On the randomized solution of initial value problems

Thomas Daun

Department of Computer Science,University of Kaiserslautern,
Postfach 3049, D-67653 Kaiserslautern,
Germany

e-mail: daun@cs.uni-kl.de

We study the randomized solution of initial value problems for systems of ordinary differential equations

$$y'(x) = f(x, y(x)),\ x \in [a, b],\ y(a) = y_0 \ \in \mathbb{R}^d.$$

Recently S. Heinrich and B. Milla presented an order optimal randomized algorithm solving this problem for γ-smooth input data (i.e. $\gamma = r + \rho$: the r-th derivatives of f satisfy a ρ-Hölder condition). This algorithm uses function values and values of derivatives of f. In this paper we present an order optimal randomized algorithm for the class of γ-smooth functions that uses only values of f. For this purpose we show how to obtain an order optimal randomized algorithm from an order (sub)optimal deterministic one.

1 Introduction

We consider algorithms for initial value problems for systems of ordinary differential equations

$$y'(x) = f(x, y(x)) \qquad (x \in [a, b]), \tag{1}$$
$$y(a) = y_0, \tag{2}$$

where $y_0 \in \mathbb{R}^d,\ -\infty < a < b < \infty,\ f : [a,b] \times \mathbb{R}^d \to \mathbb{R}^d$.
In particular we are interested in determining the optimal order of randomized algorithms, which solve (1) − (2) for specific classes of information.

In the deterministic case the problem was studied in the framework of information based complexity (IBC) by Kacewicz in [4]. In his work he provided an order optimal algorithm for the deterministic setting. Later he considered the randomized case in [5] and presented an almost optimal algorithm for the problem. That means, he reaches the optimal order only up to an arbitrarily small $\varepsilon > 0$. Heinrich and Milla introduced a new algorithm for the randomized setting and proved its optimal order in [1]. Both, the algorithm of Kacewicz and the algorithm of Heinrich and Milla are based on Taylor approximation. Thus these algorithms need knowledge about the derivatives of f.

In this paper we present a generalized approach for the considered problem class. We will show that we obtain the same order of convergence, utilizing a smaller class of information. For this purpose we start with a family of deterministic algorithms. Based on these algorithms we develop a randomized algorithm using polynomial interpolation and prove its optimal order using results and methods from [1]. By choosing a suitable family of deterministic algorithms that use only values of f, we obtain a randomized method which uses only values of f, as well. Furthermore, if we choose the samples of the randomized algorithm in a deterministic way, we obtain an order optimal deterministic algorithm using only values of f, too.

The paper is structured as follows: In Section 2 we define the considered initial value problems and present the basic notions of the IBC framework associated with the problem. In the third section we define the conditions for the family of deterministic algorithms and the randomized algorithm itself. The fourth section is dedicated to the analysis of the algorithm and numerical results for the new algorithm will be provided in Section 5.

2 Preliminaries

We start with defining the IBC framework and introduce some necessary definitions.

Definition 1. Let $r \in \mathbb{N}_0 := \{0,1,\ldots\}, \rho \in [0,1], d \in \mathbb{N} := \{1,2,\ldots\}, \kappa, L > 0, -\infty < a < b < \infty$. Then we denote by $C_d^{r,\rho}(a,b,\kappa,L)$ the set of functions $f : [a,b] \times \mathbb{R}^d \to \mathbb{R}^d$, satisfying

$$D^\alpha f(x,z) \qquad \text{continuous} \qquad (|\alpha| \le r), \tag{3}$$

$$|D^\alpha f(x,z)| \le \kappa \qquad (|\alpha| \le r), \tag{4}$$

$$|D^r f(x,z) - D^r f(t,v)| \le \kappa(|x-t|^\rho + |z-v|^\rho), \tag{5}$$

$$|f(x,z) - f(x,v)| \le L|z-v|, \tag{6}$$

where $x,t \in [a,b], z,v \in \mathbb{R}^d$ and for $\alpha = (\alpha_0,\dots,\alpha_d) \in \mathbb{N}_0^{d+1}, |\alpha| = \alpha_0 + \ldots + \alpha_d$ and

$$D^\alpha f(x,z) := \frac{\partial^{|\alpha|} f(x,z)}{\partial x^{\alpha_0} \partial z_1^{\alpha_1} \cdots \partial z_d^{\alpha_d}}.$$

$|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^d$. Statement (5) says, that the r-th derivatives of f satisfy a ρ-Hölder condition and (6) states the Lipschitz continuity of f in the second argument.

An initial value problem (1) – (2) is given by the right-hand function f and the initial value y_0. Therefore we define the set of problem instances as follows. Fix $\sigma > 0$ and put

$$F = \{(f, y_0) : f \in C_d^{r,\rho}(a,b,\kappa,L), y_0 \in \mathbb{R}^d, |y_0| \le \sigma\}. \tag{7}$$

Because of condition (6) there exists a unique solution y for the initial value problem (1) – (2) given by $(f,y_0) \in F$, thus we define the solution operator S as

$$\begin{aligned} S: \qquad F &\longrightarrow G, \\ (f, y_0) &\longmapsto y, \end{aligned}$$

where

$$G = B([a,b], \mathbb{R}^d)$$

is the linear space of bounded functions $g : [a,b] \to \mathbb{R}^d$, equipped with the norm $||g||_\infty = \sup_{x \in [a,b]} |g(x)|$.

In the framework of IBC, we have to specify the information the algorithm is allowed to use for calculating an approximate solution. In our case we define two sets of information functionals:

$$\begin{aligned} \bar\Lambda^{\mathrm{st}} &= \{\delta_{i,s}^\alpha : 1 \le i \le d, s \in [a,b] \times \mathbb{R}^d, \alpha \in \mathbb{N}_0^{d+1}, |\alpha| \le r\} \cup \{\delta_i : 1 \le i \le d\}, \\ \Lambda^{\mathrm{st}} &= \{\delta_{i,s} : 1 \le i \le d, s \in [a,b] \times \mathbb{R}^d\} \cup \{\delta_i : 1 \le i \le d\}, \end{aligned}$$

where

$$\delta_{i,s}^\alpha(f,y_0) = D^\alpha f_i(s), \quad \delta_{i,s}(f,y_0) = f_i(s), \quad \delta_i(f,y_0) = y_{0,i}.$$

The index i indicates the i-th component of f and y_0. If an algorithm is allowed to do information calls with respect to $\bar\Lambda^{\mathrm{st}}$, we are allowed to use evaluations of f and evaluations of derivatives of f. In the case of Λ^{st} we are only allowed to evaluate f. Thus Λ^{st} is a proper subset of $\bar\Lambda^{\mathrm{st}}$.

We are especially interested in the so called m-th minimal error of an algorithm. To specify this error we have to define the different kinds of algorithms, being feasible for solving numerical problems. In our case the class of randomized algorithms is of particular interest. A randomized algorithm A is a family of deterministic algorithms $A_\omega : F \to G$ depending on a randomized parameter $\omega \in \Omega$ of the associated probability space $(\Omega, \Sigma, \mathbb{P})$. In our case a deterministic algorithm takes as input an initial value problem defined by (f, y_0) and calculates an approximate solution $\bar y(x)$ to the exact solution $y(x)$ for $x \in [a, b]$. For the calculation, the algorithm is allowed to do information calls as defined above in an adaptive way. Adaptive means that the algorithm uses knowledge of previous information calls to decide which information call he uses next. For a non-adaptive algorithm the information calls are fixed before calculation. The randomized method we will define in the next section is an example of an adaptive algorithm.

In the case where the set Ω has only one single element the family A_ω describes a deterministic algorithm. Therefore we consider a deterministic algorithm as a special case of a randomized algorithm. To distinguish deterministic and randomized algorithms we use the superscripts "det" and "ran".

The error of a randomized algorithm A with respect to S, F, is defined as

$$e(S, A, F) = \sup_{(f,y_0)\in F} \left(\mathbb{E}\, ||S(f, y_0) - A_\omega(f, y_0)||_\infty^2\right)^{1/2}.$$

Then we define the m-th minimal error as

$$e_m^{\text{ran}}(S, F, \Lambda) = \inf_{\text{card}(A,F)\leq \text{m}} e(S, A, F),$$

where $\text{card}(A, F)$ describes the average number of information functionals that A needs for the calculation and Λ is the admissible information for A. Formal definitions and more details on these notions can be found in [1],[2],[3] as well as in the monographs [8] and [9].

Based on these definitions we give a short survey of important results for the considered problem. In [4], B. Kacewicz proved

$$c_1 m^{-r-\rho} \leq e_m^{\text{det}}(S, F, \bar{\Lambda}^{\text{st}}) \leq c_2 m^{-r-\rho}.$$

Later he considered the randomized case and showed in [5] for every $\varepsilon > 0$:

$$c_3 m^{-r-\rho-1/2} \leq e_m^{\text{ran}}(S, F, \bar{\Lambda}^{\text{st}}) \leq c_4 m^{-r-\rho-1/2+\varepsilon}.$$

Recently Heinrich and Milla presented an order optimal algorithm for the problem and proved in [1] that

$$c_5 m^{-r-\rho-1/2} \leq e_m^{\text{ran}}(S, F, \bar{\Lambda}^{\text{st}}) \leq c_6 m^{-r-\rho-1/2}.$$

Here and below we use the symbols c, c_0, c_1, etc. to denote positive real valued constants not depending on $m, n, k, (f, y_0) \in F$. If the specific value of the constant is not important, the same symbol may be used for different values.

In the next sections we consider the m-th minimal error in the deterministic and in the randomized case and we will show that the same orders hold even for the weaker information class Λ^{st}. To prove this we introduce a randomized algorithm in the next section and show its optimal order in Section 4.

3 The Algorithm

We define a randomized algorithm based on a certain family of deterministic algorithms:

For $n \in \mathbb{N}$ let $h = (b-a)/n$ and for $k \in \{0, 1, \ldots, n\}$ we set $x_k = a + kh$. Let $0 \le \theta \le r + \rho + 1$ and $(\mathcal{D}_{n,k})_{n\in\mathbb{N},k\in\{0,\ldots,n-1\}}$ be an arbitrary family of deterministic algorithms $\mathcal{D}_{n,k} : C_d^{r,\rho}(a, b, \kappa, L) \times \mathbb{R}^d \to B([x_k, x_{k+1}], \mathbb{R}^d)$ having the property that there exists a constant $c > 0$ such that for all $n \in \mathbb{N}, k \in \{0, \ldots, n-1\}, f \in C_d^{r,\rho}(a, b, \kappa, L)$ and $v_0 \in \mathbb{R}^d$:

$$||v - \mathcal{D}_{n,k}(f, v_0)||_{B([x_k,x_{k+1}],\mathbb{R}^d)} \le ch^\theta, \tag{8}$$

where v is the solution of

$$v'(x) = f(x, v(x)) \qquad (x \in [x_k, x_{k+1}]), \tag{9}$$
$$v(x_k) = v_0. \tag{10}$$

Put $\mathcal{D}_n = (\mathcal{D}_{n,k})_{k=0}^{n-1}$. Based on such a family of deterministic (local) algorithms we continue by defining a randomized algorithm $A_{\mathcal{D}_n}$ for the global solution.

Let $n \ge 2$ and $(f, y_0) \in F$. We calculate vectors $y_{k+1} \in \mathbb{R}^d$ $(k = 0, \ldots, n-2)$ and $\mathbb{R}^d$-valued polynomials p_k $(k = 0, \ldots, n-1)$ inductively, beginning with $k = 0$. Let u_k be the solution of the k-th local initial value problem

$$u_k'(x) = f(x, u_k(x)) \qquad (x \in [x_k, x_{k+1}]), \tag{11}$$
$$u_k(x_k) = y_k. \tag{12}$$

We calculate approximate values to u_k by

$$u_{k,i} := \mathcal{D}_{n,k}(f, y_k)(x_{k,i}) \approx u_k(x_{k,i}), \tag{13}$$

where $x_{k,i} := x_k + \frac{i}{r}h,\ i \in \{0, 1, \ldots, r\}$. Then we obtain an approximation to u_k' by the interpolation polynomial q_k of degree at most r satisfying

$$q_k(x_{k,i}) = f(x_{k,i}, u_{k,i}) \qquad (i = 0, 1, \ldots, r). \tag{14}$$

Integrating q_k yields an approximation p_k of u_k by choosing $p_k(x_k) = y_k$. If $k = n-1$, we stop here. If $k < n-1$, we let ξ_{k+1} be uniformly distributed in $[x_k, x_{k+1}]$. Then we calculate

$$y_{k+1} = p_k(x_{k+1}) + h(f(\xi_{k+1}, p_k(\xi_{k+1})) - p_k'(\xi_{k+1})). \tag{15}$$

The output of the algorithm is the following function $\bar{y} \in B([a,b],\mathbb{R}^d)$:

$$\bar{y}(x) = \begin{cases} p_k(x) & \text{if } \ x \in [x_k, x_{k+1}) \ \text{ and } \ 0 \le k < n-1, \\ p_{n-1}(x) & \text{if } \ x \in [x_{n-1}, x_n]. \end{cases} \tag{16}$$

Thus we put $A_{\mathcal{D}_n}(f, y_0) = \bar{y}$.

Remark. We explain (15) in the definition of the randomized algorithm. We have

$$y(x_{k+1}) = y(x_k) + \int_{x_k}^{x_{k+1}} f(t, y(t))\ dt. \tag{17}$$

Since $y(x_k)$ and $y(t)$ are not known itself we use the approximation p_k. Therefore we conclude

$$y(x_{k+1}) \approx y_k + \int_{x_k}^{x_{k+1}} f(t, p_k(t))\ dt. \tag{18}$$

To calculate an approximation of the integral we use Monte-Carlo Integration with separation of the main part as in [1]. As a control variate we choose $p_k'(x)$. Thus

$$\begin{aligned} y_k + \int_{x_k}^{x_{k+1}} f(t, p_k(t))\ dt = &\ y_k + \int_{x_k}^{x_{k+1}} p_k'(t)\ dt + \int_{x_k}^{x_{k+1}} (f(t, p_k(t)) - p_k'(t))\ dt \\ \approx &\ p_k(x_{k+1}) + h(f(\xi_{k+1}, p_k(\xi_{k+1})) - p_k'(\xi_{k+1})). \end{aligned}$$

4 Analysis

Our aim in this section is to prove the optimal order of the algorithm defined in Section 3 for admissible information Λ^{st} in the deterministic and in the randomized case.

For analyzing our algorithm we need an estimate for the interpolation error of r-times continuously differentiable functions whose r-th derivative satisfies a ρ-Hölder condition.

Lemma 2. *Let $r \in \mathbb{N}_0, \rho \in [0,1], \kappa_1 \in \mathbb{R}$. Then there are constants $c_1, c_2 > 0$ such that for all $a_1, b_1 \in \mathbb{R}$ with $-\infty < a_1 < b_1 < \infty$ and for all $g \in C^r([a_1, b_1])$ satisfying*

$$|g^{(r)}(x) - g^{(r)}(t)| \le \kappa_1 |x - t|^\rho \qquad (x, t \in [a_1, b_1]), \tag{19}$$

the following holds: Let p be the interpolation polynomial of degree at most r with

$$p(t_i) = g(t_i),\ t_i = a_1 + \frac{i}{r}(b_1 - a_1) \qquad (i = 0, 1, \ldots, r),$$

then

$$\sup_{x \in [a_1, b_1]} |g(x) - p(x)| \le c_1 (b_1 - a_1)^{r+\rho}. \tag{20}$$

Moreover, for any polynomial q of degree at most r:

$$\sup_{x \in [a_1, b_1]} |g(x) - q(x)| \le c_1 (b_1 - a_1)^{r+\rho} + c_2 \max_{0 \le i \le r} |g(t_i) - q(t_i)|. \tag{21}$$

This is a well-known result. However, since our formulation involves estimates with constants independent of the interval limits a_1, b_1, we include the short and elementary proof for the sake of completeness.

Proof. For a function $f \in C([a_1, b_1])$ let

$$(Pf)(t) = \sum_{i=0}^{r} f(t_i) l_i(t)$$

be the interpolating polynomial, where

$$l_i(t) = \begin{cases} \prod_{j=0, j\neq i}^{r} \frac{t-t_i}{t_i-t_j} & r \geq 1 \\ 1 & r = 0 \end{cases}$$

are the Lagrange polynomials. We have

$$\sup_{t\in[a_1,b_1]} |l_i(t)| \leq c \tag{22}$$

with a constant c independent of a_1, b_1. Furthermore, by Taylor's formula, for $r \geq 1$,

$$g(x) = \sum_{i=0}^{r} \frac{(x-a_1)^i}{i!} g^{(i)}(a_1) + \int_{a_1}^{x} \frac{(x-t)^{r-1}}{(r-1)!} (g^{(r)}(t) - g^{(r)}(a_1))\, dt. \tag{23}$$

Then (22), (23) and (19) imply (20) in the case $r \geq 1$. If $r = 0$, (20) follows directly from (19). Moreover, we have

$$\sup_{x\in[a_1,b_1]} |g(x) - q(x)| \leq \sup_{x\in[a_1,b_1]} |g(x) - (Pg)(x)| + \sup_{x\in[a_1,b_1]} |(P(g-q))(x)|,$$

which combined with (20) and (22) gives (21). □

The values, used for interpolation in the randomized algorithm, contain errors. Thus we show, using the second part of Lemma 2, that the uncertainty does not affect the error rate of the interpolation.

Lemma 3. *There are constants $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$, $k \in \{0, \ldots, n-1\}, f \in C_d^{r,\rho}(a, b, \kappa, L), v_0 \in \mathbb{R}^d$ the solution v of the initial value problem (9) − (10) satisfies*

$$|v^{(j)}(x)| \leq c_1 \qquad (x \in [x_k, x_{k+1}] \wedge j \in \{1, \ldots, r+1\}), \tag{24}$$

$$|v^{(r+1)}(x) - v^{(r+1)}(t)| \leq c_2 |x-t|^\rho \qquad (x, t \in [x_k, x_{k+1}]). \tag{25}$$

This is well-known. For a proof see e.g. [1].

Next we show the existence of a family of deterministic algorithms satisfying condition (8) for $\theta = r + \rho + 1$.

Proposition 4. *There exists a family of deterministic algorithms* $\left(\mathcal{D}^0_{n,k}\right)_{n\in\mathbb{N},k\in\{0,\ldots,n-1\}}$, $\mathcal{D}^0_{n,k} : C^{r,\rho}_d(a,b,\kappa,L)\times\mathbb{R}^d \to B([x_k,x_{k+1}],\mathbb{R}^d)$ *and a constant* $c>0$ *such that for all* $n\in\mathbb{N}$,
$k\in\{0,\ldots,n-1\}, f\in C^{r,\rho}_d(a,b,\kappa,L), v_0\in\mathbb{R}^d$ *the algorithm* $\mathcal{D}^0_{n,k}$ *uses not more than* $d[(r+1)(r+2)/2+1]$ *information functionals from* Λ^{st} *and satisfies*

$$||v-\mathcal{D}^0_{n,k}(f,v_0)||_{B([x_k,x_{k+1}],\mathbb{R}^d)} \le ch^{r+\rho+1}, \tag{26}$$

with v *defined by* $(9)-(10)$.

Proof. We define an algorithm satisfying the claimed properties.
Algorithm: Let $n\in\mathbb{N}, k\in\{0,1,\ldots,n-1\}, f\in C^{r,\rho}_d(a,b,\kappa,L), v_0\in\mathbb{R}^d$.

Step 0**:** Put

$$\bar{p}_0(x) := v_0 + (x-x_k)f(x_k,v_0). \tag{27}$$

If $r=0$ we stop and set $\mathcal{D}^0_{n,k}(f,v_0) := \bar{p}_0$. Else define $\bar{p}_{l+1}$ inductively for $l\in\{0,1,\ldots,r-1\}$:

Step $l+1$**:** Let $x_{l+1,i} := x_k + \frac{i}{l+1}h$ for $i\in\{0,1,\ldots,l+1\}$ and let $\bar{q}_{l+1}(x)$ be the $\mathbb{R}^d$-valued interpolation polynomial of degree $\le l+1$ satisfying

$$\bar{q}_{l+1}(x_{l+1,i}) = f(x_{l+1,i},\bar{p}_l(x_{l+1,i})) \qquad (i=0,1,\ldots,l+1). \tag{28}$$

Then we set

$$\bar{p}_{l+1}(x) := v_0 + \int_{x_k}^{x} \bar{q}_{l+1}(t)\, dt. \tag{29}$$

Finally $\mathcal{D}^0_{n,k}(f,v_0) := \bar{p}_r$.

Now we show that there is a constant $c>0$ such that for all n,k,f,v_0 and $l\in\{0,1,\ldots,r\}$:

$$\sup_{x\in[x_k,x_{k+1}]} |v(x)-\bar{p}_l(x)| \le ch^{l+\rho+1}. \tag{30}$$

We argue by induction: For $l = 0$ we conclude with $(9), (5), (6)$ and (24):

$$\begin{aligned}
\sup_{x\in[x_k,x_{k+1}]} |v(x) - \bar{p}_0(x)| &= \sup_{x\in[x_k,x_{k+1}]} \left| v_0 + \int_{x_k}^{x} v'(t)\ dt - v_0 - \int_{x_k}^{x} f(x_k, v_0)\ dt \right| \\
&\le h \sup_{x\in[x_k,x_{k+1}]} |f(x, v(x)) - f(x_k, v_0)| \\
&\le h \sup_{x\in[x_k,x_{k+1}]} (|f(x, v(x)) - f(x, v_0)| + |f(x, v_0) - f(x_k, v_0)|) \\
&\le\ h \sup_{x\in[x_k,x_{k+1}]} (L|v(x) - v_0| + \kappa |x - x_k|^\rho) \\
&\le h \sup_{x\in[x_k,x_{k+1}]} \left(L \left| \int_{x_k}^{x} v'(x) \right| + \kappa h^\rho \right) \\
&\le \kappa L h^2 + \kappa h^{\rho+1} \le c h^{\rho+1}.
\end{aligned}$$

Let $l \in \{0, 1, \dots, r-1\}$. In step $l+1$ we conclude with (28), Lipschitz continuity of f and the induction assumption (30), that for $0 \le i \le l+1$

$$\begin{aligned}
|v'(x_{l+1,i}) - \bar{q}_{l+1}(x_{l+1,i})| &= |f(x_{l+1,i}, v(x_{l+1,i})) - f(x_{l+1,i}, \bar{p}_l(x_{l+1,i}))| \\
&\le L|v(x_{l+1,i}) - \bar{p}_l(x_{l+1,i})| \\
&\le c h^{l+\rho+1}
\end{aligned}$$

and because of (25) and Lemma 2 applied componentwise to $g = v', a_1 = x_k, b_1 = x_{k+1}$, it follows that

$$\sup_{x\in[x_k,x_{k+1}]} |v'(x) - \bar{q}_{l+1}(x)| \le c_1 h^{l+\rho+1} + c_2 \max_{0\le i\le l+1} |v'(x_{l+1,i}) - \bar{q}_{l+1}(x_{l+1,i})| \le c h^{l+\rho+1}.$$

Integration yields

$$\begin{aligned}
\sup_{x\in[x_k,x_{k+1}]} |v(x) - \bar{p}_{l+1}(x)| &\le \sup_{x\in[x_k,x_{k+1}]} \int_{x_k}^{x} |v'(t) - \bar{q}_{l+1}(t)|\ dt \\
&\le c h^{l+\rho+2},
\end{aligned}$$

which proves (30). The algorithm uses the following information functionals from Λ^{st}: the d components of v_0 and $d(r+1)(r+2)/2$ function values of f. $\square$

Lemma 5. *There are constants $c_3, c_4 > 0$ such that for all $(f, y_0) \in F, n \in \mathbb{N}, n \ge 2, k \in \{0, \dots, n-1\}$ the following holds:*

$$\begin{aligned}
|y_k| &\le c_3 \qquad \textit{and} \\
|u_k(x)| &\le c_4 \qquad (x \in [x_k, x_{k+1}]),
\end{aligned}$$

where u_k and y_k were defined in $(11) - (12)$ *and* (15).

Proof. According to (13) we calculate $u_{k,i}$ using $\mathcal{D}_{n,k}$ in every step k. For every $i \in \{0, 1, \ldots, r\}$ we calculate

$$f(x_{k,i}, u_{k,i})$$

and obtain with condition (4)

$$|f(x_{k,i}, u_{k,i})| \leq \kappa. \tag{31}$$

We conclude using Lagrangian polynomials

$$\sup_{x \in [x_k, x_{k+1}]} |p_k'(x)| \leq c. \tag{32}$$

The algorithm calculates in every step k

$$y_{k+1} = \; y_k + \int_{x_k}^{x_{k+1}} p_k'(t) \; dt + h(f(\xi_{k+1}, p_k(\xi_{k+1})) - p_k'(\xi_{k+1})),$$

hence with (4) and (32)

$$\begin{aligned} |y_{k+1} - y_k| &\leq h|p_k'(x)| + h|(f(\xi_{k+1}, p_k(\xi_{k+1})) - p_k'(\xi_{k+1}))| \\ &\leq ch \end{aligned}$$

and because of (7) we conclude

$$\max_{i \in \{0, \ldots, n-1\}} |y_i| \leq |y_0| + n \max_{i \in \{0, \ldots, n-1\}} |y_{i+1} - y_i| \leq c.$$

Since $|u_k(x)| = |y_k + \int_{x_k}^{x} u_k'(t) \; dt|$ and by (4) and (11), $|u_k'(x)| \leq c$ for $x \in [x_k, x_{k+1}]$, the second statement follows. □

Proposition 6. *Let $0 \leq \theta \leq r + \rho + 1$ and let $(\mathcal{D}_{n,k})_{n \in \mathbb{N}, k \in \{0, \ldots, n-1\}}$ be any family as described before (see (8) and above). Then there is a constant $c > 0$ such that for all $(f, y_0) \in F, n \in \mathbb{N}, n \geq 2$ the error of the randomized algorithm satisfies*

$$\sqrt{\mathbb{E}||S(f, y_0) - A_{\mathcal{D}_n}(f, y_0)||_\infty^2} \leq c h^{\min(r+\rho, \theta) + 1/2}.$$

Proof. Let $n \in \mathbb{N}, n \geq 2, k \in \{0, 1, \ldots, n-1\}$. By (8) and (13) we have

$$|u_k(x_{k,i}) - u_{k,i}| \leq c h^\theta,$$

which together with (14) and (6) gives

$$|u_k'(x_{k,i}) - q_k(x_{k,i})| = |f(x_{k,i}, u_k(x_{k,i})) - f(x_{k,i}, u_{k,i})| \leq c h^\theta.$$

By (25) of Lemma 3 we have

$$|u_k^{(r+1)}(x) - u_k^{(r+1)}(t)| \leq c|x - t|^\rho \qquad (x, t \in [x_k, x_{k+1}]).$$

Therefore we can apply Lemma 2 with $g = u_k', a_1 = x_k, b_1 = x_{k+1}$ and get

$$\sup_{x\in[x_k,x_{k+1}]} |u_k'(x) - q_k(x)| \le c_1 h^{r+\rho} + c_2 \max_{0\le i\le r} |u_k'(x_{k,i}) - q_k(x_{k,i})| \le ch^{\min(r+\rho,\theta)}.$$

By the definition of p_k, this gives

$$\sup_{x\in[x_k,x_{k+1}]} |u_k'(x) - p_k'(x)| \le ch^{\min(r+\rho,\theta)}. \tag{33}$$

Exploiting $p_k(x_k) = u_k(x_k) = y_k$ yields

$$\begin{aligned}\mu_k := &\sup_{x\in[x_k,x_{k+1}]} |u_k(x) - p_k(x)| \\ = &\sup_{x\in[x_k,x_{k+1}]} \left|\int_{x_k}^x u_k'(t)\ dt - \int_{x_k}^x p_k'(t)\ dt\right| \\ \le &\ ch^{\min(r+\rho,\theta)+1}. \end{aligned} \tag{34}$$

Since (34) and (33) correspond to (23) and (25) of [1], the rest of the proof of Proposition 1 in [1] goes through literally, replacing γ by $\min(r+\rho,\theta)$. We do not repeat it here. □

Theorem 7. *Let $r \in \mathbb{N}_0, \rho \in [0,1], \gamma = r + \rho$ and $S, F, \Lambda^{\mathrm{st}}$ be as in Section 2. Then there are constants $c_1, c_2, c_3, c_4 > 0$ such that for every $m \in \mathbb{N}$:*

$$\begin{aligned} c_1 m^{-\gamma-1/2} \ &\le\ e_m^{\mathrm{ran}}(S, F, \Lambda^{\mathrm{st}}) \ \le\ c_2 m^{-\gamma-1/2}, \\ c_3 m^{-\gamma} \ &\le\ e_m^{\mathrm{det}}(S, F, \Lambda^{\mathrm{st}}) \ \le\ c_4 m^{-\gamma}, \end{aligned}$$

holds.

Proof. Let $C = d((r+2)(r+3)/2+1)$. For any $m \in \mathbb{N}$ with $m \ge 2C$ we put

$$n = \left\lfloor \frac{m}{C} \right\rfloor \ge 2$$

and apply the randomized algorithm $A_{\mathcal{D}_n^0}$. This algorithm uses not more than

$$d\left(1 + \frac{(r+1)(r+2)}{2} + r + 2\right) n = d\left(\frac{(r+2)(r+3)}{2} + 1\right) n = Cn \le m$$

information functionals (see Proposition 4 and relations (14) and (15)). Together with Proposition 6, this implies

$$e_m^{\mathrm{ran}}(S, F, \Lambda^{\mathrm{st}}) \le cn^{-r-\rho-1/2} \le c_1 m^{-r-\rho-1/2} \quad (m \ge 2C).$$

For $m < 2C$ we use the zero algorithm just giving zero for all algorithm calls. By assumption $|y_0| \le \sigma$, thus $||y||_\infty \le \sigma + (b-a)\kappa$ and therefore $||y - 0||_\infty \le \sigma + (b-a)\kappa$.

For the deterministic case we note that each realization of the randomized algorithm (that is, $\omega \in \Omega$ is fixed, meaning any realizations of $\xi_1, \ldots, \xi_n$ are fixed) is a deterministic algorithm with optimal deterministic error. Indeed, we use the same argument as above, except that we need one modification in the proof of Proposition 6, or more precisely in the part contained in [1]: We estimate the middle term of the last line of (40) in [1] by

$$\max_{1 \le j \le n-1} \left| \sum_{i=1}^{j} \eta_i \right| \le \sum_{i=1}^{n-1} |\eta_i| \le c h^{r+\rho}$$

and obtain

$$||y - \bar{y}||_\infty \le c h^{r+\rho}. \tag{35}$$

A similar observation was made in [1]. The lower bounds are immediate consequences of the results from [1] and [4], since $\Lambda^{\mathrm{st}} \subset \bar{\Lambda}^{\mathrm{st}}$. $\square$

Remark on the sampling region. Note that $A_{\mathcal{D}_n^0}$ samples the function f only in a neighborhood of the true solution. Let us make this more precise. First, there is a constant $c > 0$ such that the algorithm $\mathcal{D}_{n,k}^0$ uses function values of f only in points $(t, z) \in [x_k, x_{k+1}] \times \mathbb{R}^d$ with the property

$$|v_0 - z| \le ch.$$

This follows from (9) – (10) and (27) – (30) . Based on this and (32) it follows that the resulting randomized algorithm samples f for all $k \in \{0, \ldots, n-1\}$ only in points $(t, z) \in [x_k, x_{k+1}] \times \mathbb{R}^d$ with

$$|p_k(t) - z| \le |p_k(t) - \underbrace{p_k(x_k)}_{y_k}| + |y_k - z| \le ch,$$

which by (35) implies for all sample points $(t, z) \in [a, b] \times \mathbb{R}^d$

$$|y(t) - z| \le |y(t) - \bar{y}(t)| + |\bar{y}(t) - z| \le c h^{\min(r+\rho,1)}. \tag{36}$$

For the case $r + \rho = 0$ we use Proposition 2 of [1], which carries over to the situation of our paper and asserts that there are constants $\tilde{c}_1, \tilde{c}_2 > 0$ such that for all $\tau \ge \tilde{c}_1$ and all $(f, y_0) \in F$

$$\mathbb{P}\left\{||S(f, y_0) - A_{\mathcal{D}_n^0}(f, y_0)||_\infty \ge \tau n^{-1/2}\right\} = \mathbb{P}\left\{||y - \bar{y}||_\infty \ge \tau n^{-1/2}\right\} \le \exp(-\tilde{c}_2 \tau^2).$$

Thus, for $r + \rho = 0$ and any $\tau \ge \tilde{c}_1$ the algorithm samples f with probability $\ge 1 - \exp(-\tilde{c}_2 \tau^2)$ only in points $(t, z) \in [a, b] \times \mathbb{R}^d$ with

$$|y(t) - z| \le \tilde{c}_3 n^{-1} + \tau n^{-1/2}. \tag{37}$$

Remark on scalar equations of higher order. Let $r \in \mathbb{N}_0, \rho \in [0, 1], \ell \in \mathbb{N}, \kappa, L > 0, -\infty < a < b < \infty$. Then we denote by $\hat{C}_\ell^{r,\rho}(a, b, \kappa, L)$ the set of functions f :

$[a,b] \times \mathbb{R}^\ell \to \mathbb{R}$ satisfying

$$\begin{aligned}
D^\alpha f(x,z) \quad &\text{continuous} && (|\alpha| \le r),\\
|D^\alpha f(x,z)| &\le \kappa && (|\alpha| \le r),\\
|D^r f(x,z) - D^r f(t,v)| &\le \kappa(|x-t|^\rho + |z-v|^\rho),\\
|f(x,z) - f(x,v)| &\le L|z-v|,
\end{aligned}$$

for $x,t \in [a,b], z,v \in \mathbb{R}^\ell$. Let $\sigma > 0$ be fixed, $f \in \hat{C}_\ell^{r,\rho}(a,b,\kappa,L)$ and $w_0 = (w_{i,0})_{i=0}^{\ell-1} \in \mathbb{R}^\ell$ with $|w_0| \le \sigma$, then

$$w^{(\ell)}(x) = f(x, w(x), w'(x), \dots, w^{(\ell-1)}(x)), \tag{38}$$

$$w(a) = w_{0,0},\ w'(a) = w_{1,0}, \dots, w^{(\ell-1)}(a) = w_{\ell-1,0}, \tag{39}$$

defines an ordinary differential equation of order ℓ. The complexity of such equations was considered before in [6] and [7]. In [7] the order of the m-th minimal error for the randomized setting with $\bar{\Lambda}^{\mathrm{st}}$ was determined up to a gap of an arbitrarily small $\varepsilon > 0$ in the exponent of m. As an immediate consequence of our results we can close this gap and give the sharp order of the m-th minimal error both for $\bar{\Lambda}^{\mathrm{st}}$ and Λ^{st}.

We reduce problem (38) – (39) to a system in the standard way. Let $y_0(x) := w(x), y_1(x) := w'(x), \dots, y_{\ell-1}(x) = w^{(\ell-1)}(x)$, then

$$\begin{aligned}
y_0'(x) &= y_1(x), & y_0(a) &= w_{0,0},\\
y_1'(x) &= y_2(x), & y_1(a) &= w_{1,0},\\
&\ \vdots & &\ \vdots\\
y_{\ell-2}'(x) &= y_{\ell-1}(x), & y_{\ell-2}(a) &= w_{\ell-2,0},\\
y_{\ell-1}'(x) &= f(x, y_0(x), \dots, y_{\ell-1}(x)), & y_{\ell-1}(a) &= w_{\ell-1,0},
\end{aligned} \tag{40}$$

defines an equivalent system of ordinary differential equations of order 1 and dimension ℓ. We denote the right-hand side function of system (40) by f^{sys}, that is, $f^{\mathrm{sys}} : [a,b] \times \mathbb{R}^\ell \to \mathbb{R}^\ell$,

$$f^{\mathrm{sys}}(x, z_0, z_1, \dots, z_{\ell-1}) = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_{\ell-1} \\ f(x, z_0, z_1, \dots, z_{\ell-1}) \end{pmatrix}.$$

For this system we cannot apply Theorem 7 directly, since f^{sys} does not satisfy condition (4) for $\alpha = 0$. But by (38)

$$|w^{(\ell)}(x)| = |f(x, w(x), \dots, w^{(\ell-1)}(x))| \le \kappa \qquad (x \in [a,b]),$$

which together with $|w_0| \le \sigma$ and (39) gives

$$|y_i(x)| = |w^{(i)}(x)| \le c_0 \qquad (x \in [a,b], i = \ell-1, \dots, 0) \tag{41}$$

for some $c_0 > 0$. Let $\psi \in C^{r+1}(\mathbb{R})$ be such that $\psi(t) = t$ for $|t| \le 2c_0$ and $\psi(t) = 0$ for $|t| \ge 3c_0$. Define $g^{\mathrm{sys}} : [a,b] \times \mathbb{R}^\ell \to \mathbb{R}^\ell$, by

$$g^{\mathrm{sys}}(x, z_0, \dots, z_{\ell-1}) = \begin{pmatrix} \psi(z_1) \\ \psi(z_2) \\ \vdots \\ \psi(z_{\ell-1}) \\ f(x, z_0, \dots, z_{\ell-1}) \end{pmatrix}.$$

Since $\psi \in C^{r+1}(\mathbb{R})$, there are constants $\kappa_1, L_1 > 0$ such that for all $f \in \hat{C}_\ell^{r,\rho}(a,b,\kappa,L)$ we have $g^{\mathrm{sys}} \in C_\ell^{r,\rho}(a,b,\kappa_1,L_1)$. It follows from (41) that the solution $y(x) = (y_i(x))_{i=0}^{\ell-1}$ of system (40) satisfies

$$\begin{aligned} y'(x) &= f^{\mathrm{sys}}(x, y(x)) = g^{\mathrm{sys}}(x, y(x)) \qquad (x \in [a,b]), \\ y(a) &= w_0. \end{aligned} \tag{42}$$

Thus we can apply the randomized algorithm from above to g^{sys} and obtain the same convergence as in Theorem 7 for ordinary differential equations of order ℓ. The corresponding lower bound is contained in [7].

Note that the resulting order-optimal randomized algorithm formally depends on the choice of ψ and thus, on the smoothness constants of the class. In the sequel we show that this can be avoided. By (36) there is a constant $c_1 > 0$ such that $A_{\mathcal{D}_n^0}$ samples g^{sys} only in points $(t,z) \in [a,b] \times \mathbb{R}^\ell$ with

$$|y_i(t) - z_i| \le c_1 n^{-\min(r+\rho,1)} \qquad (i = 0, \dots, \ell-1).$$

Thus, for $r + \rho > 0$ and

$$n \ge \left(\frac{c_1}{c_0}\right)^{\min(r+\rho,1)^{-1}}$$

we have $|z_i| \le 2c_0$, hence $\psi(z_i) = z_i \quad (i = 0, \dots, z_{\ell-1})$ and therefore

$$g^{\mathrm{sys}}(t, z_0, \dots, z_{\ell-1}) = f^{\mathrm{sys}}(t, z_0, \dots, z_{\ell-1}).$$

This implies that the algorithm, if applied to f^{sys}, produces the same result as if applied to g^{sys}. Consequently, we obtain the same convergence rate as in Theorem 7 for the algorithm applied to f^{sys} directly.

If $r+\rho=0$, we use (37) and obtain for

$$n \geq \max\left\{2\frac{\tilde{c}_3}{c_0}, 4\left(\frac{\tilde{c}_1}{c_0}\right)^2\right\}, \quad \tau = \frac{c_0 n^{1/2}}{2},$$

that with probability

$$\geq 1 - \exp\left(-\frac{\tilde{c}_2 c_0^2 n}{4}\right)$$

all sampling points satisfy $|y(t)-z| \leq c_0$. Arguing as above, it follows that the rate of Theorem 7 holds with high probability for the algorithm applied to f^{sys} directly.

5 Numerical Results

In this section we present some numerical results. In the first two examples we compare our algorithm based on a 3-stage Runge-Kutta method with the 3-stage Runge-Kutta method itself for different test functions f. The maximum of all errors in the sample points of the interval $[a,b]$ will be displayed in the right-hand graph of the figures.

In Figure 1 we present the error for $f(x,y) := g(x)y^2$, where $g \in C^1(\mathbb{R})$ as shown in the left graph of the figure. For this example we observe a gain, with respect to the convergence of the error, by the randomized method.

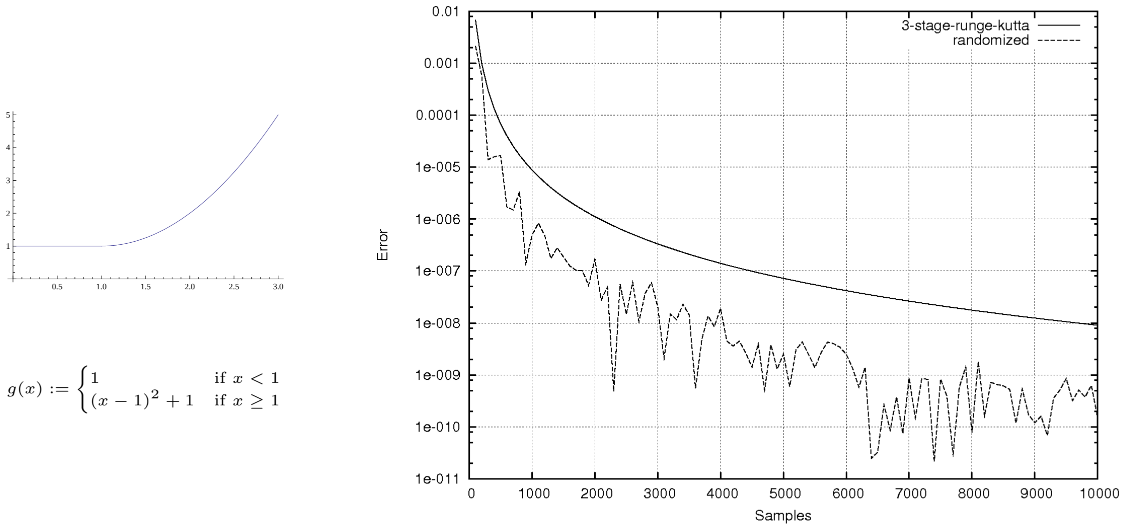
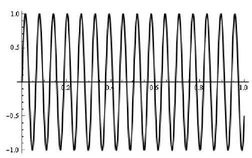


Figure 1: Plot of the error for $f(x,y) = g(x)y^2$ in log scale for the y-axis

For another example we chose a more complicated test function g, the highly oscillatory function $g(x) := \sin(100x)$. In Figure 2 we see, that the gain is even bigger than in the first example.



$g(x) := \sin(100x)$

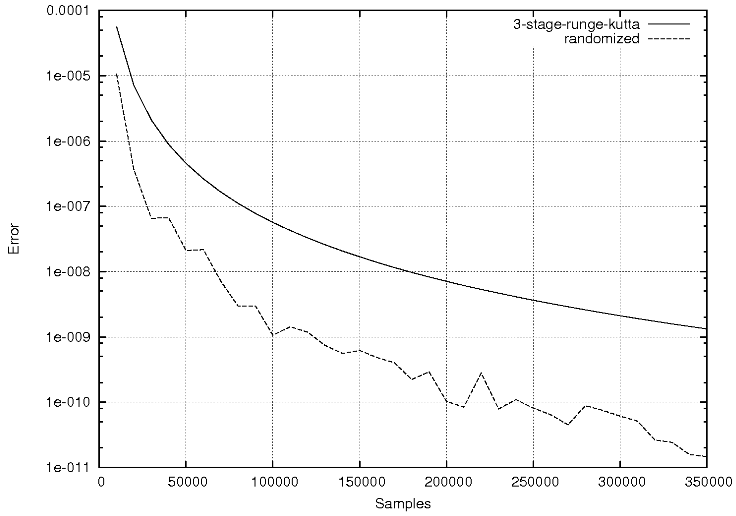
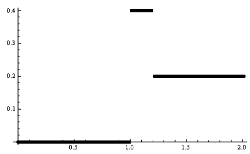


Figure 2: Plot of the error for $f(x, y) = g(x)y^2$ in log scale for the y-axis

For a last example we chose a piecewise constant function as g. Here we compared the randomized method based on the Euler method with the Euler method itself. Our results shown in Figure 3 confirm that in particular for functions f with low degree of smoothness a randomized method can be better than a deterministic one.

Acknowledgment

I would like to thank my supervisor Prof. Stefan Heinrich for helpful discussions and support.



$g(x)$ piecewise constant

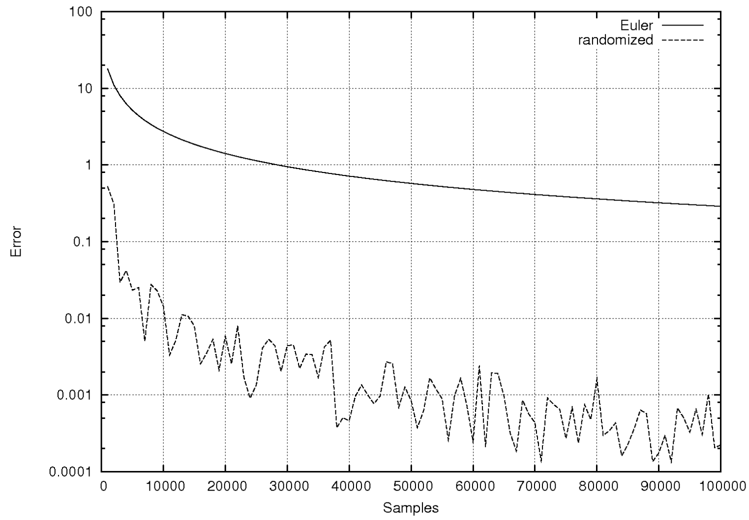


Figure 3: Plot of the error for $f(x, y) = g(x)y^2$ in log scale for the y-axis

References

[1] S. Heinrich, B. Milla, The randomized complexity of initial value problems, J. Complexity 24 (2008), 77–88.

[2] S. Heinrich, Monte Carlo approximation of weakly singular integral operators, J. Complexity 22 (2006), 192–219.

[3] S. Heinrich, The randomized information complexity of elliptic PDE, J. Complexity 22 (2006), 220–249.

[4] B. Kacewicz, How to increase the order to get minimal-error algorithms for systems of ODE, Numer. Math. 45 (1984), 93–104.

[5] B. Kacewicz, Almost optimal solution of initial-value problems by randomized and quantum algorithms, J. Complexity 22 (2006), 676–690, see also http://arXiv.org/abs/quant-ph/0510045.

[6] M. Szczęsny, Complexity of initial-value problems for ordinary differential equations of order k, J. Complexity 22 (2006), 514–532

[7] M. Goćwin, M. Szczęsny, Randomized and quantum algorithms for solving initial-value problems in ordinary differential equations of order k, Opuscula Mathematica 28 (2008), 247–277

[8] E. Novak, Deterministic and Stochastic Error Bounds in Numerical Analysis, Lecture Notes in Mathematics 1349, Springer, 1988.

[9] J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski, Information-Based Complexity, Academic Press, New York, 1988.