solution of integral equations of the second kind*, Cambridge Mono. Appl. Comput. Math. **4**, Cambridge University Press, Cambridge, 1997.

4. A.R. Barron, *Universal approximation bounds for superpositions of a sigmoidal function*, IEEE Trans. Infor. Theor. **39** (1993), 930–945.

5. K.S. Berenhaut, D.C. Morton and P.T. Fletcher, *Bounds for inverses of triangular Toeplitz matrices*, SIAM. J. Matrix Anal. Appl. **27** (2005), 212–217.

6. F. Brauer and C. Castillo Chavez, *Mathematical models in population biology and epidemiology*, Springer, New York, 2001.

7. H. Brunner, *Collocation methods for Volterra integral and related functional equations*, Cambr. Mono. Appl. Comput. Math. **15**, Cambridge University Press, Cambridge, 2004.

8. Debao Chen, *Degree of approximation by superpositions of a sigmoidal function*, Approx. Theory Appl. **9** (1993), 17–28.

9. H. Chen, T. Chen and R. Liu, *A constructive proof of Cybenko's approximation theorem and its extension*, in *Computing science and statistics*, C. Page and R. LePage, eds., 1990.

10. ———, *A constructive proof and an extension of Cybenko's approximation theorem*, in *Computing science and statistics*, Springer-Verlag, New York, 1992.

11. E.S. Chumerina, *Choice of optimal strategy of tumor chemotherapy in Gompertz model*, J. Comput. Systems Sci. Inter. **48** (2009), 325–331.

12. D. Costarelli and R. Spigler, *Constructive approximation by superposition of sigmoidal functions*, Anal. Theor. Appl. **29** (2013), 169–196.

13. G. Cybenko, *Approximation by superpositions of a sigmoidal function*, Math. Contr. Sig. Syst. **2** (1989), 303–314.

14. B. Gao and Y. Xu, *Univariant approximation by superpositions of a sigmoidal function*, J. Math. Anal. Appl. **178** (1993), 221–226.

15. N. Hahm and B. Hong, *Approximation order to a function in* $\overline{C}(\mathbf{R})$ *by superposition of a sigmoidal function*, Appl. Math. Lett. **15** (2002), 591–597.

16. N. Hritonenko and Y. Yatsenko, *Mathematical modelling in economics, ecology and the environment*, Science Press, Beijing, 2006.

17. A.J. Jerry, *Introduction to integral equations with applications*, *A series of monographs and textbooks*, Pure Appl. Math., Marcel Dekker, Inc., New York, 1985.

18. L.K. Jones, *Constructive approximations for neural networks by sigmoidal functions*, Proc. IEEE **78** (1990), 1586–1589.

19. G. Lewicki and G. Marino, *Approximation by superpositions of a sigmoidal function*, Z. Anal. Anwend. **22** (2003), 463–470.

20. ———, *Approximation of functions of finite variation by superpositions of a sigmoidal function*, Appl. Math. Lett. **17** (2004), 1147–1152.

21. K. Maleknejad, E. Hashemizadeh and R. Ezzati, *A new approach to the numerical solution of Volterra integral equations by using Bernstein's approximation*, Comm. Nonlin. Sci. Numer. Simul. **16** (2011), 647–655.

22. K. Maleknejad and E. Najafi, *Numerical solution of nonlinear Volterra integral equations using the idea of quasilinearization*, Comm. Nonlin. Sci. Numer. Simul. **16** (2011), 93–100.

23. E.A. Mantysaari, M.L. Sevon Aimonen, I. Stranden and K. Vuori, *Estimation of non-linear growth models by linearization*: *A simulation study using a Gompertz function*, Genet. Sel. Evol. **38** (2006), 343–358.

24. H.N. Mhaskar and C.A. Micchelli, *Approximation by superposition of sigmoidal and radial basis functions*, Adv. Appl. Math. **13** (1992), 350–373.

25. A. Vecchio, *A bound for the inverse of a lower triangular Toeplitz matrix*, SIAM J. Matrix Anal. Appl. **24** (2003), 1167–1174.

26. S.A. Yousefi, *Numerical solution of Abel's integral equation by using Legendre wavelets*, Appl. Math. Comput. **175** (2006), 574–580.

Dipartimento di Matematica, Università "Roma Tre" 1, Largo S. Leonardo Murialdo, 00146 Rome, Italy
Email address: costarel@mat.uniroma3.it

Dipartimento di Matematica, Università "Roma Tre" 1, Largo S. Leonardo Murialdo, 00146 Rome, Italy
Email address: spigler@mat.uniroma3.it