IMA Journal of Numerical Analysis (2018) **00**, 1–20
doi:10.1093/imanum/dry042

An ultraspherical spectral method for linear Fredholm and Volterra integro-differential equations of convolution type

NICHOLAS HALE
Department of Mathematical Sciences, Stellenbosch University, Stellenbosch, 7602, South Africa
nickhale@sun.ac.za

[Received on 1 December 2017; revised on 24 May 2018]

The Legendre-based ultraspherical spectral method for ordinary differential equations (Olver, S. & Townsend, A. (2013) A fast and well-conditioned spectral method. *SIAM Rev.*, **55**, 462–489.) is combined with a formula for the convolution of two Legendre series (Hale, N. & Townsend, A. (2014a) An algorithm for the convolution of Legendre series. *SIAM J. Sci. Comput.*, **36**, A1207–A1220.) to produce a new technique for solving linear Fredholm and Volterra integro-differential equations with convolution-type kernels. When the kernel and coefficient functions are sufficiently smooth, then the method is spectrally accurate and the resulting almost-banded linear systems can be solved with linear complexity.

Keywords: Fredholm; Volterra; convolution; Legendre; ultraspherical; spectral; integro-differential.

1. Introduction

Fredholm and Volterra integral and integro-differential equations (IDEs) arise often in mathematical models of physical phenomena, such as renewal theory (Cox, 1962), neural networks (Jackiewicz *et al.*, 2006), population dynamics (Kuang, 1993), thermodynamics (MacCamy, 1977), the spread of disease (Medlock & Kot, 2003) and financial mathematics (Cont & Tankov, 2004). This paper concerns the numerical solution of linear Fredholm and Volterra IDEs of the form

$$(\mathscr{L}^r y)(t) \;=\; f(t) + g(t)\int_0^T k(t-s)h(s)y(s)\,\mathrm{d}s, \qquad 0 \leqslant t \leqslant T, \tag{1.1}$$

and

$$(\mathscr{L}^r y)(t) \;=\; f(t) + g(t)\int_0^t k(t-s)h(s)y(s)\,\mathrm{d}s, \qquad 0 \leqslant t \leqslant T, \tag{1.2}$$

respectively, where the given functions $g(t)$ and $h(t)$ and the *convolution* (or *difference*) kernel k are smooth, and $(\mathscr{L}^r y)(t) = a_r(t)y^{(r)}(t) + a_{r-1}(t)y^{(r-1)}(t) + \ldots + a_0(t)y(t)$ is a linear differential operator with smooth coefficients. When $r = 0$ and $g(t) = h(s) = 1$, then (1.2) reduces to the well-known *convolution equation* (Linz, 1985, Chapter 6). When $r > 0$ then (1.1) and (1.2) are known as Fredholm and Volterra IDEs *of convolution type* (FIDECTs and VIDECTs), respectively, and are augmented with an auxiliary equation, $(\mathscr{B}y)(t) = \gamma$, where $\mathscr{B}$ is an $r \times \infty$ linear functional representing initial conditions, boundary conditions or other constraints.

In practical applications, solutions to such equations can not usually be obtained analytically (i.e., in closed form) and one relies on numerical approximation. Expanding on ideas introduced by (El-gendi, 1969/1970), both Driscoll (2010) and Tang *et al.* (2008) have recently shown that discretization of a broad class of integral and IDEs by collocation at Chebyshev or Legendre points can yield exponential convergence and high-accuracy solutions. However, the drawback of global collocation is that such discretizations result in dense matrices that are typically time consuming to factorize/solve. Furthermore, the matrices representing the differential operator on the left-hand side of the equations are ill-conditioned, particularly when the degree of differentiation, r, is large (Trefethen & Trummer, 1987).

Olver's ApproxFun.jl package (Olver *et al.*, 2017) uses the ultraspherical spectral (US) method (Olver & Townsend, 2013) for discretization of differential operators, resulting in sparse (almost-banded) well-conditioned matrices and spectral convergence when solving linear ordinary differential equation (ODEs). When the kernel k in (1.1) or (1.2) is separable, degenerate or well-approximated by a low-rank function (i.e., $k(s,t) \approx \sum_{i=1}^{R} \phi_i(t)\psi_i(s)$ and R is small), then equations of the form (1.1)/(1.2) can be solved by replacing the Fredholm/Volterra operator by a linear combination of multiplication and definite/indefinite integral operators. Slevinsky & Olver (2017) use such an approach, typically referred to as *degenerate kernel approximation* (Kress, 1999), to solve a broad class of singular integral equations. However, when k is not of low rank (i.e., R is large) this can potentially become computationally expensive.

The approach we take in this work is similar, but does not rely on the kernel being separable or of low rank. In particular, we continue to use the US method (in this case, with a Legendre basis) for discretization of $\mathscr{L}^r$, but to discretize the Fredholm and Volterra operators we use the formula introduced in Hale & Townsend (2014a) for computing the convolution of two Legendre series expansions. We demonstrate that when the convolution-type kernel $k(t-s)$ in (1.1) or (1.2) is sufficiently smooth, this results in banded linear operators that can be efficiently incorporated into the US method framework, obtaining spectral accuracy with linear complexity for many problems.

The outline of this paper is as follows. In Section 2 we recall the US method and describe the algorithm used in (Olver *et al.*, 2017) for solving FIDEs and VIDEs when the kernel k is of low rank (or well approximated by a low-rank function). In Section 3 we recall the formulae introduced in Hale & Townsend (2014a) for the convolution of two Legendre series, and show how this may be combined with US methods to solve FIDECTs and VIDECTs. Some examples and numerical results are given in Section 4 before we conclude in Section 5.

REMARK 1.1 It is worthwhile to note that the method described in this paper is not directly applicable to fractional integral and differential equations (FIEs and FDEs). Although FIEs take the form of (1.2), the kernel in such cases is not smooth (it has end-point singularities), and hence its Legendre series expansion will converge too slowly. For an ultraspherical method for FIEs and FDEs, see Hale & Olver (2017).

REMARK 1.2 We use the following conventions in our notation throughout: calligraphic fonts (e.g., $\mathscr{L}$ and $\mathscr{D}$) represent continuous linear operators acting on functions. Uppercase characters (e.g., L and D) represent (possibly infinite-dimensional) matrices. Bold face uppercase mathematical fonts that take an argument (e.g., $\mathbf{P}(x)$) represent *quasimatrices*, i.e., matrices whose columns are continuous functions. Underlined variables represent (possibly infinite-dimensional) vectors containing coefficients of mapped-Legendre series expansions of a function on the interval $[0, T]$. Indices for vectors and matrices begin at zero. MATLAB code to reproduce all of the numerical results in this paper can be found in the Git repository (Hale, 2017).

2. Preliminaries

2.1 *Legendre polynomials and Legendre series*

The Legendre polynomials $\widehat{P}_0, \widehat{P}_1, \ldots$ are orthogonal with respect to the L^2 inner product on the interval $[-1,1]$, with the conventional normalization $||\widehat{P}_n||_2^2 = 2/(2n+1)$.[1] We denote by $\widehat{\mathbf{P}}(x)$ the Legendre quasimatrix, whose jth column is the degree $j-1$ Legendre polynomial, $\widehat{P}_{j-1}(x)$, i.e., $\widehat{\mathbf{P}}(x) = [\widehat{P}_0(x), \widehat{P}_1(x), \ldots]$. Since (1.1) and (1.2) are defined, not on the interval $[-1,1]$, but rather on $[0,T]$, we introduce the mapped-Legendre polynomials and their associated quasimatrix

$$P_n(t) = \left(\widehat{P}_n \circ \psi_{[0,T]}\right)(t), \qquad \mathbf{P}(t) = \left[P_0(t), P_1(t), \ldots\right], \tag{2.1}$$

where $\psi_{[0,T]}(t) = 2t/T - 1$ is the linear map from $[0,T]$ to $[-1,1]$. The columns of $\mathbf{P}(t)$ then form a basis for Lipschitz continuous functions on $[0,T]$. We associate with this basis a space of coefficients, $\mathbf{P} \cong \mathbb{C}^\infty$, so that $\underline{v} = (v_0, v_1, \ldots)^\top \in \mathbf{P}$ if and only if

$$\sum_{n=0}^{\infty} |v_n| \sup_{0\leqslant t\leqslant T} \left|\widehat{P}_n(t)\right| = \sum_{n=0}^{\infty} |v_n| \sup_{-1\leqslant x\leqslant 1} |P_n(x)| = \sum_{k=0}^{\infty} |v_n| < \infty, \tag{2.2}$$

thus if $\underline{v} \in \mathbf{P}$ then $\mathbf{P}(t)\underline{v}$ defines a continuous function on $[0,T]$. If all but a finite number of the v_n are nonzero, then $p_N(t) = \mathbf{P}(t)\underline{v}$ is a polynomial of degree N, where v_N is the final nonzero coefficient in $\underline{v}$.

Conversely, it follows directly from the orthogonality of the Legendre polynomials that, given a continuous function $p(x)$ on $[-1,1]$, the best-L^2 degree N polynomial approximation of p is obtained by truncation of its Legendre series. Formally, the kth coefficient in the expansion of p is given by

$$v_n = \frac{2n+1}{2} \int_{-1}^{1} \widehat{P}_n(x) f(x)\, \mathrm{d}x, \tag{2.3}$$

but obtaining the coefficients in this manner is computationally expensive, particularly if N is large (see discussion below). The connection to best-L^2 approximation suggests that the approximation of a smooth function by a Legendre series will converge quickly as the degree of the expansion increases. For analytic functions, exactly how fast the series converges is made precise by the following result:

THEOREM 2.1 (Wang & Xiang (2012)). Denote by $v_N(x)$ the truncated Legendre series expansion of $v(x)$, i.e., $v_N(x) = \sum_{n=0}^{N} v_n P_n(x)$. If $v(x)$ is analytic inside and on a Bernstein ellipse $\mathscr{E}_\rho$, then for each $N \geqslant 0$,

$$\max_{x\in[-1,1]} \left|v(x) - v_N(x)\right| \leqslant \frac{(2N\rho + 3\rho - 2N - 1)\ell(\mathscr{E}_\rho)M}{\pi\rho^{N+1}(\rho-1)^2(1-\rho^{-2})} = \mathscr{O}(\rho^{-N-3}), \tag{2.4}$$

where $\mathscr{E}_\rho = \left\{z \in \mathbb{C} \,:\, z = \frac{1}{2}(\rho\, e^{i\theta} + \rho^{-1} e^{-i\theta}),\ \rho \geqslant 1,\ 0 \leqslant \theta \leqslant 2\pi\right\}$ and $M = \max_{z\in\mathscr{E}_\rho} |v(z)|$.

Note in particular that this convergence is geometric as N is increased. Wang & Xiang (2012) provide a similar result for when $v(x)$ is smooth, but not analytic. In such cases convergence is algebraic in N with the degree roughly the number of continuous derivatives of $v(x)$ on $[-1,1]$. However, as mentioned above, computing Legendre coefficients from the definition (2.3) is usually impractical.

[1] We include the nonstandard $\widehat{\ }$ notation here to distinguish from the mapped-Legendre polynomials introduced momentarily.

Instead, a more efficient paradigm is to first compute approximate Chebyshev coefficients of p by interpolating on a Chebyshev grid and applying a discrete cosine transform using the Fast Fourier transform (FFT) (Mason & Handscomb, 2003, Section 4.7). From these, the corresponding approximate Legendre coefficients can be rapidly computed by any of the algorithms presented in (Alpert & Rokhlin, 1991; Hale & Townsend, 2014b; Townsend *et al.*, 2018). In such a situation it is rather the approximation properties of the Chebyshev interpolant that are required (see, for example, Trefethen, 2013, Chapters 7 and 8), but since (a) these are asymptotically very similar to that of the truncated Legendre series and (b) a detailed analysis of the convergence of the presented algorithm is beyond the scope of this paper, we omit the details.

2.2 *Ultraspherical polynomials and banded operators*

The Legendre polynomials are a particular case of the more general ultraspherical (or *Gegenbauer*) polynomials, $\widehat{C}_n^{(\lambda)}(x)$, $\lambda > 0$, orthogonal on the interval $[-1, 1]$ with respect to the weight function $(1 - x^2)^{\lambda-1/2}$. As above, one can define a quasimatrix of mapped ultraspherical polynomials, $\mathbf{C}^{(\lambda)}(t)$, on the interval $[0, T]$ via an affine transformation and associate with them a space of coefficients $\mathbf{C}^{(\lambda)}$.

If $\underline{v} \in \mathbf{C}^{(\lambda)}$, then linear operators that can be applied to $v(t) = \mathbf{C}^{(\lambda)}(t)\underline{v}$ induce infinite-dimensional matrices that can be viewed as acting on $\underline{v}$. For example, given a continuous linear operator $\mathscr{L}$ satisfying

$$\left(\mathscr{L}C_n^{(\lambda)}\right)(t) = \sum_{j=n-m}^{n+m} L_{jn}C_n^{(\ell)}(t), \tag{2.5}$$

we can associate it with an *m-banded* (i.e., banded with bandwidth m) infinite-dimensional matrix

$$L := \begin{bmatrix} L_{00} & \cdots & L_{0m} & & \\ \vdots & \ddots & L_{1m} & L_{1,m+1} & \\ L_{m0} & L_{m1} & \ddots & L_{mm} & \ddots \\ & L_{m+1,1} & \ddots & \ddots & \ddots \\ & & \ddots & \ddots & \ddots \end{bmatrix}. \tag{2.6}$$

Since L is banded, multiplication is well defined on $\mathbb{C}^\infty$ and (2.6) can be viewed as an operator $L : \mathbf{C}^{(\lambda)} \to \mathbf{C}^{(\ell)}$. To relate L and $\mathscr{L}$ we note that (by construction) we have

$$\left(\mathscr{L}C_n^{(\lambda)}\right)(t) = \left(\mathscr{L}\mathbf{C}^{(\lambda)}\right)(t)\underline{e}_n = \mathbf{C}^{(\ell)}(t)L\underline{e}_n. \tag{2.7}$$

If $v(t)$ is Lipschitz continuous on $[0, T]$, then there exists $\mathbf{v} \in \mathbf{C}^{(\lambda)}$ so that $\mathbf{C}^{(\lambda)}(t)\underline{v} = v(t)$ and, because $\mathscr{L}$ is continuous, we have

$$(\mathscr{L}v)(t) = \left(\mathscr{L}\mathbf{C}^{(\lambda)}\right)(t)\underline{v} = \mathbf{C}^{(\ell)}(t)L\underline{v}. \tag{2.8}$$

Therefore, applying $\mathscr{L}$ to $v(t)$ is equivalent to applying L to $\underline{v}$.

Such banded operators form the basis of the US method, which we describe below. In Section 3 we will show that Fredholm and Volterra operators also give rise to banded operators when the kernel is sufficiently smooth.

2.3 *The US method*

Traditional Chebyshev–Galerkin spectral methods (as introduced by Lanczos, 1956, and later developed by Ortiz, 1969 and Orszag, 1971a,b) expand both the unknown solution and the right-hand side of an ODE as Chebyshev series of the first kind. One then applies the differential operator to the Chebyshev series and equates coefficients (or minimizes the residual in some suitable norm) to obtain the approximate solution. Although resulting in rapidly converging approximations (often geometric in the number of degrees of freedom), such an approach typically leads to dense and ill-conditioned matrices (Trefethen & Trummer, 1987).

The key observation of the US method introduced by Olver & Townsend (2013) is that if one is willing to map between Chebyshev and higher-order ultraspherical bases (i.e., a Petrov–Galerkin approach), then differentiation becomes a banded operator. After a diagonal preconditioner the discretized linear systems have a condition number that is bounded by a constant that is independent of the discretization size. This allows a broad class of ODEs (including some variable coefficient problems) to be solved with spectral accuracy and linear complexity, and overcomes the ill-conditioning associated with spectral differentiation matrices. Although the US method as described in Olver & Townsend (2013) expands the solution as a Chebyshev series of the first kind (which is usually the most convenient to work with), this is not necessary. Indeed, Hale & Olver (2017) recently used Legendre expansions as part of an ultraspherical method for fractional differential equations.

The Legendre-based US method for IDEs requires four banded operators that act on Legendre and related ultraspherical polynomials: differentiation, integration, conversion and multiplication. Derivation of these operators on the interval $[-1, 1]$ can be found in Hale & Olver (2017, Section 2), but for convenience, we give examples of the first three acting on $\mathbf{P}$ (i.e., on the interval $[0, T]$) in Table 1. Multiplication operators, $M_\lambda[f]$ from $\mathbf{C}^{(\lambda)} \times \mathbf{C}^{(\lambda)}$ to $\mathbf{C}^{(\lambda)}$, are less easy to write down (since they depend on the coefficients of the function f), but they are readily generated using the procedure described in Hale & Olver (2017, Section 2.3) and we omit the details here. MATLAB code for computing these matrices for a given value of N can be found (Hale, 2017).

Using operators from Table 1 we demonstrate the Legendre-based US method on the first-order IDE

$$y'(t) + ay(t) = e^t \int_0^t e^{-s} y(s)\,\mathrm{d}s, \quad 0 \leqslant t \leqslant 1, \qquad y(0) = 1. \tag{2.9}$$

To proceed we assume that the solution to (2.9) is sufficiently smooth so that we may consider its mapped-Legendre series expansion $y(t) = \mathbf{P}(t)\underline{y}$. Discretization of the left-hand side of (2.9) proceeds almost exactly as described in Olver & Townsend (2013, Section 2), but with $\mathscr{D}_0$ and $\mathscr{S}_0$ replaced by $D_{1/2}$ and $S_{1/2}$ from Table 1, respectively, so that

$$(\mathscr{L}y)(t) := y'(t) + ay(t) = \mathbf{C}^{(3/2)}(t) \underbrace{(D_{1/2} + aS_{1/2})}_{L} \underline{y}. \tag{2.10}$$

Notice in particular that the result is now expressed in the ultraspherical basis, $\mathbf{C}^{(3/2)}$, and that L is banded (since it is a linear combination of two banded matrices).

TABLE 1 *The banded infinite-dimensional matrices representing differentiation, integration and conversion operators when acting on the spaces of mapped-Legendre and ultraspherical coefficients on the interval* $[0, T]$. *One should interpret this table as follows: if* $v(t) \in \mathbf{P}(t)$ *so that* $v(u) = \mathbf{P}\underline{v}$ *, then, for example,* $\frac{\mathrm{d}}{\mathrm{d}t}v(t) = (\mathscr{D}\mathbf{P})(t)\underline{v} = \mathbf{C}^{(3/2)}(t)D_{1/2}\underline{v}$ *and* $v(t) = \mathbf{P}(t)\underline{v} = \mathbf{C}^{(3/2)}(t)S_{1/2}\underline{v}$. *These relationships may be used to solve linear differential, integral, or IDEs as outlined below (see, for example, equation (2.10)). For derivations see Hale & Olver (2017) or Olver & Townsend (2013). Multiplication operators,* $M_\lambda[f] : \mathbf{C}^{(\lambda)} \times \mathbf{C}^{(\lambda)} \mapsto \mathbf{C}^{(\lambda)}$, *are omitted from the table, but can be generated using the procedure described in Hale & Olver (2017, Section 2.3). MATLAB code for computing these, and each of the matrices below, can be found in Hale (2017).*[†]

Continuous operator	Discrete operator	Matrix representation	Range	Image	Continuous operator	Discrete operator	Matrix representation	Range	Image
$\mathscr{D} = \frac{\mathrm{d}}{\mathrm{d}t}$	$D_{1/2}$	$\frac{2}{T}\begin{bmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & 0 & 1 & & \\ & & & \ddots & \ddots & \end{bmatrix}$	$\mathbf{P}$	$\mathbf{C}^{(3/2)}$	$\mathscr{D}^2 = \frac{\mathrm{d}^2}{\mathrm{d}t^2}$	$D^2_{1/2}$	$\frac{12}{T^2}\begin{bmatrix} 0 & 0 & 1 & & \\ & 0 & 0 & 1 & \\ & & 0 & 0 & \ddots \\ & & & \ddots & \ddots \end{bmatrix}$	$\mathbf{P}$	$\mathbf{C}^{(5/2)}$
$\int_0^t \cdot\,\mathrm{d}t$	$Q_{1/2}$	$\frac{T}{2}\begin{bmatrix} 1 & -\frac{1}{3} & & & \\ 1 & 0 & -\frac{1}{5} & & \\ & \frac{1}{3} & 0 & -\frac{1}{7} & \\ & & \ddots & \ddots & \ddots \end{bmatrix}$	$\mathbf{P}$	$\mathbf{P}$	$\int_0^T \cdot\,\mathrm{d}t$	$Q^{def}_{1/2}$	$T\begin{bmatrix} 1 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix}$	$\mathbf{P}$	$\mathbf{P}$
$\mathscr{I}$	$S_{1/2}$	$\begin{bmatrix} 1 & 0 & -\frac{1}{5} & & \\ & \frac{1}{3} & 0 & -\frac{1}{7} & \\ & & \frac{1}{5} & 0 & \ddots \\ & & & \ddots & \ddots \end{bmatrix}$	$\mathbf{P}$	$\mathbf{C}^{(3/2)}$	$\mathscr{I}$	$S_{3/2}$	$3\begin{bmatrix} \frac{1}{3} & 0 & -\frac{1}{7} & & \\ & \frac{1}{5} & 0 & -\frac{1}{9} & \\ & & \frac{1}{7} & 0 & \ddots \\ & & & \ddots & \ddots \end{bmatrix}$	$\mathbf{C}^{(3/2)}$	$\mathbf{C}^{(5/2)}$

[†]Note that whilst technically the indefinite integral operator $\int_0^T \cdot\,\mathrm{d}t$ is a linear functional mapping $\mathbf{P}$ to $\mathbf{C}$, for convenience of notation we consider it as a mapping from $\mathbf{P}$ to $\mathbf{P}$.

For the right-hand side we require multiplication operators $M_{1/2}[\cdot]$ and the indefinite integral operator $Q_{1/2}$ so that

$$e^t \int_0^t e^{-s} y(s)\,\mathrm{d}s = \mathbf{P}(t) \underbrace{M_{1/2}[\underline{e}^t] Q_{1/2} M_{1/2}[\underline{e}^{-s}]}_{V} \underline{y} = \mathbf{C}^{(3/2)}(t) S_{1/2} V \underline{y}, \tag{2.11}$$

where in the second equality we have multiplied by $S_{1/2}$ to make the bases of (2.10) and (2.11) compatible. Therefore, (2.9) is equivalent to

$$\underbrace{(L - S_{1/2}V)}_{A} \underline{y} = 0. \tag{2.12}$$

If the coefficients of e^t and e^{-s} are truncated at machine precision, then each of the operators that make up A are banded, and hence A is banded (with bandwidth around 15 in this example). The initial condition $y(0) = 1$ can be expressed as $B\underline{y} := \mathbf{P}(0)\underline{y} = [1, -1, 1, -1, 1, -1, \ldots]\underline{y} = 1$, and hence we arrive at the set of linear equations

$$\begin{bmatrix} B \\ L - S_{1/2}V \end{bmatrix} \underline{y} = \begin{bmatrix} 1 \\ \underline{0} \end{bmatrix}. \tag{2.13}$$

The infinite-dimensional matrix on the left-hand side of the above is now *almost-banded*, that is, banded with one dense row (see the left panel of Fig. 1). One can solve for the approximate coefficients of the solution by truncating the matrix equation at a suitable size N. Using a Schur-complement factorization of the 1-1 block, or the Woodbury formula, this can be done in $\mathcal{O}(N)$ operations. Alternatively, one can employ the adaptive-QR method, which does not truncate, but rather 'solves the infinite-dimensional problem' until the coefficients of the solution fall below a specified tolerance (Olver & Townsend, 2013, 2014). In either case, since the kernel is smooth (in this case entire) fast convergence to the underlying solution is usually observed.

Remark 2.2 As described in Section 2.1, mapped Legendre and Chebyshev series converge geometrically fast for functions analytic in a neighbourhood of $[0, T]$ (and super geometrically for entire functions), and hence one typically requires $\mathcal{O}(\log \varepsilon)$ coefficients to resolve an analytic function to an accuracy of ε (once any oscillatory nature has been resolved). The bandwidth of the matrix A will therefore grow proportionally to $\log \varepsilon$ and require $\mathcal{O}(\log(\varepsilon)N)$ operations to factorize/solve. However, in this manuscript we will fix ε at machine double precision and absorb this term in the big-$\mathcal{O}$.

2.4 *FIDECTS and VIDECTS with low-rank kernels*

A bivariate function $k(s, t)$ is *separable* (or *degenerate*) if it can be expressed as

$$k(s, t) = \sum_{i=1}^{R} \phi_i(t)\psi_i(s), \qquad R < \infty. \tag{2.14}$$

If (2.14) holds, then we say that k has *rank* at most R, and if R is small we say that k is of *low rank*. Even if k is not of finite rank, it can often be well approximated (in some appropriate norm) by a finite-

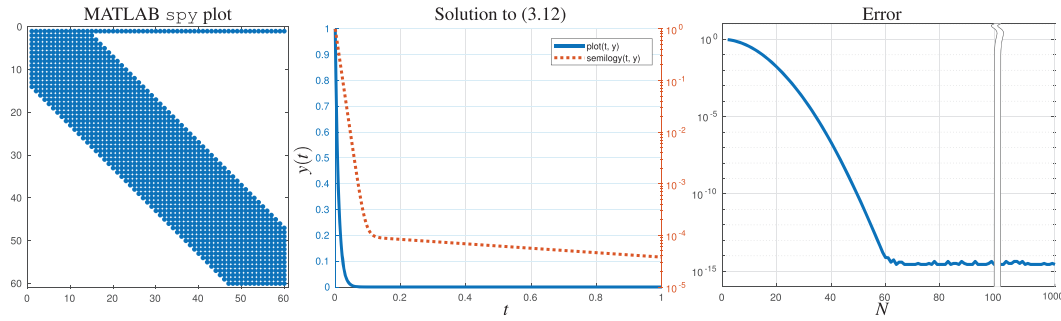

FIG. 1. Left: MATLAB spy plot of the $N = 60$ discretization of the operator in (3.14). The operator is 'almost-banded' (with one dense row at the top due to the boundary condition) and the bandwidth is determined by the decay in the Legendre coefficients of the kernel. Middle: Solution (solid line = linear scale, dashed line = $\log_{10}$ in y) of (3.12) with $a = 100$, evaluated using Clenshaw's algorithm. Right: Error (approx. infinity norm at 1000 equally-spaced points) for various values of N. As the solution $y(t)$ is entire we observe super-geometric convergence. The well-conditioned discretization of the differentiation operator means there is no growth in the error even for very large values of N.

rank function so that the equality (2.14) can be replaced by $\approx$ (Townsend, 2014, Chapter 3). Adaptive low-rank approximations of bivariate functions can be efficiently computed by Chebfun2 in Chebfun (Townsend & Trefethen, 2013; Driscoll *et al.*, 2014).

Substituting the expansion of the kernel $k(s,t)$ in (2.14) to a Volterra operator yields

$$\int_0^t k(s,t)y(s)\,\mathrm{d}s = \sum_{i=1}^{R} \phi_i(t) \int_0^t \psi_i(s)y(s)\,\mathrm{d}s, \tag{2.15}$$

hence, for a degenerate kernel, the Volterra operator is equivalent to a linear combination of multiplication and indefinite integral operators (Kress, 1999, Chapter 11). To discretize the operator on the left-hand side of (2.15) using the Legendre-US method we proceed as above, i.e., expand y, the ϕ_i, and the ψ_i as Legendre series so that (2.15) becomes

$$\int_0^t k(s,t)y(s)\,\mathrm{d}s = \widehat{\mathbf{P}}(t) \underbrace{\left[\sum_{i=1}^{R} M[\underline{\phi}_i]Q_{1/2}M[\underline{\psi}_i]\right]}_{V} \underline{y}. \tag{2.16}$$

To solve VIDEs of the form

$$(\mathscr{L}^r y)(t) = f(t) + \int_0^t k(s,t)y(s)\,\mathrm{d}s, \tag{2.17}$$

subject to $\mathscr{B}y(t) = \gamma$, one forms Legendre-based US discretizations L of the differential operator $\mathscr{L}^r$ and B of the functional $\mathscr{B}$, and then solves

$$\begin{bmatrix} B \\ L - S_{[r]}V \end{bmatrix} \underline{y} = \begin{bmatrix} \gamma \\ S_{[r]}\underline{f} \end{bmatrix}, \tag{2.18}$$

where $S_{[r]} = S_{r+1/2}S_{r-1/2}\dots S_{3/2}S_{1/2} : \mathbf{P} \to \mathbf{C}^{(r+1/2)}$. In fact, the VIDE solved in Section 2.3 is precisely the Volterra equation (2.17) with $(\mathscr{L}^r y)(t) = y'(t)+ay(t), f(t) = 0$, the rank-1 kernel $k(s,t) = e^{-(t-s)}$ and initial condition $\mathscr{B}y = y(0) = 1$.

Fredholm operators, where the upper limit of integration in (2.15) becomes T rather than t, can be represented similarly, with the matrix $Q_{1/2}$ in (2.16) replaced by $Q^{def}_{1/2}$ from Table 1. The approach can also be extended to equations involving combinations of multiple Fredholm and Volterra operators, systems of linear IDEs (where the almost-banded structure of the operators can be maintained by interlacing the coefficients of the variables) and partial IDEs.

The algorithm described above is essentially that used by the ApproxFun (Olver *et al.*, 2017), but one drawback is that it requires a separable decomposition of $k(s,t)$ (or an approximation thereof). Furthermore, if R is large, then constructing the operator V in this manner may be time consuming. In the next section we propose a modification to the approach which avoids the need to compute low-rank approximations when the kernel is of convolution type, i.e., $k(s,t) = k(t-s)$.

3. New method: a spectral convolution operator

3.1 *A banded convolution operator*

Consider $\underline{k}$ and $\underline{y} \in \mathbf{P}$ so that $k(x) = \mathbf{P}(x)\underline{k}$ and $y(x) = \mathbf{P}(x)\underline{y}$ are Lipschitz continuous functions on $[0,T]$.

Proposition 3.1 Suppose $\underline{k}$ has a finite number of nonzero entries so that $k(t) = \mathbf{P}(t)\underline{k}$ is a polynomial of degree m, then the convolution operator

$$\big(\mathscr{V}[k]y\big)(t) := \int_0^t k(t-s)y(s)\,\mathrm{d}s, \qquad 0 \leqslant t \leqslant T, \tag{3.1}$$

corresponds to an $(m+2)$-banded infinite dimensional matrix $V[\underline{k}] : \mathbf{P} \to \mathbf{P}$, i.e.,

$$(\mathscr{V}[k]y)(t) = \mathbf{P}(t)V[\underline{k}]\underline{y}. \tag{3.2}$$

We refer to $V[\underline{k}]$ as the *discrete Legendre convolution operator* with kernel k.

Proof. Since $k(t)$ is Lipschitz continuous on $[0,T]$ then $\mathscr{V}[k]$ is a compact operator, and therefore there exists $\underline{u} \in \mathbf{P}$ such that $u(t) = \mathbf{P}(t)\underline{u} = (\mathscr{V}[k]y)(t)$. Modifying the derivation in (Hale & Townsend, 2014a, Section 4) to account for the fact that we consider the interval $[0,T]$, we have from (2.3) that for $j = 0, 1, \dots$ the mapped-Legendre coefficients of $u(t)$ satisfy

$$u_j = \frac{2j+1}{T}\int_0^T P_j(t)\int_0^t k(t-s)y(s)\,\mathrm{d}s\,\mathrm{d}t, \tag{3.3}$$

$$= \sum_{n=0}^{\infty} \underbrace{\left[\frac{2j+1}{T}\int_0^T P_j(t)\int_0^t k(t-s)P_n(s)\,\mathrm{d}s\,\mathrm{d}t\right]}_{V[\underline{k}]_{j,n}} y_n = \big[V[k]\underline{y}\big]_j, \tag{3.4}$$

where in the second equality we have used the absolute convergence of the coefficients $\underline{y}$. Hence, we have that $\underline{u} = V[\underline{k}]\underline{y}$. It remains to show that $V[\underline{k}]$ is banded.

It was shown in (Hale & Townsend, 2014a, Theorem 4.1) that the entries of $V[\underline{k}]$ satisfy the recurrence relation

$$V[\underline{k}]_{j,n+1} = -\frac{2n+1}{2j+3}V[\underline{k}]_{j+1,n} + \frac{2n+1}{2j-1}V[\underline{k}]_{j-1,n} + V[\underline{k}]_{j,n-1}, \quad n \geqslant 1, \tag{3.5}$$

with starting values (again, modified to account for the fact we are on the interval $[0, T]$)

$$V[\underline{k}]_{j,0} = \begin{cases} \frac{T}{2}[k_{j-1}/(2j-1) - k_{j+1}/(2j+3)], & j \neq 0, \\ \frac{T}{2}[k_0 - k_1/3], & j = 0, \end{cases} \tag{3.6}$$

$$V[\underline{k}]_{j,1} = \begin{cases} V[\underline{k}]_{j-1,0}/(2j-1) - V[\underline{k}]_{j,0} - V[\underline{k}]_{j+1,0}/(2j+3), & j \neq 0, \\ -V[\underline{k}]_{1,0}/3, & j = 0, \end{cases} \tag{3.7}$$

$$V[\underline{k}]_{0,n} = \ (-1)^n V[\underline{k}]_{n,0}/(2n+1), \ \ n \geqslant 0. \tag{3.8}$$

Since $k(t)$ is a polynomial of degree m, we have $k_j = 0$ for $j > m$, and hence from (3.6) that $V[\underline{k}]_{j,0} = 0$ for $j > m+1$. Subsequently, from (3.7), we also have that $V[\underline{k}]_{j,1} = 0$ for $j > m+2$, and from the recurrence (3.5) it follows by induction that $V[\underline{k}]$ is lower-banded with bandwidth $m+2$. The fact that $V[\underline{k}]$ is also upper-banded, and hence banded, follows from the scaled-symmetry relation (Hale & Townsend, 2014a, (4.8)):

$$V[\underline{k}]_{j,n} = (-1)^{n+j}\frac{2j+1}{2n+1}V[\underline{k}]_{n,j} \qquad j, n \geqslant 0. \tag{3.9}$$

□

3.2 *Solving Volterra IDEs*

We are now in a position to solve Volterra equations of the form (1.2). Let $f(t)$, $g(t)$, $h(t)$ and $k(t)$ be Lipschitz continuous functions on $[0, T]$. Rearranging (1.2) and expressing the result in operator notation so that $\mathscr{A} = (\mathscr{L}^r - g\mathscr{V}[k]h)$, then we have

$$(\mathscr{A}y)(t) = f(t), \qquad 0 \leqslant t \leqslant T, \tag{3.10}$$

which is equivalent to

$$A\underline{y} := \Big(L - S_{[r]}M[\underline{g}]V[\underline{k}]M[\underline{h}]\Big)\,\underline{y} = S_{[r]}\underline{f}, \tag{3.11}$$

where L is the Legendre-US discretization of $\mathscr{L}^r$ and $V[\underline{k}]$ is as in Section 3.1.

Example 3.2 Consider the following example from (Ma & Brunner, 2006),

$$y'(t) + ay(t) = \int_0^t e^{-(t-s)}y(s)\,\mathrm{d}s, \quad 0 \leqslant t \leqslant 1, \quad y(0) = 1, \tag{3.12}$$

where a is a given constant. The exact solution is given by

$$y(t) = e^{-\frac{a+1}{2}t}\Big(\cosh(bt) + \tfrac{1-a}{2b}\sinh(bt)\Big), \tag{3.13}$$

where $b = \frac{1}{2}\sqrt{a^2 - 2a + 5}$. Here $g(t) = h(s) = 1$, so we need only compute the Legendre coefficients of e^{-t} on $[0, T]$ in order to construct the Volterra operator as described above. These can be computed in quasi-linear time using recently developed fast algorithms (Alpert & Rokhlin, 1991; Hale & Townsend, 2014b; Townsend *et al.*, 2018), or in linear time by observing that $k(t)$ satisfies $k' + k = 0$, $k(0) = 1$ and solving this simple ODE on $[0, 1]$ using the US-Legendre method. In either case, the resulting discretization is given by

$$\begin{bmatrix} \begin{bmatrix} 1, & -1, & 1, & -1, & \dots \end{bmatrix} \\ D_{1/2} + aS_{1/2}(I - V[\underline{e}^{-t}]) \end{bmatrix} \underline{y} = \begin{bmatrix} 1 \\ \underline{0} \end{bmatrix}. \tag{3.14}$$

Figure 1 shows the discretization, solution and convergence of the Legendre US method for this VIDECT when $a = 100$. The first panel verifies that the operator in (3.14) is almost-banded, with one dense row at the top due to the boundary condition $y(0) = 1$, and the bandwidth of the remainder determined by the decay in the Legendre coefficients of the kernel. It takes around 13 terms in a Legendre series expansion to represent e^{-t} to machine precision on $[0, 1]$, so here the bandwidth of the operator is around $13+2 = 15$. This drops to around 10 if only 8 digits of precision are required. The centre panel shows a plot of the obtained solution. The approximated Legendre series returned upon solving (3.11) is evaluated using Clenshaw's algorithm. Since the solution here has such a rapid initial decay, we show also the logarithm (base-10) of the solution. The third panel shows the error (approx. infinity norm error at 1000 equally-spaced points, again evaluated using Clenshaw's algorithm) in the computed solution as N is increased. Because the solution $y(t)$ is entire we observe super-geometric convergence until this plateaus at around 10^{-15} due to rounding error in evaluating the computed Legendre series. The well-conditioned discretization of the differentiation operator means there is no growth in the error even for very large values of N.

3.3 *Solving Fredholm IDEs*

We saw in the previous section how a US method for VIDEs might be derived. However, VIDEs are typically initial value problems, for which local or time-stepping discretizations are more commonly used. Fredholm operators, on the other hand, are themselves global in nature, and hence a global spectral discretization of FIDEs is more natural. To this end, consider now the basic Fredholm operator with a convolution kernel, which we may write as

$$\big(\mathscr{F}[k]y\big)(t) := \int_0^T k(t-s)y(s)\,\mathrm{d}s = \big(\mathscr{V}[k]y\big)(t) + \big(\widetilde{\mathscr{V}}[k]y\big)(t), \tag{3.15}$$

where

$$\big(\widetilde{\mathscr{V}}[k]y\big)(t) := \int_t^T k(t-s)y(s)\,\mathrm{d}s. \tag{3.16}$$

Making the substitution $\tau = T - t$, $\sigma = T - s$ gives

$$\big(\widetilde{\mathscr{V}}[k]y\big)(t) = \int_0^\tau k(\sigma-\tau)y(T-\sigma)\,\mathrm{d}\sigma = \int_0^\tau \widetilde{k}(\tau-\sigma)\widetilde{y}(\sigma)\,\mathrm{d}\sigma = \big(\mathscr{V}[\widetilde{k}]\widetilde{y}\big)(\tau), \quad 0 \leqslant t, \tau \leqslant T, \tag{3.17}$$

where $\widetilde{y}(\tau) = y(T - t)$ and $\widetilde{k}(t) = k(-t)$. Since $\widehat{P}_n(-x) = (-1)^n\widehat{P}_n(x)$ for Legendre polynomials on $[-1, 1]$, we have that $P_n(T - t) = (-1)^n P_n(t)$ for mapped-Legendre polynomials on $[0, T]$, and hence

that $\widetilde{\underline{y}} = \widetilde{I}\underline{y}$ and $\underline{y} = \widetilde{I}\widetilde{\underline{y}}$ where

$$\widetilde{I} := \begin{bmatrix} 1 & & & & & \\ & -1 & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \end{bmatrix}. \tag{3.18}$$

Therefore, the discrete version of $\widetilde{\mathscr{V}}[k]$ may be written as

$$\widetilde{V}[\underline{k}]\underline{y} = \widetilde{I}V[\widetilde{\underline{k}}]\widetilde{I}\underline{y}, \tag{3.19}$$

where $\widetilde{\underline{k}}$ are the Legendre coefficients of $k(-t)$ in $[0,T]$, and we may express the Fredholm operator (3.15) as

$$\big(\mathscr{F}[k]y\big)(t) = \mathbf{P}(t)F[\underline{k},\widetilde{\underline{k}}]\underline{y} := \mathbf{P}(t)\left(V[\underline{k}] + \widetilde{I}V[\widetilde{\underline{k}}]\widetilde{I}\right)\underline{y}. \tag{3.20}$$

Like the discrete Volterra operator in the previous section, $F[\underline{k},\widetilde{\underline{k}}]$ is banded when $k(t)$ is a polynomial (or sufficiently well-approximated by a polynomial on $[-T,0]$ and $[0,T]$) and can be combined with the Legendre-based US method to solve Fredholm IDEs. The derivation above is readily extended to Fredholm operators with nonconstant coefficients, which may be written as

$$(\mathscr{F}[k]y)(t) = g(t)\int_0^T k(t-s)h(s)y(s)\,\mathrm{d}s = \mathbf{P}(t)M_{1/2}[\underline{g}]\left(V[\underline{k}] + \widetilde{I}V[\widetilde{\underline{k}}]\widetilde{I}\right)M_{1/2}[\underline{h}]\underline{y}.$$

Example 3.3 We modify the example from the previous section so that the solution (3.13) to (3.12) is also the solution to the second-order Fredholm IDE

$$y''(t) + ay'(t) - y(t) = f(t) - \int_0^T e^{-(t-s)}y(s)\,\mathrm{d}s, \quad 0 \leqslant t \leqslant 1, \quad y(0) = 1,\ y(T) = \gamma_T, \tag{3.21}$$

where $f(t) = \exp(-t)\big(\exp\left(\frac{1-a}{2}T\right)\sinh(bT) - \exp\left(\frac{1-a}{2}t\right)\sinh(bt)\big)/b$, a is a given constant and γ_T is obtained by evaluating (3.13) at $t = T$. Using the approach described above, our discretization is given by

$$\begin{bmatrix} \begin{bmatrix} 1, & -1, & 1, & -1, & 1 & -1, & 1 & -1, & \dots \\ 1, & 1, & 1, & 1, & 1, & 1, & 1, & 1, & \dots \end{bmatrix} \\ D_{1/2}^2 + S_{3/2}(aD_{1/2} + S_{1/2}(-I + F[\underline{e}^{-t}, \underline{e}^{t}])) \end{bmatrix} \underline{y} = \begin{bmatrix} \begin{bmatrix} 1 \\ \gamma_T \end{bmatrix} \\ S_{3/2}S_{1/2}\underline{f} \end{bmatrix}. \tag{3.22}$$

Figure 2 shows the resulting discretization, solution and convergence when $a = 100$. There is very little difference to the corresponding results for (3.12), other than that the two boundary conditions now give rise to two dense rows at the top of the matrix. As before, the discretization is almost-banded and the convergence super-geometric.

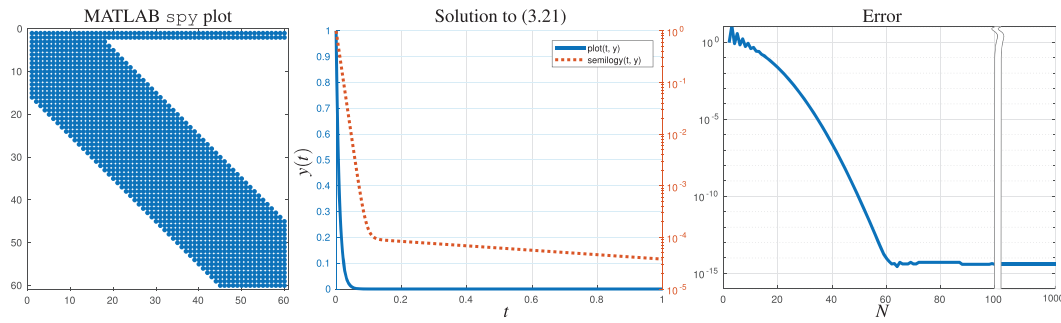

FIG. 2. As in Fig. 1, but here for the FIDECT (3.21) with $a = 100$. The only significant difference in the discretizations is that the operator now has two dense rows at the top, because of the two boundary conditions.

REMARK 3.4 Note that the splitting in (3.15) means we require only that the kernel is *piecewise* smooth on $[-T, 0]$ and $[0, T]$, not necessarily smooth on $[-T, T]$. In particular, this allows kernels of the form $k(|t - s|)$ since here

$$\int_0^T k(|t-s|)y(s)\,\mathrm{d}s = \underbrace{\int_0^t k(t-s)y(s)\,\mathrm{d}s}_{(\mathscr{V}[k]y)(t)} + \underbrace{\int_t^T k(s-t)y(s)\,\mathrm{d}s}_{(\mathscr{V}[\widehat{k}]y)(t)}, \tag{3.23}$$

from which it follows that the discrete version of the operator is given by $F[\underline{k}, \underline{k}] = V[\underline{k}] + \widetilde{I}V[\underline{k}]\widetilde{I}$.

4. Further examples

The two examples considered in the previous section were intentionally elementary. In fact, in both cases the rank-1 kernel, $k(t - s) = e^{-(t-s)}$, could be easily dealt with by a variant of the degenerate kernel approach described in Section 2.3. In this section we consider two more challenging examples and make comparisons with some existing spectrally-accurate methods for IDEs from the literature.

EXAMPLE 4.1 Gaussian kernels arise often in Volterra and Fredholm integral equations with applications in filtering and scattering. These functions are not degenerate and such equations are therefore not so easily solved with standard techniques as those with, say, exponential kernels in the examples above. However, these kernels are smooth and of convolution type, so the method of this paper is readily applicable. To this end, consider the second-order Fredholm IDE given by

$$\xi^2 y''(t) + ty'(t) + y(t) + \int_0^1 e^{-\frac{(t-s)^2}{2\sigma^2}} y(s)\,\mathrm{d}s = f(t), \quad 0 \leqslant t \leqslant 1. \tag{4.1}$$

For convenience we choose the function $f(t)$ on the right-hand side of (4.1) as

$$f(t) = \frac{\xi\sigma}{r}\sqrt{\frac{\pi}{2}}e^{-\frac{t^2}{2r^2}}\left(\operatorname{erf}\left(\frac{\xi t}{\sigma r\sqrt{2}}\right) + \operatorname{erf}\left(\frac{r^2 - \xi^2 t}{\xi\sigma r\sqrt{2}}\right)\right), \tag{4.2}$$

where $r = \sqrt{\sigma^2 + \xi^2}$ and enforce the additional constraints

$$y(0) = 1 \quad \text{and} \quad \int_0^1 y(t)\,\mathrm{d}t = \sqrt{\frac{\pi}{2}}\xi \operatorname{erf}\left(\frac{1}{\sqrt{2}\xi}\right) =: \gamma_\xi, \tag{4.3}$$

so that the solution to (4.1) is given by

$$y(t) = e^{-\frac{t^2}{\xi^2}}. \tag{4.4}$$

To discretize this FIDE, we first note that the mean-value constraint in (4.3) can be enforced by observing that if $P_n(t)$ is a mapped-Legendre polynomial on $[0, T]$, then

$$\int_0^T P_n(t)\,\mathrm{d}t = \begin{cases} T, & n = 0, \\ 0, & \text{otherwise}, \end{cases} \tag{4.5}$$

so that here we require $[1, 0, 0, \ldots]\underline{y} = \gamma_\xi$. The additional complication in this example is the non-constant coefficient in front of the $y'(t)$ term. Since $D_{1/2} : \mathbf{P} \to \mathbf{C}^{3/2}$, the multiplication must take place in $\mathbf{C}^{3/2}$. To this end, we construct the mapped-Legendre coefficients of t on $[0, 1]$ (i.e., $\underline{t} = [\frac{1}{2}, \frac{1}{2}, 0, \ldots]$) and convert these to coefficients in $\mathbf{C}^{3/2}$ by multiplying with the conversion operator $S_{1/2}$ (i.e., $S_{1/2}\underline{t} = [\frac{1}{2}, \frac{1}{6}, 0, \ldots]$). The resulting multiplication operator $M_{3/2}[S_{1/2}\underline{t}]$ can then be constructed using the technique described in Hale & Olver (2017, Section 2.3). In this simple case it is given by

$$M_{3/2}[S_{1/2}\underline{t}] = \frac{1}{2}\begin{bmatrix} 1 & \frac{3}{5} & & & \\ \frac{1}{3} & 1 & \frac{4}{7} & & \\ & \frac{2}{5} & 1 & \frac{5}{9} & \\ & & \frac{3}{7} & 1 & \ddots \\ & & & \ddots & \ddots \end{bmatrix} = \frac{1}{2}(I + J_{3/2}), \tag{4.6}$$

where $J_{3/2}$ is given by Hale & Olver (2017, (2.15)) with $\lambda = 3/2$.[2] It remains then to compute the mapped-Legendre coefficients $\underline{f}$ of $f(t)$, $\underline{k}$ of the kernel $k(t) = e^{-t^2/2\sigma^2}$ and $\widehat{\underline{k}}$ of the flipped kernel $\widehat{k}(t) = e^{t^2/2\sigma^2}$ on the interval $[0, 1]$ so that the discretized version of (4.1) and (4.3) is given by

$$\begin{bmatrix} \begin{bmatrix} 1, & -1, & 1, & -1, & 1, & -1, & 1, & -1, & \ldots \\ 1, & 0, & 0, & 0, & 0, & 0, & 0, & 0, & \ldots \end{bmatrix} \\ D_{1/2}^2 + S_{3/2}\big(M_{3/2}[S_{1/2}\underline{t}]D_{1/2} + S_{1/2}(I + F[\underline{k}, \widehat{\underline{k}}])\big) \end{bmatrix} \underline{y} = \begin{bmatrix} \begin{bmatrix} 1 \\ \gamma_\xi \end{bmatrix} \\ S_{3/2}S_{1/2}\underline{f} \end{bmatrix}. \tag{4.7}$$

Figure 3 shows the result of solving this system when $\xi = 1/10$ and $\sigma = 1$. As usual, in the first panel we see that the discretization is almost-banded, here with only one dense row since the mean value constraint in (4.3) is sparse in $\mathbf{P}$. For this example, the bandwidth of the operator is determined by σ in the Gaussian kernel; making σ larger will result in a wider bandwidth, whilst decreasing it will make the

[2] Although in this instance we could have taken advantage of the relationship $\frac{\mathrm{d}}{\mathrm{d}t}(ty) = ty' + y$, resulting in $D_{1/2}M_{1/2}[\underline{t}]$ rather than $M_{3/2}[S_{1/2}\underline{t}]D_{1/2} + S_{1/2}$ in (4.7), we instead choose to multiply in $\mathbf{C}^{3/2}$ to demonstrate the more general situation.

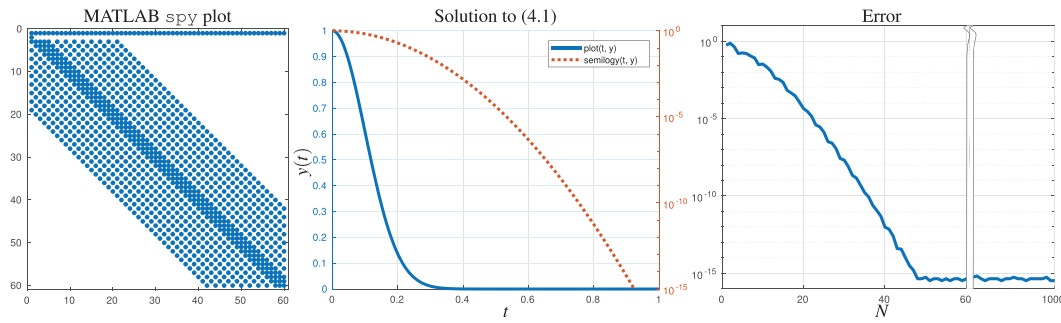

FIG. 3. Discretization, solution and convergence for the second-order FIDECT (4.1) with $\xi = 1/10$ and $\sigma = 1$. As in the previous examples, the almost-banded structure and exponential convergence of the solution are still obtained, even in the presence of nonconstant coefficients and more general boundary conditions.

bandwidth narrower. Conversely, the rate of convergence (right-most figure) depends only on the value of ξ; smaller ξ will require a higher N to resolve. We see in the rightmost panel that the convergence of the solution is again exponential, and that the conditioning of the linear system does not appear to degrade as the discretization size is increased.

Figure 4 shows the result of solving the same problem, but now using the methods of Driscoll (2010).[3] In particular, the first panel shows a MATLAB spy plot of the discretization using standard Chebyshev collocation and the approach given in (Driscoll, 2010, Section 2)—which is effectively the implementation used in Chebfun (Driscoll *et al.*, 2014). As expected, the discretized system is dense and unstructured, and hence solving the system will become computationally expensive when a large discretization is required. Furthermore, the plot of the error in the third panel (red line) shows that the error creeps upwards as the discretization size increases, indicative of the well-established ill-conditioning in the second order Chebyshev spectral collocation matrix. In the same paper Driscoll describes how this ill-conditioning can be reduced by reformulating the original IDE as an integral equation. As seen again in the third panel (yellow line) this certainly improves the conditioning, however the discretization is still dense (centre panel). We are hesitant to present execution times, since these may be highly dependent on the implementation, but one would typically expect a system with the structure in the left-most panel of Fig. 3 to be solved much more quickly than those in Fig. 4, even for moderate discretization sizes.

EXAMPLE 4.2 As our final example we consider, for any $\omega \in \mathbb{R}$ and non-negative integer $\mu > 0$, the second-order VIDECT

$$y''(t) + \omega^2 y(t) = f(t) - \omega \int_0^t J_\mu\big(\omega(t-s)\big) y(s)\,\mathrm{d}s, \quad 0 \leqslant t \leqslant T, \quad y(0) = y'(0) = 0, \tag{4.8}$$

where $f(t)$ on the right-hand side is taken as

$$f(t) = J_{\mu+\eta}(\omega t) + \frac{1}{2t^2}\big((\eta-1)(\eta-2)J_{\eta-1}(\omega t) + (\eta+1)(\eta+2)J_{\eta+1}(\omega t)\big) \tag{4.9}$$

[3] It is important to note that the methods described in (Driscoll, 2010) can be readily applied to arbitrary smooth kernels, and hence are far more general than the one described in the present manuscript. Our intent here is simply to show that the present method can be advantageous in the special case of smooth convolution-type kernels, such as in (4.1).

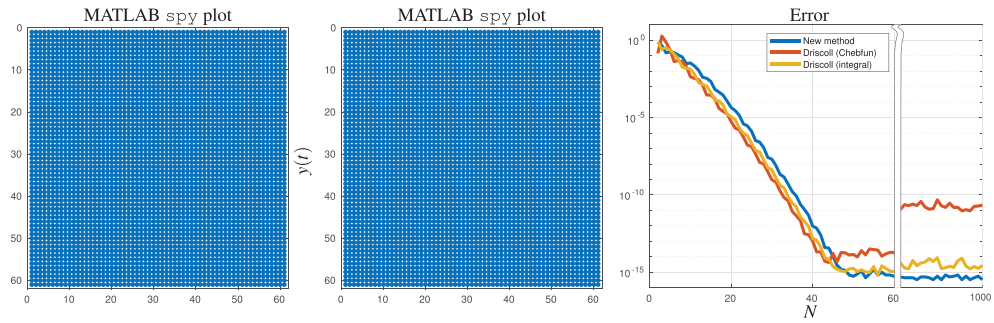

FIG. 4. Result of applying the Chebyshev spectral collocation methods described by Driscoll (2010) to (4.1). The left and centre panels show that the resulting discretizations are dense and unstructured when using standard Chebyshev collocation discretization of the second-order differentiation operator (left) and the integral reformulation (centre) suggested in (Driscoll, 2010, Section 3). The final panel shows the convergence as the discretization size is increased. The ill-conditioning of the standard collocation approach is evident, but this is mitigated by the integral reformulation.

so that for any integer $\eta \geqslant 3$ the exact solution is given by[4]

$$y(t) = \eta J_\eta(\omega t)/(\omega t). \tag{4.10}$$

Integral equations with Bessel function kernels arise often in the fields of scattering and potential theory (Xiang & Brunner, 2013), and since $J_\mu(\omega(t-s))$ is not degenerate, the Volterra operator in (4.8) cannot trivially be written as a linear combination of multiplication and integration operators.

We proceed as before, discretizing the differential operator $\mathscr{D}^2 + \omega^2\mathscr{I}$ using the Legendre US method for ODEs, and the convolution-type Volterra operator $\mathscr{V}[J_\mu(\omega t)]$ using the approach outlined in Section 3.1. The Dirichlet boundary condition $y(0) = 0$ can be enforced as in (3.14), and for the Neumann condition $y'(0) = 0$ we have that

$$y'(0) = \mathbf{P}'(0)\underline{y} = T\left[0,\ \ 2,\ \ -6,\ \ 12,\ \ -20,\ \ \dots\ \right]\underline{y} = 0. \tag{4.11}$$

The full discretization of (4.8) on the interval $[0,1]$ is then given by

$$\begin{bmatrix} \begin{bmatrix} 1, & -1, & 1, & -1, & 1, & \dots \\ 0, & 2, & -6, & 12, & -20, & \dots \end{bmatrix} \\ D^2_{1/2} + S_{3/2}S_{1/2}(\omega^2 I + \omega V[\underline{J}_{\mu,\omega}]) \end{bmatrix} \underline{y} = \begin{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \underline{f} \end{bmatrix}, \tag{4.12}$$

where as usual $\underline{f}$ and $\underline{J}_{\mu,\omega}$ are the Legendre coefficients of $f(t)$ and $J_\mu(\omega t)$ on $[0,1]$, respectively.

[4] In fact, if the initial conditions are changed to $y(0) = \delta_{1,\eta}/2$, $y'(0) = \omega\delta_{2,\eta}/4$ then the result holds also for $\eta = 1$ and $\eta = 2$. The solution to this VIDECT was reverse engineered using the differential equation satisfied by Bessel function (Olver *et al.*, 2010, 10.2.1) and the known convolution of two Bessel functions (Olver *et al.*, 2010, 10.22.34).

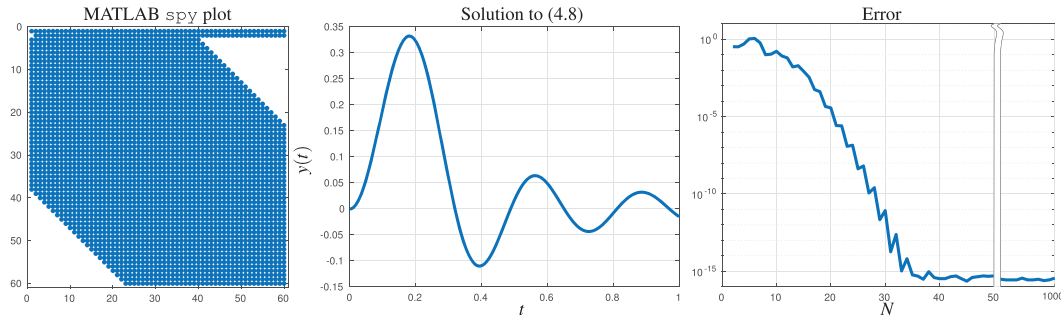

FIG. 5. Discretization, solution and convergence for the second-order VIDECT (4.8) with $\nu = 3$, $\mu = 2$ and $\omega = 20$. Exponential convergence is obtained (right-most panel) to the Bessel function-like solution (centre panel). However, although almost-banded as $N \to \infty$, this structure is not utilized in practice for this example since the frequency of the solution is the same as that kernel (left-most panel). One can construct an example for which this is not the case by adding a singular perturbation to the left-hand side (see Fig. 6).

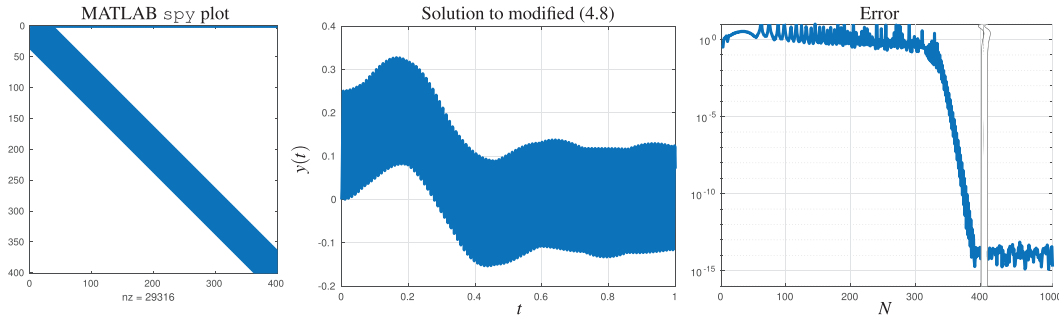

FIG. 6. As in Fig. 5, but with the left-hand side of (4.8) modified to $0.001y''(t) + \omega^2 y(t)$. The singular perturbation makes the solution far more oscillatory (centre), and the banded nature of the discretization is utilized (left). Here the underlying solution is not known, so the error is computed against a discretization with $N = 2000$. Once the oscillations are resolved, convergence is again exponential (right).

Figure 5 shows the discretization, solution and error when solving (4.8) with $\eta = 3, \mu = 2$ and $\omega = 20$. As in the previous examples, the first panel shows the almost-banded structure of the discretization, here with two dense rows arising from the two initial conditions. Here the bandwidth is determined by the frequency ω in the kernel which, by construction, is the same as the frequency of the solution, and so the banded nature of the operator (in the limit as $N \to \infty$) is not utilized in practice for this example. However, one can readily imagine related problems which require larger values of N to solve (such as introducing small perturbation parameter in front of the y'' term or adding an oscillatory or less-smooth term to $f(t)$) that will not affect the structure of the operator. Figure 6 shows such a situation, where the left-hand side of (4.8) is modified to $0.001y''(t) + \omega^2 y(t)$. This singular perturbation causes the solution to become far more oscillatory, and hence far more degrees of freedom are required to resolve it. In this case, the banded nature of the discretization is evident, and utilized. In both cases, exponential convergence and high accuracy is observed to around the level of machine precision.

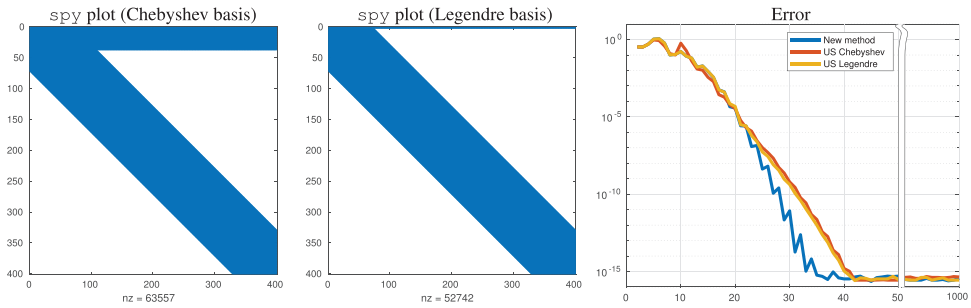

FIG. 7. Left and centre: $N = 400$ discretizations of the VIDECT (4.8) using the separable kernel approach described in Section 2.4 with a Chebyshev (left) and Legendre (centre) basis. Comparing with the leftmost panel of Fig. 6, we see that the bandwidth and number of nonzero entries in the matrix has approximately doubled in both cases. Furthermore, in the case of the Chebyshev basis, the number of dense rows at the top of the matrix has also increased significantly. Right: Convergence of these approaches when applied to (4.8). As expected, the convergence is exponential, but slightly slower than the convolution-kernel approach introduced in Section 3.

Finally, we compare the present approach to that proposed in Section 2.4 (which is essentially the approach used by ApproxFun (Olver *et al.*, 2017) as introduced in (Slevinsky & Olver, 2017) when solving (4.8).[5] In particular, we use ultraspherical discretizations of the differentiation operator and employ Chebfun2 (Townsend & Trefethen, 2013) to efficiently form a separable low-rank approximation to the kernel $k(s,t) = J_\mu(w(t-s))$. When $\mu = 2$ and $\omega = 20$, Chebfun2 determines a rank 17 approximation that is sufficient to approximate $k(s,t)$ to machine precision, and the Volterra operator in (4.8) can be then approximated as in (2.16). Figure 7 shows the resulting discretizations when using ultraspherical discretizations in the Chebyshev T basis (left) and in the Legendre basis (centre). Comparing to the discretization of the new method depicted in the leftmost panel of Fig. 6, we notice two things: First, although the discretizations are again almost banded, the bandwidth in these two cases is approximately double that of the present method. This is to be expected since the bandwidth of the two multiplication operators in (2.16) are comparable to each other (and the bandwidth of $V[\underline{k}]$ in Section 3.1), and are pre- and post-multiplying the integration operators Q_0 and $Q_{1/2}$. Secondly, we note that in the case of the Chebyshev discretization the number of dense rows at the top of the matrix has similarly increased, due to the fact that Q_0 itself is almost banded. ($Q_{1/2}$ is banded, so this effect is not present in the Legendre case.) Therefore, not only will these discretizations take longer to solve (approximately four times longer in the Legendre case, and more for the Chebyshev), but experiments suggests that the dominant cost of the computation in this approach is in actually forming the discretization of V via the sum in (2.16), rather than the cost of solving the resulting system itself. As in the previous example, we are hesitant to state precise timings, since the implementations of these methods are far from optimal, but to form and solve the systems using the low-rank kernel approximation takes around 5–10 times longer than the present method. Finally, the rightmost panel suggests that, although the rate of convergence of each of the methods considered is exponential, it appears to be

[5] As with the comparison to Driscoll's method with Example 4.1, it is important to note that the ultraspherical and low-rank kernel approximation used in Approxfun and described in (Slevinsky & Olver, 2017) is far more general than the present approach. It is not our intent to show the present method is superior to the existing approach in general, rather that the present approach may be advantageous for the special case of smooth convolution-type kernels.

marginally faster for the present method. However, we have not investigated this further. We conclude that the present method offers some advantages over the separable kernel approximation method in the case of convolution kernels.

5. Conclusion

We have shown that the linear operators representing the action of Fredholm and Volterra integral operators on Legendre series expansions can be combined with the Legendre US method for ordinary differential equations to efficiently solve IDEs with convolution-type kernels. It was demonstrated that if the kernel, coefficient functions, and right-hand side are sufficiently smooth (analytic) then the rate of convergence is very fast (exponential down to machine precision). If the kernel and coefficient functions are polynomials (or sufficiently well approximated by a polynomial) then the discretizations result in banded or almost-banded matrices which can be solved with linear complexity. In the final section we compared the approach to some existing spectral methods for IDEs from the literature and found that it performed favourably in terms of the sparsity and conditioning of the resulting discretization.

Volterra equations typically correspond to initial value problems, and hence most numerical techniques use some form of time-marching. However, the exponential convergence and linear complexity of the algorithm presented in this paper should make it competitive with such approaches in many cases. Fredholm integral operators are like boundary-value problems and a global discretization is more natural.

We presented here the simplest version of this algorithm, namely linear problems in one dimension. The approach readily extends to time-dependent problems, systems of equations, and higher-dimensions, analogously to the ultraspherical method for differential equations (Olver & Townsend, 2014). Nonlinear problems may be solved by a Newton–Raphson-like approach, but in such situations the banded structure (although not the well-conditioned) property of the discretization is lost. A linear complexity method for nonlinear problems, even for ODES, remains elusive.

Acknowledgements

I am grateful to Sheehan Olver (Imperial College London), Alex Townsend (Cornell) and André Weideman (Stellenbosch) for useful discussions and to the anonymous referees for their informative comments and suggestions.

Funding

National Research Foundation of South Africa (Grant Number 109210).

References

Alpert, B. K. & Rokhlin, V. (1991) A fast algorithm for the evaluation of Legendre expansions. *SIAM J. Sci. Statist. Comput*., **12**, 158–179.

Cont, R. & Tankov, P. (2004) *Financial modelling with jump processes*, Boca Raton, FL: Chapman & Hall/CRC.

Cox, D. R. (1962) Renewal theory. London: Methuen; New York: John Wiley.

Driscoll, T. A. (2010) Automatic spectral collocation for integral, integro-differential, and integrally reformulated differential equations. *J. Comput. Phys.*, **229**, 5980–5998.

Driscoll, T. A., Hale, N. & Trefethen, L. N. (2014) *Chebfun Guide*. Oxford, UK: Pafnuty Publications.

El-gendi, S. E. (1969/1970) Chebyshev solution of differential, integral and integro-differential equations. *Comput. J.*, **12**, 282–287.

Hale, N. (2017) An ultraspherical spectral method for linear Fredholm and Volterra integro-differential equations of convolution type. *Companion code to this paper.* Available at https://github.com/nickhale/usconv.

Hale, N. & Olver, S. (2017) A fast and spectrally convergent algorithm for rational-order fractional integral and differential equations,. *SIAM J. Sci. Comput.* (to appear).

Hale, N. & Townsend, A. (2014a) An algorithm for the convolution of Legendre series. *SIAM J. Sci. Comput.*, **36**, A1207–A1220.

Hale, N. & Townsend, A. (2014b) A fast, simple, and stable Chebyshev–Legendre transform using an asymptotic formula. *SIAM J. Sci. Comput.*, **36**, A148–A167.

Jackiewicz, Z., Rahman, M. & Welfert, B. D. (2006) Numerical solution of a Fredholm integro-differential equation modelling neural networks. *Appl. Numer. Math.*, **56**, 423–432.

Kress, R. (1999) *Linear Integral Equations*, vol. 82, 2nd edn. New York: Springer.

Kuang, Y. (1993) *Delay Differential Equations with Applications in Population Dynamics*, vol. 191. Boston, MA: Academic Press.

Lanczos, C. (1956) *Applied Analysis.* Englewood Cliffs, NJ: Prentice Hall.

Linz, P. (1985) *Analytical and Numerical Methods for Volterra Equations*, vol. 7. Philadelphia, PA: SIAM.

Ma, J. & Brunner, H. (2006) A posteriori error estimates of discontinuous Galerkin methods for non-standard Volterra integro-differential equations refpunct. *IMA J. Numer. Anal.*, **26**, 78–95.

MacCamy, R. C. (1977) An integro-differential equation with application in heat flow. *Quart. Appl. Math.*, **35**, 1–19.

Mason, J. C. & Handscomb, D. C. (2003) *Chebyshev Polynomials*. Boca Raton, FL: Chapman & Hall/CRC.

Medlock, J. & Kot, M. (2003) Spreading disease: integro-differential equations old and new. *Math. Biosci.*, **184**, 201–222.

Olver, F. W. J., Lozier, D. W., Boisvert, R. F. & Clark, C. W. (eds) (2010) *NIST Handbook of Mathematical Functions*. New York: Cambridge University Press.

Olver, S., Goretkin, G., Slevinsky, R. M. & Townsend, A. (2017) *Approx Fun.* Available at https://github.com/ApproxFun/ApproxFun.jl.

Olver, S. & Townsend, A. (2013) A fast and well-conditioned spectral method. *SIAM Rev.*, **55**, 462–489.

Olver, S. & Townsend, A. (2014) A practical framework for infinite-dimensional linear algebra. *Proceedings of the 1st First Workshop for High Performance Technical Computing in Dynamic Languages*. New Orleans, LA: IEEE Press, pp. 57–62.

Orszag, S. A. (1971a) Accurate solution of the Orr–Sommerfeld stability equation. *J. Fluid Mech.*, **50**, 689–703.

Orszag, S. A. (1971b) Numerical simulation of incompressible flows within simple boundaries. I. Galerkin (spectral) representations. *Studies Appl. Math.*, **50**, 293–327.

Ortiz, E. L. (1969) The tau method. *SIAM J. Numer. Anal.*, **6**, 480–492.

Slevinsky, R. M. & Olver, S. (2017) A fast and well-conditioned spectral method for singular integral equations. *J. Comput. Phys.*, **332**, 290–315.

Tang, T., Xu, X. & Cheng, J. (2008) On spectral methods for Volterra integral equations and the convergence analysis. *J. Comput. Math.*, 825–837.

Townsend, A. (2014) Computing with functions in two dimensions. *Ph.D. Thesis*, University of Oxford.

Townsend, A. & Trefethen, L. N. (2013) An extension of Chebfun to two dimensions. *SIAM J. Sci. Comput.*, **35**, C495–C518.

Townsend, A., Webb, M. & Olver, S. (2018) Fast polynomial transforms based on Toeplitz and Hankel matrices. *Math. Comp.*, **87**, 1913–1934.

Trefethen, L. N. (2013) *Approximation Theory and Approximation Practice*. Philadelphia, PA: SIAM.

Trefethen, L. N. & Trummer, M. R. (1987) An instability phenomenon in spectral methods. *SIAM J. Numer. Anal.*, **24**, 1008–1023.

Wang, H. & Xiang, S. (2012) On the convergence rates of Legendre approximation. *Math. Comp.*, **81**, 861–877.

Xiang, S. & Brunner, H. (2013) Efficient methods for Volterra integral equations with highly oscillatory Bessel kernels. *BIT*, **53**, 241–263.