AVERAGE CASE COMPLEXITY OF WEIGHTED APPROXIMATION AND INTEGRATION OVER $\mathbb{R}_+$

LESZEK PLASKOTA, KLAUS RITTER, AND GRZEGORZ W. WASILKOWSKI

ABSTRACT. We study weighted approximation and integration of Gaussian stochastic processes X defined over $\mathbb{R}_+$ whose rth derivatives satisfy a Hölder condition with exponent β in the quadratic mean. We assume that the algorithms use samples of X at a finite number of points. We study the average case (information) complexity, i.e., the minimal number of samples that are sufficient to approximate/integrate X with the expected error not exceeding ε. We provide sufficient conditions in terms of the weight and the parameters r and β for the weighted approximation and weighted integration problems to have finite complexity. For approximation, these conditions are necessary as well. We also provide sufficient conditions for these complexities to be proportional to the complexities of the corresponding problems defined over $[0, 1]$, i.e., proportional to $\varepsilon^{-1/\alpha}$ where $\alpha = r + \beta$ for the approximation and $\alpha = r + \beta + 1/2$ for the integration.

1. INTRODUCTION

Complexity of approximating or integrating a function defined over a bounded domain has already been a well established area. We mention only Traub, Wasilkowski, and Woźniakowski (1988), Ritter (2000), and the references therein. Complexity results include various settings such as the worst case and the average case settings. There are, however, very few results that address these problems for functions defined over unbounded domains such as $\mathbb{R}^d$.

Some progress has recently been made in the worst case setting for the approximation and integration problems over $\mathbb{R}$ and $\mathbb{R}^d$; see, respectively, Wasilkowski and Woźniakowski (2000a) and (2000b). See also Traub, Wasilkowski, and Woźniakowski (1983), Curbera (1998), and Mathé (1998). In the present paper we study complexity of approximating functions[1] $f : \mathbb{R}_+ \to \mathbb{R}$ and their integrals over $\mathbb{R}_+ = [0, \infty)$ in the average case setting, assuming that the class of functions is equipped with a probability measure. Equivalently, we assume that f is a trajectory of a stochastic process X on $\mathbb{R}_+$, and we measure the errors by the quadratic mean. These problems seem not to have been studied yet.

In contrast to processes defined on a compact interval, say $[0, 1]$, the expected squared L_2-norm of typical processes defined on $\mathbb{R}_+$ (including the fractional Brownian motion) is infinite. Furthermore, the integral over $\mathbb{R}_+$ does not exist with probability one. Hence the complexity analysis of those problems is of interest only in a weighted sense.

More specifically, let $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ be a measurable *weight function*. For a given zero mean Gaussian stochastic process $X(t)$, $t \in \mathbb{R}_+$, we want to approximate X or its

Date: November 28, 2000.

[1]We consider $\mathbb{R}_+$ instead of $\mathbb{R}$ as the domain of the functions f for simplicity only.

(weighted) integral

$$\text{Int}_\rho X = \int_0^\infty X(t) \cdot \rho(t)\, dt. \tag{1}$$

The error of an approximation $\mathcal{A}X$ of X is given as

$$e(\mathcal{A}, \text{App}_\rho) = \left(\mathbb{E} \left(\int_0^\infty (X - \mathcal{A}X)^2(t) \cdot \rho^2(t)\, dt \right) \right)^{1/2}$$

and the error of a quadrature $\mathcal{Q}X$ for $\text{Int}_\rho X$ is given as

$$e(\mathcal{Q}, \text{Int}_\rho) = \left(\mathbb{E}(\text{Int}_\rho X - \mathcal{Q}X)^2 \right)^{1/2}.$$

Here and elsewhere $\mathbb{E}$ stands for expectation.

We assume that any method, i.e., any approximation $\mathcal{A}$ or quadrature $\mathcal{Q}$, can use only samples (or observations) of X at a finite number of points $t_i \in \mathbb{R}_+$. We call this number the *cardinality* and denote it by $\text{card}(\mathcal{A})$ or $\text{card}(\mathcal{Q})$, respectively[2].

We are interested in the *(information) complexity* of weighted approximation and integration, which is the minimal number of samples needed to construct an approximation (algorithm) with error not exceeding a given $\varepsilon > 0$. That is, for the approximation,

$$\text{comp}(\varepsilon, \text{App}_\rho) = \min\{\, \text{card}(\mathcal{A}) : \mathcal{A} \text{ s.t. } e(\mathcal{A}, \text{App}_\rho) \le \varepsilon \,\},$$

and for integration $\text{comp}(\varepsilon, \text{Int}_\rho)$ is defined correspondingly.

We present results that do not depend on the particular process X but hold for classes of processes. These classes are defined by quadratic mean properties, see Section 2 for details and examples. In particular, we assume that for some $r \in \mathbb{N}_0$ the derivative $X^{(r)}$ is Hölder continuous in quadratic mean with exponent $\beta \in [0, 1]$.

It is clear that for some weight functions ρ the complexity of approximation is infinite, and the integration problem is not even well defined. Therefore, one of our first results provides a necessary and sufficient condition for the complexity of approximation to be finite for every $\varepsilon > 0$. We also provide a necessary and sufficient condition for the weighted integral to exist with probability one. This condition simultaneously gives finite complexity for the integration problem.

Approximation over $\mathbb{R}_+$ cannot have smaller complexity than the corresponding problem restricted to a compact interval. The same usually (but not always) holds for integration. Typically, the complexity on compact subintervals is $\Theta(\varepsilon^{-1/\alpha})$ with $\alpha = r + \beta$ for approximation and $\alpha = r + \beta + 1/2$ for integration. We provide sufficient conditions for the complexity of weighted problems on $\mathbb{R}_+$ to be proportional to $\varepsilon^{-1/\alpha}$ as well.

To give a flavor of the results, let $\rho(x) = \Theta(x^{-\gamma})$. Then $\gamma \ge \alpha + 1/2$ implies the complexity $\Theta(\varepsilon^{-1/\alpha})$ for both problems. On the other hand, if $\gamma < \alpha + 1/2$ then the complexity of approximation is infinite, and the integration problem may not be well defined.

---

[2]We formally consider nonadaptive observations. Note that adaptive observations with varying cardinality do not lead to essentially better approximations for problems considered in this paper, see Wasilkowski (1986).

Finally, we state that in cases where $\mathrm{comp}(\varepsilon, \rho) = \Theta(\varepsilon^{-1/\alpha})$, the upper bounds are provided by the cost of specific algorithms. For the approximation problem, these algorithms are deterministic and enjoy certain robustness properties. Indeed, they are based on a simple piecewise polynomial interpolation, and they do *not* require any specific information about X other than an upper bound for the parameter $r + \beta$. For the integration, similar deterministic algorithms are constructed only in special cases. In general case, the upper bound is given by Monte Carlo arguments.

2. Assumptions and Examples

We consider a measurable Gaussian stochastic process $X(t)$, $t \in \mathbb{R}_+$, with zero mean, i.e., $\mathbb{E}(X(t)) = 0$ for every t. The covariance kernel K of X is defined by

$$K(s,t) = \mathbb{E}(X(s) \cdot X(t))$$

for $s, t \in \mathbb{R}_+$.

Let $r \in \mathbb{N}_0$ and $\beta \in [0,1]$. The process X satisfies Hölder condition of order (r, β) if the derivatives $X^{(1)}, \ldots, X^{(r)}$ exist and are continuous in quadratic mean and if

$$\mathbb{E}\left(X^{(r)}(s) - X^{(r)}(t)\right)^2 \le C_1^2 \cdot |s-t|^{2\beta} \tag{2}$$

for all $s, t \in \mathbb{R}_+$ with a constant $C_1 > 0$. This property can be equivalently stated in terms of the covariance kernel. Namely, the partial derivatives $K^{(i,j)}$ exist and are continuous on $\mathbb{R}_+^2$ for $i, j = 0, \ldots, r$, and

$$K^{(r,r)}(s,s) - 2K^{(r,r)}(s,t) + K^{(r,r)}(t,t) \le C_1^2 \cdot |s-t|^{2\beta}. \tag{3}$$

In fact, the left-hand sides in (2) and (3) coincide.

Example 1. The fractional Brownian motion with parameter $\beta \in (0,1)$ is the zero mean Gaussian process with covariance kernel

$$K(s,t) = \tfrac{1}{2}\left(s^{2\beta} + t^{2\beta} - |s-t|^{2\beta}\right).$$

This process satisfies the Hölder condition of order $(0, \beta)$, since (3) holds with equality for $C_1 = 1$ and $r = 0$. In particular, for $\beta = \frac{1}{2}$ we get the Brownian motion with covariance kernel

$$K(s,t) = \tfrac{1}{2}\left(s + t - |s-t|\right) = \min\{s,t\}.$$

Suppose that Y is zero mean Gaussian and satisfies the Hölder condition of order $(0, \beta)$. Take $r \ge 1$. By r-fold integration,

$$X(t) = \int_0^t \frac{(t-u)^{r-1}}{(r-1)!}\, Y(u)\, du,$$

we obviously get a zero mean Gaussian process X that satisfies the Hölder condition of order (r, β). This construction yields, in particular, the r-fold integrated (fractional) Brownian motion.

Now we consider a stationary process X with spectral density φ. Such processes are naturally defined on the whole real line. It is well known that the smoothness of X is closely related to decay properties of its spectral density. By definition, φ is symmetric, nonnegative, and integrable, and the covariance kernel K of X satisfies

$$K(s,t) = \int_{-\infty}^{\infty} \exp(\imath(s-t)u) \cdot \varphi(u)\, du.$$

Assume that

$$\varphi(u) \le c \cdot |u|^{-2\gamma}$$

with constants $c > 0$ and $\gamma > \frac{1}{2}$ for $|u|$ sufficiently large. If $\gamma - \frac{1}{2} \notin \mathbb{N}$ then X satisfies the Hölder condition with

$$r = \lfloor \gamma - \tfrac{1}{2} \rfloor \qquad \text{and} \qquad \beta = \gamma - \tfrac{1}{2} - r, \tag{4}$$

see Ritter (2000, Lemma VI.5). We add that all major results in this paper hold for $\gamma - \frac{1}{2} \in \mathbb{N}$, as well, if $r + \beta$ is replaced by $\gamma - \frac{1}{2}$.

The Sacks-Ylvisaker conditions[3], see Ritter, Wasilkowski, and Woźniakowski (1995), define another class of processes that satisfy Hölder conditions of order $(r, \frac{1}{2})$.

We will use Hölder conditions to derive upper bounds for the complexity. These conditions do not imply nontrivial lower bounds neither for approximation nor for integration. To derive nontrivial lower bounds for approximation we require the following additional property.

For $a < b$ and $t \in (a,b)$ let $\widetilde{X}_{a,b}(t)$ denote the conditional expectation of $X(t)$ given $X(s)$, $s \in [0,a] \cup [b,\infty)$. Thus $\widetilde{X}_{a,b}(t)$ has minimal mean squared error among all estimators for $X(t)$ that are based on complete knowledge of X outside of (a,b). We assume that

$$\mathbb{E}\left(X(t) - \widetilde{X}_{a,b}(t)\right)^2 \ge C_2^2 \cdot \left(\frac{(b-t)\cdot(t-a)}{b-a}\right)^{2(r+\beta)} \tag{5}$$

for all $t \in (a,b)$ with a constant $C_2 > 0$ that does not depend on a and b. This property can be equivalently formulated by using the Hilbert space $H(K)$ with reproducing kernel K. Namely, for every $t \in [a,b]$,

$$\sup\{|h(t)| : h \in B(K),\ \operatorname{supp} h \subseteq [a,b]\} \ge C_2 \cdot \left(\frac{(b-t)\cdot(t-a)}{b-a}\right)^{r+\beta}, \tag{6}$$

where $B(K)$ denotes the unit ball in $H(K)$. In fact, the left-hand sides in (5) and (6) coincide, up to taking the square root.

Example 2. Let $t \in (a,b)$. For the Brownian motion X we have

$$\widetilde{X}_{a,b}(t) = \frac{X(a)\cdot(b-t) + X(b)\cdot(t-a)}{b-a}.$$

Since X has independent increments, we get

$$\mathbb{E}\left(X(t) - \widetilde{X}_{a,b}(t)\right)^2 = \frac{(b-t)\cdot(t-a)}{b-a}.$$

[3]These conditions are usually defined in the compact case $t \in [0,1]$; they may be used for $t \in \mathbb{R}_+$ in the same way.

Thus (5) holds with equality for $r = 0$, $\beta = \frac{1}{2}$, and $C_2 = 1$. This is generalized to the r-fold integrated Brownian motion in the following way. The conditional expectation $\widetilde{X}_{a,b}(t)$ is given by the polynomial of degree at most $2r+1$ that interpolates the boundary values $X^{(k)}(a)$ and $X^{(k)}(b)$ for $k = 0, \dots, r$, and (5) holds with equality for $\beta = \frac{1}{2}$ and $C_2^2 = 1/((2r+1)(r!)^2)$, see Speckman (1979). Note that $\widetilde{X}_{a,b}(t)$ only depends on the boundary values of the r-fold integrated Brownian motion. This is due to the fact that $(X^{(0)}, \dots, X^{(r)})$ is a Markov process in this case.

Using the results from Ritter, Wasilkowski, and Woźniakowski (1995), the lower bound (6) with $\beta = \frac{1}{2}$ can be verified under Sacks-Ylvisaker conditions of order $r \in \mathbb{N}_0$.

The fractional Brownian motion with $\beta \neq \frac{1}{2}$ is non-Markovian. For the corresponding reproducing kernel Hilbert space we have $h \in H(K)$ for every function $h \in C^\infty(\mathbb{R}_+)$ with compact support that does not include zero. Moreover, the norm of these functions is given by

$$\|h\|_K^2 = c \cdot \int_{-\infty}^{\infty} |u|^{2\beta+1} \cdot \left|\widehat{h}(u)\right|^2 du$$

for some constant c, see Singer (1994). Here $\widehat{h}$ denotes the Fourier transform of h. This allows us to establish (6) in the following way. Take $g \in C^\infty(\mathbb{R})$ such that $g(0) = 1$ and $g(s) = 0$ if $|s| \geq 1$, and put

$$C_2 = \left(2^\beta c^{1/2} \cdot \int_{-\infty}^{\infty} |u|^{2\beta+1} \cdot |\widehat{g}(u)|^2 \, du\right)^{-1}.$$

For $t \in (a, b)$ and $\delta = \min\{t - a, b - t\}$ define

$$h(s) = C_2 \cdot \delta^\beta \cdot g\left(2(s - t)/\delta\right)$$

for $s \in \mathbb{R}_+$. Then $h \in H(K)$ since $t > \delta/2$, and $\|h\|_K = 1$. Furthermore, $h = 0$ on $[0, a] \cup [b, \infty)$ and

$$h(t) = C_2 \cdot \delta^\beta \geq C_2 \cdot \left(\frac{(b - t) \cdot (t - a)}{b - a}\right)^\beta.$$

In a similar way, one can verify (6) for the r-fold integrated fractional Brownian motion.

Consider a stationary process X on the real line, whose spectral density φ satisfies

$$\varphi(u) \geq c \cdot \left(1 + u^2\right)^{-\gamma}$$

with constants $c > 0$ and $\gamma > \frac{1}{2}$ for every $u \in \mathbb{R}$. Then every function $h \in C^\infty(\mathbb{R})$ with compact support belongs to $H(K)$ and

$$\|h\|_K^2 \leq c' \cdot \int_{-\infty}^{\infty} \left(1 + u^2\right)^\gamma \cdot \left|\widehat{h}(u)\right|^2 du$$

for some constant $c' > 0$, see Ritter (2000, Lemma VI.7). Therefore (6), with r and β given by (4), can be verified as in the case of fractional Brownian motion with $\delta = \min\{t - a, b - t, 1\}$.

3. Weighted Approximation on $\mathbb{R}_+$

In this section we assume that X satisfies (2) and (5) with $r+\beta>0$ and, for simplicity, that

$$C_1 = 1.$$

3.1. Preliminary Results.

First, we study the error of piecewise polynomial interpolation of degree $r_0 \ge r$ on compact intervals. Put

$$\overline{r_0} = \max\{r_0, 1\}.$$

Lemma 1. *For $a<b$ and $n \in \mathbb{N}$, let $\mathcal{U}$ denote the operator of interpolation by piecewise polynomials of degree at most r_0 at the knots*

$$a + j \cdot \frac{b-a}{n\overline{r_0}}, \qquad j = 0, \dots, n\overline{r_0}.$$

There exists a constant $A_1 = A_1(r_0, r, \beta) > 0$, such that

$$\sup_{t\in[a,b]} \mathbb{E}\left(X(t) - \mathcal{U}X(t)\right)^2 \le A_1^2 \cdot \left(\frac{b-a}{n}\right)^{2(r+\beta)}.$$

Proof. As previously, let $B(K)$ denote the unit ball in the Hilbert space $H(K)$ with reproducing kernel K. We have

$$\mathbb{E}(X(t) - \mathcal{U}X(t))^2 = \sup_{h\in B(K)} |h(t) - \mathcal{U}h(t)|^2.$$

Moreover, the functions $h \in H(K)$ are r-times continuously differentiable with

$$\left|h^{(r)}(s) - h^{(r)}(t)\right| \le |s-t|^\beta,$$

see Ritter, Wasilkowski, and Woźniakowski (1993).

We give the proof of the lemma only in the case $r \ge 1$. Assume that $n=1$, at first. Write

$$h(t) = h_1(t) + h_2(t)$$

for $t \in [a,b]$, where

$$h_1(t) = \sum_{k=0}^{r} \frac{h^{(k)}(a)}{k!}\,(t-a)^k$$

and

$$h_2(t) = \int_a^t \frac{(t-u)^{r-1}}{(r-1)!}\left(h^{(r)}(u) - h^{(r)}(a)\right)\,du.$$

Observe that $h - \mathcal{U}h = h_2 - \mathcal{U}h_2$. Define

$$F(a,b) = \{h \in C^r([a,b]) : h^{(k)}(a) = 0 \text{ for } k = 0,\dots,r,\quad \left|h^{(r)}(s) - h^{(r)}(t)\right| \le |s-t|^\beta \text{ for } s,t \in [a,b]\},$$

so that $h_2 \in F(a,b)$. Moreover, let $p_0, \ldots, p_{r_0}$ denote the Lagrange polynomials for interpolation at the knots j/r_0 with $j = 0, \ldots, r_0$. Define $\widetilde{h}(z) = h_2(a + z \cdot (b-a))$ for $z \in [0,1]$. Then

$$h_2(t) - \mathcal{U}h_2(t) = \widetilde{h}\left(\frac{t-a}{b-a}\right) - \sum_{j=0}^{r_0} \widetilde{h}\left(\frac{j}{r_0}\right) \cdot p_j\left(\frac{t-a}{b-a}\right),$$

and $(b-a)^{-(r+\beta)} \cdot \widetilde{h} \in F(0,1)$ if $h_2 \in F(a,b)$. Therefore

$$\sup_{t\in[a,b]} \mathbb{E}(X(t) - \mathcal{U}X(t))^2 \leq \sup_{t\in[a,b]} \sup_{h\in F(a,b)} |h(t) - \mathcal{U}h(t)|^2 \leq (b-a)^{2(r+\beta)} \cdot A_1^2,$$

where

$$A_1 = \sup_{t\in[0,1]} \sup_{h\in F(0,1)} \left| h(t) - \sum_{j=0}^{r_0} h\left(\frac{j}{r_0}\right) \cdot p_j(t) \right|.$$

This constant is finite since every function $h \in F(0,1)$ is bounded by $(b-a)^{r+\beta}$.

For $n \geq 2$ the same arguments work on the respective subintervals of $[a,b]$ of length $(b-a)/n$. □

Next, we discuss the complexity in the classical case of unweighted L_2-approximation on a compact interval.

Theorem 1. *Let $a < b$ and*

$$\rho = 1_{[a,b]}.$$

Then

$$\mathrm{comp}(\varepsilon, \mathrm{App}_\rho) = \Theta\left(\varepsilon^{-1/(r+\beta)}\right).$$

Proof. Consider the piecewise polynomial interpolation $\mathcal{U}$ from Lemma 1. We get

$$e^2(\mathcal{U}, \mathrm{App}_\rho) \leq A_1^2 \cdot \left(\frac{b-a}{n}\right)^{2(r+\beta)} \cdot (b-a),$$

and the number of knots used by $\mathcal{U}$ is of order n. Hereby the upper bound for the complexity follows.

Consider an arbitrary method $\mathcal{A}$ that uses knots $t_1 < \cdots < t_n$. Assume without loss of generality that $a, b \in \{t_1, \ldots, t_n\}$, say $a = t_i$ and $b = t_j$. Using (5) we get

$$\begin{aligned} e^2(\mathcal{A}, \mathrm{App}_\rho) &= \sum_{k=i+1}^{j} \int_{t_{k-1}}^{t_k} \mathbb{E}\left(X(t) - \mathcal{A}X(t)\right)^2 \, dt \\ &\geq C_2^2 \cdot \sum_{k=i+1}^{j} \int_{t_{k-1}}^{t_k} \left(\frac{(t_k - t) \cdot (t - t_{k-1})}{t_k - t_{k-1}}\right)^{2(r+\beta)} dt \\ &= c \cdot \sum_{k=i+1}^{j} (t_k - t_{k-1})^{2(r+\beta)+1} \\ &\geq c \cdot (b-a)^{2(r+\beta)+1} \cdot n^{-2(r+\beta)} \end{aligned}$$

with a constant $c > 0$ that only depends on r, β, and C_2. Hence the lower bound for the complexity follows. $\square$

From Theorem 1 we conclude that $\mathrm{comp}(\varepsilon, \mathrm{App}_\rho)$ is at least of order $\varepsilon^{-1/(r+\beta)}$, if ρ is an arbitrary weight function on $\mathbb{R}_+$ that is bounded away from zero on an interval of positive length.

3.2. **Finite Complexity.** We give a necessary and sufficient condition for the complexity of approximation of X to be finite for any $\varepsilon > 0$.

Define the function $L : \mathbb{R}_+ \to \mathbb{R}_+ \cup \{\infty\}$ by

$$L(R) = \left(\int_R^\infty \rho^2(t) \cdot t^{2(r+\beta)}\, dt \right)^{1/2}.$$

Lemma 2. *We have*

$$\forall\, \varepsilon > 0 :\ \mathrm{comp}(\varepsilon, \mathrm{App}_\rho) < \infty, \tag{7}$$

iff

$$\lim_{R\to\infty} L(R) = 0 \tag{8}$$

and

$$\forall\ 0 \le a < b < \infty\ \forall\ \varepsilon > 0 :\ \mathrm{comp}(\varepsilon, \mathrm{App}_{\rho\cdot 1_{[a,b]}}) < \infty. \tag{9}$$

Proof. We claim that

$$K(t,t) \le A_2^2 \cdot t^{2(r+\beta)} \tag{10}$$

for $t \ge 1$, say, with a constant $A_2 = A_2(K, r, \beta) > 0$. To show this, define X_1 and X_2 by

$$X_1(t) = \sum_{k=0}^r \frac{X^{(k)}(0)}{k!}\, t^k$$

and $X(t) = X_1(t) + X_2(t)$. Then

$$K(t,t) \le 2 \left(\mathbb{E}(X_1(t))^2 + \mathbb{E}(X_2(t))^2 \right)$$

with $\mathbb{E}(X_1(t))^2 = O\left(t^{2r}\right)$. If $r \ge 1$ then

$$\begin{aligned} &\mathbb{E}(X_2(t))^2 \\ &\quad = \int_0^t \int_0^t \frac{(t-u)^{r-1}(t-v)^{r-1}}{((r-1)!)^2}\, \mathbb{E}\left(\left(X^{(r)}(u) - X^{(r)}(0)\right) \cdot \left(X^{(r)}(v) - X^{(r)}(0)\right) \right) du\, dv \\ &\quad \le \left(\int_0^t \frac{(t-u)^{r-1}}{(r-1)!} \cdot u^\beta\, du \right)^2 \le t^{2(r+\beta)}. \end{aligned}$$

This upper bound is obviously valid in the case $r = 0$, too, and (10) is proven.

Suppose now that (8) and (9) hold. For a given $\varepsilon > 0$, let $R_\varepsilon \ge 1$ be such that $L(R_\varepsilon) \le \varepsilon_1$ with $\varepsilon_1 = \varepsilon(1 + A_2^2)^{-1/2}$. Let $\mathcal{A}_\varepsilon$ be a method such that $e(\mathcal{A}_\varepsilon, \mathrm{App}_{\rho\cdot 1_{[0,R_\varepsilon]}}) \le \varepsilon_1$, and $\mathcal{A}_\varepsilon f$ is zero on $[R_\varepsilon, \infty)$. Then (10) yields

$$e(\mathcal{A}_\varepsilon, \mathrm{App}_\rho)^2 = e\left(\mathcal{A}_\varepsilon, \mathrm{App}_{\rho\cdot 1_{[0,R_\varepsilon]}}\right)^2 + \int_{R_\varepsilon}^\infty \rho^2(t) \cdot K(t,t)\, dt \le \varepsilon_1^2 \cdot \left(1 + A_2^2\right)^2 \le \varepsilon^2,$$

which proves (7).

Suppose that (7) holds. Then, of course, (9) holds as well, and we only need to show (8). For a given $\varepsilon > 0$, let $\mathcal{A}_\varepsilon$ be a method such that $e(\mathcal{A}_\varepsilon, \mathrm{App}_\rho) \le \varepsilon$. Let $t_{1,\varepsilon} < \cdots < t_{n,\varepsilon}$ denote the knots used by $\mathcal{A}_\varepsilon$, and put $R_\varepsilon = t_{n,\varepsilon}$. Recall the definition of $\widetilde{X}_{a,b}(t)$ from Section 2. For every $b > R_\varepsilon$ we have

$$\varepsilon^2 \ge \int_{R_\varepsilon}^b \mathbb{E}\,(X(t) - \mathcal{A}_\varepsilon X(t))^2 \cdot \rho^2(t)\,dt \ge \int_{R_\varepsilon}^b \mathbb{E}\left(X(t) - \widetilde{X}_{R_\varepsilon,b}(t)\right)^2 \cdot \rho^2(t)\,dt.$$

Using (5) we get

$$\varepsilon^2 \ge C_2^2 \cdot \int_{R_\varepsilon}^\infty \rho^2(t) \cdot 1_{[R_\varepsilon,b]}(t) \cdot \left(\frac{(b-t)\cdot(t-R_\varepsilon)}{b-R_\varepsilon}\right)^{2(r+\beta)} dt.$$

For $b \to \infty$ the integrand converges monotonically towards $\rho^2(t)\cdot(t-R_\varepsilon)^{2(r+\beta)}$. Thus

$$L^2(2R_\varepsilon) \le 2^{2(r+\beta)} \cdot \int_{2R_\varepsilon}^\infty \rho^2(t)\cdot(t-R_\varepsilon)^{2(r+\beta)}\,dt \le \frac{2^{2(r+\beta)}}{C_2^2}\cdot\varepsilon^2,$$

which proves (8). □

Lemma 3. *Assume that*

$$\int_0^\infty \rho^2(t)\,dt < \infty. \tag{11}$$

Then we have (9)*, i.e., finite complexity on compact subintervals. Moreover, we have finite complexity* (7) *iff*

$$\int_0^\infty \rho^2(t)\cdot t^{2(r+\beta)}\,dt < \infty. \tag{12}$$

Proof. We use Lemma 1 to conclude that (11) implies (9). Moreover, given (11), we have equivalence of (8) and (12). It remains to apply Lemma 2. □

3.3. **Upper Bounds.** We already know that the complexity of approximating X is at least of order $\varepsilon^{-1/(r+\beta)}$ if the weight ρ is bounded away from zero on a subinterval of positive length. In the following, we provide a method which, under some assumptions on ρ, has error ε and cardinality proportional to $\varepsilon^{-1/(r+\beta)}$. We also give a necessary condition for the complexity to be of that order.

Let $a_i = 2^i - 1$ for $i \in \mathbb{N}_0$, and define

$$\rho_i = \operatorname{ess\,sup}\{\rho(t) : t \in [a_{i-1}, a_i]\}$$

as well as

$$c_i = \rho_i^{1/(r+\beta+1/2)} \cdot 2^{i-1}$$

for $i \in \mathbb{N}$. We assume that

$$A_3 := A_3(r,\beta,\rho) := \sum_{i=1}^\infty c_i < \infty \tag{13}$$

and

$$L(a_i) \le A_4 \cdot c_{i+1}^{r+\beta+1/2} \tag{14}$$

with a constant $A_4 = A_4(r,\beta,\rho)$. To exclude trivial cases, we also assume $c_i > 0$ for all $i \in \mathbb{N}$.

The upper bound $\operatorname{comp}(\varepsilon, \mathrm{App}_\rho) = O\left(\varepsilon^{-1/(r+\beta)}\right)$ for the complexity is obtained by the following method $\mathcal{A}^*$. Let $k \in \mathbb{N}$ and $n_1, \dots, n_k \in \mathbb{N}$, and put $I_i = [a_{i-1}, a_i]$ for $i = 1, \dots, k$. On each of these intervals, $\mathcal{A}^* X$ is an interpolation of X by piecewise polynomials of degree at most $r_0 \ge r$, as in Lemma 1. The interpolation points are given by

$$a_{i-1} + j \cdot \frac{a_i - a_{i-1}}{n_i \overline{r_0}}, \qquad j = 0, \dots, n_i \overline{r_0},$$

for $i \le k$. On $I_{k+1} = [a_k, \infty)$ we use $\mathcal{A}^* X(t) = 0$. Clearly,

$$\operatorname{card}(\mathcal{A}^*) = \overline{r_0} \cdot \sum_{i=1}^{k} n_i + 1.$$

The particular choice of $k, n_1, \dots, n_k$ depends on ε, r, β, and ρ in the following way. For $\ell \in \mathbb{N}$ we define

$$G(\varepsilon, \ell) = \left(\frac{A^2}{\varepsilon^2} \cdot \sum_{i=1}^{\ell+1} c_i \right)^{1/(2(r+\beta))}$$

with

$$A = \max\{A_1,\ A_2 \cdot A_4\}.$$

We take

$$k = k(\varepsilon) = \min\left\{\ell \in \mathbb{N} : c_{\ell+1} \cdot G(\varepsilon, \ell) \le 1\right\}$$

and

$$n_i = \lceil c_i \cdot G(\varepsilon, k) \rceil .$$

Observe that k is well-defined for every ε because of (13). Moreover, $n_{k+1} = 1$.

Theorem 2. *Suppose that* (13) *and* (14) *hold. Then we have*

$$e(\mathcal{A}^*, \mathrm{App}_\rho) \le \varepsilon$$

and

$$\begin{aligned} \operatorname{comp}(\varepsilon, \mathrm{App}_\rho) &\le \operatorname{card}(\mathcal{A}^*) \\ &\le \overline{r_0} \cdot \left(2 \cdot \varepsilon^{-1/(r+\beta)} \cdot A^{1/(r+\beta)} \cdot A_3^{(r+\beta+1/2)/(r+\beta)} + 1 \right) + 1. \end{aligned}$$

Proof. We first show that the error of $\mathcal{A}^*$ is at most ε. Lemma 1 yields upper bounds for the error of $\mathcal{A}^*$ on the subintervals I_i for $i \le k$. We have

$$e_i^2 := \int_{a_{i-1}}^{a_i} \mathbb{E}(X(t) - \mathcal{A}^* X(t))^2 \cdot \rho^2(t)\, dt \le A_1^2 \cdot \frac{c_i^{2(r+\beta+1/2)}}{n_i^{2(r+\beta)}}.$$

Furthermore, by (10),

$$e_{k+1}^2 := \int_{a_k}^{\infty} \mathbb{E}(X(t) - \mathcal{A}^* X(t))^2 \cdot \rho^2(t)\, dt \le A_2^2 \cdot \int_{a_k}^{\infty} \rho^2(t) \cdot t^{2(r+\beta)}\, dt.$$

Using (14) we get

$$e_{k+1}^2 \le A_2^2\, A_4^2 \cdot c_{k+1}^{2(r+\beta+1/2)}.$$

The particular choice of n_i, $1 \le i \le k$, yields

$$e_i^2 \le A_1^2 \cdot G^{-2(r+\beta)}(\varepsilon, k) \cdot c_i.$$

The particular choice of k yields

$$e_{k+1}^2 \le A_2^2\, A_4^2 \cdot G^{-2(r+\beta)}(\varepsilon, k) \cdot c_{k+1}.$$

We therefore have

$$e^2(\mathcal{A}^*, \rho) = \sum_{i=1}^{k+1} e_i^2 \le A^2 \cdot G^{-2(r+\beta)}(\varepsilon, k) \cdot \sum_{i=1}^{k+1} c_i = \varepsilon^2.$$

Now we derive the upper bound for the cardinality of $\mathcal{U}$. For that end, we need to estimate the sum of n_i. Clearly,

$$\sum_{i=1}^{k} n_i \le k + G(\varepsilon, k) \cdot \sum_{i=1}^{k} c_i.$$

For $i = 2, \ldots, k$ we have $1 < c_i \cdot G(\varepsilon, i-1) \le c_i \cdot G(\varepsilon, k)$, and therefore

$$k \le 1 + G(\varepsilon, k) \cdot \sum_{i=1}^{k} c_i.$$

Finally

$$G(\varepsilon, k) \cdot \sum_{i=1}^{k} c_i \le 1 + G(\varepsilon, k) \cdot A_3 \le \varepsilon^{-1/(r+\beta)} \cdot A^{1/(r+\beta)} \cdot A_3^{(r+\beta+1/2)/(r+\beta)}.$$

We conclude that

$$\operatorname{card}(\mathcal{A}^*) \le \overline{r_0} \cdot \Big(2 \cdot \varepsilon^{-1/(r+\beta)} \cdot A^{1/(r+\beta)} \cdot A_3^{(r+\beta+1/2)/(r+\beta)} + 1\Big) + 1,$$

which completes the proof of the theorem. □

Let us discuss assumptions (13) and (14). First, note that (13) implies boundedness of ρ. It also implies integrability of ρ^2 and of $\rho^{1/(r+\beta+1/2)}$ over $\mathbb{R}_+$.

Suppose now that ρ is monotonically decreasing. Then

$$\frac{1}{2} \cdot \sum_{i=2}^{\infty} c_i \le \int_0^{\infty} \rho^{1/(r+\beta+1/2)}(t)\, dt \le \sum_{i=1}^{\infty} c_i,$$

so that (13) is equivalent to integrability of $\rho^{1/(r+\beta+1/2)}$ over $\mathbb{R}_+$. Furthermore, in this case, (13) implies (12). Indeed, if $\int_0^\infty \rho^{1/(r+\beta+1/2)}(t)\, dt = c < \infty$ then $\rho^{1/(r+\beta+1/2)}(t) \le c/t$, so that

$$\int_0^{\infty} \rho^2(t) \cdot t^{2(r+\beta)}\, dt \le c^{2(r+\beta)} \cdot \int_0^{\infty} \rho^{1/(r+\beta+1/2)}\, dt < \infty.$$

Thus we already get finite complexity (7) from (13) by Lemma 3.

Verifying (14) may be more complicated. The following simple observation can ease this task in some cases. Suppose

$$L(1) < \infty \qquad \text{and} \qquad \forall\, x, y \geq 1: \ \rho(xy) \leq A_5 \cdot \rho(x) \cdot \rho(y). \tag{15}$$

Then (14) holds with $A_4 = L(1) \cdot A_5$ as follows from

$$L^2(R) = R^{2r+2\beta+1} \int_1^\infty \rho^2(x \cdot R) \cdot x^{2(r+\beta)}\, dx \leq (A_5 \cdot L(1))^2 \cdot \rho^2(R) \cdot R^{2r+2\beta+1}.$$

We now illustrate assumptions (13) and (14) by the following two examples.

Example 3. Consider the weight

$$\rho(t) = (t+1)^{-\alpha}.$$

Then $\rho_i = 2^{-(i-1)\,\alpha}$ and

$$A_3 = \sum_{i=0}^\infty 2^{i\,(1-\alpha/(r+\beta+1/2))}.$$

Thus (13) holds iff

$$\alpha > r + \beta + \tfrac{1}{2}, \tag{16}$$

and in this case

$$A_3 = \frac{1}{1 - 2^{1-\tau}}$$

with $\tau = \alpha / \left(r + \beta + \frac{1}{2}\right) > 1$. Note also that

$$L^2(R) = \int_R^\infty \frac{t^{2(r+\beta)}}{(t+1)^{2\alpha}}\, dt,$$

so that (16) is necessary for finite complexity (7). Finally,

$$L^2(R) \leq \int_R^\infty (t+1)^{-2\,(\alpha-r-\beta)}\, dt = \frac{(R+1)^{-2\,(\alpha-r-\beta)+1}}{2\,(\alpha-r-\beta)-1},$$

so that (16) yields (14) with

$$A_4^2 = \frac{1}{2(\alpha - r - \beta) - 1}.$$

Example 4. Consider the weight function

$$\rho(x) = \exp\left(-\alpha_1 \cdot x^{\alpha_2}\right)$$

for positive α_1, α_2. Of course, $L(0) < \infty$ and (13) holds. Note also that (15) holds with $A_5 = \exp(\alpha_1)$. Hence (14) holds with $A_4 = L(1) \cdot e^{\alpha_2}$.

Remark 1. There are weight functions for which the complexity is finite and (13) holds; however, (14) is not satisfied. In such cases Theorem 2 is not applicable. For instance, consider

$$\rho(t) = (t+1)^{-(r+\beta+1/2)} \cdot \ln^{-\alpha}(t+e). \tag{17}$$

Then (7) is equivalent to $\alpha > \frac{1}{2}$, (13) is equivalent to $\alpha > r + \beta + 1/2$, yet (14) does not hold no matter how large α is.

In Plaskota, Ritter, and Wasilkowski (2000), we develop a different technique that allows to find the complexity for weights like (17).

In many cases (13) is also a *necessary* condition for the complexity to be of the same order as in the compact case.

Theorem 3. *Suppose that the weight function* ρ *is continuous or monotonically decreasing. Then* $\mathrm{comp}(\varepsilon, \mathrm{App}_\rho) = O\left(\varepsilon^{-1/(r+\beta)}\right)$ *implies*

$$\int_0^\infty \rho^{1/(r+\beta+1/2)}(t)\, dt < \infty.$$

Proof. Let $\mathcal{A}_\varepsilon$ be a method such that $e(\mathcal{A}_\varepsilon, \mathrm{App}_\rho) \le \varepsilon$ and $\mathrm{card}(\mathcal{A}_\varepsilon) = O\left(\varepsilon^{-1/(r+\beta)}\right)$. Let $t_1 < \cdots < t_n$ denote the knots used by $\mathcal{A}_\varepsilon$, where $n = \mathrm{card}(\mathcal{A}_\varepsilon)$. Put $R = t_n$ and $t_0 = 0$ as well as

$$\widetilde{\rho}_i = \inf\{\rho(t) : t \in [t_{i-1}, t_i]\}.$$

Using (5) we obtain

$$\begin{aligned}
\varepsilon^2 &\ge C_2^2 \cdot \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \rho^2(t) \cdot \left(\frac{(t_i - t)\cdot(t - t_{i-1})}{t_i - t_{i-1}} \right)^{2(r+\beta)} dt \\
&\ge c \cdot \sum_{i=1}^n \widetilde{\rho}_i^{\,2} \cdot (t_i - t_{i-1})^{2(r+\beta)+1} \\
&\ge c \cdot n^{-2(r+\beta)} \cdot \left(\sum_{i=1}^n \widetilde{\rho}_i^{\,1/(r+\beta+1/2)} \cdot (t_i - t_{i-1}) \right)^{2(r+\beta)+1}
\end{aligned}$$

with a constant $c > 0$ that depends only on C_2, r, and β. Thus

$$\sum_{i=1}^n \widetilde{\rho}_i^{\,1/(r+\beta+1/2)} \cdot (t_i - t_{i-1})$$

is uniformly bounded in ε and the corresponding step functions converge to $\rho^{1/(r+\beta+1/2)}$ at every point of continuity of ρ, i.e., at least almost everywhere. It remains to apply Fatou's Lemma. □

4. Weighted Integration on $\mathbb{R}_+$

In this section we assume that X satisfies (2) with $r + \beta > 0$ and, for simplicity, that

$$C_1 = 1.$$

For the integration problem to be well defined in the Lebesgue sense, it is necessary and sufficient that

$$\int_0^\infty K^{1/2}(t,t) \cdot \rho(t)\, dt < \infty. \tag{18}$$

Indeed, since X is Gaussian, we have

$$\mathbb{E}\left(\int_0^\infty |X(t)| \cdot \rho(t)\, dt\right) = \int_0^\infty \mathbb{E}(|X(t)|) \cdot \rho(t)\, dt = \sqrt{2/\pi} \cdot \int_0^\infty K^{1/2}(t,t) \cdot \rho(t)\, dt,$$

which, together with (18), implies that the weighted integral $\mathrm{Int}_\rho X$ is well defined for almost every trajectory of X. Conversely, assume that the latter holds true. Then, by Fernique's inequality,

$$\mathbb{E}\left(\int_0^\infty |X(t)| \cdot \rho(t)\, dt\right) < \infty,$$

which implies (18). We add that

$$\int_0^\infty \rho(t) \cdot \max\{1,t\}^{r+\beta}\, dt < \infty \tag{19}$$

is a sufficient condition for (18) to hold, see (10). Moreover, for processes that satisfy (5), the condition (19) is only slightly stronger than (18), since $K(t,t) \geq c \cdot t^{2(r+\beta)}$ for $t \in \mathbb{R}_+$ in this case.

We use the general technique from Wasilkowski (1994) to derive upper bounds for the complexity of the integration problem. In this approach one analyzes suitable randomized (Monte Carlo) methods. By a mean value argument, a Monte Carlo method with average error at most ε yields the existence of a deterministic method with the same error bound and the same number of samples.

4.1. **Preliminary Results.** First, we consider the case of a bounded weight function with compact support. Of course, this includes the classical case of unweighted integration on a compact interval.

Theorem 4. *Let $a < b$ be such that*

$$\mathrm{supp}\,\rho \subseteq [a,b] \qquad \textit{and} \qquad \operatorname*{ess\,sup}_{t\in[a,b]} \rho(t) < \infty.$$

Then

$$\mathrm{comp}(\varepsilon, \mathrm{Int}_\rho) = O\left(\varepsilon^{-1/(r+\beta+1/2)}\right).$$

Proof. Consider the piecewise polynomial interpolation $\mathcal{U}$ from Lemma 1, which uses $1 + n\overline{r_0}$ knots. Define

$$\mathcal{MC}X = \int_a^b \mathcal{U}X(t) \cdot \rho(t)\, dt + \frac{b-a}{1+n\overline{r_0}} \cdot \sum_{j=0}^{n\overline{r_0}} (X - \mathcal{U}X)(t_j) \cdot \rho(t_j), \tag{20}$$

where $t_0, \ldots, t_{n\overline{r_0}}$ are independent and uniformly distributed in $[a,b]$. We use $\mathbb{E}_t$ to denote the expectation with respect to the joint distribution of the points t_j. For every fixed trajectory of X,

$$\mathbb{E}_t(\mathcal{MC}X) = \mathrm{Int}_\rho X$$

and

$$\mathbb{E}_t(\mathrm{Int}_\rho X - \mathcal{MC}X)^2 \leq \frac{b-a}{1+n\overline{r_0}} \cdot \int_a^b (X - \mathcal{U}X)^2(t) \cdot \rho^2(t)\, dt.$$

Therefore

$$\mathbb{E}_t\left(\mathbb{E}(\mathrm{Int}_\rho X - \mathcal{MC}X)^2\right) = \mathbb{E}\left(\mathbb{E}_t(\mathrm{Int}_\rho X - \mathcal{MC}X)^2\right) \le \frac{b-a}{1+n\overline{r_0}} \cdot e^2(\mathcal{U}, \mathrm{App}_\rho).$$

Hence there exists a choice of deterministic points t_i such that the quadrature formula $\mathcal{Q}$ defined by the right-hand side of (20) satisfies

$$e(\mathcal{Q}, \mathrm{Int}_\rho) \le \left(\frac{b-a}{1+n\overline{r_0}}\right)^{1/2} \cdot e(\mathcal{U}, \mathrm{App}_\rho).$$

We apply Lemma 1 to obtain

$$e(\mathcal{Q}, \mathrm{Int}_\rho) \le A_1 \cdot (b-a)^{r+\beta+1} \cdot \operatorname*{ess\,sup}_{t\in[a,b]} \rho(t) \cdot n^{-(r+\beta+1/2)},$$

and the upper bound on the complexity follows, since $\mathrm{card}(\mathcal{Q}) = 2(1+n\overline{r_0}) = O(n)$. □

Under Hölder conditions (2), we are able only to provide upper bounds since without additional restrictions, the complexity of the integration problem could be 1 independently of ε. Indeed, this holds when, e.g., $K(s,t) = g(s)\cdot g(t)$ for a suitable nonzero function g, since one sample determines a trajectory of X precisely.

However, the upper bound from Theorem 4 cannot be improved in general. Indeed, for the processes of Example 2 we have

$$\mathrm{comp}(\varepsilon, \mathrm{Int}_\rho) = \Theta\left(\varepsilon^{-1/(r+\beta+1/2)}\right),$$

and simple constructions of almost optimal quadrature formulas are known. Moreover, $\mathrm{comp}(\varepsilon, \mathrm{Int}_\rho)$ is at least of order $\varepsilon^{-1/(r+\beta+1/2)}$ for those processes and arbitrary weight functions on $\mathbb{R}_+$ that are bounded away from zero on a compact interval of positive length. See Ritter, Wasilkowski, and Woźniakowski (1995), and Ritter (2000, Sec. VI.1.2, VI.1.4).

We also mention that (18) implies finite complexity on compact intervals, i.e.,

$$\forall\ 0 \le a < b < \infty\ \forall\ \varepsilon > 0 :\ \mathrm{comp}(\varepsilon, \mathrm{Int}_{\rho\cdot 1_{[a,b]}}) < \infty,$$

and finite complexity on $\mathbb{R}_+$, i.e.,

$$\forall\ \varepsilon > 0 :\ \mathrm{comp}(\varepsilon, \mathrm{Int}_\rho) < \infty.$$

4.2. **Upper Bounds.** We present two different approaches that yield an upper bound of order $\varepsilon^{-1/(r+\beta+1/2)}$ for $\mathrm{comp}(\varepsilon, \mathrm{Int}_\rho)$.

In the first approach, we apply randomization to compact subintervals of $\mathbb{R}_+$. To this end, let $a_i = 2^i - 1$ and

$$\rho_i = \operatorname*{ess\,sup}_{t\in[a_{i-1},a_i]} \rho(t)$$

as in Section 3.3. Redefine

$$c_i = \rho_i^{1/(r+\beta+1)} \cdot 2^{i-1}$$

and

$$L(R) = \left(\int_R^\infty \int_R^\infty K(s,t)\cdot \rho(s)\,\rho(t)\,ds\,dt\right)^{1/2}.$$

We assume that

(21) $$A_3 := A_3(r,\beta,\rho) := \sum_{i=1}^{\infty} c_i < \infty$$

and

(22) $$L(a_i) \le A_4 \cdot c_{i+1}^{r+\beta+1}$$

with a constant $A_4 = A_4(r,\beta,\rho)$.

Theorem 5. *Suppose that* (21) *and* (22) *hold. Then we have*

$$\mathrm{comp}(\varepsilon, \mathrm{Int}_\rho) \le \overline{r_0} \cdot \Big(3 \cdot \varepsilon^{-1/(r+\beta+1/2)} \cdot A^{1/(r+\beta+1/2)} \cdot A_4^{(r+\beta+1)/(r+\beta+1/2)} + 2\Big) + 1.$$

Proof. We study suitable linear combinations of the Monte Carlo methods from the proof of Theorem 4. Given $\varepsilon > 0$, let

$$G(\varepsilon,\ell) = \left(\frac{A^2}{\varepsilon^2} \cdot \sum_{i=1}^{\ell+1} c_i\right)^{1/(2r+2\beta+1)}$$

with

$$A = \max\{A_1, A_4\}.$$

We take

$$k = k(\varepsilon) = \min\{\ell \in \mathbb{N} \,:\, c_{\ell+1} \cdot G(\varepsilon,\ell) \le 1\}$$

and

$$n_i = n_i(\varepsilon) = \lceil c_i \cdot G(\varepsilon,k) \rceil \,.$$

For every interval $[a,b] = I_i = [a_{i-1}, a_i]$ and $n = n_i$ with $i = 1,\dots,k$ let $\mathcal{U}_i$ be the piecewise polynomial interpolation from Lemma 1 and let $\mathcal{MC}_i$ be defined by the right-hand side of (20). Assuming that all the Monte Carlo points are chosen independently, $\mathcal{MC}_1 X, \dots, \mathcal{MC}_k X$ are independent random variables for every fixed trajectory of X. We define

$$\mathcal{MC}X = \sum_{i=1}^{k} \mathcal{MC}_i X.$$

Then

$$\mathbb{E}_t\,(\mathrm{Int}_\rho X - \mathcal{MC}X)^2 = \left(\int_{a_k}^{\infty} X(t)\cdot\rho(t)\,dt\right)^2 + \sum_{i=1}^{k} \mathbb{E}_t \left(\mathrm{Int}_{\rho\cdot 1_{I_i}} - \mathcal{MC}_i X\right)^2,$$

and hereby

$$\mathbb{E}_t\,(\mathbb{E}(\mathrm{Int}_\rho X - \mathcal{MC}X)^2) \le L^2(a_k) + \sum_{i=1}^{k} \frac{2^{i-1}}{1 - n_i \overline{r_0}} \cdot e^2(\mathcal{U}_i, \mathrm{App}_{\rho\cdot 1_{I_i}}).$$

The rest of the proof is very similar to the proof of Theorem 2, and we omit it. $\square$

Note that (21) is in general a stronger assumption than (13). For instance, $\rho(t) = (t+1)^{-\alpha}$ satisfies (13) iff $\alpha > r + \beta + 1/2$, whereas α must be greater than $r + \beta + 1$ for (21) to hold. Hence, although the complexity of weighted integration is smaller than the complexity of weighted approximation, we need a stronger assumption on the weight for the complexity to be of minimal order.

In a second approach, we apply randomization directly to the half-line $\mathbb{R}_+$.

Theorem 6. *Suppose there exists $\delta \in (0,2)$ such that*

$$\int_0^\infty \rho^\delta(t)\,dt < \infty, \tag{23}$$

and $\rho_\delta = \rho^{1-\delta/2}$ satisfies the assumptions (13) *and* (14), *i.e.,*

$$\sum_{i=1}^\infty 2^{i-1} \cdot \rho_i^{(1-\delta/2)/(r+\beta+1/2)} < \infty \tag{24}$$

and

$$\int_{a_i}^\infty \rho^{2-\delta}(t) \cdot t^{2(r+\beta)}\,dt \le A_4^2 \cdot 2^{i(2r+2\beta+1)} \cdot \rho_{i+1}^{2-\delta}. \tag{25}$$

Then

$$\mathrm{comp}(\varepsilon, \mathrm{Int}_\rho) = O\left(\varepsilon^{-1/(r+\beta+1/2)}\right).$$

Proof. Given ε, let $\mathcal{A}_\varepsilon^*$ be the method from Theorem 2 for the weight function ρ replaced by ρ_δ. Let $n = n(\varepsilon)$ be the cardinality of $\mathcal{A}_\varepsilon^*$. Consider the following randomized method

$$\mathcal{MC}_n X = \int_0^\infty \mathcal{A}_\varepsilon^*(X)(t) \cdot \rho(t)\,dt + \frac{a}{n}\sum_{i=1}^n (X - \mathcal{A}_\varepsilon^*(X))(t_i) \cdot \rho^{1-\delta}(t_i)$$

where the points t_i are chosen independently according to the probability distribution whose density equals ρ^δ/a with $a = \int_0^\infty \rho^\delta(t)\,dt$. Note that the cardinality of $\mathcal{MC}_n$ equals $2n$. It is easy to check that for every trajectory of X,

$$\mathbb{E}_t \left(\mathrm{Int}_\rho(X) - \mathcal{MC}_n(X)\right)^2 \le \frac{1}{n} \cdot \int_0^\infty \rho^\delta(t)\,dt \cdot \int_0^\infty \left((X - \mathcal{A}_{\varepsilon^*}(X))(t) \cdot \rho_\delta(t)\right)^2\,dt,$$

where $\mathbb{E}_t$ denotes the expectation with respect to the points t_i. Finally we use Theorem 2 and proceed as in the proof of Theorem 4. □

We illustrate the assumptions of Theorem 6 for $\rho(t) = (t+1)^{-\alpha}$ as before. For ρ^δ to be integrable, we need $\delta\alpha > 1$. For the other assumptions, we need $\alpha(1-\delta/2) > r+\beta+1/2$. Equivalently, we need $1/\alpha < \delta < 2 - 2(r+\beta+1/2)/\alpha$. This means that such a δ exists iff $\alpha > r+\beta+1$ which is exactly the same condition as the condition for satisfying the assumptions of Theorem 5.

For every monotone function ρ, (21) is equivalent to the existence of $\delta \in (0,2)$ with (23) and (24). Furthermore, (25) is not needed in this case, see Plaskota, Ritter, and Wasilkowski (2000). On the other hand, there exist nonmonotonic weight functions such that Theorem 6 yields the upper bound, while Theorem 5 is not applicable.

4.3. **A Special Case: Sacks-Ylvisaker conditions.** We now discuss a special case. We assume that X satisfies the Sacks-Ylvisaker conditions of order $r \in \mathbb{N}_0$. As shown in Ritter, Wasilkowski, and Woźniakowski (1995), the corresponding reproducing kernel Hilbert space is, essentially, equal to the Sobolev space $W_2^{r+1}(\mathbb{R}_+)$. It is also well known, see, e.g., Traub, Wasilkowski and Woźniakowski (1988), Ritter (2000), that for integration the average case complexity is equal to the worst case complexity with respect to the unit ball in the reproducing kernel Hilbert space. Moreover, (almost) optimal methods in one of the setting are also (almost) optimal in the other. Thus, for X satisfying Sacks-Ylvisaker conditions, the average complexity of the weighted integration reduces to the worst case complexity with respect to the unit ball in $W_2^{r+1}(\mathbb{R}_+)$. The latter problem, among others, was considered in Wasilkowski and Woźniakowski (2000a).

If (21) and (22) hold with $\beta = 1/2$ also in the definition of c_i, then there are constructions of simple methods $\mathcal{Q}_\varepsilon^*$ whose errors do not exceed ε and cardinalities are proportional to $\varepsilon^{-1/(r+1)}$. Hence they are almost optimal since the complexity of the problem also equals

$$\operatorname{comp}(\varepsilon, \operatorname{Int}_\rho) = \Theta\left(\varepsilon^{-1/(r+1)}\right).$$

For specifics concerning these methods see Wasilkowski and Woźniakowski (2000a) and Han and Wasilkowski (2000).

We sketch a possible construction. With the choice of k and n_i from Theorem 5 take

$$\mathcal{Q}_i X = \int_{a_{i-1}}^{a_i} \mathcal{U}_i X(t) \cdot \rho(t)\, dt,$$

where $\mathcal{U}_i$ is the piecewise linear interpolation from Lemma 1 on $[a_{i-1}, a_i]$ with $1 + n_i \overline{r_0}$ knots. Then $\mathcal{Q} = \sum_{i=1}^k \mathcal{Q}_i$ is an almost optimal method. A proof can be based on the following facts. Let $X - \widehat{X}$ denote the Taylor polynomial of degree r at a_{i-1}. Since $\overline{r_0} \geq r$,

$$\int_{a_{i-1}}^{a_i} X(t) \cdot \rho(t)\, dt - \mathcal{Q}_i X = \int_{a_{i-1}}^{a_i} \widehat{X}(t) \cdot \rho(t)\, dt - \mathcal{Q}_i \widehat{X}.$$

Moreover, note that these random variables are independent for $i = 1, \dots k$, if X is the r-fold integrated Brownian motion.

5. Concluding Remarks

We discuss possible improvements to the proposed methods. We will do this only for the approximation problem; however, the same comments pertain to the integration problem. Due to the lower bounds the improvements can only lead to better constants in the estimates for the error or the cardinality.

Remark 2. The method $\mathcal{A}^*$ is based on piecewise polynomial interpolation. Instead one could use error-optimal algorithms. The latter are given by the means of the corresponding conditional process, or, equivalently, by interpolating K-splines. In view of the lower bounds we have decided to work with piecewise polynomials, since they are easy to implement and do not depend on specific type of the process X.

Recall that $\mathcal{A}^*X(t)$ vanishes for $t > a_k$. Alternatively, we could define $\mathcal{A}^*X|_{[a_k,\infty)}$ by extrapolation, using a few values X in a neighbourhood of a_k.

Remark 3. In the definition of $\mathcal{A}^*$, the parameters k and n_i are chosen based on an upper bound on error of an interpolating piecewise polynomial. Specifically, we use the following inequality

$$e_i^2 = \mathbb{E}\left(\int_{a_{i-1}}^{a_i} |X(t) - \mathcal{A}^*X(t)|^2 \cdot \rho^2(t)\, dt\right) \leq \rho_i^2 \cdot \ (a_i - a_{i-1}) \cdot \max_t \mathbb{E}(X(t) - \mathcal{A}^*X(t))^2.$$

This could be improved by using

$$e_i^2 = \mathbb{E}\left(\int_{a_{i-1}}^{a_i} |X(t) - \mathcal{A}^*X(t)|^2 \cdot \rho^2(t)\, dt\right)$$

if the above expectation are easy to compute, or by using

$$e_i^2 \leq \rho_i^2 \cdot \mathbb{E}\left(\int_{a_{i-1}}^{a_i} |X(t) - \mathcal{A}^*X(t)|^2\, dt\right)$$

that in many cases is not difficult to compute. Any such improvement would require a new definition of c_i; the rest of the method would remain unchanged.

Remark 4. The method $\mathcal{A}^*$ uses the values of the suprema ρ_i. This could result in a very high combinatorial cost for a number of weights ρ. Of course, this does not concern monotonic weights ρ since then the numbers ρ_i are given explicitly by $\rho_i = \rho(a_{i-1})$.

Remark 5. The sample points used by $\mathcal{A}^*$ are equally spaced in each subinterval $[a_{i-1}, a_i]$. Instead, one could use the sampling similar to the one proposed in Han and Wasilkowski (2000), a paper that deals with the worst case setting.

Remark 6. Suppose we only know an upper bound σ for the Hölder smoothness $r+\beta$ of X. Then we can also achieve an error of order ε at cost of order $\varepsilon^{-1/(r+\beta)}$ by the following modification of the method $\mathcal{A}^*$. Take piecewise polynomial interpolation of degree $r_0 = \lceil \sigma \rceil$. Redefine

$$c_i = \rho_i^{1/(\sigma+1/2)} \cdot 2^{i-1}$$

and assume that $A_3 := \sum_{i=1}^{\infty} c_i < \infty$. Moreover, assume that

$$\int_{a_i}^{\infty} \rho^2(t) \cdot t^{2\sigma}\, dt \leq A_4^2 \cdot c_{i+1}^{2\sigma+1}$$

with a constant $A_4 > 0$. Finally, redefine

$$G(\varepsilon, \ell) = \left(\frac{1}{\varepsilon^2} \cdot \sum_{i=1}^{\ell+1} c_i\right)^{1/(2\sigma)},$$

and take k and n_i as previously.

Acknowledgments

We are grateful to S. Kwapień for his remarks concerning (18).

The first author was partially supported by the State Committee for Scientific Research of Poland (KBN). The second author was supported in part by the Alexander von Humboldt Foundation. The third author was partially supported by the National Science Foundation under the Grant CCR-9729971.

References

1. F. Curbera, Optimal integration of Lipschitz functions with a Gaussian weight, *Journal of Complexity*, **14** (1998), 122–149.
2. L. Han and G.W. Wasilkowski, A new optimal algorithm for weighted approximation and integration over $\mathbb{R}$, *Numerical Algorithms*, **23** (2000), 393–406.
3. P. Mathé, Asymptotically optimal weighted numerical integration, *Journal of Complexity*, **14** (1998), 34–48.
4. L. Plaskota, K. Ritter, and G.W. Wasilkowski, (2000), *in progress.*
5. K. Ritter, "Average-Case Analysis of Numerical Problems," Lecture Notes in Mathematics **1733**, Springer, Berlin, 2000
6. K. Ritter, G.W. Wasilkowski, and H. Woźniakowski, On multivariate integration of stochastic processes, *International Series of Numerical Mathematics,* **112** (1993), 331–347.
7. K. Ritter, G.W. Wasilkowski, and H. Woźniakowski, Multivariate integration and approximation for random fields satisfying Sacks-Ylvisaker conditions, *Annals of Applied Probability*, **5** (1995), 518–540.
8. P. Singer, An integrated fractional Fourier transform, *J. Comput. Appl. Math.* **54** (1994), 221–237.
9. P. Speckman, L_p approximation of autoregressive Gaussian processes, Tech. Rep., Dept. of Statistics, Univ. of Oregon, Eugene, 1979.
10. J.F. Traub, G.W. Wasilkowski, and H. Woźniakowski, "Information, Uncertainty, Complexity", Addison-Wesley, Reading, 1983.
11. J.F. Traub, G.W. Wasilkowski, and H. Woźniakowski, "Information-Based Complexity," Academic Press, New York, 1988.
12. G.W. Wasilkowski, Information of varying cardinality, *J. Complexity* **2** (1986), 204–228.
13. G.W. Wasilkowski, Integration and approximation of multivariate functions: average case complexity with isotropic Wiener measure, *J. Approx. Theory* **77** (1994), 212–227.
14. G.W. Wasilkowski and H. Woźniakowski, Complexity of weighted approximation over $\mathbb{R}^1$, *J. Approx. Theory* **103** (2000a), 223–251.
15. G.W. Wasilkowski and H. Woźniakowski, Complexity of weighted approximation over $\mathbb{R}^d$, *J. Complexity* (2000b), *to appear.*

Department of Mathematics, Informatics, and Mechanics, Warsaw University, ul. Banacha 2, 02-097 Warsaw, Poland
E-mail address: leszekp@mimuw.edu.pl

Fachbereich Mathematik, Technische Universität Darmstadt, Schlossgartenstr. 7, 64289 Darmstadt, Germany
E-mail address: ritter@mathematik.tu-darmstadt.de

Department of Computer Science, University of Kentucky, 773 Anderson Hall, Lexington, KY 40506-0046, USA
E-mail address: greg@cs.uky.edu