

RELIABLE APPROXIMATE SOLUTION OF SYSTEMS OF VOLTERRA INTEGRO-DIFFERENTIAL EQUATIONS WITH TIME-DEPENDENT DELAYS

M. SHAKOURIFAR [†] AND W. H. ENRIGHT [†]

Abstract. Volterra integro-differential equations with time-dependent delay arguments (DVIDEs) can provide us with realistic models of many real-world phenomena. Delayed Lokta-Volterra predator-prey systems arise in Ecology and are well-known examples of DVIDEs first introduced by Volterra in 1928. We investigate the numerical solution of systems of DVIDEs using an adaptive stepsize selection strategy. We will present a generic variable stepsize approach for solving systems of neutral DVIDEs based on an explicit continuous Runge-Kutta method using defect error control and study the convergence of the resulting numerical method for various kind of delay arguments. We will show that the global error of the numerical solution can be effectively and reliably controlled by monitoring the size of the defect of the approximate solution and adjusting the stepsize on each step of the integration. Numerical results will be presented to demonstrate the effectiveness of this approach.

Key words. Continuous Runge-Kutta Methods, Delay Volterra integro-differential equations, Global Error, Defect Control, Adaptive Stepsize Selection

AMS subject classifications. 65R20 (65L06)

1. Introduction. Significant progress has been achieved in the qualitative understanding and the numerical solution of delay differential equations in recent years. This has increased the use of this class of equations for simulations in various fields such as Biology, Medicine, Chemistry, Financial Mathematics and Engineering.

Ordinary and partial differential equations are often derived as a first approximation to model a real-world situation, where the state of the system depends not only on the present time, but also on the history of the system. In this situation delay Volterra integro-differential equations (denoted DVIDEs) can provide a more realistic model. In particular, they play an important role in mathematical modeling of population dynamics phenomena. Delayed Lokta-Volterra predator-prey systems are the basis of many models in population biology. The dynamic of two interacting species was first modeled by Volterra [1] as,

$$(1) \quad \begin{aligned} N_1'(t) &= N_1(t) \left(\epsilon_1 - \gamma_1 N_2(t) - \int_{t-\tau}^t F_1(t-s) N_2(s) ds \right), \\ N_2'(t) &= N_2(t) \left(-\epsilon_2 + \gamma_2 N_1(t) + \int_{t-\tau}^t F_2(t-s) N_1(s) ds \right), \end{aligned}$$

where $t \in I := [0, T]$, $\epsilon_i > 0$, $\gamma_i \geq 0$, $F_i(t) \geq 0$ is continuous, and $N_1(t) = \phi_1$, $N_2(t) = \phi_2$ for $t \leq 0$. $N_1(t)$ and $N_2(t)$ represent two populations (prey and predator) at time t (see for example [2] for more details). For a comprehensive bibliography see [3]. Also, De Gaetano and Arino [4] introduced a delay integro-differential equation for use in the glucose-insulin regulatory system in relation to diabetes,

$$\begin{aligned} G'(t) &= -b_1 G(t) - b_4 I(t) G(t) + b_7, \\ I_2'(t) &= -b_2 I(t) + \frac{b_6}{b_5} \int_{t-b_5}^t G(s) ds, \end{aligned}$$

[†]Department of Computer Science, University of Toronto, Toronto, ON, Canada M5S 3G4 (shakouri@cs.toronto.edu, enright@cs.toronto.edu).

where $G(t)$ denotes blood glucose concentration at time t and $I(t)$ is representing insulin blood concentration. For more details see [5].

This paper is concerned with designing, analyzing and implementing an efficient method to approximate the solution of a general system of neutral Volterra integro-differential equations with time dependent delay arguments using a continuous Runge-Kutta (CRK) method. CRK methods were introduced for initial value problems (IVPs), and they determine an approximation to the solution of an IVP for any t in the interval of interest. The accuracy of the continuous approximation is consistent with the accuracy of the underlying discrete Runge-Kutta formula which generates approximations only at the discrete mesh points. The numerical stability of CRK methods for delay differential equations has been investigated in [6, 7].

We consider neutral VIDEs with a time dependent delay $\tau(t)$,

$$(2) \quad y'(t) = f(t, y(t)) + \int_{t-\tau(t)}^t K(t, s, y(s), y'(s)) ds,$$

for $t \in [t_0, T]$, $f : R \times R^m \rightarrow R^m$ and $K : R \times R \times R^m \times R^m \rightarrow R^m$. To make the problem well defined, a unique solution $y(t)$ is usually identified by specifying an initial function $\phi(t)$ for $t \leq t_0$, with $\phi(t) : R \rightarrow R^m$.

The approximate solution of DVIDEs has been studied by several authors. Linear multistep methods for DVIDEs were studied in [8]. Collocation methods were investigated in [9, 10] for strictly increasing delay arguments and block-by-block methods investigated in [11] for DVIDEs with smooth solutions using a fixed stepsize strategy. Brunner investigated the numerical solution of nonlinear VIDEs with infinite delay in [12] and neutral VIDEs with constant delay in [13] and showed that the order of accuracy at the discrete mesh points can be $2\bar{m}$ when using a method based on collocation at the \bar{m} Gauss (-Legendre) points. Numerical methods have been investigated on the basis of both uniform and adaptive stepsize selection. In the adaptive stepsize case the nature of the solution can have an important effect on the performance. Collocation approximate solutions of nonlinear systems of VIDEs with delay arguments of the general type,

$$(3) \quad \begin{aligned} N_1'(t) &= f_1(N_1(t), N_2(t)) + \int_{t-\tau}^t F_1(t-s)G_1(N_1(t), N_2(s))ds, \\ N_2'(t) &= f_2(N_1(t), N_2(t)) + \int_{t-\tau}^t F_2(t-s)G_2(N_2(t), N_1(s))ds, \end{aligned}$$

which includes system (1) as a special case were analyzed in [14] based on a constant stepsize strategy. Discrete RK methods and their numerical stability for VIDEs with constant delay have been investigated in [15, 16, 17] assuming a fixed stepsize strategy. These methods are mainly extensions of the classical Pouzet and Belyukov RK methods where the discretization is a part of the method and no continuous approximation is produced.

In contrast to local error control strategy which attempts to estimate and bound the local error on each step, defect control is defined for methods such as those based on a CRK formula, and attempts to estimate and bound the magnitude of the defect of the associated approximation (for all t) on each step of the integration. Regardless of the particular numerical method used for approximating the solution of (2), the achieved accuracy of a reliable numerical method should depend primarily on the prescribed user defined tolerance and the conditioning of the problem. Enright introduced and analyzed the error and stepsize control for CRK methods applied to IVPs

based on defect monitoring [18, 19]. With his colleagues he has recently identified a new class of explicit CRK formulas in which the defect over a timestep has a consistent shape which is problem independent and easy to estimate [20, 21].

In this investigation, we extend this approach to VIDEs with time dependent delay arguments. It is well known that discontinuities may occur in various orders of the derivative of the solution associated with delay problems (if the solution is C^p at point ξ , then the discontinuity point ξ has order p). The sequence of discontinuity points $\{\xi_\mu : \mu \geq 0\}$ can be characterized in general by a recursion,

$$(4) \quad \xi_{\mu+1} - \tau(\xi_{\mu+1}) = \xi_r \quad \text{for some } 0 \leq r \leq \mu, \quad \xi_0 = t_0.$$

It is worth mentioning that the delay argument is not necessarily increasing in our investigation and can be of either vanishing or non-vanishing type. Brunner and Zhang [22] have thoroughly investigated the regularity and smoothing properties of scalar DVIDEs for integral and integro-differential equations with various kind of delays. It is shown that, unlike the solutions of neutral delay differential equations, smoothing indeed does happen for the solutions of neutral DVIDEs. This increase in regularity results in avoiding the clustering of low order discontinuity points in case of vanishing delays (see problem 2). In addition, Baker and Willé [23], investigated the propagation of discontinuities in the solutions of scalar and systems of DVIDEs. Guglielmi and Hairer have discussed how to accurately compute these breaking points in [24]. The use of arbitrary meshes will in general result in a reduction in the order of accuracy due to the presence of these discontinuities. Therefore, we attempt to detect discontinuity points during the integration process and ensure these points are forced to be mesh points so as not to contaminate the order of the underlying CRK formula. As we do not restrict the delay argument to be increasing, we must keep track of all discontinuities in the integration. One of the main results that we establish is that the global error of the numerical solution can be effectively and reliably controlled by directly monitoring the magnitude of the associated defect. Since the definite integrals arising in the underlying equations defining our approximate solution can not in general be calculated analytically, we will explore the convergence properties of an approximate solution to (2) that uses a numerical quadrature scheme to approximate these integrals.

In the next section we will present an overview of how a generic CRK method can be applied to approximate the solution of (2) and describe different strategies that can be used to control the size of the defect of the associated interpolants. In section §3 we investigate the global convergence properties of this Runge-Kutta approach when applied to (2). We show that the global error is proportional to the norm of the defect. In the fourth section, numerical results for our implementations will be presented which illustrate the effectiveness of the approach. In the final section we discuss the significance and limitations of this approach.

2. Continuous Runge-Kutta Methods. Discrete numerical methods for IVPs introduce a set of approximations $\{y_0, y_1, \dots, y_N\}$ corresponding to the mesh points $a = t_0 < t_1 < \dots < t_N = b$. Obtaining these approximations is done using an underlying discrete approximation formula. An s -stage, p^{th} order, explicit Runge-Kutta formula when applied to the standard IVP,

$$(5) \quad y' = f(t, y), \quad y(a) = y_0, \quad \text{for } t \in [a, b],$$

determines,

$$(6) \quad y_{n+1} = y_n + h_n \sum_{i=1}^s w_i k_i,$$

where,

$$(7) \quad k_i = f(t_n + c_i h_n, y_n + h_n \sum_{j=1}^{i-1} a_{i,j} k_j), \quad i = 1, \dots, s.$$

We compute the approximation y_{n+1} , if (t_n, y_n) is known, as an explicit computation requiring only s evaluations of the differential equation. To derive an optimal order CRK associated with the discrete formula (6) additional stages, $k_{s+1} = f(t_{n+1}, y_{n+1})$, $k_{s+2}, \dots, k_{\bar{s}+1}$ are used to obtain an approximation for any $t \in (t_n, t_{n+1})$ as,

$$(8) \quad u_n(t) = y_n + h_n \sum_{j=1}^{\bar{s}+1} b_j(\tau) k_j,$$

where $b_j(\tau)$ is a polynomial of degree at most $p+1$, $\tau = \frac{t-t_n}{h_n}$, $b_j(\tau) = \sum_{k=1}^{p+1} \beta_{j,k} \tau^k$, and $u_n(t)$ agrees with $y_n(t)$ to $O(h_n^{p+1})$, where $y_n(t)$ is the solution of the local IVP,

$$(9) \quad y'_n(t) = f(t, y_n(t)), \quad y_n(t_n) = u_n(t_n).$$

A *Butcher Tableau* (Table 2.1) can be used to represent the resulting CRK formula. Different criteria have been introduced to define a suitable continuous extension,

TABLE 2.1
Butcher Tableau for a CRK formula

0	0							
c_2	$a_{2,1}$	0						
\vdots								
c_s	$a_{s,1}$	$a_{s,2}$	\dots	$a_{s,s-1}$	0			
1	w_1	w_2	\dots	w_{s-1}	w_s	0		
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots		
$c_{\bar{s}+1}$	$a_{\bar{s}+1,1}$	$a_{\bar{s}+1,2}$	\dots	\dots	\dots	\dots	$a_{\bar{s}+1,\bar{s}}$	0
	w_1	w_2	\dots	w_{s-1}	w_s	0	\dots	0
	$b_1(\tau)$	$b_2(\tau)$	\dots	\dots	\dots	\dots	$b_{\bar{s}}(\tau)$	$b_{\bar{s}+1}(\tau)$

and the parameters defining such a CRK (represented in Table 2.1) are usually chosen to ensure that the piecewise polynomial $u(t)$, defined by the mesh $t_0 < t_1 < \dots < t_N$ and the associated polynomials $u_n(t)$ $n = 0, 1, \dots, (N-1)$, is in $C^1[a, b]$. Enright and Hu [25] introduced a compatible definition for each k_i and used the associated CRK to define an approximation formula that can be applied to (2) with constant delay and smooth solutions. We will follow the same approach to define k_i for (2) with time-dependent delay arguments. To advance from t_n to t_{n+1} after $u(t)$ has been

defined for $t \leq t_n$, $u_n(t)$ is defined for $t \in [t_n, t_{n+1}]$ by (8) where the k_i 's are defined not by (7), but by the following system of equations,

$$(10) \quad k_i = f(t_n + c_i h_n, y_n + h_n \sum_{j=1}^{i-1} a_{i,j} k_j) + \int_{t_n + c_i h_n - \tau(t_n + c_i h_n)}^{t_n + c_i h_n} K(t_n + c_i h_n, s, u(s), u'(s)) ds,$$

for $i = 1, 2, \dots, \bar{s} + 1$.

Note that regardless of the location of the lower bound in the integral in (10), for any CRK, we have a system of $(\bar{s} + 1)m$ coupled implicit equations (defining the k_i 's), since $u_n(t)$ is defined in terms of all the k_i 's introduced on this step. It is worth mentioning that if the k_i 's satisfy a CRK formula given by Table 2.1, k_{s+1} on step n is the same as k_1 on step $n + 1$. This reduces the size of the implicit set of equations to be solved on each step.

For $t \in [t_n, t_{n+1}]$ the CRK interpolant $u_n(t)$ (defined by (8) and (10)) has an associated defect or residual,

$$\delta_n(t) = u'_n(t) - f(t, u_n(t)) - \int_{t-\tau(t)}^t K(t, s, u(s), u'(s)) ds.$$

To analyze the accuracy and convergence of this piecewise polynomial, $u(t)$, it is convenient to introduce the local error associated with each step. Consider $z_n(t)$ to be the exact solution of the 'local' IVP on step n ,

$$(11) \quad z'_n(t) = f(t, z_n(t)) + \int_{t-\tau(t)}^t K(t, s, u(s), u'(s)) ds \quad , \quad z_n(t_n) = y_n.$$

where $u(t)$ is the piecewise polynomial already computed for $t \leq t_n$ and $u(t) \equiv u_n(t)$ (which is well-defined in terms of $k_1, k_2, \dots, k_{\bar{s}+1}$, but not necessarily computable) for $t \in [t_n, t_{n+1}]$. If the standard ODE CRK formula (6),(7) and (8) is applied to this local IVP problem, we will get the same piecewise polynomial $u(t)$ (as its continuous approximate solution); therefore, we can conclude from the theory of optimal order CRK methods applied to IVPs and the global order of convergence of the approximate solution that the associated defect satisfies [20],

$$\delta_n(t) = G(\tau)h_n^p + O(h_n^{p+1}),$$

where

$$(12) \quad G(\tau) = q_1(\tau)F_1 + q_2(\tau)F_2 + \dots + q_K(\tau)F_K.$$

The q_i 's are known polynomials in τ that depend only on the CRK formula, and the F_i 's are constants depending on both the CRK formula and the problem (Each of the F_i 's are expressions involving the derivatives of the problem evaluated at the point t_n and are referred to as *elementary differentials* . See [19] for more details). Note that we assume that h_n is small enough to ensure sufficient differentiability for the right hand side of the local IVP (11). Methods for IVPs have been designed and implemented which attempt to ensure that $\|\delta_n(t)\| \leq TOL$ for $t_n \leq t \leq t_{n+1}$, for some norm $\|\cdot\|$, on each step of the integration. It is seen from the above relation that as $h_n \rightarrow 0$ the defect will behave like a linear combination of the q_i 's over the subinterval $[t_n, t_{n+1}]$. This property allows the defect to be monitored and controlled

in an efficient and reliable manner.

Enright and Yan [21] introduced a new, more restrictive, class of CRK interpolants for the standard IVP (5) for which (12) had the special form,

$$\delta_n(t) = u'_n(t) - f(t, u_n(t)) = q_1(\tau)F_1 h_n^p + O(h_n^{p+1}) \quad , \quad t \in [t_n, t_{n+1}].$$

In this case, $u_n(t)$ is generally a more accurate interpolant in approximating the local solution $y_n(t)$ in (9). In addition, the shape of the defect (as $h_n \rightarrow 0$) will be independent of both problem and the step, and the maximum defect should occur (as $h_n \rightarrow 0$) at the location in $[0, 1]$ of the local maximum of $|q_1(\tau)|$ (known *à priori*). They referred to the use of this class of new interpolants and the reliable estimate of the maximum defect on a step as a strict defect control (SDC) strategy. They also introduced a validity check that reflects the credibility of this estimate. The resulting defect control strategy is an improvement of earlier defect control strategies investigated in Enright and Hayes and referred to as relaxed defect control (RDC) schemes. We will investigate and implement an SDC and an RDC CRK method for (2). Since the shape and magnitude of the defect depends on the algorithm used to compute the approximate solution $u_n(t)$, the shape of the defect for DVIDEs is generally different from that of the ODE cases. This is because of various additional sources of error involved with the numerical solution of DVIDEs that will be discussed in subsequent sections. One should note that the solution of the implicit system of equations arising on each step of the integration, the detection and inclusion of the propagated discontinuities as mesh points, and the approximation of the quadratures that arise are important components of the method which affect the overall accuracy and efficiency of the approach. We will discuss these components and how they each contribute to the observed defect and global error in the following sections. Although we have specifically investigated the generalization of a new class of explicit CRK methods to DVIDEs, the same approach can be carried out for implicit CRK methods as well. While implicit methods are generally thought to be inefficient for non-stiff ODEs, in the case of VIDEs, we eventually end up with an implicit system of equations on each step of the integration (even starting with an explicit CRK method for ODEs). As a result, implicit methods may result in a suitable VIDE method. We are planning to consider this idea in a future investigation.

3. Global Error Bound and Convergence Results. Suppose that the CRK formulae (6), (8) and (10) define a continuous piecewise polynomial interpolant $u(t) \in C^1[a, b]$. We are interested in the optimal order of accuracy associated with $u(t)$, assuming that the CRK is of order p when applied to standard IVPs. In fact, we are looking for the largest value of $q \geq 1$ for which the following error bound is satisfied,

$$\sup_{t \in [a, b]} \|y(t) - u(t)\| = O(H^q),$$

where $y(t)$ is the exact solution of (2) and $H := \max_n h_n$. In order to establish our convergence results, we assume that $y(t)$ exists and that f and K are sufficiently smooth functions over their respective domains and satisfy the following Lipschitz conditions,

$$(13) \quad \begin{aligned} \|f(t, y_1) - f(t, y_2)\| &\leq L_f \|y_1 - y_2\|, \\ \|K(t, s, y_1, z_1) - K(t, s, y_2, z_2)\| &\leq L_y \|y_1 - y_2\| + L_{y'} \|z_1 - z_2\|. \end{aligned}$$

Note that our assumption of global Lipschitz condition (13) is very strong and will not likely be satisfied by problems arising in practice. The assumptions can be replaced

by locally Lipschitz assumptions (where it is assumed that each inequality of (13) is satisfied for all $y_1(t)$ and $y_2(t)$ in a neighborhood of the solution $y(t)$ and all $z_1(t)$ and $z_2(t)$ in a neighborhood of $y'(t)$). This is analogous to the standard approach used to analyze the convergence and accuracy of IV methods, where a rigorous analysis is presented for problems satisfying a global Lipschitz condition with an understanding that the results can be established for problems satisfying a local Lipschitz condition. Enright and Hu [25] proved that the method converges assuming that the delay argument τ is constant, the exact solution is smooth throughout the integration interval, the integrals are computed analytically, and $\tau L_{y'} < 1$. This latter requirement is quite strong but is typical of a condition required to guarantee that (2) has a unique solution (see, for example, [17, 26]). We will present a convergence theorem which applies to the more general time-dependent class of problems (with arbitrary delay arguments) and also accounts for discontinuity points and the quadrature errors associated with the evaluation of (10) and $\delta(t)$. We will assume that the delay argument is continuous over $[a, b]$ and $\bar{B}L_{y'} < 1$, where $\bar{B} := \max_{t \in [a, b]} |\tau(t)|$.

THEOREM 3.1. *Let f and K have sufficient continuous derivatives on their respective domains and be such that for a given initial function $\phi(t)$, (2) has unique solution $y(t)$. Assume*

- (a) $u(t) \in C^1[a, b]$ denotes an optimal order CRK specified by Table 2.1 and defined by (8) and (10),
- (b) all discontinuity points of order ≤ 1 are included in the set of mesh points,
- (c) $\bar{B}L_{y'} < 1$.

Then

$$\sup_{t \in [a, b]} \|y^{(l)}(t) - u^{(l)}(t)\| = O(H), \quad l = 0, 1.$$

Proof. This result follows directly from the theory of CRKs applied to IVPs and the observation that $u(t)$ satisfies the associated IVP (10), (11). \square

The assumption of this theorem, which is not likely to be satisfied when this approach is implemented, is that the quadratures appearing in (10) can be evaluated analytically. These definite integrals will have to be approximated numerically by a suitable numerical quadrature rule when this approach is applied to DVIDES. We know that the CRK approach applied to the local IVP (11) defines the piecewise interpolant $u(t)$. If, in the definition of k_i 's in (10), we replace the integral operator \int by a suitable numerical approximation, symbolically denoted by \sum , we will obtain a *fully discretized CRK interpolant*, denoted $\tilde{u}(t)$, defined for $t \in [t_n, t_{n+1}]$ by $\tilde{u}_n(t)$. By the triangular inequality we have,

$$\|y(t) - \tilde{u}(t)\| \leq \|y(t) - u(t)\| + \|u(t) - \tilde{u}(t)\|$$

We already know from the theory of perturbed IVPs and Theorem 3.1 that $\|y(t) - u(t)\| = O(H)$. Therefore, all we need is to investigate the convergence and accuracy of $\|u(t) - \tilde{u}(t)\|$. We investigate the classical notion of convergence (as $H \rightarrow 0$) as well as the adaptive (variable step) interpretation (as $TOL \rightarrow 0$). We begin with establishing an $H \rightarrow 0$ result by deriving a bound on the global error and showing that this bound tends to zero as $H \rightarrow 0$. This bound makes no assumption about the local error and stepsize control and can be very pessimistic as a predictor of the achieved global error. We will also establish an optimal order of convergence result (where we show that the global error is $O(H^p)$ as $H \rightarrow 0$ under some mild assumptions). Our

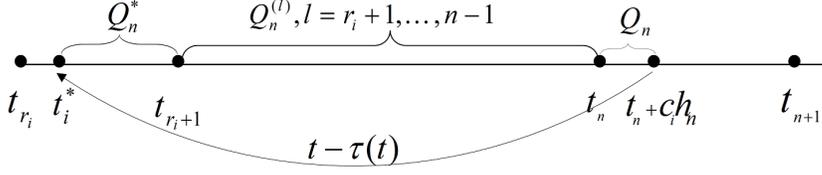


FIG. 3.1. Schematic representation of the partitioning associated with the quadrature

$TOL \rightarrow 0$ convergence result assumes that the defect on each step is monitored, and the method adjust the stepsize in an attempt to bound the magnitude of the defect by TOL on each step.

Consider a mesh point t_n and the parameter $0 \leq c_i \leq 1$ associated with the definition of k_i on the step from t_n to t_{n+1} . To simplify the notation for specifying the upper and lower bounds of the integral term appearing in (10) we introduce (for each $i = 1, 2, \dots, \bar{s} + 1$), $t_i^* = t_n + c_i h_n - \tau(t_n + c_i h_n)$:

$$t_{r_i} \leq t_i^* < t_{r_i+1}.$$

That is, the delayed value associated with the definition of k_i lies in the interval between t_{r_i} and t_{r_i+1} . We assume that $0 \leq r_i \leq n - 1$ (special cases where $t_i^* \leq 0$ or t_i^* falls within the current interval can be easily handled). The variable t_i^* can be written as $t_{r_i} + \bar{c}_i h_{r_i}$ for some value of $0 \leq \bar{c}_i \leq 1$. The following definitions will help us simplify the statement of the convergence results. Let N be the number of subintervals partitioning the interval $[a, b]$, $t_{n,i} := t_n + c_i h_n$ for $0 \leq n \leq N - 1$. For a function $g(t) \in C^1[a, b]$ we define (see figure 3.1),

$$\begin{aligned} Q_n^* g(t_{n,i}) &:= \int_{\bar{c}_i}^1 K(t_{n,i}, t_{r_i} + \theta h_{r_i}, g(t_{r_i} + \theta h_{r_i}), g'(t_{r_i} + \theta h_{r_i})) d\theta, \\ Q_n g(t_{n,i}) &:= \int_0^{c_i} K(t_{n,i}, t_n + \theta h_n, g(t_n + \theta h_n), g'(t_n + \theta h_n)) d\theta, \\ Q_n^{(l)} g(t_{n,i}) &:= \int_0^1 K(t_{n,i}, t_l + \theta h_l, g(t_l + \theta h_l), g'(t_l + \theta h_l)) d\theta, \end{aligned} \quad (14) \quad (r_i + 1 \leq l \leq n - 1).$$

We define the discretized form of the integrals in (14) using suitable quadrature formulae denoted by $\tilde{Q}_n^* g(t_{n,i})$, $\tilde{Q}_n g(t_{n,i})$ and $\tilde{Q}_n^{(l)} g(t_{n,i})$ respectively.

We next introduce a lemma from the theory of difference inequalities which is based on a Gronwall inequality and often used in convergence analysis of DDEs (see for example [27, 28]).

LEMMA 3.2. *If the sequence $\{e_n\}$ satisfies the difference inequality,*

$$e_n \leq (1 + Lh_n)e_{n-1} + d_n h_n \quad , \quad n \geq 0,$$

where $\{e_n\}$, $\{d_n\}$, and $\{h_n\}$ are nonnegative sequences, and L is a nonnegative constant, then

$$e_n \leq \left(e_{-1} + \sum_{i=0}^n d_i h_i \right) \exp \left(\sum_{j=0}^n h_j L \right).$$

THEOREM 3.3. *Let f and K have sufficient continuous derivatives over their respective domains and be such that for a given initial function $\phi(t)$, (2) has a unique solution $y(t)$. Assume*

- (a) $u(t) \in C^1[a, b]$ denotes the CRK interpolant defined by (6), (8) and (10),
- (b) $\tilde{u}(t) \in C^1[a, b]$ denotes the fully discretized CRK interpolant where quadrature formulae $\tilde{Q}_n^* g(t_{n,i})$ and $\tilde{Q}_n^{(l)} g(t_{n,i})$ are accurate to at least $O(H)$ in their approximation of the integrals in (14) for $c_1 = 0$,
- (c) all discontinuity points of order ≤ 1 are included in the set of mesh points,
- (d) $\bar{B}L_{y'} < 1$.

Then the interpolant $\tilde{u}(t)$ satisfies

$$\sup_{t \in [a, b]} \|y^{(l)}(t) - \tilde{u}^{(l)}(t)\| = O(H), \quad l = 0, 1.$$

Proof. It suffices to show that $\|u^{(l)}(t) - \tilde{u}^{(l)}(t)\|$ is $O(H)$. Based on the analytical calculation of the integral terms, the CRK interpolant $u(t)$ can be characterized as follows,

$$(15) \quad \begin{aligned} u_n(t_n + \theta h_n) &= y_n + h_n \sum_{i=1}^{\bar{s}+1} b_i(\theta) k_{n,i}, \\ k_{n,i} &= f(t_{n,i}, y_n + h_n \sum_{j=1}^{i-1} a_{i,j} k_{n,j}) + \int_{t_{n,i}-\tau(t_{n,i})}^{t_{n,i}} K(t_{n,i}, \theta, u(\theta), u'(\theta)) d\theta. \end{aligned}$$

Also, using the symbolic operator \sum , the fully discretized CRK interpolant $\tilde{u}(t)$ can be represented as,

$$(16) \quad \begin{aligned} \tilde{u}_n(t_n + \theta h_n) &= \tilde{y}_n + h_n \sum_{i=1}^{\bar{s}+1} b_i(\theta) \tilde{k}_{n,i}, \\ \tilde{k}_{n,i} &= f(t_{n,i}, \tilde{y}_n + h_n \sum_{j=1}^{i-1} a_{i,j} \tilde{k}_{n,j}) + \sum_{t_{n,i}-\tau(t_{n,i})}^{t_{n,i}} K(t_{n,i}, \theta, \tilde{u}(\theta), \tilde{u}'(\theta)) d\theta. \end{aligned}$$

We can define, $e_n := \max_{t_0 \leq t \leq t_{n+1}} \|u(t) - \tilde{u}(t)\|$ and $e'_n := \max_{t_0 \leq t \leq t_{n+1}} \|u'(t) - \tilde{u}'(t)\|$. Using a Taylor series expansion of $u_n(t)$ and $\tilde{u}_n(t)$ about $t = t_n$ we have,

$$(17) \quad u_n(t_n + \theta h_n) = u_n(t_n) + \theta h_n \left[f(t_n, u_n(t_n)) + \int_{t_n - \tau(t_n)}^{t_n} K(t_n, s, u(s), u'(s)) ds \right] + O(h_n^2),$$

$$(18) \quad \begin{aligned} \tilde{u}_n(t_n + \theta h_n) &= \tilde{u}_n(t_n) + \theta h_n \left[f(t_n, \tilde{u}_n(t_n)) + h_{r_1} (\tilde{Q}_n^* \tilde{u}_{r_1}(t_{n,1})) \right. \\ &\quad \left. + \sum_{l=r_1+1}^{n-1} h_l (\tilde{Q}_n^{(l)} \tilde{u}_l(t_{n,1})) \right] + O(h_n^2). \end{aligned}$$

Regarding the assumption (b) in the theorem the following relations hold for the quadrature formulae,

$$(19) \quad \begin{aligned} \tilde{Q}_n^* \tilde{u}_{r_1}(t_n) &= Q_n^* \tilde{u}_{r_1}(t_n) - \xi_n^*(t_n), \\ \tilde{Q}_n^{(l)} \tilde{u}_l(t_n) &= Q_n^{(l)} \tilde{u}_l(t_n) - \xi_n^{(l)}(t_n) \end{aligned}$$

where all error terms are at worst $O(H)$. Substituting these relations into (18) leads to,

$$(20) \quad \tilde{u}_n(t_n + \theta h_n) = \tilde{u}_n(t_n) + \theta h_n \left[f(t_n, \tilde{u}_n(t_n)) + \int_{t_n - \tau(t_n)}^{t_n} K(t_n, s, \tilde{u}(s), \tilde{u}'(s)) ds - h_{r_1} \xi_n^*(t_n) - \sum_{l=r_1+1}^{n-1} h_l \xi_n^{(l)}(t_n) \right] + O(h_n^2).$$

Subtracting (20) from (17), using the assumed Lipschitz conditions and choosing appropriate bounds we obtain,

$$(21) \quad \|u_n(t_n + \theta h_n) - \tilde{u}_n(t_n + \theta h_n)\| \leq (1 + h_n L_f + h_n \bar{B} L_y) e_{n-1} + h_n \bar{B} L_{y'} e'_{n-1} + C_1 h_n H + C_2 h_n^2.$$

Now, if $\max_{t_0 \leq t \leq t_{n+1}} \|u(t) - \tilde{u}(t)\|$ occurs at some t^* , $t^* \leq t_n$, then trivially $e_n = e_{n-1}$. Otherwise, $t^* = t_n + \theta^* h_n$ and $e_n = \|u_n(t_n + \theta^* h_n) - \tilde{u}_n(t_n + \theta^* h_n)\|$. In either case we see from (21) that,

$$(22) \quad e_n \leq (1 + h_n L_f + h_n \bar{B} L_y) e_{n-1} + h_n \bar{B} L_{y'} e'_{n-1} + C_1 h_n H + C_2 h_n^2.$$

Using a similar approach, for $u'_n(t)$ and $\tilde{u}'(t)$ we have,

$$e'_n \leq (L_f + \bar{B} L_y) e_{n-1} + \bar{B} L_{y'} e'_n + C_3 H + C_4 H.$$

Since $0 \leq \bar{B} L_{y'} < 1$ we have,

$$(23) \quad e'_n \leq \left(\frac{L_f + \bar{B} L_y}{1 - \bar{B} L_{y'}} \right) e_n + \left(\frac{C_3 + C_4}{1 - \bar{B} L_{y'}} \right) H.$$

Substituting (23) in (22) leads to,

$$(24) \quad e_n \leq \left(1 + h_n L_f + h_n \bar{B} L_y + h_n \bar{B} L_{y'} \left(\frac{L_f + \bar{B} L_y}{1 - \bar{B} L_{y'}} \right) \right) e_{n-1} + \left(\frac{h_n \bar{B} L_{y'}}{1 - \bar{B} L_{y'}} \right) (C_3 + C_4) H + C_1 h_n H + C_2 h_n^2.$$

Now, applying Lemma (3.2) with e_n defined as above, $e_{-1} = 0$,

$$L := L_f + \bar{B} L_y + \bar{B} L_{y'} \left(\frac{L_f + \bar{B} L_y}{1 - \bar{B} L_{y'}} \right), \text{ and}$$

$$d_n := (C_1 + C_2 + \left(\frac{\bar{B} L_{y'}}{1 - \bar{B} L_{y'}} \right) (C_3 + C_4)) H,$$

we observe that

$$\sum_{i=0}^n d_i h_i \leq H(b-a) \left(C_1 + C_2 + \left(\frac{\bar{B} L_{y'}}{1 - \bar{B} L_{y'}} \right) (C_3 + C_4) \right),$$

and we obtain the bound,

$$e_n \leq H(b-a) \left(C_1 + C_2 + \left(\frac{\bar{B} L_{y'}}{1 - \bar{B} L_{y'}} \right) (C_3 + C_4) \right) \exp(L(b-a)).$$

Also, from (23) we readily see that $e'_n = O(H)$. \square

Although we have proved that $\max_{t_0 \leq t \leq t_N} \|y(t) - \tilde{u}(t)\| \leq CH$, our bound for C is not sharp and will in general be pessimistic. In order to establish the optimal order of convergence (for arbitrary delay arguments), we have to put more restrictions on the conditions (b) and (d) in Theorem 3.3. These complications are due to the presence of y' in the kernel of *neutral* DVIDE (2).

THEOREM 3.4. *Let f and K have sufficient continuous derivatives over their respective domains and be such that for a given initial function $\phi(t)$, the delay system (2) has unique solution $y(t)$. Assume*

(a) $u(t) \in C^1[a, b]$ denotes the CRK interpolant defined by (6), (8) and (10).

(b) $\tilde{u}(t) \in C^1[a, b]$ denotes the fully discretized CRK interpolant where quadrature formulae $\tilde{Q}_n^* g(t_{n,i})$, $\tilde{Q}_n g(t_{n,i})$ and $\tilde{Q}_n^{(l)} g(t_{n,i})$ have degree of precision at least $p - 1$ where p is the order of the underlying Runge-Kutta method.

(c) all discontinuity points of order $\leq p$ are included in the set of mesh points.

(d) $\bar{B}L_{y'} < \frac{1 - \bar{h}L_f\alpha}{\beta'}$ where $\alpha := \max_{i=1,2,\dots,\bar{s}+1} \sum_j |a_{ij}|$, $\beta' := \max_{0 \leq \theta \leq 1} \sum_{j=1}^{\bar{s}+1} |b'_j(\theta)|$ and $H < \bar{h} < \frac{1}{\alpha L_f}$.

Then for sufficiently small H , the interpolant $\tilde{u}(t)$ satisfies

$$\sup_{t \in [a, b]} \|y^{(l)}(t) - \tilde{u}^{(l)}(t)\| = O(H^p), \quad l = 0, 1.$$

Proof. Over each subinterval $[t_n, t_{n+1}]$, the approximation we compute is equivalent to applying the CRK method to the following standard (local) ODE:

$$\begin{aligned} w'_n(t) &= f(t, w_n(t)) + \sum_{t-\tau(t)}^t K(t, s, \tilde{u}(s), \tilde{u}'(s)) ds, \\ (25) \quad w_n(t_n) &= \tilde{y}_n = \tilde{u}(t_n), \quad t_n \leq t \leq t_{n+1}. \end{aligned}$$

whose associated interpolant is $\tilde{u}_n(t)$. Since $\tilde{u}(t)$ is only C^1 continuous (for our CRK method), we do not have enough regularity for $w_n(t)$ over $[t_n, t_{n+1}]$. To address this difficulty (and for our analysis only) we introduce the following artificial ODE problem (similar artificial problems have been used in the convergence analysis of DDEs in [28]):

$$\begin{aligned} y'_n(t) &= f(t, y_n(t)) + \int_{t-\tau(t)}^t K(t, s, y(s), y'(s)) ds, \\ (26) \quad y_n(t_n) &= y(t_n), \quad t_n \leq t \leq t_{n+1}. \end{aligned}$$

Including all discontinuity points of order $\leq p$ as mesh points we have,

$$\|y_n(t) - \rho_n(t)\| \leq O(h_n^{p+1}), \quad \|z_{n+1} - y(t_{n+1})\| \leq O(h_n^{p+1}),$$

where $\rho_n(t)$ and z_{n+1} are the continuous and discrete approximations resulting from applying the CRK method to (26). Note that, because of uniqueness of the solution of (26), we observe that $y_n(t)$ will be equal to $y(t)$ for all n . We have:

$$\begin{aligned} (27) \quad \|y_{n+1} - y(t_{n+1})\| &\leq \|y_{n+1} - z_{n+1}\| + \|z_{n+1} - y(t_{n+1})\| \\ &\leq \|y_{n+1} - z_{n+1}\| + O(h_n^{p+1}), \end{aligned}$$

and

$$(28) \quad \begin{aligned} \|y(t) - u_n(t)\| &\leq \|y(t) - \rho_n(t)\| + \|\rho_n(t) - u_n(t)\| \\ &\leq \|\rho_n(t) - u_n(t)\| + O(h_n^{p+1}). \end{aligned}$$

For $\rho_n(t)$ and $\tilde{u}_n(t)$ we have:

$$\begin{aligned} \rho_n(t_n + \theta h_n) &= y(t_n) + h_n \sum_{j=1}^{\bar{s}+1} b_j(\theta) k_{n,j}, \\ \tilde{u}_n(t_n + \theta h_n) &= \tilde{y}_n + h_n \sum_{j=1}^{\bar{s}+1} b_j(\theta) \tilde{k}_{n,j}. \end{aligned}$$

Therefore,

$$(29) \quad \|\rho_n(t_n + \theta h_n) - \tilde{u}_n(t_n + \theta h_n)\| \leq \|y(t_n) - \tilde{y}_n\| + h_n \beta \max_{j=1,2,\dots,\bar{s}+1} \|k_{n,j} - \tilde{k}_{n,j}\|,$$

and

$$(30) \quad \|y_{n+1} - z_{n+1}\| \leq \|y(t_n) - \tilde{y}_n\| + h_n \bar{\beta} \max_{j=1,2,\dots,s} \|k_{n,j} - \tilde{k}_{n,j}\|,$$

where $\beta = \max_{0 \leq \theta \leq 1} \sum_{j=1}^{\bar{s}+1} |b_j(\theta)|$ and $\bar{\beta} = \sum_{j=1}^s |b_j|$. Now, we establish an upper bound for $\max_j \|k_{n,j} - \tilde{k}_{n,j}\|$. From definition (10) we have:

$$k_{n,j} = f(t_n + c_j h_n, y(t_n)) + h_n \sum_{i=1}^{j-1} a_{j,i} k_{n,i} + \int_{t_n + c_j h_n - \tau(t_n + c_j h_n)}^{t_n + c_j h_n} K(t_n + c_j h_n, s, y(s), y'(s)) ds,$$

and

$$\begin{aligned} \tilde{k}_{n,j} &= f(t_n + c_j h_n, \tilde{y}_n) + h_n \sum_{i=1}^{j-1} a_{j,i} \tilde{k}_{n,i} + \int_{t_n + c_j h_n - \tau(t_n + c_j h_n)}^{t_n + c_j h_n} K(t_n + c_j h_n, s, \tilde{u}(s), \tilde{u}'(s)), \\ &= f(t_n + c_j h_n, \tilde{y}_n) + h_n \sum_{i=1}^{j-1} a_{j,i} \tilde{k}_{n,i} + \int_{t_n + c_j h_n - \tau(t_n + c_j h_n)}^{t_n + c_j h_n} K(t_n + c_j h_n, s, \tilde{u}(s), \tilde{u}'(s)) ds \\ &\quad + \varepsilon_n(t_{n,i}) \end{aligned}$$

where (see (14))

$$\begin{aligned} \varepsilon_n(t_{n,i}) &:= h_{r_i} \xi_n^*(t_{n,i}) + \sum_{l=r_i+1}^{n-1} h_l \xi_n^{(l)}(t_{n+1}) + h_n \xi_n(t_{n,i}), \\ \tilde{Q}_n^* \tilde{u}_{r_i}(t_{n,i}) &= Q_n^* \tilde{u}_{r_i}(t_{n,i}) + \xi_n^*(t_{n,i}), \\ \tilde{Q}_n \tilde{u}_n(t_{n,i}) &= Q_n \tilde{u}_n(t_{n,i}) + \xi_n(t_{n,i}), \\ \tilde{Q}_n^{(l)} \tilde{u}_l(t_{n,i}) &= Q_n^{(l)} \tilde{u}_l(t_{n,i}) + \xi_n^{(l)}(t_{n,i}). \end{aligned}$$

According to the assumption of the theorem and the Peano Theorem [29] we know that $\|\varepsilon_n(t_{n,i})\| = O(H^p)$. Therefore, using the above relations we get:

$$\begin{aligned} \|k_{n,j} - \tilde{k}_{n,j}\| &\leq L_f \|y(t_n) - \tilde{y}_n\| + h_n L_f \alpha \max_{j=1,2,\dots,\bar{s}+1} \|k_{n,j} - \tilde{k}_{n,j}\| + \bar{B} L_y \max_{t \leq t_{n+1}} \|\tilde{u}(t) - y(t)\| \\ &\quad + \bar{B} L_{y'} \max_{t \leq t_{n+1}} \|\tilde{u}'(t) - y'(t)\| + O(H^p) \end{aligned}$$

So, for h_n sufficiently small (say $h_n < \bar{h} < \frac{1}{L_f\alpha}$) we arrive at the following bound:

$$\begin{aligned} \max_{j=1,2,\dots,\bar{s}+1} \|k_{n,j} - \tilde{k}_{n,j}\| &\leq \frac{L_f}{1 - h_n L_f \alpha} \|y(t_n) - \tilde{y}_n\| + \frac{\bar{B}L_y}{1 - h_n L_f \alpha} \max_{t \leq t_{n+1}} \|\tilde{u}(t) - y(t)\| \\ &+ \frac{\bar{B}L_{y'}}{1 - h_n L_f \alpha} \max_{t \leq t_{n+1}} \|\tilde{u}'(t) - y'(t)\| + O(H^p) \end{aligned}$$

Substituting the above relation, (29) and (30) in (27) and (28) and after some manipulation we obtain respectively,

$$\begin{aligned} \|y(t_{n+1}) - y_{n+1}\| &\leq (1 + h_n \text{const}) \|y(t_n) - \tilde{y}_n\| + h_n \text{const} \max_{t \leq t_{n+1}} \|\tilde{u}(t) - y(t)\| \\ (31) \quad &+ h_n \max_{t \leq t_{n+1}} \|\tilde{u}'(t) - y'(t)\| + h_n (\text{const} H^p + O(H^{p+1})), \end{aligned}$$

and

$$\begin{aligned} \|y(t) - \tilde{u}_n(t)\| &\leq (1 + h_n \text{const}) \|y(t_n) - \tilde{y}_n\| + h_n \text{const} \max_{t \leq t_{n+1}} \|\tilde{u}(t) - y(t)\| \\ (32) \quad &+ h_n \max_{t \leq t_{n+1}} \|\tilde{u}'(t) - y'(t)\| + h_n (\text{const} H^p + O(H^{p+1})), \end{aligned}$$

where "const" stands for different constant values. Also, since

$$\begin{aligned} \rho'_n(t_n + \theta h_n) &= \sum_{j=1}^{\bar{s}+1} b'_j(\theta) k_{n,j}, \\ \tilde{u}'_n(t_n + \theta h_n) &= \sum_{j=1}^{\bar{s}+1} b'_j(\theta) \tilde{k}_{n,j}. \end{aligned}$$

a similar inequality to (32) exists for the derivatives,

$$\begin{aligned} \|y'(t) - \tilde{u}'_n(t)\| &\leq \frac{\beta' L_f}{1 - h_n L_f \alpha} \|y(t_n) - \tilde{y}_n\| + \frac{\beta' \bar{B}L_y}{1 - h_n L_f \alpha} \max_{t \leq t_{n+1}} \|\tilde{u}(t) - y(t)\| \\ (33) \quad &+ \frac{\beta' \bar{B}L_{y'}}{1 - h_n L_f \alpha} \max_{t \leq t_{n+1}} \|\tilde{u}'(t) - y'(t)\| + O(H^p). \end{aligned}$$

Introducing

$$e_{n+1} := \max_{i \leq n+1} \|y(t_i) - \tilde{y}_i\|, E_{n+1} := \max_{t \leq t_{n+1}} \|y(t) - \tilde{u}(t)\|, E'_{n+1} := \max_{t \leq t_{n+1}} \|y'(t) - \tilde{u}'(t)\|,$$

from (33) we obtain (notice that E'_{n+1} might not necessarily happen in the interval $[t_n, t_{n+1}]$. So, we replace h_n by \bar{h}):

$$E'_{n+1} \leq \frac{\beta' L_f}{1 - \bar{h} L_f \alpha} e_n + \frac{\beta' \bar{B}L_y}{1 - \bar{h} L_f \alpha} E_{n+1} + \frac{\beta' \bar{B}L_{y'}}{1 - \bar{h} L_f \alpha} E'_{n+1} + O(H^p).$$

So, for $1 - \frac{\beta' \bar{B}L_{y'}}{1 - \bar{h} L_f \alpha} > 0$ (condition (d) in the theorem) we have:

$$(34) \quad E'_{n+1} \leq \text{const} e_n + \text{const} E_{n+1} + O(H^p).$$

It is worth mentioning that $\beta' \geq 1$ in our numerical method. Therefore, $\frac{1-\bar{h}L_f\alpha}{\beta'} \leq 1$ in the condition (d) of the theorem (compare with the condition (d) in the Theorem 3.3). Substituting (34) in (32) results in,

$$(35) \quad E_{n+1} \leq \text{const}e_n + O(H^p)$$

and then (34) yields:

$$(36) \quad E'_{n+1} \leq \text{const}e_n + O(H^p)$$

From (31) we get,

$$(37) \quad \begin{aligned} e_{n+1} &\leq (1 + h_n \text{const})e_n + h_n \text{const}E_{n+1} + h_n \text{const}E'_{n+1} \\ &+ h_n(\text{const}H^p + O(H^{p+1})) \end{aligned}$$

Substituting (35) and (36) in (37) we obtain,

$$e_{n+1} \leq (1 + h_n \text{const})e_n + h_n(\text{const}H^p + O(H^{p+1}))$$

which is the Gronwall inequality that allows us (from Lemma 3.2) to conclude that $e_n = O(H^p)$. Then, from (35) and (36) we obtain $E_n = O(H^p)$ and $E'_n = O(H^p)$. \square

As mentioned earlier, the defect $\delta(t)$ is the amount by which the CRK interpolant $u(t)$ fails to satisfy the VIDEs (2). The method we have implemented attempts to ensure that some measure of the size of this defect is bounded by the user-defined TOL. We now show that the global error associated with our numerical solution $\tilde{u}(t)$ is bounded on the whole interval by a multiple of the tolerance if the magnitude of this defect is bounded by TOL. The following Lemma [30] helps us establish a relationship between the maximum magnitude of the defect and a global error bound for delay VIDEs.

LEMMA 3.5. *Let K_1 and K_2 be nonnegative constants and $\rho(t)$ be a continuous nonnegative function on an interval $a \leq t \leq b$ satisfying the inequality $\rho(t) \leq K_1 + K_2 \int_a^t \rho(s)ds$ for $a \leq t \leq b$. Then $\rho(t) \leq K_1 \exp[K_2(t - a)]$ for $a \leq t \leq b$.*

THEOREM 3.6. *Let $\tilde{u}(t)$ be the CRK interpolant resulting from the fully discretized (16). In addition, assume that the integral term \int is approximated by the quadrature \sum with the error $QE(t)$ (see 43). If*

- (a) *the norm of the defect, $\|\tilde{\delta}(t)\|$, is bounded by TOL on each step of the integration in $[a, b]$,*
- (b) *the functions f and K satisfy the Lipschitz conditions (13),*
- (c) *$\bar{B}L_{y'} < 1$,*
- (d) *$|QE(t)| \leq TOL$.*

Then

$$\|y(t) - \tilde{u}(t)\| \leq C \cdot TOL \quad \text{on} \quad [a, b],$$

where C is a nonnegative constant depending only on the problem.

Proof. The exact solution of (2) satisfies,

$$(38) \quad y'(t) = f(t, y(t)) + \int_{t-\tau(t)}^t K(t, s, y(s), y'(s))ds,$$

and the associated continuous interpolant $u(t)$ satisfies the perturbed Volterra differential equation,

$$(39) \quad u'(t) = f(t, u(t)) + \int_{t-\tau(t)}^t K(t, s, u(s), u'(s)) ds + \delta(t),$$

on $[a, b]$. Equation (39) reveals that the defect $\delta(t)$ is also defined in terms of a definite integral operator \int which has to be computed numerically whenever the fully discretized defect, $\tilde{\delta}(t)$, is to be evaluated. Having computed the continuous interpolant $\tilde{u}(t)$ from the fully discretized CRK method on each step, we have to calculate the associated fully discretized defect $\tilde{\delta}(t)$. The continuous interpolant, $\tilde{u}(t)$, associated with the fully discretized method satisfies,

$$(40) \quad \tilde{u}'(t) = f(t, \tilde{u}(t)) + \sum_{t-\tau(t)}^t K(t, s, \tilde{u}(s), \tilde{u}'(s)) ds + \tilde{\delta}(t),$$

where the symbol \sum represents a suitable quadrature rule. Subtracting (38) from (40) and integrating yields,

$$(41) \quad \begin{aligned} \tilde{u}(t) - y(t) &= \int_a^t [f(x, \tilde{u}(x)) - f(x, y(x))] dx + \int_a^t \sum_{x-\tau(x)}^x K(x, s, \tilde{u}(s), \tilde{u}'(s)) ds dx \\ &- \int_a^t \int_{x-\tau(x)}^x K(x, s, y(s), y'(s)) ds dx + \int_a^t \tilde{\delta}(x) dx. \end{aligned}$$

We now manipulate the right side of the above equation by adding and subtracting the integral term $\int_a^t \int_{x-\tau(x)}^x K(x, s, \tilde{u}(s), \tilde{u}'(s)) ds dx$ and then take a norm of both sides to obtain,

$$(42) \quad \|\tilde{u}(t) - y(t)\| \leq (L_f + \bar{B}L_y) \int_a^t \tilde{\beta}(x) dx + L_{y'} \bar{B} \int_a^t \tilde{\gamma}(x) dx + \int_a^t QE(x) dx + \int_a^t \|\tilde{\delta}(x)\| dx,$$

where we denote the quadrature error $QE(x)$ by,

$$(43) \quad QE(x) := \left\| \sum_{x-\tau(x)}^x K(x, s, \tilde{u}(s), \tilde{u}'(s)) - \int_{x-\tau(x)}^x K(x, s, \tilde{u}(s), \tilde{u}'(s)) ds \right\|,$$

and

$$\tilde{\beta}(t) := \max_{a \leq s \leq t} \|\tilde{u}(s) - y(s)\|, \quad \tilde{\gamma}(t) := \max_{a \leq s \leq t} \|\tilde{u}'(s) - y'(s)\|.$$

We first have to establish an upper bound for the term $\tilde{\gamma}(x)$ in (42). Subtracting (38) from (40), in addition to adding and subtracting an integral term will result in,

$$\begin{aligned} \tilde{u}'(x) - y'(x) &= f(x, \tilde{u}(x)) - f(x, y(x)) + \sum_{x-\tau(x)}^x K(x, s, \tilde{u}(s), \tilde{u}'(s)) \\ &- \int_{x-\tau(x)}^x K(x, s, \tilde{u}(s), \tilde{u}'(s)) ds + \int_{x-\tau(x)}^x K(x, s, \tilde{u}(s), \tilde{u}'(s)) ds \\ &- \int_{x-\tau(x)}^x K(x, s, y(s), y'(s)) ds + \tilde{\delta}(x). \end{aligned}$$

Taking a norm of both sides of the above equation and using our Lipschitz assumptions with $\bar{B}L_{y'} < 1$ we obtain,

$$(44) \quad \tilde{\gamma}(x) \leq \frac{(L_f + L_y \bar{B})}{(1 - L_{y'} \bar{B})} \tilde{\beta}(x) + \frac{\|\tilde{\delta}(x)\| + QE(x)}{(1 - L_{y'} \bar{B})}.$$

Substituting this inequality into (42) and using $\|\tilde{\delta}(t)\| \leq TOL$ leads to,

$$\begin{aligned} \tilde{\beta}(t) &\leq (L_f + \bar{B}L_y + \bar{B}L_{y'} \frac{(L_f + L_y \bar{B})}{(1 - L_{y'} \bar{B})}) \int_a^t \tilde{\beta}(x) dx + TOL(b - a) \left(1 + \frac{\bar{B}L_{y'}}{1 - L_{y'} \bar{B}}\right) \\ &\quad + \left(1 + \frac{\bar{B}L_{y'}}{1 - L_{y'} \bar{B}}\right) \int_a^t QE(x) dx. \end{aligned}$$

It is readily seen from the above inequality that if the error in the quadrature approximation $QE(t)$ is bounded by TOL , then from Lemma (3.5) with $K_1 := 2 TOL(b - a) \left(1 + \frac{\bar{B}L_{y'}}{1 - L_{y'} \bar{B}}\right)$, $K_2 := (L_f + \bar{B}L_y + \bar{B}L_{y'} \frac{(L_f + L_y \bar{B})}{(1 - L_{y'} \bar{B})})$ and $\rho(s) = \tilde{\beta}(s)$ we will achieve the bound,

$$\|\tilde{u}(t) - y(t)\| \leq C \cdot TOL,$$

where $C := \left(2(b - a) \left(1 + \frac{\bar{B}L_{y'}}{1 - L_{y'} \bar{B}}\right) \right) \cdot \exp \left((b - a) \left(L_f + \bar{B}L_y + \bar{B}L_{y'} \frac{(L_f + L_y \bar{B})}{(1 - L_{y'} \bar{B})}\right) \right)$.

□

4. Implementation and Numerical Results. In this section, we illustrate the performance and reliability of a fully discretized CRK method on a selection of problems and justify the strategies and heuristics we have adopted in our implementation. At a very general level, an attempted step from t_n to t_{n+1} can be viewed as consisting of the following five stages:

1. Determine a trial stepsize h_n .
2. Check whether there exists a discontinuity point within $(t_n, t_n + h_n)$. If there exist any, change the attempted step to $(t_n, t_n + h^*)$ where $t_n + h^*$ is the leftmost discontinuity point detected.
3. Form the interpolant $u_n(t)$.
4. Evaluate some measure, Δ , of the size of the defect, $\tilde{\delta}_n(t)$.
5. Accept the step if $\Delta \leq TOL$.

The experimental code we have developed calls a version of a 6th order CRK identified and investigated in [21]. It is based on an underlying discrete 6th order, 7– stage explicit RK formula developed by Verner [31] for IVPs. It should be noted that other formulas and orders can be used as well. Three additional stages are computed to form an RDC CRK of $O(h^7)$ accuracy, and a total of 15 stages are used overall to define the associated SDC CRK. A validity check is also used to ensure a more reliable defect estimate. This strategy is called SDCV. Point 4 deserves further comment. As mentioned earlier, the assumption that the solution is sufficiently smooth results in a credible estimate of the size of the maximum defect in the case of ODEs, i.e., all we need is to evaluate the defect at a predetermined single point. In addition, it is shown in [21] that there is a direct asymptotic relationship between the local error of the underlying discrete RK formula and the continuous defect of any associated SDC CRK.

Shampine extends the concept of the local error to the span of the step and obtains an asymptotically correct estimate of the maximum local error by a single evaluation of the residual [32]. But, this may not be the case when we deal with delay equations since the asymptotic results are not necessarily directly applicable. So, to estimate the maximum defect when it is not sufficiently smooth or the stepsize is fairly large, one could sample it at more than one point. Our objective is to show that the maximum defect will be proportional to the maximum global error. If the maximum defect on a step is dominated by the contribution from the truncation error associated with the CRK formula, sampling the defect at a single point will deliver reliable results. We have implemented this estimate, and our numerical results (described below) verify that this is generally the case. Nevertheless, there are problems where iteration errors or quadrature errors can significantly contribute to the observed maximum defect and developing more reliable estimates when this is the case is a topic we are currently investigating.

As mentioned earlier, an implicit system of equations has to be solved on each step of the integration to determine $\tilde{k}_2, \tilde{k}_3, \dots, \tilde{k}_{\bar{s}+1}$. In order to take advantage of the underlying explicit RK formula, a convergent Gauss-Seidel type iterative procedure (similar to that used in a Picard iteration or in waveform relaxation) is used to approximate the solution of the implicit system of equations as follows:

$$\begin{aligned}
 \tilde{k}_i^{[m]} &= f(t_n + c_i h_n, \tilde{y}_n + h_n \sum_{j=1}^{i-1} a_{i,j} \tilde{k}_j^{[m]}) \\
 &+ \int_{\min\{t_n, t_n + c_i h_n - \tau(t_n + c_i h_n)\}}^{t_n} K(t_n + c_i h_n, s, \tilde{u}(s), \tilde{u}'(s)) ds \\
 (45) \quad &+ \int_{\max\{t_n, t_n + c_i h_n - \tau(t_n + c_i h_n)\}}^{t_n + c_i h_n} K(t_n + c_i h_n, s, \tilde{u}_{[m-1]}(s), \tilde{u}'_{[m-1]}(s)) ds,
 \end{aligned}$$

where

$$\tilde{u}_{[m]}(t_n + \theta h_n) = \tilde{y}_n + h_n \sum_{i=1}^{\bar{s}+1} b_i(\theta) \tilde{k}_i^{[m]} \quad , \quad \tilde{u}'_{[m]}(t_n + \theta h_n) = \sum_{i=1}^{\bar{s}+1} b'_i(\theta) \tilde{k}_i^{[m]} .$$

In our experimental implementation, we allow a maximum number of iterations, I_{max} , on each step, and we define I_{av} to be the average (over all accepted steps) of the number of iterations required before an acceptable approximation $\tilde{u}_{[m]}(t)$ is determined. For each step we force at least I_{min} iterations before accepting the step, and we accept the first iterate, $\tilde{u}_{[m]}(t)$, such that $m \geq I_{min}$ and the estimate of the maximum defect for $\tilde{u}_{[m]}(t)$ is less than TOL . In the reported numerical experiments we use $I_{max} = 4$ and $I_{min} = 1$, but in order to minimize the effect of iteration errors (on the accepted steps) on the results, we have run examples with $I_{min} = I_{max} = 8$ (see Table 5.1). Note that we also monitor the convergence of the iteration. We reject the step whenever divergence is detected and adjust the next trial stepsize accordingly. The larger the value of I_{max} , the more likely it will be that a larger step, h_n , will be accepted, and the number of steps will therefore be smaller. However, the number of kernel evaluations might be significantly larger (depending on the problem). In order to start the iteration (45) we need an initial guess $\tilde{u}_{[0]}(t)$. We determine $\tilde{u}_{[0]}(t)$ by a natural extension of the approximate solution $\tilde{u}_{n-1}(t)$ on the last step. For the first interval, the initial function is evaluated for $t \geq t_0$ (whenever possible) to provide this

initial guess.

The CRK approach we have developed ensures a $C^1[a, b]$ interpolant provided the iteration (45) converges to \tilde{k}_i which satisfies (16), in which case $\tilde{k}_{n,s+1} = \tilde{k}_{n+1,1} = f(t_{n+1}, \tilde{y}_{n+1}) + \int_{t_{n+1}-\tau(t_{n+1})}^{t_{n+1}} K(t_{n+1}, s, \tilde{u}(s), \tilde{u}'(s))ds$. However, the iteration (45) is halted before (16) is exactly satisfied. As a consequence, the derivative at t_{n+1} on step n (which is the $(s+1)^{th}$ stage value $\tilde{k}_{n,s+1}^{[m]}$ defined in terms of the interpolant $\tilde{u}_{[m-1]}(t)$) is not necessarily equal to the derivative at t_{n+1} on the next attempted step $(n+1)$ (which is the first stage value defined in terms of the updated interpolant $\tilde{u}_{[m]}(t)$). Therefore, the interpolant $\tilde{u}(t)$ may no longer be C^1 continuous. Although there are a number of ways to overcome this difficulty, we found it the most helpful to update the stage value \tilde{k}_{s+1} on step n one more time having updated all stages contributing to the definition of the interpolant associated with step n . In other words, we have set $\tilde{k}_{n,s+1}^{[m+1]}$ on step n to be the same as $\tilde{k}_{n+1,1}$ on step $(n+1)$. The difference between the two most recently updated stage values $\tilde{k}_{n,s+1}^{[m]}$ and $\tilde{k}_{n,s+1}^{[m+1]}$ can be viewed as an indication of the size of the derivative discontinuity in the approximate solution at t_{n+1} arising from the iteration error.

We demonstrate the performance of the fully discretized CRK methods developed in this paper on several examples. The statistics we report are:

- **GEMAX** : maximum absolute error over the integration interval (determined by evaluating the reference solution at 10000 equally-spaced points in the integration interval).
- **GE** : maximum absolute error over the mesh points.
- **NSTP** : total number of successful steps, N , required to solve the problem.
- **NFCN** : total number of function evaluations, $f(t, y(t))$, required to solve the problem.
- **NKER** : total number of kernel evaluations, $K(t, s, y(s), y'(s))$, required to solve the problem.
- **DMAX** : ratio of the maximum observed defect over all steps to the tolerance.
- **Frac-D** : fraction of steps where the maximum defect exceeds the tolerance.
- **R-Max** : maximum ratio over all steps of the true maximum defect (over the entire step) to its estimate.
- **Frac-G** : fraction of steps where the ratio of the true maximum defect to its estimate was bounded by 1.1. That is, the estimate was within 10% of the true maximum.

GEMAX and **GE** are two measures of the accuracy of the numerical solution. **NSTP**, **NFCN** and **NKER** are three measures of the cost. **DMAX** and **Frac-D** are two measures of the reliability of the method demonstrating how well the method controls the maximum magnitude of the defect. **R-Max** and **Frac-G** are two measures of the reliability of the estimate helping us assess how well the estimate of the maximum defect reflects its true value. One can estimate I_{av} , the average number of iterations used on each attempted step (assuming a small fraction of rejected steps), for the following problems by $\frac{\text{NFCN}}{(\bar{s}+1)(\text{NSTP})}$ (where $\bar{s}+1$ equals 11 for RDC, 15 for SDC and 17 for SDCV). Calculating a similar quantity in terms of the number of kernel evaluations is not straightforward because the delay is time-dependent and also our stepsize selection strategy is adaptive. However, we can calculate a rough bound for **NKER** by considering the following two typical cases assuming that the percentage of rejected steps is small:

Large delays (on step n , the associated delay $\tau(t) \simeq t - t_0$) :

TABLE 4.1
Results for the first example for three defect control strategies

TOL/CRK	GEMAX	GE	NSTP	NFCN	NKER	DMAX	Frac-D	R-Max	Frac-G
10^{-4}									
RDC	2.52×10^{-5}	2.52×10^{-5}	75	3169	55692	3.27	0.67	8.47	0.0
SDC	1.66×10^{-5}	1.61×10^{-5}	36	2641	27674	3.36	0.25	4.98	0.02
SDCV	1.18×10^{-5}	1.12×10^{-5}	34	3473	43104	1.36	0.06	1.59	0.15
10^{-6}									
RDC	3.61×10^{-7}	2.26×10^{-7}	57	3205	46164	3.40	0.46	8.78	0.0
SDC	1.88×10^{-7}	1.88×10^{-7}	49	3425	41082	2.03	0.22	5.37	0.02
SDCV	9.96×10^{-8}	9.95×10^{-8}	47	4085	57132	1.45	0.13	3.67	0.13
10^{-8}									
RDC	4.82×10^{-9}	4.29×10^{-9}	72	3325	59279	6.67	0.64	18.18	0.0
SDC	3.06×10^{-9}	3.06×10^{-9}	49	3665	43951	1.28	0.15	10.33	0.04
SDCV	1.77×10^{-9}	1.77×10^{-9}	52	4247	67117	1.30	0.19	5.13	0.08
10^{-10}									
RDC	3.26×10^{-11}	9.04×10^{-12}	62	3493	53091	10.14	0.37	67.45	0.0
SDC	2.52×10^{-11}	2.52×10^{-11}	53	3393	49925	4.59	0.17	6.37	0.12
SDCV	2.51×10^{-11}	2.50×10^{-11}	43	3881	64515	1.48	0.05	4.09	0.19

$$\mathbf{NKER} \simeq M \cdot I_{av} \cdot \mathbf{NSTP} \cdot (\bar{s} + 1) + M \cdot \mathbf{NSTP}^2 \cdot (\bar{s} + 1).$$

Small delays (on step n , the associated delay $\tau(t) \simeq t - t_n$) :

$$\mathbf{NKER} \simeq M \cdot I_{av} \cdot \mathbf{NSTP} \cdot (\bar{s} + 1).$$

M stands for the number of abscissas at which the integrands associated with the integral terms are evaluated (which is 4 in our experiments). Note that, for the case of large delays, the above observations indicate that the number of kernel evaluations can be quite large. As seen in the reported results for problem 3, the number of kernel evaluations exceeds even the above bounds, because the delay argument is decreasing and maps the interval $[0, 2]$ to the interval $[-1, -5.38]$. We are investigating alternative quadrature strategies to address this issue.

Problem 1: A DVIDE with *vanishing* delay argument,

$$y'(t) = -\sin(t) + t^2(\cos(t) - \cos(t - \cos(t) - 1)) + \int_{t - \cos(t) - 1}^t t^2 \sin(s) (y'(s)^2 + y(s)^2) ds,$$

$$0 \leq t \leq 4, \quad y(t) = \begin{cases} 1, & t \leq 0, \\ \cos(t), & t \geq 0. \end{cases}$$

In this example we only have C^1 continuity at the starting point. The recursion (4) generates a cluster of discontinuity points in the left neighborhood of the vanishing point π . However, due to smoothing, we only need to keep track of the discontinuities of order $\leq p = 6$.

Problem 2 : A mathematical model of the human immune response with delays [33],

$$\begin{aligned} X'_U(t) &= s_U - \alpha_1 X_U(t) B(t) - \mu_{X_U} X_U(t), \\ X'_I(t) &= \alpha_1 X_U(t) B(t) - \alpha_2 X_I(t) A_R(t) - X_I(t), \\ B'(t) &= \alpha_{20} B(t) \left(1 - \frac{B(t)}{\sigma}\right) - \alpha_3 B(t) I_R(t) - \alpha_4 B(t) A_R(t), \end{aligned}$$

TABLE 4.2
Results for the second example for three defect control strategies

TOL/CRK	GEMAX	GE	NSTP	NFCN	NKER	DMAX	Frac-D	R-Max	Frac-G
10^{-2}									
RDC	0.85×10^{-2}	0.74×10^{-2}	69	1549	51784	2.50	0.33	38.40	0.35
SDC	1.00×10^{-2}	0.97×10^{-2}	77	2449	82816	0.99	0.0	2.67	0.97
SDCV	1.00×10^{-2}	0.97×10^{-2}	77	2759	97436	0.99	0.0	1.89	0.97
10^{-4}									
RDC	8.37×10^{-5}	8.06×10^{-5}	129	2077	137600	2.58	0.35	72.53	0.39
SDC	1.62×10^{-4}	1.13×10^{-4}	142	2913	220528	0.98	0.0	1.51	0.99
SDCV	1.16×10^{-4}	1.13×10^{-4}	142	3279	251640	0.98	0.0	1.51	0.99
10^{-6}									
RDC	1.11×10^{-6}	1.01×10^{-6}	265	3661	515912	4.14	0.37	89.33	0.42
SDC	1.79×10^{-6}	1.76×10^{-6}	291	5233	852572	0.92	0.0	6.70	0.99
SDCV	1.79×10^{-6}	1.76×10^{-6}	291	5889	966186	0.92	0.0	6.70	0.99
10^{-8}									
RDC	1.14×10^{-8}	1.12×10^{-8}	553	7117	2116464	8.47	0.38	67.93	0.42
SDC	1.62×10^{-8}	1.62×10^{-8}	610	10353	3585012	1.27	0.00	10.00	0.98
SDCV	1.56×10^{-8}	1.54×10^{-8}	610	11667	4052092	1.55	0.0	10.40	0.99

$$I'_R(t) = s_{I_R} + \int_{t-\tau_1}^t \omega_1(\theta - t)B(\theta)d\theta - \mu_{I_R}I_R(t),$$

$$A'_R(t) = s_{A_R} + \int_{t-\tau_2}^t \omega_2(\theta - t)B(\theta)d\theta - \mu_{A_R}A_R(t), \quad 0 \leq t \leq 50,$$

which consists of five components: uninfected target cells (X_U), infected cells (X_I), bacteria (B), and phenomenological variables capturing innate (I_R) and adaptive (A_R) immunity. $\omega_1(\theta)$ and $\omega_2(\theta)$ are both chosen to be $A \exp(K\theta)$ for $A = K = \frac{\ln 2}{6}$. Also, we set $\tau_1 = \tau_2 = 6$. For a detailed description of the model, the selection of parameter values as well as initial conditions see [33]. For this example, we use $TOL = 10^{-10}$ and a very fine mesh to generate a reference solution.

Problem 3: A DVIDE with *decreasing* delay argument,

$$y'(t) = te^t - e^{-t} + \int_{t-e^t}^t te^{2s}y(s)y'(s)ds, \quad 0 \leq t \leq 2, \quad y(t) = e^{-t}.$$

It is worth pointing that there is no propagated discontinuity for this example. In addition, the delay argument is decreasing over the integration interval which leads to a very large number of kernel evaluations relative to the number of steps.

Problem 4: Predator-Prey system (1) with the following set of parameters [14],

$$\epsilon_1 = 0.02, \epsilon_2 = 1, \gamma_1 = 1, \gamma_2 = 1, \tau = 0.2, T = 2.$$

The initial functions are $\phi_1(t) = \phi_2(t) = 3$, $t \in [-\tau, 0]$. Also, we have set $F_1(t) = F_2(t) = \frac{1}{2!}t^3e^{-3t}$. Predator-Prey systems are simple cases of the general system (3). Table 4.4 reports the results of our method on this problem. This problem does not have a known closed form solution. We computed a reference exact solution using an adaptive mesh with 410 subintervals. Many DVIDEs arising as mathematical models in population dynamics, rheology, etc. have the non-standard form (3) which is a

TABLE 4.3
Results for the third example for three defect control strategies

TOL/CRK	GEMAX	GE	NSTP	NFCN	NKER	DMAX	Frac-D	R-Max	Frac-G
10^{-4}									
RDC	9.18×10^{-4}	9.18×10^{-4}	17	1105	79755	2.20	0.56	9.15	0.0
SDC	1.98×10^{-3}	1.98×10^{-3}	11	865	55467	3.47	0.20	8.32	0.10
SDCV	5.34×10^{-4}	5.34×10^{-4}	12	999	88777	0.81	0.0	3.07	0.18
10^{-6}									
RDC	2.91×10^{-5}	2.91×10^{-5}	19	1213	85961	6.49	0.72	98.22	0.0
SDC	5.4×10^{-6}	5.4×10^{-6}	17	1585	108251	1.18	0.062	3.82	0.062
SDCV	1.16×10^{-5}	1.16×10^{-5}	17	1979	174229	1.19	0.062	1.82	0.12
10^{-8}									
RDC	6.46×10^{-8}	6.46×10^{-8}	25	1501	113291	2.26	0.66	6.30	0.0
SDC	3.43×10^{-7}	3.43×10^{-7}	18	1585	108273	1.84	0.11	13.8	0.058
SDCV	8.81×10^{-8}	8.81×10^{-8}	18	2059	186779	1.50	0.058	1.82	0.058
10^{-10}									
RDC	7.74×10^{-10}	7.74×10^{-10}	26	1489	119191	4.82	0.60	345.44	0.0
SDC	1.40×10^{-9}	1.40×10^{-9}	20	1585	114423	1.98	0.10	13.82	0.26
SDCV	5.07×10^{-11}	5.07×10^{-11}	20	2135	201027	0.91	0.0	3.53	0.26

particular case of the more general DVIDE

$$(46) \quad y'(t) = f(t, y(t)) + \int_{t-\tau(t)}^t K(t, s, \underline{y}(t), y(s), y'(s)) ds,$$

Our variable stepsize approach and the analytical framework we have established can be extended in a straightforward way to this class of DVIDEs. It would require a slight modification to the definition of k_i in (10) (and to the corresponding iteration scheme),

$$k_i = f(t_n + c_i h_n, Y_i) + \int_{t_n + c_i h_n - \tau(t_n + c_i h_n)}^{t_n + c_i h_n} K(t_n + c_i h_n, s, Y_i, u(s), u'(s)) ds,$$

$$Y_i := y_n + h_n \sum_{j=1}^{i-1} a_{i,j} k_j$$

for $i = 1, 2, \dots, \bar{s} + 1$. A more general equation might also include a state dependent delay argument as $t - \tau(t, y(t))$. The same standard discretization approach can be used here to define a CRK method. However, a reliable iterative approach has to be used to obtain approximate solutions within the required accuracy. We are currently investigating how this is best done.

From the reported results in the tables it is clear that the achieved accuracy is consistent with the user-defined tolerance at all tolerances, and the accuracy at mesh points is consistent with the accuracy of off-mesh points. In addition, the maximum defect across the entire timestep can be reliably controlled using the new SDC or SDCV strategies. SDC with validity check is often the most reliable strategy at all tolerances. It is also clear that the RDC delivers acceptable reliability.

5. Observations and Conclusion. There are four potential sources of error associated with our numerical method:

TABLE 4.4
Results for the fourth example for three defect control strategies

TOL/CRK	GEMAX	GE	NSTP	NFCN	NKER	DMAX	Frac-D	R-Max	Frac-G
10^{-4}									
RDC	5.33×10^{-6}	3.99×10^{-6}	15	217	4420	0.72	0.0	1.81	0.78
SDC	4.35×10^{-6}	3.93×10^{-6}	15	353	6388	0.62	0.0	1.02	1.0
SDCV	4.35×10^{-6}	3.93×10^{-6}	15	397	7308	0.62	0.0	1.02	1.0
10^{-6}									
RDC	9.40×10^{-8}	8.72×10^{-8}	25	577	16052	1.28	0.04	16.04	0.83
SDC	7.37×10^{-8}	6.93×10^{-8}	24	689	15372	0.71	0.0	1.37	0.91
SDCV	7.37×10^{-8}	6.93×10^{-8}	24	777	18516	0.71	0.0	1.37	0.91
10^{-8}									
RDC	9.89×10^{-10}	9.28×10^{-10}	49	1201	41680	0.95	0.0	160.27	0.79
SDC	7.73×10^{-10}	7.47×10^{-10}	48	1585	50624	0.91	0.0	15.41	0.89
SDCV	7.70×10^{-10}	7.45×10^{-10}	48	1793	61260	0.93	0.0	10.03	0.89
10^{-10}									
RDC	8.96×10^{-12}	1.14×10^{-11}	95	1717	92800	2.58	0.02	40.67	0.78
SDC	9.94×10^{-12}	9.72×10^{-12}	91	2161	112948	0.97	0.0	21.35	0.94
SDCV	9.94×10^{-12}	9.72×10^{-12}	91	2453	132976	0.97	0.0	9.41	0.94

- **Truncation error** : that arises when approximating the true solution of the DVIDEs.
- **Quadrature error** : arising both in defining the fully discretized $\tilde{u}(t)$ and in computing $\tilde{\delta}(t)$. This error can be reduced by employing quadratures with a higher degree of precision which increases the number of kernel evaluations.
- **Iteration error** : which is associated with the nonlinear systems of equations arising on each step of the integration. The larger the number of iterations associated with the accepted steps, the smaller this component of the error is likely to be. However, increasing this number may result in a significant increase in the number of function and kernel evaluations although the number of steps will only be slightly reduced.
- **Round-off error**

We have implicitly assumed (in our analysis and implementation) that truncation error dominates the other sources of error. The effect of other errors when this assumption is not valid, such as situations where the accuracy requests are stringent and round-off effects are significant or the formula requires many floating-point operations per step (due to a large delay or a high accurate quadrature), is of future interest. The reliability of both the methods we have implemented and the associated defect estimate can be significantly improved if we reduce the effects of other sources of error (at an increase in cost). Table 5.1 shows the effects of increasing the degree of precision of the quadrature formula to 13 as well as increasing the number of iterations to be fixed and equal to 8 on the accepted steps in integrating the third problem for SDC and SDCV strategies. These changes reduce the effects of the quadrature errors and the iteration errors. One can see the significant improvement in the reliability measures **Frac-D** and **Frac-G** compared to those reported in Table 4.3. It is also noticeable that there is a decrease in the number of steps which naturally results in a fewer number of function and kernel evaluations. However, this is not always the case. We ran similar experiments using the human immune system example as well and found that the number of steps remains almost the same while there is a significant increase in the cost measures **NFCN** and **NKER** as expected. For both

TABLE 5.1
 Results for the third problem, with $I_{min} = I_{max} = 8$ and a quadrature of degree 13

TOL/CRK	GEMAX	GE	NSTP	NFCN	NKER	DMAX	Frac-D	R-Max	Frac-G
10^{-4}									
SDC	2.14×10^{-6}	2.14×10^{-6}	5	625	26943	0.04	0.0	3.75	0.50
SDCV	2.14×10^{-6}	2.14×10^{-6}	5	721	47821	0.04	0.0	1.30	0.75
10^{-6}									
SDC	2.17×10^{-6}	2.17×10^{-6}	7	897	40817	0.23	0.0	2.0	0.50
SDCV	2.17×10^{-6}	2.17×10^{-6}	7	1089	71269	0.23	0.0	1.62	0.50
10^{-8}									
SDC	8.98×10^{-8}	8.93×10^{-8}	10	1345	69301	0.35	0.0	2.63	0.55
SDCV	8.82×10^{-8}	8.82×10^{-8}	10	1627	120319	0.35	0.0	2.19	0.55
10^{-10}									
SDC	1.43×10^{-9}	1.43×10^{-9}	13	1665	84673	0.44	0.0	13.82	0.75
SDCV	1.40×10^{-9}	1.40×10^{-9}	13	2005	155891	0.44	0.0	3.53	0.75

problems the reliability measures improve significantly. We are currently investigating the use of asymptotically correct estimates of the iteration and quadrature errors so that the number of iterations (on each step) and the quadrature formulas used can be adaptively chosen without a significant increase in the cost.

These sources of error affect the asymptotic behavior of the defect function corresponding to the approximate solution as well. In case of ODEs, for a given problem and any numerical method, as $H \rightarrow 0$, the (scaled) plots of $\delta_n(\tau)$ vs. τ approach a unique polynomial. This allows us to estimate the maximum defect by a single evaluation of the defect at run-time at a predetermined point. However, it is difficult to choose a fixed sample point for DVIDES, since the shape of the defect might be different for different problems partly due to the extra sources of error arising in approximating the solution. The plot of $|\delta(\tau)|$ vs. τ for all steps required by our CRK method to solve a typical problem (Tol = 10^{-6} , $H \rightarrow 0$ and with the normal error control) is presented in figure (5.1). The figure reveals that the limiting polynomial is asymptotically approached (as $H \rightarrow 0$) at least for the steps where the maximum defect is close to the requested tolerance. Furthermore, $\tau^* = 0.92$ seems to be a good choice at which we can evaluate a reliable estimate of the maximum defect. The figure also contains the plot of $\delta(\tau)/\delta(\tau^*)$ for the majority of steps where round-off error is not significant. We are currently investigating the characteristics of the defect functions corresponding to the DVIDES.

In the present work, we have investigated CRK methods applied to a general class of VIDEs with arbitrary time-dependent delay arguments. We analyzed the convergence properties of the fully discretized CRK methods (using a variable stepsize approach) and considered the effects of the numerical quadratures when approximating the solutions of the DVIDES. We demonstrated that the global error associated with the introduced continuous approximation will be bounded by a multiple of the prescribed tolerance if the magnitude of the defect is bounded by the tolerance. We also included the propagated discontinuities in the set of the mesh points using an automatic discontinuity detection strategy. We have implemented our approach as an experimental Fortran code and carried out numerical experiments over various kinds of DVIDES and reported the statistics. Various kind of delays including non-vanishing, vanishing, proportional and etc. are handled by the solver.

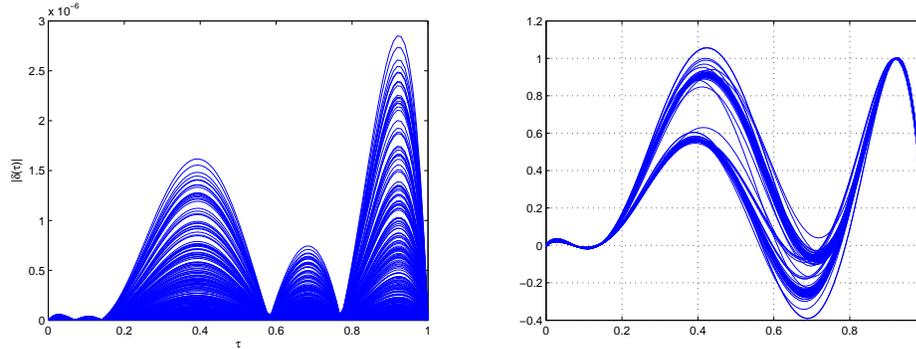


FIG. 5.1. Plots of $|\delta(\tau)|$ vs. τ (left) and $\delta(\tau)/\delta(\tau^*)$ vs. τ (right).

Acknowledgments. We would like to thank the referees for their helpful suggestions and comments. These led to improvements in the paper. This research was supported in part by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

- [1] V. VOLTERRA, *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*, *Memorie del R. Comitato talassografico italiano*, Mem. CXXXI, 1927.
- [2] J. M. CUSHING, *Integro-differential Equations and Delay Models in Population Dynamics*, Lecture Notes in Biomathematics 20, Springer, Berlin, 1977.
- [3] G. A. BOCHAROV, F. A. RIHAN, *Numerical modelling in biosciences using delay differential equations*, *J. Comput. Appl. Math.*, 125(2000), pp. 183-199.
- [4] A. DE GAETANO, O. ARINO, *Mathematical modelling of the intravenous glucose tolerance test*, *J. Math. Biol.*, 40 (2000), pp. 136-168.
- [5] A. MAKROGLOU, J. LI, Y. KUANG, *Mathematical models and software tools for the glucose-insulin regulatory system and diabetes: an overview*, *Appl. Numer. Math.*, 56 (2006), pp. 559-573.
- [6] A. BELLEN, R. VERMIGLIO, *Some applications of continuous Runge-Kutta methods*, *Appl. Numer. Math.*, 22(1996), pp. 63-80.
- [7] A. BELLEN, N. GUGLIELMI, M. ZENNARO, *Numerical stability of nonlinear delay differential equations of neutral type*, *J. Comput. Appl. Math.*, 125(2000), pp. 251-263.
- [8] C. T. H. BAKER, N. J. FORD, *Asymptotic error expansions for linear multistep methods for a class of delay integro-differential equations*, *Bull. Greek Math. Soc.*, 31(1990), pp. 5-10.
- [9] H. BRUNNER, *Collocation Methods for Volterra Integral and Related Functional Differential Equations*, Cambridge University Press, Cambridge, 2004.
- [10] H. BRUNNER, *Recent advances in the numerical analysis of Volterra functional differential equations with variable delays*, *J. Comput. Appl. Math.*, 228(2009), pp. 524-537.
- [11] A. MAKROGLOU, *A block-by-block method for the numerical solution of Volterra delay integro-differential equation*, *Computing*, 30(1983), pp. 49-62.
- [12] H. BRUNNER, *Collocation methods for nonlinear Volterra Integro-differential equations with infinite delay*, *Math. Comp.*, 53(1989), pp. 571-587.
- [13] H. BRUNNER, *The numerical solution of neutral Volterra integro-differential equations with delay arguments*, *Ann. Numer. Math.*, 1(1994), pp. 309-322.
- [14] M. SHAKOURIFAR, M. DEGHAN, *On the numerical solution of nonlinear systems of Volterra integro-differential equations with delay arguments*, *Computing*, 82(2008), pp. 241-260.
- [15] Y. YU, L. WEN, S. LI, *Nonlinear stability of RungeKutta methods for neutral delay integro-differential equations*, *Appl. Math. Comput.*, 191(2007), pp. 543-549.
- [16] C. ZHANG, S. VANDEWALLE, *Stability analysis of RungeKutta methods for nonlinear Volterra delay-integro-differential equations*, *IMA J. Numer. Anal.*, 24(2004), pp. 193-214.
- [17] W. S. WANG, S. F. LI, *Convergence of Runge-Kutta Methods for Neutral Volterra Delay-*

- Integro-Differential Equations*, Front. Math. China., 4(2009), pp. 195216.
- [18] W. H. ENRIGHT, *Analysis of error control strategies for continuous Runge-Kutta methods*, SIAM J. Numer. Anal., 26(1989), pp. 588-599.
 - [19] W. H. ENRIGHT, *A new error-control for initial value solvers*. Appl. Math. Comput., 31(1989), pp. 288-301.
 - [20] W. H. ENRIGHT, W. B. HAYES, *Robust and reliable defect control for Runge-Kutta methods*. ACM Trans. Math. Software, 33(2007), pp. 1-19.
 - [21] W. H. ENRIGHT, L. YAN, *The reliability/cost trade-off for a class of ODE solvers*, Numer. Algorithms., 53(2009), pp. 239-260.
 - [22] H. BRUNNER, W. ZHANG, *Primary discontinuities in solutions for delay integro-differential equations*, Methods. Appl. Anal., 6(1999), pp. 525-534.
 - [23] C. T. H. BAKER, D. WILLÉ, *On the propagation of derivative discontinuities in Volterra retarded integro-differential equations*, New Zeland Journal of Mathematics, 29(2000), pp. 103-113.
 - [24] N. GUGLIELMI, E. HAIRER, *Computing breaking points in implicit delay differential equations*, Adv. Comput. Math., 29(2008), pp. 229-247.
 - [25] W. H. ENRIGHT, M. HU, *Continuous Runge-Kutta methods for neutral Volterra integro-differential equations with delay*, Appl. Numer. Math., 24(1997), pp. 175-190.
 - [26] Z. KAMONT, M. KWAPISZ, *On the Cauchy problem for differential-delay equations in a Banach space*, Math. Nachr., 74(1976), pp. 173-190.
 - [27] C. T. H. BAKER, C. A. H. PAUL, *Parallel continuous Runge-Kutta methods and vanishing lag delay differential equations*, Adv. Comput. Math., 1(1993), pp. 367-394.
 - [28] A. BELLEN, M. ZENNARO, *Numerical Methods for Delay Differential Equations*, Oxford University Press, Oxford, 2003.
 - [29] A. H. STROUD, *Numerical Quadrature and Solution of Ordinary Differential Equations*, Springer-Verlag, New York, 1974.
 - [30] F. BRAUER, J. A. NOHEL, *The Qualitative Theory of Ordinary Differential Equations: An Introduction*, Dover, New York, 1989.
 - [31] J. H. VERNER, *Explicit Runge-Kutta Methods with Estimates of the Local Truncation Error*, SIAM J. Numer. Anal., 15(1978), pp. 772-790.
 - [32] L. F. SHAMPINE, *Solving ODEs and DDEs with Residual Control*, Appl. Numer. Math., 52(2005), pp. 113-127.
 - [33] S. MARINO, E. BERETTA, D. E. KIRSCHNER, *The Role of Delays in Innate and Adaptive Immunity to Intracellular Bacterial Infection*, Math. Biosci. Eng., 4(2007), pp. 261-88.