Complexity of initial value problems in Banach spaces

Stefan Heinrich
Department of Computer Science
University of Kaiserslautern
D-67653 Kaiserslautern, Germany

Dedicated to the memory of Mikhail Iosifovich Kadets

Abstract

We study the complexity of initial value problems for Banach space valued ordinary differential equations in the randomized setting. The right-hand side is assumed to be r-smooth, the r-th derivatives being ϱ-Hölder continuous. We develop and analyze a randomized algorithm. Furthermore, we prove lower bounds and thus obtain complexity estimates. They are related to the type of the underlying Banach space. We also consider the deterministic setting. The results extend previous ones for the finite dimensional case from [10, 9, 2].

Keywords: ordinary differential equation, initial value problem, Banach space, Monte Carlo algorithm, information-based complexity, lower bounds

MSC2010: 65C05, 65J15, 65L05, 68Q17, 46N40

1 Introduction

Randomized solution of initial value problems for ordinary differential equations (ODEs) has been studied in various papers [16, 17, 11, 12, 9, 2], all of them dealing with the $\mathbb{R}^n$-valued case. In this paper we study initial value problems for Banach space valued ODEs. We develop a randomized algorithm and analyze its convergence, extending results from [9, 2].

We also prove lower bounds and consider the complexity. It turns out that the complexity is connected with the type of the underlying Banach space. For general Banach spaces upper and lower bounds are almost matching in the sense that an arbitrarily small gap in the exponent remains. For special spaces, including the L_p spaces, the bounds are matching and the algorithm is of optimal order. Furthermore, we show that for arbitrary Banach spaces and for any fixed choice of the random parameters the resulting deterministic algorithm is of optimal order in the deterministic setting. This way we generalize complexity results from [10].

The results of the Banach space valued case are a crucial tool for the complexity analysis of parameter dependent ODEs, a topic which will be treated in a subsequent paper [4].

2 Preliminaries and the problem

Let $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. For a Banach space X the closed unit ball is denoted by B_X, the identity operator by I_X, and the dual space by X^*. Given $k \in \mathbb{N}_0$ and another Banach space Y, we set $\mathscr{L}_0(X, Y) = Y$, while for $k \geq 1$ we let $\mathscr{L}_k(X, Y)$ be the space of bounded multilinear mappings $T : X^k \to Y$ endowed with the canonical norm

$$\|T\|_{\mathscr{L}_k(X,Y)} = \sup_{x_1,\dots,x_k \in B_X} \|T(x_1, \dots, x_k)\|.$$

If $k = 1$, this is the space of bounded linear operators, for which we write $\mathscr{L}(X, Y)$. If $X = Y$, we write $\mathscr{L}_k(X)$ instead of $\mathscr{L}_k(X, X)$ and $\mathscr{L}(X)$ instead of $\mathscr{L}(X, X)$.

Let $1 \leq p \leq 2$. A Banach space X is said to be of (Rademacher) type p (see [14, 13]), if there is a constant $c > 0$ such that for all $n \in \mathbb{N}$ and $x_1, \dots, x_n \in X$

$$\mathbb{E} \Big\| \sum_{i=1}^{n} \varepsilon_i x_i \Big\|^p \leq c^p \sum_{k=1}^{n} \|x_i\|^p, \tag{1}$$

where $(\varepsilon_i)_{i=1}^n$ is a sequence of independent Bernoulli random variables with $\mathbb{P}\{\varepsilon_i = -1\} = \mathbb{P}\{\varepsilon_i = +1\} = 1/2$. The smallest constant satisfying (1) is called the type p constant if X and is denoted by $\tau_p(X)$. If there is no such $c > 0$, we put $\tau_p(X) = \infty$. The space $L_{p_1}(\mathcal{N}, \nu)$ with $(\mathcal{N}, \nu)$ an arbitrary measure space and $p_1 < \infty$ is of type p with $p = \min(p_1, 2)$. We will use the following result (see [13], Prop. 9.11).

Lemma 2.1. *Let $1 \leq p \leq 2$, let X be a Banach space, $n \in \mathbb{N}$ and $(\theta_i)_{i=1}^n$ be a sequence of independent X-valued random variables with $E\|\theta_i\|^p < \infty$ and $\mathbb{E}\,\theta_i = 0$ $(i = 1, \dots, n)$. Then*

$$\left(\mathbb{E} \Big\| \sum_{i=1}^{n} \theta_i \Big\|^p \right)^{1/p} \leq 2\tau_p(X) \left(\sum_{k=1}^{n} \mathbb{E}\, \|\theta_i\|^p \right)^{1/p}.$$

We will work in the setting of information-based complexity theory (IBC), see [18, 15]. For the notation used here we also refer to [7, 8]. Let us first explain the general approach, later we specify everything for the initial value problems.

An abstract numerical problem is given by a tuple $\mathcal{P} = (F, G, S, K, \Lambda)$, where F is a non-empty set – the set of inputs, G a normed space, $S : F \to G$ an arbitrary mapping – the solution operator, which maps the input $\psi \in F$ to the exact solution of the problem $S(\psi) \in G$. Furthermore, K is another non-empty set and Λ is any set of mappings from F to K – the set of information functionals.

Next we define classes of algorithms for $\mathcal{P}$. In this paper we consider adaptive deterministic and randomized algorithms of fixed cardinality (all algorithms developed later on will be of this type). For the respective notions of algorithms with varying cardinality see [7, 8].

First we consider the deterministic case and introduce the class $\mathcal{A}_n^{\det}(F, G)$ of deterministic algorithms for $\mathcal{P}$ which use n information functionals, where $n \in \mathbb{N}$. An element $A \in \mathcal{A}_n^{\det}(F, G)$ is a tuple $A = ((\mu_i)_{i=1}^n, \varphi)$, where $\mu_1 \in \Lambda$ and

$$\begin{aligned} \mu_i :&\quad K^{i-1} \to \Lambda \qquad (i = 2, 3, \dots, n) \\ \varphi :&\quad K^n \to G \end{aligned}$$

are arbitrary mappings. Given $\psi \in F$, we define $\lambda_1 = \mu_1$,

$$\lambda_i \;=\; \mu_i(\lambda_1(\psi), \dots, \lambda_{i-1}(\psi)) \quad (i = 2, 3, \dots, n),$$

and the output $A(\psi)$ of algorithm A at input ψ by

$$A(\psi) = \varphi(\lambda_1(\psi), \dots, \lambda_n(\psi)).$$

The error of A is defined as

$$e(S, A, F) = \sup_{\psi \in F} \|S(\psi) - A(\psi)\|_G.$$

Thus we measure the error in the norm of G. The central notion of IBC is the n-th minimal error, which is defined for $n \in \mathbb{N}$ as

$$e_n^{\det}(S, F) = \inf_{A \in \mathcal{A}_n^{\det}(F,G)} e(S, A, F).$$

So $e_n^{\det}(S, F)$ is the minimal possible error among all deterministic algorithms that use n information functionals.

Next we introduce the respective class of randomized algorithms. An element $A \in \mathcal{A}_n^{\mathrm{ran}}(F, G)$ is a tuple $A = ((\Omega, \Sigma, \mathbb{P}), (A_\omega)_{\omega \in \Omega})$, where $(\Omega, \Sigma, \mathbb{P})$ is a probability space,

$$A_\omega \in \mathcal{A}_n^{\det}(F, G) \quad (\omega \in \Omega),$$

and for each $\psi \in F$ the mapping

$$\omega \in \Omega \to A_\omega(\psi) \in G$$

is Σ-to-Borel measurable and essentially separably valued, that is, there is a separable subspace $G_0 \subseteq G$ such that

$$A_\omega(\psi) \in G_0 \quad \text{for } \mathbb{P}\text{-almost all } \omega \in \Omega.$$

Thus, the output $A(\psi)$ of algorithm A at input $\psi \in F$ is the G-valued random variable $A_\omega(\psi)$ on $(\Omega, \Sigma, \mathbb{P})$. The error of A is given by

$$e(S, A, F) = \sup_{\psi \in F} \mathbb{E} \, \|S(\psi) - A_\omega(\psi)\|_G$$

and the n-th minimal error for $n \in \mathbb{N}$ by

$$e_n^{\mathrm{ran}}(S, F) = \inf_{A \in \mathcal{A}_n^{\mathrm{ran}}(F,G)} e(S, A, F).$$

Consequently, $e_n^{\mathrm{ran}}(S, F)$ is the minimal possible error among all randomized algorithms that use n information functionals.

The minimal errors $e_n^{\mathrm{det}}(S, F)$ and $e_n^{\mathrm{ran}}(S, F)$ describe the intrinsic difficulty of approximating the solution of problem $\mathcal{P}$ in the deterministic and randomized setting, respectively. In this connection let us mention closely related quantities. The information complexity in the deterministic setting (set = det) and in the randomized setting (set = ran) is defined for $\varepsilon > 0$ by

$$\begin{aligned} &\mathrm{comp}_\varepsilon^{\mathrm{set}}(S, F) \\ &\quad = \min\{n \in \mathbb{N} : \text{there is an } A \in \mathcal{A}_n^{\mathrm{set}}(F, G) \text{ with } e(S, A, F) \le \varepsilon\}, \end{aligned}$$

where we put $\mathrm{comp}_\varepsilon^{\mathrm{set}}(S, F) = +\infty$ if there is no such $n \in \mathbb{N}$. So $\mathrm{comp}_\varepsilon^{\mathrm{set}}(S, F)$ is the minimal number of information functionals needed to reach an error $\le \varepsilon$, and thus, is a way of assessing the complexity of problem $\mathcal{P}$. It is readily checked that $e_n^{\mathrm{set}}(S, F)$ and $\mathrm{comp}_\varepsilon^{\mathrm{set}}(S, F)$ are inverse to each other in the following sense: For all $n \in \mathbb{N}$ and $\varepsilon > 0$, $e_n^{\mathrm{set}}(S, F) \le \varepsilon$ if and only if $\mathrm{comp}_{\varepsilon_1}^{\mathrm{set}}(S, F) \le n$ for all $\varepsilon_1 > \varepsilon$. Hence it suffices to determine one of them. We shall study minimal errors.

Now we describe the Banach space valued initial value problems and specify the abstract notions. Let X be a Banach space over the reals. (We make this assumption since below we consider only real differentiation. The results can also be applied to complex Banach spaces by just regarding them as spaces over the reals.) Throughout the paper $\|\cdot\|$ denotes the norm of X. Other norms are distinguished by subscripts. For $-\infty < a < b < +\infty$, $U \subseteq [a, b] \times X$ open, $\kappa, L > 0$, $r \in \mathbb{N}_0$, $0 \le \varrho \le 1$ we consider the following class

$$\mathscr{C}^{r,\varrho}(U, \kappa, L) \quad \text{of continuous functions} \quad f : U \to X$$

having continuous partial (Fréchet-)derivatives D^α with $\alpha = (\alpha_0, \alpha_1) \in \mathbb{N}_0^2$ of order $|\alpha| = \alpha_0 + \alpha_1 \le r$

$$D^\alpha f(t,x) = \frac{\partial^{|\alpha|} f(t,x)}{\partial t^{\alpha_0} \partial x^{\alpha_1}} \in \mathscr{L}_{\alpha_1}(X)$$

satisfying for $(s,x),(t,y),(t,z) \in U$

$$\begin{aligned} \|D^\alpha f(s,x)\|_{\mathscr{L}_{\alpha_1}(X)} &\le \kappa \qquad (0 \le |\alpha| \le r), &(2)\\ \|D^\alpha f(s,x) - D^\alpha f(t,y)\|_{\mathscr{L}_{\alpha_1}(X)} &\le \kappa(|s-t|^\varrho + \|x-y\|^\varrho) \qquad (|\alpha| = r), &(3)\\ \|f(t,y) - f(t,z)\| &\le L\|y-z\|. &(4) \end{aligned}$$

We consider initial value problems for ODEs with values in X

$$u'(t) = f(t,u(t)) \quad (t \in [a,b]), \qquad u(a) = u_0, \tag{5}$$

with $f \in \mathscr{C}^{r,\varrho}(U,\kappa,L)$ and $u_0 \in X$. A function $u : [a,b] \to X$ is called a solution, if u is continuously differentiable and satisfies (5). For background on Banach space valued differential calculus and ODE we refer to [1]. Let $U_0 \subseteq X$, $V \subseteq U$ be any subsets. For $t \in [a,b]$ denote $V(t) = \{x \in X : (t,x) \in V\}$ and define

$$\begin{aligned} F = \mathscr{F}^{r,\varrho}(U,\kappa,L,U_0,V) = \{(f,u_0) :\ & f \in \mathscr{C}^{r,\varrho}(U,\kappa,L), u_0 \in U_0, \text{ and} \\ & \text{there is a solution } u \text{ of (5) with } u(t) \in V(t)\ (t \in [a,b])\}. \end{aligned} \tag{6}$$

To avoid trivial cases, throughout the paper we assume

$$F = \mathscr{F}^{r,\varrho}(U,\kappa,L,U_0,V) \ne \emptyset. \tag{7}$$

Note that due to (4), the solution u is unique. The solution operator $S : F \to G$ is defined for $(f,u_0) \in F$ by $S(f,u_0) = u$, where u is the solution of the initial value problem (5) and $G = B([a,b],X)$ is the space of all X-valued, bounded on $[a,b]$ functions, equipped with the supremum norm

$$\|g\|_{B([a,b],X)} = \sup_{x \in [a,b]} \|g(t)\|.$$

Observe that, if

$$U_0 + \kappa(t-a)B_X \subseteq V(t) \quad (t \in [a,b]) \tag{8}$$

(which is satisfied, in particular, if $V = [a,b] \times X$), then it follows from (2) that for all $f \in \mathscr{C}^{r,\varrho}(U,\kappa,L)$ and $u_0 \in U_0$ there is a solution $u(t)$ of (5) with $u(t) \in V(t)$ $(t \in [a,b])$, and consequently,

$$\mathscr{F}^{r,\varrho}(U,\kappa,L,U_0,V) = \mathscr{C}^{r,\varrho}(U,\kappa,L) \times U_0. \tag{9}$$

We have chosen this type (6) of F to cover two typical situations: In the first case we demand that for all $f \in \mathscr{C}^{r,\varrho}(U,\kappa,L)$, $u_0 \in U_0$ the solution of (5) exists on $[a,b]$, which usually amounts to restricing the size of $b-a$. This is the local situation (8) treated in Theorem 3.2. In the second case we assume that we have some solution u of (5) (or a set of solutions) and the function f satisfies (2–4) in some neighbourhood of the solution. This is the global approach with no restriction on $b-a$, which is considered in Theorem 3.3.

The class of information functionals Λ is defined as

$$\Lambda = \{\delta_{(t,x)} : \ (t,x) \in U\} \cup \{\delta\},$$

where $\delta_{(t,x)}(f,u_0) = f(t,x)$ and $\delta(f,u_0) = u_0$ for $(f,u_0) \in F$. So here we consider X-valued information functionals, that is $K = X$. This defines our problem $\mathcal{P} = (F,G,S,K,\Lambda)$.

Previous results on the complexity of the initial value problem (5) were all concerned with the case $X = \mathbb{R}^d$ and $U = V = [a,b] \times \mathbb{R}^d$. For the deterministic setting Kacewicz showed in [10] (see also the comments on p. 827 of [11]) that there are constants $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$

$$c_1 n^{-r-\varrho} \le e_n^{\mathrm{det}}(S,F) \le c_2 n^{-r-\varrho}.$$

For the randomized setting it is proved in [9, 2] that there are constants $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$

$$c_1 n^{-r-\varrho-1/2} \le e_n^{\mathrm{ran}}(S,F) \le c_2 n^{-r-\varrho-1/2}$$

(with an aditional arbitrarily small $\varepsilon > 0$ in the exponent of the upper estimate this was already shown in [12]).

It is the goal of this paper to prove appropriate generalizations of these results for the case of Banach space valued ODEs. Moreover, we consider more general set U, V than just $U = V = [a,b] \times X$. In the case of $X = \mathbb{R}^d$ this can be done in a standard fashion by using sufficiently smooth bump functions on $X = \mathbb{R}^d$. For arbitrary Banach spaces this requires a different approach since such functions, in general, do not exist, see [6]. In Section 3 we define the algorithm (which slightly extends that in [2]) and present error estimates. In Section 4 we prove lower bounds.

Constants $c, c_1, c_2, \ldots$ appearing in the paper may depend on the class $\mathscr{F}$ and related parameters like r, ϱ, κ, L, etc., but are independent of the discretization parameters n, k, randomness $\omega \in \Omega$, and the input (f, u_0). Moreover, in Section 3 constants are even independent of X and the related sets U, U_0, V. In all basic statements like theorems etc. this is made clear anyway by the order of quantifiers. Note also that the same symbol may denote different constants, even in a sequence of relations.

3 The algorithm and its analysis

Let $r_1 \in \mathbb{N}_0$, $n \in \mathbb{N}$, put $h = (b-a)/n$ and $t_k = a + kh$ $(k = 0, 1, \ldots, n)$. To define the needed random variables, let $\Omega = [t_0, t_1] \times \cdots \times [t_{n-1}, t_n]$, let Σ be the sigma-algebra of Lebesgue measurable subsets of Ω, and $\mathbb{P}$ the normalized Lebesgue measure on Ω. Define $\xi_k : \Omega \to [t_{k-1}, t_k]$ $(k = 1, \ldots, n)$ by setting $\xi_k(\omega) = s_k$ for $\omega = (s_1, \ldots, s_n) \in \Omega$. Then $(\xi_k)_{k=1}^n$ are independent random variables on $(\Omega, \Sigma, \mathbb{P})$ such that each ξ_k is uniformly distributed on $[t_{k-1}, t_k]$.

Given $f \in \mathscr{C}^{r,\varrho}(U, \kappa, L)$ and $u_0 \in U_0$, we inductively define $(u_k)_{k=1}^n \subset X$ and X-valued polynomials $p_{kj}(t)$ for $k = 0, \ldots, n-1$, $j = 0, \ldots, r_1$ as follows. Suppose u_k is already defined and satisfies

$$u_k \in U(t_k) \tag{10}$$

(note that u_0 is the initial value). Then we set for $t \in [t_k, t_{k+1}]$

$$p_{k0}(t) \quad = \quad u_k + f(t_k, u_k)(t - t_k). \tag{11}$$

Furthermore, suppose $r_1 \geq 1$, $0 \leq j < r_1$, and p_{kj} is already defined. Let $P_{k,j+1}$ be the Lagrange interpolation operator of degree $j+1$ on the equidistant grid $t_{k,j+1,i} = t_k + ih/(j+1)$ $(i = 0, \ldots, j+1)$ on $[t_k, t_{k+1}]$. If

$$p_{kj}(t_{k,j+1,i}) \in U(t_{k,j+1,i}) \quad (i = 0, \ldots, j+1), \tag{12}$$

we put

$$q_{kj} \quad = \quad (f(t_{k,j+1,i}, p_{kj}(t_{k,j+1,i})))_{i=0}^{j+1} \tag{13}$$

and define $p_{k,j+1}$ by setting

$$p_{k,j+1}(t) \quad = \quad u_k + \int_{t_k}^{t} (P_{k,j+1} q_{kj})(s) ds. \tag{14}$$

Finally, if p_{kr_1} is defined and

$$p_{kr_1}(t) \in U(t) \quad (t \in [t_k, t_{k+1}]), \tag{15}$$

we set

$$u_{k+1} = p_{kr_1}(t_{k+1}) + h\left(f(\xi_{k+1}, p_{kr_1}(\xi_{k+1})) - p'_{kr_1}(\xi_{k+1})\right). \tag{16}$$

The latter choice is motivated by the following. First we approximate

$$u(t_{k+1}) = u(t_k) + \int_{t_k}^{t_{k+1}} f(s, u(s)) ds \approx u_k + \int_{t_k}^{t_{k+1}} f(s, p_{kr_1}(s)) ds.$$

Then the integral is approximated by the Monte Carlo method with one sample and with variance reduction by separation of the main part. As such we use the function $p'_{kr_1}(s)$, which is close to $u'(s) = f(s, u(s))$, the latter, in turn, being near to the integrand $f(s, p_{kr_1}(s))$. This gives

$$\begin{aligned} & u_k + \int_{t_k}^{t_{k+1}} f(s, p_{kr_1}(s))ds \\ \approx \quad & u_k + \int_{t_k}^{t_{k+1}} p'_{kr_1}(s)ds + h\big(f(\xi_{k+1}, p_{kr_1}(\xi_{k+1})) - p'_{kr_1}(\xi_{k+1})\big) \\ = \quad & p_{kr_1}(t_{k+1}) + h\big(f(\xi_{k+1}, p_{kr_1}(\xi_{k+1})) - p'_{kr_1}(\xi_{k+1})\big), \end{aligned}$$

explaining (16).

The full approximate solution $v(t)$ on $[a, b]$ is defined as

$$v(t) = \begin{cases} p_{kr_1}(t) & \text{if} \quad t \in [t_k, t_{k+1}) \quad \text{and} \quad 0 \le k \le n-1, \\ u_n & \text{if} \quad t = t_n, \end{cases} \tag{17}$$

so we put

$$A^{r_1}_{n,\omega}(f, u_0) = v \in B([a, b], X) \tag{18}$$

and $A^{r_1}_n = (A^{r_1}_{n,\omega})_{\omega\in\Omega}$. We say that $A^{r_1}_n(f, u_0)$ is defined (or, more precisely, defined on U) if for all $\omega \in \Omega$ this definition goes through till (18), that is, (10), (12), (15) are satisfied at all stages. If for some ω, at some stage k, any of the conditions (10), (12), (15) is violated (in particular, if $u_0 \notin U(a)$), we leave $A^{r_1}_n(f, u_0)$ undefined.

Note that, as far as the definition of $A^{r_1}_{n,\omega}$ is concerned, fixing any $\omega \in \Omega$ is the same as fixing any values of $\xi_k \in [t_{k-1}, t_k]$ ($k = 1, \dots, n$). This way the algorithm becomes deterministic.

First we show that $A^{r_1}_n$ is indeed a randomized algorithm in the sense of the general notion introduced in Section 2.

Lemma 3.1. *Let $(f, u_0) \in F$. If $A^{r_1}_n(f, u_0)$ is defined, then $A^{r_1}_n(f, u_0)$ is a $B([a, b], X)$-valued random variable.*

Proof. For the purposes of this proof we include the dependence on $\xi_1, \dots, \xi_k$ into the notation and write $u_k(\xi_1, \dots, \xi_k)$, $p_{kj}(\xi_1, \dots, \xi_k, t)$, etc. We shall show that $u_k(s_1, \dots, s_k)$ $(0 \le k \le n)$ and $p_{kj}(s_1, \dots, s_k, t)$ $(0 \le k \le n-1,\ t \in [t_k, t_{k+1}])$ depend continuously on $(s_1, \dots, s_k) \in [t_0, t_1] \times \dots \times [t_{k-1}, t_k]$.

First note that $A^{r_1}_n(f, u_0)$ being defined means, in particular, that these functions are defined for all possible values of the s_i. To prove continuity, we argue by induction. For $k = 0$ the statement is trivial, since there is no dependence. Now assume the statement holds for some k with $0 \le k < n$. Then by (16), with $\bar{s} = (s_1, \dots, s_k)$ and $s_{k+1} \in [t_k, t_{k+1}]$

$$u_{k+1}(\bar{s}, s_{k+1}) \quad = \quad p_{kr_1}(\bar{s}, t_{k+1}) + h\left(f(s_{k+1}, p_{kr_1}(\bar{s}, s_{k+1})) - p'_{kr_1}(\bar{s}, s_{k+1})\right).$$

Since $p_{kr_1}(\bar{s},t)$ is a polynomial in t, continuity with respect to $\bar{s}$ for each $t \in [t_k, t_{k+1}]$ implies continuity with respect to $(\bar{s},t)$. Therefore u_{k+1} depends continuously on $(\bar{s}, s_{k+1})$. Now assume $k < n-1$. To prove the statement about $p_{k+1,j}$ we argue by induction over j. For $j=0$ we have by (11), with $\tilde{s} = (s_1, \dots, s_k, s_{k+1})$

$$p_{k+1,0}(\tilde{s},t) = u_{k+1}(\tilde{s}) + f(t_{k+1}, u_{k+1}(\tilde{s}))(t - t_{k+1}),$$

which shows the continuous dependence. Now assume that $p_{k+1,j}(\tilde{s},t)$ depends continuously on $\tilde{s}$ and hence, also on $(\tilde{s},t)$. Then by (13–14)

$$q_{k+1,j}(\tilde{s}) \quad = \quad (f(t_{k+1,j+1,i}, p_{k+1,j}(\tilde{s}, t_{k+1,j+1,i})))_{i=0}^{j+1}$$

and therefore also

$$p_{k+1,j+1}(\tilde{s},t) \quad = \quad u_{k+1}(\tilde{s}) + \int_{t_{k+1}}^{t} \left(P_{k+1,j+1} q_{k+1,j}(\tilde{s})\right)(\tau) d\tau$$

have the required continuous dependencies. This completes the induction over j and also that over k. From (17) we infer

$$z(s_1, \dots, s_n, t) = \begin{cases} p_{kr_1}(s_1, \dots, s_k, t) & \text{if} \quad t \in [t_k, t_{k+1}) \text{ and } 0 \le k \le n-1, \\ u_n(s_1, \dots, s_n) & \text{if} \quad t = t_n, \end{cases}$$

hence $z(s_1, \dots, s_n) \in B([a,b], X)$ depends continuously on $s_1, \dots, s_n$. This implies that the mapping $\omega \to A^{r_1}_{n,\omega}(f, u_0)$ is Σ-to-Borel measurable and essentially separably valued.

□

In the rest of this section all constants $c, c_1, \dots$ are even independent of X, U_0, V, and U, which is also made clear by the order of quantifiers in the statements. Moreover, let us introduce the following constants depending only on $r_1 \in \mathbb{N}_0$

$$\begin{aligned} c_0(0) &= 1, \quad c_0(r_1) = \max_{1 \le j \le r_1} \|P_j\|_{\mathscr{L}(C([0,1],\mathbb{R}))} \ge 1 \quad (r_1 \ge 1), \\ c(r_1) &= 2c_0(r_1) + 1, \end{aligned}$$

where P_j is the operator of Lagrange interpolation of degree j on $[0,1]$ and $C([0,1], X)$ denotes the space of continuous, X-valued functions, endowed with the supremum norm. Observe that if we consider P_j (without change of notation) as interpolation operator of X-valued functions in $C([0,1], X)$, then the constant remains the same for an arbitrary Banach space X, i.e., we have

$$\max_{1 \le j \le r_1} \|P_j\|_{\mathscr{L}(C([0,1],X))} = c_0(r_1). \tag{19}$$

Indeed, the upper bound follows by taking linear functionals, the lower bound by the fact that, trivially, X contains an isometric copy of $\mathbb{R}$.

Now we estimate the error of the algorithm.

Theorem 3.2. *Given $r, r_1, \varrho, a, b, \kappa, L$ as above and $1 \le p \le 2$, there are constants $c_1, c_2 > 0$ such that the following holds. Let X be a Banach space, $\emptyset \ne U_0 \subseteq X$, $V \subseteq U \subseteq [a,b] \times X$, U open, satisfying (8) and*

$$U_0 + c(r_1)\kappa(t-a)B_X \subseteq U(t) \quad (t \in [a,b]) \tag{20}$$

(which holds, in particular, if $U = [a,b] \times X$). Then $F \ne \emptyset$ and for all $n \in \mathbb{N}$ and $(f, u_0) \in F$, $A_n^{r_1}(f, u_0)$ is defined on U and satisfies

$$\sup_{(f,u_0)\in F} \|S(f,u_0) - A_{n,\omega}^{r_1}(f,u_0)\|_{B([a,b],X)} \le c_1 n^{-\min(r+\varrho, r_1+1)} \quad (\omega \in \Omega) \tag{21}$$

and

$$\sup_{(f,u_0)\in F} \left(\mathbb{E}\, \|S(f,u_0) - A_{n,\omega}^{r_1}(f,u_0)\|_{B([a,b],X)}^p\right)^{1/p} \le c_2 \tau_p(X) n^{-\min(r+\varrho, r_1+1)-1+1/p}. \tag{22}$$

In the previous statement we imposed a certain relationship on κ, a, b, U_0, U and V. Now we drop this assumption while assuming something on V and U as well as a minimal smoothness ($r + \varrho > 0$).

Theorem 3.3. *Let $r, r_1, \varrho, a, b, \kappa, L$ be as above, $1 \le p \le 2$, $\delta_0 > 0$ and assume that $r + \varrho > 0$. Then there are constants $c_1, c_2 > 0$, $n_0 \in \mathbb{N}$, such that the following holds. Let X be a Banach space, $U_0 \subseteq X$, $V \subseteq U \subseteq [a,b] \times X$, U open, such that $F \ne \emptyset$ and*

$$V(t) + \delta_0 B_X \subseteq U(t) \quad (t \in [a,b]). \tag{23}$$

Then for all $n \in \mathbb{N}$ with $n \ge n_0$ and $(f, u_0) \in F$, $A_n^{r_1}(f, u_0)$ is defined on U and the estimates (21) and (22) hold.

We prove these theorems in the following way. With Proposition 3.4 below we show a respective statement under a stronger assumption. This is the key part of the proof, which is different from the analysis in [9] and [2]. The latter would require a martingale type property instead of the type assumption. Then the proofs of Theorems 3.2 and 3.3 will be a reduction to this proposition.

For $n \in \mathbb{N}$ we define

$$\begin{aligned} &\mathscr{H}^{r,\varrho}(U, \kappa, L, U_0, V, r_1, n) \\ &= \{(f, u_0) \in \mathscr{F}^{r,\varrho}(U, \kappa, L, U_0, V) : \text{ If } u = S(f, u_0), \text{ then} \\ &\quad u(t_k) + c_0(r_1)\kappa(t - t_k)B_X \subseteq U(t) \ \ (t \in [t_k, t_{k+1}], 0 \le k \le n-1) \\ &\quad \text{and } A_n^{r_1}(f, u_0) \text{ is defined on } U\}. \end{aligned} \tag{24}$$

Let us note that, if $U = [a,b] \times X$, then for all $r_1 \in \mathbb{N}_0$ and $n \in \mathbb{N}$ the sets defined in (6) and (24) coincide:

$$\mathscr{H}^{r,\varrho}(U, \kappa, L, U_0, V, r_1, n) = \mathscr{F}^{r,\varrho}(U, \kappa, L, U_0, V). \tag{25}$$

Proposition 3.4. *Given $r, r_1, \varrho, a, b, \kappa, L$ as above and $1 \le p \le 2$, there are constants $c_1, c_2 > 0$ such that the following holds. Let X be a Banach space, $U_0 \subseteq X$, $V \subseteq U \subseteq [a,b] \times X$, U open, $n \in \mathbb{N}$ be such that $\mathscr{H}^{r,\varrho}(U, \kappa, L, U_0, V, r_1, n) \neq \emptyset$. Then for all $(f, u_0) \in \mathscr{H}^{r,\varrho}(U, \kappa, L, U_0, V, r_1, n)$*

$$\|S(f,u_0) - A^{r_1}_{n,\omega}(f,u_0)\|_{B([a,b],X)} \quad \le \quad c_1 n^{-\min(r+\varrho, r_1+1)} \qquad (\omega \in \Omega) \tag{26}$$

and

$$\begin{aligned} &\Big(\mathbb{E}\,\|S(f,u_0) - A^{r_1}_{n,\omega}(f,u_0)\|^p_{B([a,b],X)}\Big)^{1/p} \\ &\le \quad c_2 \tau_p(X) n^{-\min(r+\varrho, r_1+1)-1+1/p}. \end{aligned} \tag{27}$$

Proof. Differentiating equation (5) and using the assumptions on f it follows that $u = S(f, u_0)$ is $(r+1)$-times continuously differentiable and there are constants $c_1, c_2 > 0$ such that

$$\|u^{(j)}(t)\| \le c_1 \qquad (t \in [t_k, t_{k+1}],\ 1 \le j \le r+1)\,, \tag{28}$$

$$\|u^{(r+1)}(s) - u^{(r+1)}(t)\| \le c_2 |t-t|^{\varrho} \qquad (s, t \in [t_k, t_{k+1}])\,. \tag{29}$$

Indeed, we show by induction over j that for $0 \le j \le r$ there exist constants κ_j, L_j such that for each $f \in \mathscr{C}^{r,\varrho}(U, \kappa, L)$ there is an

$$f_j \in \mathscr{C}^{r-j,\varrho}(U, \kappa_j, L_j) \tag{30}$$

such that

$$u^{(j+1)}(t) = f_j(t, u(t)) \quad (t \in [a,b]). \tag{31}$$

For $j = 0$ we just set $f_0 = f$, $\kappa_0 = \kappa$, $L_0 = L$. Now assume (30–31) hold for some j with $0 \le j < r$. Differentiating (31), we obtain for $t \in [a,b]$

$$\begin{aligned} u^{(j+2)}(t) &= \frac{\partial f_j}{\partial t}(t, u(t)) + \frac{\partial f_j}{\partial x}(t, u(t)) u'(t) \\ &= \frac{\partial f_j}{\partial t}(t, u(t)) + \frac{\partial f_j}{\partial x}(t, u(t)) f(t, u(t)). \end{aligned}$$

Setting for $(t, x) \in U$

$$f_{j+1}(t,x) = \frac{\partial f_j}{\partial t}(t,x) + \frac{\partial f_j}{\partial x}(t,x) f(t,x),$$

it follows that (31) holds for $j+1$. Moreover, from (30) we conclude that $f_j \in \mathscr{C}^{r-j-1,\varrho}(U, \kappa_{j+1}, L_{j+1})$ for some constants $\kappa_{j+1}, L_{j+1} > 0$. This completes the induction and proves (30–31). Relation (28) follows directly from (30–31), while (29) is a consequence of (3), (30–31), and the Lipschitz continuity of u implied by (28).

Observe that for $j \geq 1$

$$\|P_{kj}\|_{\mathscr{L}(C([t_k,t_{k+1}],X))} = \|P_j\|_{\mathscr{L}(C([0,1],X))} \leq c_0(r_1). \tag{32}$$

For $k = 0, \dots, n-1$, $j = 0, \dots, r_1$ let $\tilde{p}_{kj}$ and $\tilde{q}_{kj}$ be defined by (10–14) with u_k replaced by $u(t_k)$. We show by induction over j that there are constants $c_{1-4,j} > 0$ such that for all j with $0 \leq j \leq r_1$ the following hold:

$$\tilde{p}_{kj} \text{ is defined and } \|\tilde{p}_{kj}(t) - u(t_k)\| \leq c_0(r_1)\kappa(t - t_k) \quad (t \in [t_{k-1}, t_k]), \tag{33}$$

$$\sup_{t\in[t_k,t_{k+1}]} \|u(t) - \tilde{p}_{kj}(t)\| \leq c_{1,j} h^{\min(r+\varrho,j+1)+1}, \tag{34}$$

$$\sup_{t\in[t_k,t_{k+1}]} \|u'(t) - \tilde{p}'_{kj}(t)\| \leq c_{2,j} h^{\min(r+\varrho,j+1)}, \tag{35}$$

$$\sup_{t\in[t_k,t_{k+1}]} \|\tilde{p}_{kj}(t) - p_{kj}(t)\| \leq c_{3,j}\|u(t_k) - u_k\|, \tag{36}$$

$$\sup_{t\in[t_k,t_{k+1}]} \|\tilde{p}'_{kj}(t) - p'_{kj}(t)\| \leq c_{4,j}\|u(t_k) - u_k\|. \tag{37}$$

Note that, since $u(t_k) = \tilde{p}_{kj}(t_k)$, (34) follows with $c_{1,j} = c_{2,j}$ from (35) by integration. Similarly, because of $u_k = p_{kj}(t_k)$, (36) is a consequence of (37), with $c_{3,j} \leq 1 + hc_{4,j} \leq 1 + (b-a)c_{4,j}$.

Let $j = 0$. By (6), $u(t_k) \in U(t_k)$, so $\tilde{p}_{k0}$ is defined. Moreover, for $t \in [t_k, t_{k+1}]$

$$\|\tilde{p}_{k0}(t) - u(t_k)\| = \|f(t_k, u(t_k))(t - t_k)\| \leq \kappa(t - t_k),$$

which gives (33). Furthermore,

$$\|u'(t) - \tilde{p}'_{k0}(t)\| = \|f(t, u(t)) - f(t_k, u(t_k))\|. \tag{38}$$

If $r = 0$, it follows from (2–3) and (5) that

$$\|f(t, u(t)) - f(t_k, u(t_k))\| \leq \kappa(|t - t_k|^\varrho + \|u(t) - u(t_k)\|^\varrho) \leq \kappa(1 + \kappa^\varrho)h^\varrho.$$

This together with (38) gives (35). If $r \geq 1$, we have

$$\begin{aligned} &\|f(t, u(t)) - f(t_k, u(t_k))\| \\ &= \left\| \int_{t_k}^t \left(\frac{\partial f}{\partial t}(s, u(s)) + \frac{\partial f}{\partial x}(s, u(s)) f(s, u(s)) \right) ds \right\| \leq (\kappa + \kappa^2)h, \end{aligned}$$

which, combined with (38) gives (35) also for this case. Finally, we have by (4)

$$\|\tilde{p}'_{k0}(t) - p'_{k0}(t)\| = \|f(t_k, u(t_k)) - f(t_k, u_k)\| \leq L\|u(t_k) - u_k\|,$$

showing (37) and completing the proof of (33–37) for $j = 0$.

Now we assume that $r_1 \geq 1$ and (33–37) hold for some j with $0 \leq j < r_1$. It follows from (33) and (24) that $\tilde{p}_{kj}(t_{k,j+1,i}) \in U(t_{k,j+1,i})$, so $\tilde{p}_{k,j+1}$ is defined. Furthermore, using (32), we get for $t \in [t_k, t_{k+1}]$

$$\|\tilde{p}_{k,j+1}(t) - u(t_k)\| = \left\| \int_{t_k}^{t} \left(P_{k,j+1}\tilde{q}_{kj}\right)(s)ds \right\| \leq c_0(r_1)\kappa(t - t_k),$$

which shows (33) for $j+1$. We have

$$\begin{aligned} &\|u'(t) - \tilde{p}'_{k,j+1}(t)\| \\ &\quad = \|u'(t) - (P_{k,j+1}\tilde{q}_{kj})(t)\| \\ &\quad \leq \|u'(t) - (P_{k,j+1}u')(t)\| + \|(P_{k,j+1}u')(t) - (P_{k,j+1}\tilde{q}_{kj})(t)\|. \end{aligned} \tag{39}$$

Setting $l = \min(r, j+1)$, (28) and the Taylor series with integral remainder term give

$$u'(t) = \sum_{i=0}^{l} \frac{u^{(i+1)}(t_k)}{i!}(t - t_k)^i + R(t) \quad (t \in [t_k, t_{k+1}]) \tag{40}$$

with

$$R(t) = \begin{cases} \dfrac{1}{(l-1)!}\displaystyle\int_{t_k}^{t}(t-s)^{l-1}\left(u^{(l+1)}(s) - u^{(l+1)}(t_k)\right)ds & \text{if} \quad l \geq 1 \\ u'(t) - u'(t_k) & \text{if} \quad l = 0. \end{cases} \tag{41}$$

Since $P_{k,j+1}$ is exact on (X-valued) polynomials of degree $\leq j+1$, (32) and (40–41) give

$$\begin{aligned} &\sup_{t \in [t_k, t_{k+1}]} \|u'(t) - (P_{k,j+1}u')(t)\| \\ &\quad \leq \frac{c_0(r_1)+1}{l!} h^l \sup_{t \in [t_k, t_{k+1}]} \|u^{(l+1)}(t) - u^{(l+1)}(t_k)\|. \end{aligned} \tag{42}$$

For $j+1 \geq r$ we have $l = r$ and by (29),

$$\sup_{t \in [t_k, t_{k+1}]} \|u^{(l+1)}(t) - u^{(l+1)}(t_k)\| \leq ch^{\varrho}. \tag{43}$$

If $j+1 < r$, meaning that $l = j+1$ and $l+2 \leq r+1$, we have by (28)

$$\sup_{t \in [t_k, t_{k+1}]} \|u^{(l+1)}(t) - u^{(l+1)}(t_k)\| \leq \int_{t_k}^{t_{k+1}} \|u^{(l+2)}(s)\|ds \leq ch. \tag{44}$$

Using (32) and the induction assumption (34), we obtain

$$\begin{aligned}
&\sup_{t\in[t_k,t_{k+1}]} \|(P_{k,j+1}u')(t) - (P_{k,j+1}\tilde{q}_{kj})(t)\| \\
&\quad\le\; c_0(r_1) \max_{0\le i\le j+1} \|u'(t_{k,j+1,i}) - f(t_{k,j+1,i}, \tilde{p}_{kj}(t_{k,j+1,i}))\| \\
&\quad=\; c_0(r_1) \max_{0\le i\le j+1} \|f(t_{k,j+1,i}, u(t_{k,j+1,i})) - f(t_{k,j+1,i}, \tilde{p}_{kj}(t_{k,j+1,i}))\| \\
&\quad\le\; c_0(r_1) L \sup_{t\in[t_k,t_{k+1}]} \|u(t) - \tilde{p}_{kj}(t)\| \le c_0(r_1) L c_{1,j} h^{\min(r+\varrho+1,j+2)}. \qquad (45)
\end{aligned}$$

Combining (39) and (42–45) proves (35) for $j+1$. Finally, we have for $t \in [t_k, t_{k+1}]$

$$\begin{aligned}
&\|\tilde{p}'_{k,j+1}(t) - p'_{k,j+1}(t)\| \\
&\quad=\; \|(P_{k,j+1}\tilde{q}_{kj})(t) - (P_{k,j+1}q_{kj})(t)\| \\
&\quad\le\; c_0(r_1) \max_{0\le i\le j+1} \|f(t_{k,j+1,i}, \tilde{p}_{kj}(t_{k,j+1,i})) - f(t_{k,j+1,i}, p_{kj}(t_{k,j+1,i}))\| \\
&\quad\le\; c_0(r_1) L \max_{0\le i\le j+1} \|\tilde{p}_{kj}(t_{k,j+1,i})) - p_{kj}(t_{k,j+1,i})\| \\
&\quad\le\; c_0(r_1) L c_{3,j} \|u(t_k) - u_k\|,
\end{aligned}$$

where in the last step we used the induction assumption (36). This shows (37) for $j+1$, completes the induction step and thus the proof of (33–37).

Note that (24) and (33) ensure $\tilde{p}_{kr_1}(t) \in U(t)$ ($t \in [t_k, t_{k+1}]$). From (4), (34), and (35) we get

$$\begin{aligned}
&\sup_{t\in[t_k,t_{k+1}]} \|f(t, \tilde{p}_{kr_1}(t)) - \tilde{p}'_{kr_1}(t)\| \\
&\quad=\; \sup_{t\in[t_k,t_{k+1}]} \big(\|f(t, \tilde{p}_{kr_1}(t)) - f(t, u(t))\| + \|u'(t) - \tilde{p}'_{kr_1}(t)\|\big) \\
&\quad\le\; c h^{\min(r+\varrho, r_1+1)}. \qquad (46)
\end{aligned}$$

We define for $0 \le k \le n-1$

$$w_{k+1} = \tilde{p}_{kr_1}(t_{k+1}) + h(f(\xi_{k+1}, \tilde{p}_{kr_1}(\xi_{k+1})) - \tilde{p}'_{kr_1}(\xi_{k+1})). \qquad (47)$$

Let

$$e_k = u(t_k) - u_k$$

denote the error at point t_k. Then we have $e_0 = 0$ and for $0 \le k \le n-1$

$$e_{k+1} - e_k = u(t_{k+1}) - u(t_k) - (u_{k+1} - u_k) = d_{k+1} + g_{k+1} + \eta_{k+1}, \qquad (48)$$

where

$$\begin{aligned}
d_{k+1} &= \int_{t_k}^{t_{k+1}} (f(t, u(t)) - f(t, \tilde{p}_{kr_1}(t)))\, dt, &(49)\\
g_{k+1} &= w_{k+1} - u(t_k) - (u_{k+1} - u_k), &(50)\\
\eta_{k+1} &= \int_{t_k}^{t_{k+1}} f(t, \tilde{p}_{kr_1}(t)) dt - (w_{k+1} - u(t_k)). &(51)
\end{aligned}$$

We have by (16)

$$u_{k+1} - u_k = \int_{t_k}^{t_{k+1}} p'_{kr_1}(t)dt + h(f(\xi_{k+1}, p_{kr_1}(\xi_{k+1})) - p'_{kr_1}(\xi_{k+1})), \quad (52)$$

and, similarly, from (47)

$$w_{k+1} - u(t_k) = \int_{t_k}^{t_{k+1}} \tilde{p}'_{kr_1}(t)dt + h(f(\xi_{k+1}, \tilde{p}_{kr_1}(\xi_{k+1})) - \tilde{p}'_{kr_1}(\xi_{k+1})). \quad (53)$$

Since $\tilde{p}_{kr_1}$ does not depend on ω, taking the expectation in (53) gives

$$\mathbb{E}w_{k+1} - u(t_k) = \int_{t_k}^{t_{k+1}} f(t, \tilde{p}_{kr_1}(t))dt. \quad (54)$$

By (51), (53), and (54),

$$\begin{aligned} \eta_{k+1} &= \int_{t_k}^{t_{k+1}} \left(f(t, \tilde{p}_{kr_1}(t)) - \tilde{p}'_{kr_1}(t)\right) dt \\ &\quad -h(f(\xi_{k+1}, \tilde{p}_{kr_1}(\xi_{k+1})) - \tilde{p}'_{kr_1}(\xi_{k+1})) \quad (55) \\ &= \mathbb{E}w_{k+1} - w_{k+1}. \quad (56) \end{aligned}$$

Relations (46) and (55) imply

$$\|\eta_{k+1}\| \leq ch^{\min(r+\varrho, r_1+1)+1}. \quad (57)$$

Moreover, from (4), (34), and (49) we get

$$\|d_{k+1}\| \leq ch^{\min(r+\varrho, r_1+1)+2}. \quad (58)$$

Subtracting (52) from (53) and using (4), (36), and (37) gives

$$\|g_{k+1}\| \leq ch\|e_k\|. \quad (59)$$

Denote

$$\theta = \sum_{j=1}^{n} \|d_j\| + \max_{1\leq j\leq n} \Big\| \sum_{i=1}^{j} \eta_i \Big\|. \quad (60)$$

Since $e_0 = 0$, we obtain from (48) and (59) for $1 \leq k \leq n$

$$\|e_k\| = \left\| \sum_{j=1}^{k} (g_j + d_j + \eta_j) \right\| \leq \sum_{j=1}^{k} \|g_j\| + \theta \leq ch \sum_{j=1}^{k-1} \|e_j\| + \theta. \quad (61)$$

Using (61) recursively, we get

$$\begin{aligned}\sum_{j=1}^{k-1}\|e_j\| &\le \sum_{j=1}^{k-2}\|e_j\| + ch\sum_{j=1}^{k-2}\|e_j\| + \theta = (1+ch)\sum_{j=1}^{k-2}\|e_j\| + \theta \\ &\le (1+ch)^2\sum_{j=1}^{k-3}\|e_j\| + \theta + (1+ch)\theta \le \cdots \le \frac{(1+ch)^{k-1}-1}{ch}\theta.\end{aligned}$$

Inserting this into (61) yields

$$\|e_k\| \le (1+ch)^{k-1}\theta \le e^{c(k-1)h}\theta \le e^{c(b-a)}\theta,$$

and, with (58) and (60),

$$\max_{0\le k\le n}\|e_k\| \quad \le \quad ch^{\min(r+\varrho,r_1+1)+1} + c\max_{1\le k\le n}\Big\|\sum_{i=1}^{k}\eta_i\Big\|. \tag{62}$$

With v given by (17) we have

$$\begin{aligned}&\sup_{t\in[a,b]}\|u(t)-v(t)\| \\ &\quad = \max\Big(\max_{0\le k\le n-1}\sup_{t\in[t_k,t_{k+1}]}\|u(t)-p_{kr_1}(t)\|, \|u(t_n)-u_n\|\Big).\end{aligned} \tag{63}$$

Moreover, using (34) and (36),

$$\begin{aligned}\sup_{t\in[t_k,t_{k+1}]}\|u(t)-p_{kr_1}(t)\| &\le \sup_{t\in[t_k,t_{k+1}]}(\|u(t)-\tilde{p}_{kr_1}(t)\| + \|\tilde{p}_{kr_1}(t)-p_{kr_1}(t)\|) \\ &\le ch^{\min(r+\varrho,r_1+1)+1} + c\|e_k\|,\end{aligned}$$

which together with (62) and (63) gives

$$\|S(f,u_0)-A_{n,\omega}^{r_1}(f,u_0)\|_{B([a,b],X)} \le ch^{\min(r+\varrho,r_1+1)+1} + c\max_{1\le k\le n}\Big\|\sum_{i=1}^{k}\eta_i\Big\|. \tag{64}$$

Now (26) is a consequence of (57) and (64). The case $p=1$ of (27), in turn, follows from (26). Now we assume $1<p\le 2$. From (64) we get

$$\begin{aligned}&\mathbb{E}\,\|S(f,u_0)-A_{n,\omega}^{r_1}(f,u_0)\|_{B([a,b],X)}^p \\ &\quad \le ch^{p\min(r+\varrho,r_1+1)+p} + c\,\mathbb{E}\max_{1\le k\le n}\Big\|\sum_{i=1}^{k}\eta_i\Big\|^p.\end{aligned} \tag{65}$$

By (55–56), $(\eta_k)_{k=1}^n$ is a sequence of independent X-valued random variables of mean zero. Consequently, $\left(\left\|\sum_{i=1}^k \eta_i\right\|\right)_{k=1}^n$ is a non-negative submartingale. From Doob's inequality ([5], Ch. VII, Th. 3.4) we obtain

$$\mathbb{E} \max_{1\le k\le n} \Big\| \sum_{i=1}^k \eta_i \Big\|^p \le \frac{p^p}{(p-1)^p} \mathbb{E} \Big\| \sum_{i=1}^n \eta_i \Big\|^p . \tag{66}$$

On the other hand, from Lemma 2.1 we get

$$\mathbb{E} \Big\| \sum_{i=1}^n \eta_i \Big\|^p \le 2^p \tau_p(X)^p \sum_{i=1}^n \mathbb{E}\|\eta_i\|^p . \tag{67}$$

Combining (57), (66), and (67) gives

$$\mathbb{E} \max_{1\le k\le n} \Big\| \sum_{i=1}^k \eta_i \Big\|^p \le c\tau_p(X)^p h^{p\min(r+\varrho, r_1+1)+p-1}.$$

Inserting this into (65) gives (27) for $1 < p \le 2$. □

Proof of Theorem 3.2. By assumption, (8) and hence (9) is valid, so $F \ne \emptyset$. Let $(f, u_0) \in F$. Using (10–16) and (20), we show by induction that for $0 \le k \le n$ the following hold:

$$u_k \text{ is defined and } \|u_k - u_0\| \le c(r_1)\kappa k h, \tag{68}$$

and, if $k \le n-1$, then for all j with $0 \le j \le r_1$

$$p_{kj} \text{ is defined and } \|p_{kj}(t) - u_k\| \le c_0(r_1)\kappa(t - t_k) \quad (t \in [t_k, t_{k+1}]). \tag{69}$$

First we show that for $0 \le k \le n-1$ (68) implies (69). So suppose (68) holds for some $0 \le k \le n-1$. To derive (69), we argue by induction over j. Let $j = 0$. By (20) and (68), $u_k \in U(t_k)$, so p_{k0} is defined and we have

$$\|p_{k0}(t) - u_k\| = \|f(t_k, u_k)(t - t_k)\| \le \kappa(t - t_k) \quad (t \in [t_k, t_{k+1}]).$$

Now we assume that (69) holds for some j with $0 \le j < r_1$. It follows that for $t \in [t_k, t_{k+1}]$

$$\|p_{kj}(t) - u_0\| \le \|p_{kj}(t) - u_k\| + \|u_k - u_0\| \le c(r_1)\kappa(t - a),$$

hence $p_{kj}(t_{k,j+1,i}) \in U(t_{k,j+1,i})$ for $i = 0, \ldots, j+1$, so $p_{k,j+1}$ is defined. Furthermore, using (32), we get

$$\|p_{k,j+1}(t) - u_k\| = \left\| \int_{t_k}^t (P_{k,j+1} q_{kj})(s) ds \right\| \le c_0(r_1)\kappa(t - t_k),$$

showing (69) for $j+1$, completing the induction over j and thus, the proof that (68) implies (69).

Next we prove (68) by induction over k. For $k = 0$, (68) holds by definition. Now suppose (68) and hence (69) hold for some k with $0 \le k \le n-1$. It follows that for $t \in [t_k, t_{k+1}]$

$$\|p_{kr_1}(t) - u_0\| \le \|p_{kr_1}(t) - u_k\| + \|u_k - u_0\| \le c(r_1)\kappa(t-a),$$

and therefore, $p_{kr_1}(t) \in U(t)$. This shows that u_{k+1} is defined. Note that

$$\|p'_{kr_1}(t)\| \le c_0(r_1)\kappa \quad (t \in [t_k, t_{k+1}]),$$

which is a consequence of (11) if $r_1 = 0$ and of (14) if $r_1 \ge 1$. We have

$$\begin{aligned} \|u_{k+1} - u_k\| &\le \|p_{kr_1}(t_{k+1}) - u_k\| + h\left\|f(\xi_{k+1}, p_{kr_1}(\xi_{k+1})) - p'_{kr_1}(\xi_{k+1})\right\| \\ &\le c_0(r_1)\kappa h + (c_0(r_1)+1)\kappa h = c(r_1)\kappa h, \end{aligned}$$

hence

$$\|u_{k+1} - u_0\| \le c(r_1)\kappa(k+1)h.$$

This shows (68) for $k+1$, completes the induction over k and the proof of (68–69).

It follows that $A_n^{r_1}(f, u_0)$ is defined. Moreover, with $u = S(f, u_0)$ we get from (5)

$$u(t_k) \in U_0 + \kappa(t_k - a)B_X,$$

and therefore, using (20), for $t \in [t_k, t_{k+1}]$, $0 \le k \le n-1$

$$u(t_k) + c_0(r_1)\kappa(t - t_k)B_X \subseteq U_0 + c_0(r_1)\kappa(t-a)B_X \subseteq U(t).$$

This shows that $(f, u_0) \in \mathscr{H}^{r,\varrho}(U, \kappa, L, U_0, V, r_1, n)$, therefore

$$\mathscr{F}^{r,\varrho}(U, \kappa, L, U_0, V) = \mathscr{H}^{r,\varrho}(U, \kappa, L, U_0, V, r_1, n),$$

and the result follows from Proposition 3.4. □

Proof of Theorem 3.3. We set

$$W = \{(t, u(t)) : t \in [a,b],\ u = S(f, u_0) \text{ for some } (f, u_0) \in F\}. \tag{70}$$

Then $W \subseteq V$ and $W(t) \ne \emptyset$ $(t \in [a,b])$. Define the function $\psi : [0, +\infty) \to [0,1]$ by

$$\psi(t) = \begin{cases} 1 & \text{if} \quad 0 \le t \le 2\delta_0/3 \\ 3(1 - t/\delta_0) & \text{if} \quad 2\delta_0/3 < t < \delta_0 \\ 0 & \text{if} \quad \delta_0 \le t. \end{cases}$$

Given $(f, u_0) \in \mathscr{F}^{r,\varrho}(U, \kappa, L, U_0, V)$ we define $\tilde{f} : [a, b] \times X \to X$ by

$$\tilde{f}(t, x) = \begin{cases} \psi(d(x, W(t)))f(t, x) & \text{if} \quad x \in U(t) \\ 0 & \text{otherwise,} \end{cases}$$

where $d(x, W(t)) = \inf\{\|x - y\| : y \in W(t)\}$ is the distance of x to $W(t)$. Consequently,

$$\tilde{f}(t, x) = f(t, x) \qquad \left(t \in [a, b],\, x \in X,\, d(x, W(t)) \le \frac{2\delta_0}{3} \right). \tag{71}$$

Put

$$\tilde{\varrho} = \begin{cases} 1 & \text{if} \quad r \ge 1 \\ \varrho & \text{if} \quad r = 0, \end{cases}$$

thus, $0 < \tilde{\varrho} \le 1$. We show that

$$\tilde{f} \in \mathscr{C}^{0,\tilde{\varrho}}([a, b] \times X, \kappa_1, L_1) \tag{72}$$

for some $\kappa_1, L_1 > 0$ depending only on κ, L and δ_0.

By assumption, $d(x, W(t)) < \delta_0$ implies $x \in U(t)$, therefore

$$\sup_{t \in [a,b], x \in X} \|\tilde{f}(t, x)\| \le \sup_{(t,x) \in U} \|f(t, x)\| \le \kappa. \tag{73}$$

Now fix $s, t \in [a, b]$, $x, y \in X$. Let $v = S(g, v_0)$ for some $(g, v_0) \in \mathscr{F}^{r,\varrho}(U, \kappa, L, U_0, V)$. Then

$$\begin{aligned} d(x, W(s)) &\le d(x, v(s)) \le d(x, v(t)) + \|v(s) - v(t)\| \\ &\le d(x, v(t)) + \kappa|s - t|. \end{aligned}$$

Taking the infimum over v and using (70), we get

$$d(x, W(s)) \le d(x, W(t)) + \kappa|s - t|,$$

and, exchanging s and t and combining both estimates, we arrive at

$$|d(x, W(s)) - d(x, W(t))| \le \kappa|s - t|. \tag{74}$$

Using this, we derive

$$\begin{aligned} &|\psi(d(x, W(s))) - \psi(d(y, W(t)))| \\ &\quad \le |\psi(d(x, W(s))) - \psi(d(x, W(t)))| + |\psi(d(x, W(t))) - \psi(d(y, W(t)))| \\ &\quad \le 3\delta_0^{-1}\big(|d(x, W(s)) - d(x, W(t))| + |d(x, W(t)) - d(y, W(t))|\big) \\ &\quad \le 3\delta_0^{-1}(\kappa|s - t| + \|x - y\|). \end{aligned} \tag{75}$$

Now we verify that $\tilde{f}$ satisfies the $\tilde{\varrho}$-Hölder condition. We can assume $d(x, W(s)) \leq d(y, W(t))$. If $d(x, W(s)) \geq \delta_0$, we have

$$\tilde{f}(s, x) = \tilde{f}(t, y) = 0. \tag{76}$$

If $d(x, W(s)) < \delta_0 \leq d(y, W(t))$, then $\psi(d(y, W(t))) = 0$, and therefore $\tilde{f}(t, y) = 0$. Taking into account (75), we conclude

$$\begin{aligned} \|\tilde{f}(s, x) - \tilde{f}(t, y)\| &= \big\|\psi(d(x, W(s)))f(s, x)\big\| \\ &= \big|\psi(d(x, W(s))) - \psi(d(y, W(t)))\big|\|f(s, x)\| \\ &\leq 3\delta_0^{-1}\kappa(\kappa|s - t| + \|x - y\|). \end{aligned} \tag{77}$$

Finally, we assume $d(x, W(s)) < \delta_0$ and $d(y, W(t)) < \delta_0$. Then, using again (75),

$$\begin{aligned} &\|\tilde{f}(s, x) - \tilde{f}(t, y)\| \\ &\quad\leq |\psi(d(x, W(s))) - \psi(d(y, W(t)))|\|f(s, x)\| \\ &\qquad + \psi(d(y, W(t)))\|f(s, x) - f(t, y)\| \\ &\quad\leq 3\delta_0^{-1}\kappa(\kappa|s - t| + \|x - y\|) + \|f(s, x) - f(t, y)\|. \end{aligned} \tag{78}$$

Now (3), (73), and (76–78) imply the $\tilde{\varrho}$-Hölder condition for $\tilde{f}$. Finally, Lipschitz continuity of $\tilde{f}$ follows from (4) and (76–78) with $s = t$, which completes the proof of (72).

Let $u = S(f, u_0)$. Then, by (70), for all $t \in [a, b]$ we have $u(t) \in W(t)$ and therefore, by (71),

$$u'(t) = f(t, u(t)) = \tilde{f}(t, u(t)).$$

It follows that

$$S(\tilde{f}, u_0) = S(f, u_0). \tag{79}$$

Let $\omega \in \Omega$, $n \in \mathbb{N}$, and let $\tilde{u}_k$ $(0 \leq k \leq n)$, $\tilde{p}_{kj}$ and $\tilde{q}_{kj}$ $(0 \leq k \leq n - 1, 0 \leq j \leq r_1)$ be the resulting sequences from the definition (10–16) of $A_{n,\omega}^{r_1}(\tilde{f}, u_0)$. By (25), (72), and Proposition 3.4,

$$\|S(\tilde{f}, u_0) - A_{n,\omega}^{r_1}(\tilde{f}, u_0)\|_{B([a,b],X)} \leq c_1 n^{-\tilde{\varrho}} \leq \delta_0/3, \tag{80}$$

provided $n \geq n_1$, where $n_1 = \lceil (3c_1/\delta_0)^{1/\tilde{\varrho}} \rceil$. Taking into account (70) and (79), it follows from (80) that for $0 \leq k \leq n - 1$, $t \in [t_k, t_{k+1}]$

$$\begin{aligned} d(\tilde{p}_{kr_1}(t), W(t)) &\leq \|\tilde{p}_{kr_1}(t) - \big(S(f, u_0)\big)(t)\| \\ &= \|\tilde{p}_{kr_1}(t) - \big(S(\tilde{f}, u_0)\big)(t)\| \leq \frac{\delta_0}{3}, \end{aligned}$$

and hence, also

$$d(\tilde{u}_k, W(t_k)) \leq \frac{\delta_0}{3}. \tag{81}$$

By (11) and (13–14), for $0 \leq j \leq r_1$

$$\|\tilde{p}_{kj}(t) - \tilde{u}_k\| \leq c_0(r_1)\kappa_1 h. \tag{82}$$

Using (74), (81), and (82), we conclude

$$d(\tilde{p}_{kj}(t), W(t)) \leq c_0(r_1)\kappa_1 h + \frac{\delta_0}{3} + \kappa h \leq \frac{2\delta_0}{3} \quad (t \in [t_k, t_{k+1}]), \tag{83}$$

provided $n \geq n_2$, with a suitably chosen $n_2 \geq n_1$.

Next we consider algorithm $A^{r_1}_{n,\omega}$ for (f, u_0). Let u_k $(0 \leq k \leq n)$, p_{kj}, and q_{kj} $(0 \leq k \leq n-1,\ 0 \leq j \leq r_1)$ be the corresponding sequences from the definition (10–16), as far as they are defined (compare conditions (10), (12), and (15)). We show by induction that for $0 \leq k \leq n$ the following holds:

$$u_k \text{ is defined and } u_k = \tilde{u}_k, \tag{84}$$

and, if $k \leq n-1$, then for all j with $0 \leq j \leq r_1$

$$p_{kj} \text{ is defined and } p_{kj} = \tilde{p}_{kj}. \tag{85}$$

First we prove that (84) implies (85). So assume that (84) holds for some $0 \leq k \leq n-1$. We show (85) by induction over j. Let $j = 0$. By (84), (81), and (71), $u_k \in U(t_k)$ and

$$f(t_k, u_k) = f(t_k, \tilde{u}_k) = \tilde{f}(t_k, \tilde{u}_k),$$

therefore p_{k0} is defined and

$$p_{k0}(t) = u_k + f(t_k, u_k)(t - t_k) = \tilde{u}_k + \tilde{f}(t_k, \tilde{u}_k)(t - t_k) = \tilde{p}_{k0}(t).$$

Now we assume that (85) holds for some j with $0 \leq j < r_1$. Then

$$p_{kj}(t_{k,j+1,i}) = \tilde{p}(t_{k,j+1,i}) \quad (i = 0, \ldots, j+1), \tag{86}$$

hence, by (83),

$$d(p_{kj}(t_{k,j+1,i}), W(t_{k,j+1,i})) \leq \frac{2\delta_0}{3}. \tag{87}$$

In particular, $p_{kj}(t_{k,j+1,i}) \in U(t_{k,j+1,i})$ for $i = 0, \ldots, j+1$, so $p_{k,j+1}$ is defined. Moreover, by (86) and (87)

$$f(t_{k,j+1,i}, p_{kj}(t_{k,j+1,i})) = \tilde{f}(t_{k,j+1,i}, \tilde{p}_{kj}(t_{k,j+1,i})),$$

so, compare (13), $q_{kj} = \tilde{q}_{kj}$, and therefore,

$$\begin{aligned} p_{k,j+1}(t) &= u_k + \int_{t_k}^{t} (P_{k,j+1} q_{k,j+1})(s)ds \\ &= \tilde{u}_k + \int_{t_k}^{t} (P_{k,j+1} \tilde{q}_{k,j+1})(s)ds = \tilde{p}_{k,j+1}(t). \end{aligned}$$

This completes the induction over j and the proof that (84) implies (85).

Next we show (84) by induction over k. For $k = 0$ it holds by definition. Now suppose (84) and thus (85) hold for some k with $0 \le k \le n-1$. Since by (83) and (85)

$$d(p_{kr_1}(t), W(t)) = d(\tilde{p}_{kr_1}(t), W(t)) \le \frac{2\delta_0}{3} \quad (t \in [t_k, t_{k+1}]),$$

it follows that u_{k+1} is defined and

$$\begin{aligned} u_{k+1} &= p_{kr_1}(t_{k+1}) + h\big(f(\xi_{k+1}, p_{kr_1}(\xi_{k+1})) - p'_{kr_1}(\xi_{k+1})\big) \\ &= \tilde{p}_{kr_1}(t_{k+1}) + h\big(\tilde{f}(\xi_{k+1}, \tilde{p}_{kr_1}(\xi_{k+1})) - \tilde{p}'_{kr_1}(\xi_{k+1})\big) = \tilde{u}_{k+1}. \end{aligned}$$

This shows (84) for $k+1$, completes the induction over k, and proves (84–85).

It follows that $A_n^{r_1}(f, u_0)$ is defined. Setting $S(f, u_0) = u$, we have for $0 \le k \le n-1$ and $t \in [t_k, t_{k+1}]$

$$\|u(t) - u(t_k)\| \le \kappa(t - t_k)$$

and therefore, using (23) and (19)

$$u(t_k) + c_0(r_1)\kappa(t - t_k)B_X \subseteq u(t) + (c_0(r_1) + 1)\kappa(t - t_k)B_X \subseteq U(t)$$

provided $n \ge n_0$, with $n_0 \ge n_2$ suitably chosen. Recalling (24), it follows that for $n \ge n_0$ we have $(f, u_0) \in \mathscr{H}^{r,\varrho}(U, \kappa, L, U_0, V, r_1, n)$, consequently

$$\mathscr{F}^{r,\varrho}(U, \kappa, L, U_0, V) = \mathscr{H}^{r,\varrho}(U, \kappa, L, U_0, V, r_1, n).$$

Now Theorem 3.3 follows from Proposition 3.4.

□

4 Lower bounds, Banach space valued integration, and complexity

To prove lower bounds we shall exploit that Banach space valued integration is a special case of the initial value problem. While complexity of integration in the scalar case is well-studied, the Banach space case has been investigated only recently in [3]. This paper covered the case C^r, while here we need the case $C^{r,\varrho}$. The (short) proof of the lower bound is analogous, we include it though for the sake of completeness. Since algorithms for the initial value problem lead to algorithms for integration, we can use the algorithm from above to get upper bounds. Furthermore, it is informative to see what this algorithm means for integration.

For $-\infty < a < b < \infty$, $r \in \mathbb{N}_0$, $0 \le \varrho \le 1$ let $F_0 = \mathscr{C}^{r,\varrho}([a,b],1)$ be the set of all r-times continuously differentiable functions $f : [a,b] \to X$ satisfying for $s, t \in [a,b]$

$$\begin{aligned} \|f^{(j)}(t)\| &\le 1 \qquad (0 \le j \le r), \\ \|f^{(r)}(s) - f^{(r)}(t)\| &\le |s-t|^{\varrho}, \end{aligned}$$

let $G_0 = X$ and define $S_0 : F_0 \to G_0$ by

$$S_0(f) = \int_a^b f(t)dt.$$

Moreover, let $K_0 = X$ and $\Lambda_0 = \{\delta_s : \ s \in [a,b]\}$ with $\delta_s(f) = f(s)$. This defines the integration problem $\mathcal{P}_0 = (F_0, G_0, S_0, K_0, \Lambda_0)$.

Let, as before, $U_0 \subseteq X$, $V \subseteq U \subseteq [a,b] \times X$, U open, $\kappa, L > 0$. We assume that there is a $u_0 \in U_0$ and a $\delta_1 > 0$ such that

$$u_0 + \delta_1(t-a)B_X \subseteq V(t) \quad (t \in [a,b]). \tag{88}$$

In particular, these conditions are satisfied if

$$U_0 \ne \emptyset \quad \text{and} \quad V = U = [a,b] \times X. \tag{89}$$

Then we can reduce the integration problem to the initial value problem. For this purpose, set $\sigma_0 = \min(\kappa, \delta_1)$ and define

$$R : F_0 \to \mathscr{C}^{r,\varrho}(U, \kappa, L) \times U_0, \quad R(f) = (\sigma_0 \bar{f}, u_0),$$

where $\bar{f}$ is given by $\bar{f}(t,x) = f(t)$ $((t,x) \in U)$. Then the solution of the system

$$u'(t) = \sigma_0 \bar{f}(t, u(t)) \quad (t \in [a,b]), \qquad u(a) = u_0,$$

is

$$u(t) = u_0 + \sigma_0 \int_a^t f(s)ds,$$

which, by (88), satisfies $u(t) \in V(t)$ $(t \in [a,b])$. Therefore by (6),

$$R(f) \in F = \mathscr{F}^{r,\varrho}(U, \kappa, L, U_0, V) \quad (f \in F_0).$$

Define

$$\Psi : B([a,b], X) \to X, \quad \Psi(g) = \sigma_0^{-1}(g(b) - u_0). \tag{90}$$

Obviously, we have

$$S_0 = \Psi \circ S \circ R.$$

This shows that the integration problem $\mathcal{P}_0 = (F_0, G_0, S_0, K_0, \Lambda_0)$ reduces to $\mathcal{P}$ (see [8] for the formal definition and additional details like the requirements on R, which are easily seen to be satisfied here). Consequently, with $e_n^{\rm set}$ standing for $e_n^{\rm det}$ or $e_n^{\rm ran}$, we have for all n

$$e_n^{\rm set}(S_0, F_0) \quad \le \quad \|\Psi\|_{\rm Lip} e_n^{\rm set}(S, F) = \sigma_0^{-1} e_n^{\rm set}(S, F), \tag{91}$$

where $\|\Psi\|_{\rm Lip}$ denotes the Lipschitz constant of Ψ. Next let us see how algorithm $A_{n,\omega}^{r_1}$ transforms into an algorithm for S_0. Considering $A_{n,\omega}^{r_1}$, applied to $(\sigma_0 \bar{f}, u_0)$, we have by (11) and (13–14) for $t \in [t_k, t_{k+1}]$

$$p_{kr_1}(t) \quad = \quad u_k + \sigma_0 \int_{t_k}^{t} (P_{kr_1} f)(s) ds,$$

Here $P_{kr_1} f$ stands for $P_{kr_1}(f(t_{k,r_1,i}))_{i=0}^{r_1}$ with $t_{k,r_1,i} = t_k + ih/r_1$ if $r_1 \ge 1$, while $P_{k0} f$ is given by $(P_{k0} f)(t) \equiv f(t_k)$. It follows that

$$u_{k+1} = u_k + \sigma_0 \int_{t_k}^{t_{k+1}} (P_{kr_1} f)(t) dt + \sigma_0 h \left(f(\xi_{k+1}) - (P_{kr_1} f)(\xi_{k+1}) \right),$$

and therefore

$$u_n = u_0 + \sigma_0 \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (P_{kr_1} f)(t) dt + \sigma_0 h \sum_{k=0}^{n-1} \left(f(\xi_{k+1}) - (P_{kr_1} f)(\xi_{k+1}) \right).$$

Together with (17) and (90) this gives

$$\begin{aligned} A_{n,\omega}^{(0)}(f) &:= \Psi \circ A_{n,\omega}^{r_1} \circ R(f) = \sigma_0^{-1}(u_n - u_0) \\ &= \sum_{k=0}^{n-1} \left(\int_{t_k}^{t_{k+1}} (P_{kr_1} f)(t) dt + \frac{b-a}{n} \left(f(\xi_{k+1}) - (P_{kr_1} f)(\xi_{k+1}) \right) \right), \end{aligned} \tag{92}$$

which is Monte Carlo integration with stratified sampling and separation of the main part. Moreover, for $f \in F_0$

$$\begin{aligned} \|S_0(f) - A_{n,\omega}^{(0)}(f)\| \quad &= \quad \|\Psi \circ S \circ R(f) - \Psi \circ A_{n,\omega}^{r_1} \circ R(f)\| \\ &\le \quad \sigma_0^{-1} \|S(\sigma_0 \bar{f}, u_0) - A_{n,\omega}^{r_1}(\sigma_0 \bar{f}, u_0)\|_{B([a,b],X)}. \end{aligned} \tag{93}$$

Taking, e.g., the choice (89), the conditions of Theorem 3.2 are satisfied, which together with (93) yields that for any $\omega \in \Omega$

$$\begin{aligned} \sup_{f \in F_0} \|S_0(f) - A_{n,\omega}^{(0)}(f)\| \quad &\le \quad \sigma_0^{-1} \sup_{f \in F_0} \|S(\sigma_0 \bar{f}, u_0) - A_{n,\omega}^{r_1}(\sigma_0 \bar{f}, u_0)\|_{B([a,b],X)} \\ &\le \quad \sigma_0^{-1} \sup_{(g,v_0) \in F} \|S(g, v_0) - A_{n,\omega}^{r_1}(g, v_0)\|_{B([a,b],X)} \\ &\le \quad c n^{-\min(r+\varrho, r_1+1)} \end{aligned} \tag{94}$$

and similarly

$$\sup_{f\in F_0}\left(\mathbb{E}\,\|S_0(f)-A^{(0)}_{n,\omega}(f)\|^p\right)^{1/p} \quad\leq\quad c\tau_p(X)n^{-\min(r+\varrho,r_1+1)-1+1/p}. \qquad (95)$$

Let us mention that (95) could also be derived directly from (92), Lemma 2.1, and

$$\sup_{t\in[t_k,t_{k+1}]}\|f(t)-(P_{kr_1}f)(t)\|\leq cn^{-\min(r+\varrho,r_1+1)},$$

which is just (35) for the present situation.

Proposition 4.1. *Let r,ϱ,a,b,X be as above. Then there are constants $c_1,c_2>0$ such that for all $n\in\mathbb{N}$ the deterministic n-th minimal error of the integration problem satisfies*

$$c_1n^{-r-\varrho}\leq e_n^{\det}(S_0,F_0)\leq c_2n^{-r-\varrho}.$$

Let, in addition, $1\leq p\leq 2$ and assume that X is of type p. Let p_X be the supremum of all p_1 such that X is of type p_1. Then there are constants $c_3,c_4>0$ such that for all $n\in\mathbb{N}$ the randomized n-th minimal error fulfills

$$c_3n^{-r-\varrho-1+1/p_X}\leq e_n^{\mathrm{ran}}(S_0,F_0)\leq c_4n^{-r-\varrho-1+1/p}.$$

Proof. Choosing $r_1\geq r$, the upper bounds follow from (94) and (95), since algorithm $A^{(0)}_{n,\omega}$ uses not more than cn values of f. Let us turn to the lower bounds. Since every Banach space X contains an isometric copy X_0 of $\mathbb{R}$, scalar problems reduce to the Banach case by considering problems such that all values of f are in X_0. Therefore the lower bounds in the deterministic setting and in the randomized setting with $p_X=2$ follow from the respective scalar results.

Now we assume $p_X<2$ and consider the randomized setting. Since a finite dimensional space Z always satisfies $p_Z=2$, the space X must be infinite dimensional. By the Maurey-Pisier Theorem (see [14], Th. 2.3) for every $n\in\mathbb{N}$ there is a subspace $E_n\subset X$ of dimension $8n$ and an isomorphism $T:\ell_{p_X}^{8n}\to E_n$ with $\|T\|\leq 1$ and $\|T^{-1}\|\leq 2$. Let $x_i=Te_i$, where $(e_i)_{i=0}^{8n-1}$ is the unit vector basis of $\ell_{p_X}^{8n}$. Let $\psi\in C^\infty(\mathbb{R})$ be such that $\psi(t)>0$ for $t\in(0,1)$ and $\operatorname{supp}\psi\subset[0,1]$, let $n\in\mathbb{N}$ and define for $t\in\mathbb{R}$, $i=0,\dots,8n-1$

$$\psi_i(t)=\psi(8n(t-t_i)),\quad t_i=a+i\frac{(b-a)}{8n}.$$

There is a constant $c_0>0$ such that for all $n\in\mathbb{N}$ and $(\alpha_i)_{i=0}^{8n-1}\subset[-1,1]^{8n}$

$$c_0n^{-r-\varrho}\sum_{i=0}^{8n-1}\alpha_i\psi_ix_i\in F_0.$$

Setting $f_i = c_0 n^{-r-\varrho}\psi_i x_i$ and $\sigma_1 = \int_0^1 \psi(t)dt$, we get for all $(\alpha_i)_{i=0}^{8n-1} \subset \mathbb{R}^{8n}$

$$\begin{aligned}
&\Big\| \sum_{i=0}^{8n-1} \alpha_i S_0 f_i \Big\| = c_0 n^{-r-\varrho} \Big\| \sum_{i=0}^{8n-1} \alpha_i x_i \int_a^b \psi_i(t)dt \Big\| \\
= \;& \frac{1}{8} c_0 (b-a)\sigma_1 n^{-r-\varrho-1} \Big\| \sum_{i=0}^{8n-1} \alpha_i x_i \Big\| \ge c n^{-r-\varrho-1} \left(\sum_{i=0}^{8n-1} |\alpha_i|^{p_X} \right)^{1/p_X} .
\end{aligned}$$

Using Lemma 5 and 6 of [7] with $K = X$ (Lemma 6 is formulated for $K = \mathbb{R}$, but is easily seen to hold also for $K = X$), we conclude

$$e_n^{\mathrm{ran}}(S_0, F_0) \;\ge\; \frac{1}{4} \min_{I \subseteq \{0,\dots,4n-1\}, |I| \ge 4n} \mathbb{E} \Big\| \sum_{i \in I} \varepsilon_i S_0 f_i \Big\| \ge c n^{-r-\varrho-1+1/p_X},$$

with $(\varepsilon_i)_{i=0}^{8n-1}$ being a sequence of independent Bernoulli random variables with $\mathbb{P}\{\varepsilon_i = -1\} = \mathbb{P}\{\varepsilon_i = +1\} = 1/2$. □

Now we consider the initial value problem.

Theorem 4.2. *Let $r, \varrho, a, b, \kappa, L$ be as above, let X be a Banach space, $\emptyset \ne U_0 \subseteq X$, $V \subseteq U \subseteq [a,b] \times X$, and U open. Assume that one of the following is fulfilled:*
1. Conditions (8) and (20) hold or
2. $r + \varrho > 0$ and there are $\delta_0, \delta_1 > 0$ and $u_0 \in U_0$ such that conditions (23) and (88) hold.

Then there are constants $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$ the deterministic n-th minimal error of the initial value problem satisfies

$$c_1 n^{-r-\varrho} \le e_n^{\det}(S, F) \le c_2 n^{-r-\varrho}.$$

Let, moreover, $1 \le p \le 2$ and assume that X is of type p. Let p_X be the supremum of all p_1 such that X is of type p_1. Then there are constants $c_3, c_4 > 0$ such that for all $n \in \mathbb{N}$ the randomized n-th minimal error fulfills

$$c_3 n^{-r-\varrho-1+1/p_X} \le e_n^{\mathrm{ran}}(S, F) \le c_4 n^{-r-\varrho-1+1/p}.$$

Proof. The upper bounds follow directly from Theorems 3.2 and 3.3 and the fact that the algorithm needs not more than cn values of f. The lower bounds follow from Proposition 4.1 and (91). □

Note that the bounds in the randomized cases of Proposition 4.1 and Theorem 4.2 are matching up to an arbitrarily small gap in the exponent. In some cases, they are even of matching order.

Corollary 4.3. *Assume that the conditions of Theorem 4.2 hold. Let p_X be the supremum of all p_1 such that X is of type p_1. Then for each $\varepsilon > 0$ there are constants $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$*

$$c_1 n^{-r-\varrho-1+1/p_X} \le e_n^{\mathrm{ran}}(S, F) \le c_2 n^{-r-\varrho-1+1/p_X+\varepsilon}.$$

If, moreover, the supremum of types is attained, that is, X is of type p_X, then there are constants $c_3, c_4 > 0$ such that for all $n \in \mathbb{N}$

$$c_3 n^{-r-\varrho-1+1/p_X} \le e_n^{\mathrm{ran}}(S, F) \le c_4 n^{-r-\varrho-1+1/p_X}.$$

The latter holds, in particular, for spaces of type 2 and, if $1 \le p_1 < \infty$, for spaces $X = L_{p_1}(\mathcal{N}, \nu)$, where $(\mathcal{N}, \nu)$ is some measure space.

Under the conditions of Proposition 4.1 the same results hold with $e_n^{\mathrm{ran}}(S, F)$ replaced by $e_n^{\mathrm{ran}}(S_0, F_0)$.

The lower bounds in Proposition 4.1, Theorem 4.2, and Corollary 4.3 remain true for algorithms of varying cardinality (see [7, 8] for the definition), since the results from [7] used in proof of Proposition 4.1 are valid for this class of algorithms, as well.

For general Banach spaces X upper and lower bounds of matching order of $e_n^{\mathrm{ran}}(S, F)$ and $e_n^{\mathrm{ran}}(S_0, F_0)$ remain an open problem.

References

[1] H. Cartan, Cours de calcul différentiel, Hermann, Paris, 1977.

[2] Th. Daun, On the randomized solution of initial value problems, J. Complexity 27 (2011), 300–311.

[3] Th. Daun, S. Heinrich, Complexity of Banach space valued and parametric integration, paper submitted to the Proceedings of MCQMC 2012.

[4] Th. Daun, S. Heinrich, Complexity of parametric initial value problems, paper in preparation.

[5] J. L. Doob, Stochastic Processes, Wiley, New York, 1953.

[6] G. Godefroy, Renorming of Banach spaces, in: W. B. Johnson, J. Lindenstrauss, eds., Handbook of the Geometry of Banach Spaces, Vol.1, Elsevier, Amsterdam 2001, pp. 781–835 .

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

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

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

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

[11] B. Kacewicz, Randomized and quantum algorithms yield a speed-up for initial-value problems, J. Complexity 20 (2004), 821–834.

[12] B. Kacewicz, Almost optimal solution of initial-value problems by randomized and quantum algorithms, J. Complexity 22 (2006), 676–690.

[13] M. Ledoux, M. Talagrand, Probability in Banach Spaces, Springer, 1991.

[14] B. Maurey, G. Pisier, Series de variables aléatoires vectorielles independantes et propriétés geométriques des espaces de Banach, Stud. Math. 58, 45–90 (1976).

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

[16] G. Stengle, Numerical methods for systems with measurable coefficients, Appl. Math. Lett. 3 (1990), 25–29.

[17] G. Stengle, Error analysis of a randomized numerical method, Numer. Math. 70 (1995), 119–128.

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