Contemporary Mathematics
http://ojs.wiserpub.com/index.php/CM/

UNIVERSAL WISER PUBLISHER

Research Article

On the Power of Standard Information for Tractability for L_2-Approximation in the Randomized Setting

Wanting Lu, Heping Wang*

School of Mathematical Sciences, Capital Normal University, Beijing, China
E-mail: wanghp@cnu.edu.cn

Received: 2 November 2021; **Revised:** 20 December 2021; **Accepted:** 28 December 2021

Abstract: We study the approximation of multivariate functions from a separable Hilbert space in the randomized setting with the error measured in the weighted L_2 norm. We consider algorithms that use standard information Λ^{std} consisting of function values or general linear information Λ^{all} consisting of arbitrary linear functionals. We investigate the equivalences of various notions of algebraic and exponential tractability in the randomized setting for Λ^{std} and Λ^{all} for the normalized or absolute error criterion. For the normalized error criterion, we show that the power of Λ^{std} is the same as that of Λ^{all} for all notions of exponential tractability and some notions of algebraic tractability without any condition. For the absolute error criterion, we show that the power of Λ^{std} is the same as that of Λ^{all} for all notions of algebraic and exponential tractability without any condition. Specifically, we solve Open Problems 98, 101, 102 and almost solve Open Problem 100 as posed by E.Novak and H.Woźniakowski in the book: Tractability of Multivariate Problems, Volume III: Standard Information for Operators, EMS Tracts in Mathematics, Zürich, 2012.

***Keywords*:** tractability, standard information, general linear information, randomized setting

MSC: 41A63, 65C05, 65D15, 65Y20

1. Introduction

We study multivariate approximation $\text{APP} = \{\text{APP}_d\}_{d\in\mathbb{N}}$, where

$$\text{APP}_d : F_d \to G_d \quad \text{with} \quad \text{APP}_d(f) = f$$

is the compact embedding operator, F_d is a separable Hilbert function space on D_d, G_d is a weighted L_2 space on D_d, $D_d \subset \mathbb{R}^d$, and the dimension d is large or even huge. We consider algorithms that use finitely many information evaluations. Here information evaluation means linear functional on F_d (general linear information) or function value at some point (standard information). We use Λ^{all} and Λ^{std} to denote the extended class of all linear functionals (not necessarily continuous) and the extended class of all function values (defined only almost everywhere), respectively.

For a given error threshold $\varepsilon \in (0, 1)$, the information complexity $n(\varepsilon, d)$ is defined to be the minimal number of information evaluations for which the approximation error of some algorithm is at most ε. Tractability is aimed

DOI: https://doi.org/10.37256/cm.3120221229

at studying how the information complexity $n(\varepsilon, d)$ depends on ε and d. There are two kinds of tractability based on polynomial convergence and exponential convergence. The Algebraic tractability (ALG-tractability) describes how the information complexity $n(\varepsilon, d)$ behaves as a function of d and ε^{-1}, while the Exponential tractability (EXP-tractability) does as one of d and $1 + \ln\varepsilon^{-1}$. The existing notions of tractability mainly include Strong Polynomial Tractability (SPT), Polynomial Tractability (PT), Quasi-Polynomial Tractability (QPT), Weak Tractability (WT), (s, t)-Weak Tractability ((s, t)-WT), and Uniform Weak Tractability (UWT). In recent years the study of algebraic and exponential tractability has attracted much interest, and a great number of interesting results have been obtained (see [1-11] and the references therein). This paper is devoted to investigating the equivalences of various notions of algebraic and exponential tractability for Λ^{std} and Λ^{all} in the randomized setting, see Chapter 22 in [8]. The class Λ^{std} is much smaller and much more practical and is much more difficult to analyze than the class Λ^{all}. Hence, it is very important to study the power of Λ^{std} compared to Λ^{all}. There are many papers devoted to this field. For example, for the randomized setting, see [8, 12-16]; for the average case setting, see [8, 17-19]; for the worst case setting see [8, 20-26]. The authors in [8, 16] obtained the equivalences of ALG-SPT, ALG-PT, ALG-QPT, ALG-WT in the randomized setting for Λ^{std} and Λ^{all} for the normalized error criterion without any condition. Meanwhile, for the absolute error criterion under some conditions, the equivalences of ALG-SPT, ALGPT, ALG-QPT, ALG-WT were also obtained in [8].

In this paper, we obtain the remaining equivalences of all notions of algebraic and exponential tractability in the randomized setting for Λ^{std} and Λ^{all} for the normalized or absolute error criterion without any condition. Our results particularly imply that for the absolute error criterion the imposed conditions are not necessary. This solves Open Problems 98, 101, 102 as posed by Novak and Woźniakowski in [8]. We also give an almost complete solution to Open Problem 100 in [8].

This paper is organized as follows. Section 2 contains 5 subsections. In Subsections 2.1 and 2.2 we introduce the approximation problem in the worst case and randomized settings. The various notions of algebraic and exponential tractability are given in Subsection 2.3. Subsection 2.4 is devoted to giving the equivalences of tractability for Λ^{all} for the absolute or normalized error criterion in the worst case and randomized settings. Our main results, Theorems 2.2, 2.3, 2.5, and 2.6, are stated in Subsection 2.5. In Section 3, we give the proofs of Theorems 2.2 and 2.3. After that, in Section 4, we establish the equivalence results for the notions of algebraic tractability. The equivalence results for the notions of exponential tractability are proved in Section 5.

2. Preliminaries and main results

2.1 *Deterministic worst case setting*

For $d \in \mathbb{N}$, let F_d be a separable Hilbert space of d-variate functions defined on $D_d \subset \mathbb{R}^d$, $G_d = L_2(D_d, \rho_d(x)dx)$ be a weighted L_2 space, where D_d is a Borel measurable subset of $\mathbb{R}^d$ with positive Lebesgue measure, ρ_d is a probability density function on D_d. We consider the multivariate approximation problem $\text{APP} = \{\text{APP}_d\}_{d\in\mathbb{N}}$ in the deterministic worst case setting which is defined via the compact embedding operator

$$\text{APP}_d : F_d \to G_d \quad \text{with} \quad \text{APP}_d(f) = f. \tag{1}$$

We approximate $\text{APP}_d(f)$ by algorithms $A_{n,d}(f)$ of the form

$$A_{n,d}(f) = \phi_{n,d}(L_1(f), L_2(f), \ldots, L_n(f)), \tag{2}$$

where L_1, L_2, ..., L_n are general linear functionals on F_d, $\phi_{n,d} : \mathbb{R}^n \to G_d$ is an arbitrary measurable mapping, and the number n may be adaptively depend on the input. The worst case approximation error for the algorithm $A_{n,d}$ of the form (2) is defined as

$$e^{\text{wor}}(A_{n,d}) = \sup_{f\in F_d,\, \|f\|_{F_d}\le 1} \|\text{APP}_d(f) - A_{n,d}(f)\|_{G_d}.$$

The nth minimal worst case error is defined by

$$e^{\mathrm{wor}}(n,d;\Lambda^{\mathrm{all}}) = \inf_{A_{n,d} \text{ with } L_i \in \Lambda^{\mathrm{all}}} e^{\mathrm{wor}}(A_{n,d}),$$

where the infimum is taken over all algorithms of the form (2).

For $n = 0$, we use $A_{0,d} = 0$. We obtain the so-called initial error $e^{\mathrm{wor}}(0, d; \Lambda^{\mathrm{all}})$, defined by

$$e^{\mathrm{wor}}(0,d;\Lambda^{\mathrm{all}}) = \sup_{f \in F_d,\, \|f\|_{F_d} \le 1} \| \mathrm{APP}_d(f) \|_{G_d}.$$

From [6, 27] we know that $e^{\mathrm{wor}}(n, d; \Lambda^{\mathrm{all}})$ depends on the eigenpairs $\left\{(\lambda_{k,d}, e_{k,d})\right\}_{k=1}^{\infty}$ of the operator

$$W_d = \mathrm{APP}_d^* \mathrm{APP}_d : F_d \to F_d,$$

where APP_d is given by (1), APP_d^* is the adjoint operator of APP_d, and

$$\lambda_{1,d} \ge \lambda_{2,d} \ge \ldots \lambda_{n,d} \ldots \ge 0.$$

That is, $\{e_{j,d}\}_{j \in \mathbb{N}}$ is an orthonormal basis in F_d, and

$$W_d e_{j,d} = \lambda_{j,d} e_{j,d}.$$

From [6] we get that the nth minimal worst case error is

$$e^{\mathrm{wor}}(n,d;\Lambda^{\mathrm{all}}) = (\lambda_{n+1,d})^{1/2},$$

and it is achieved by the optimal algorithm

$$S_{n,d}^*(f) = \sum_{k=1}^{n} \langle f, e_{k,d} \rangle_{F_d} e_{k,d},$$

that is,

$$e^{\mathrm{wor}}(n,d;\Lambda^{\mathrm{all}}) = \sup_{f \in F_d,\, \|f\|_{F_d} \le 1} \| f - S_{n,d}^*(f) \|_{G_d} = (\lambda_{n+1,d})^{1/2}. \tag{3}$$

Without loss of generality, we may assume that all the eigenvalues are positive. We set

$$\eta_{k,d} = \lambda_{k,d}^{-1/2} e_{k,d}, \ \ k \in \mathbb{N}.$$

We remark that $\{e_{k,d}\}$ is an orthonormal basis in F_d, $\{\eta_{k,d}\}$ is an orthonormal system in G_d, and for $f \in F_d$,

$$\langle f, e_{k,d} \rangle_{F_d} e_{k,d} = \langle f, \eta_{k,d} \rangle_{G_d} \eta_{k,d},$$

and

$$S^*_{n,d}(f) = \sum_{k=1}^{n} \langle f, \eta_{k,d} \rangle_{G_d} \eta_{k,d}. \tag{4}$$

2.2 *Randomized setting*

In the randomized setting, we consider randomized algorithms $A^{\omega}_{n,d}$ of the form

$$A^{\omega}_{n,d}(f) = \phi_{n,d,\omega}(L_{1,\omega}(f), \cdots, L_{n,\omega}(f)), L_{j,\omega} \in \Lambda,\ 1 \le j \le n, \tag{5}$$

where $\Lambda \in \{\Lambda^{\text{all}}, \Lambda^{\text{std}}\}$, $\phi_{n,d,\omega}$ and $L_{j,\omega}$ could be randomly selected according to some probability space $(\Omega, \Sigma, \mathcal{P})$, for any fixed $\omega \in \Omega$, $A^{\omega}_{n,d}$ is a deterministic method with cardinality $n = n(f, \omega)$, the number $n = n(f, \omega)$ may be randomized and adaptively depend on the input, and the cardinality of $A^{\omega}_{n,d}$ is then defined by

$$\text{Card}(A^{\omega}_{n,d}) = \sup_{f \in F_d, \|f\|_{F_d} \le 1} \mathbb{E}_{\omega} n(f, \omega).$$

The randomized approximation error for the algorithm $A^{\omega}_{n,d}$ of the form (5) is defined as

$$e^{\text{ran}}(A^{\omega}_{n,d}) = \sup_{f \in F_d, \|f\|_{F_d} \le 1} \left(\mathbb{E}_{\omega} \left\| \text{APP}_d(f) - A^{\omega}_{n,d}(f) \right\|^2_{G_d} \right)^{1/2}.$$

The nth minimal randomized error for $\Lambda \in \{\Lambda^{\text{all}}, \Lambda^{\text{std}}\}$ is defined by

$$e^{\text{ran}}(n, d; \Lambda) = \inf_{A^{\omega}_{n,d} \text{ with } L_{i,\omega} \in \Lambda} e^{\text{ran}}(A^{\omega}_{n,d}),$$

where the infimum is taken over all randomized algorithms $A^{\omega}_{n,d}$ of the form (5) with $\text{Card}(A^{\omega}_{n,d}) \le n$.

For $n = 0$, we use $A^{\omega}_{0,d} = 0$. We have

$$e^{\text{ran}}(0, d; \Lambda) = e^{\text{wor}}(0, d; \Lambda^{\text{all}}) = (\lambda_{1,d})^{1/2}.$$

There are many papers devoted to studying randomized approximation and relations of $e^{\text{ran}}(n, d; \Lambda)$ and $e^{\text{wor}}(n, d; \Lambda)$ for $\Lambda \in \{\Lambda^{\text{all}}, \Lambda^{\text{std}}\}$ (see [6, 8, 14-16, 28-33]).

This paper is devoted to discussing the equivalence of tractability for Λ^{all} and Λ^{std} in the randomized settings. For Λ^{std} the authors in [8, 14, 16] used simplified randomized algorithms of the form

$$A_{n,\vec{t}}(f) = \sum_{j=1}^{n} f(t_j) g_{j,\vec{t}}, \tag{6}$$

where $\vec{t} = [t_1, ..., t_n]$ for some random points $t_1, ..., t_n$ from D_d, which are independent, and each t_j is distributed according to some probability. The functions $g_{j,\vec{t}} \in G_d$ may depend on the selected points t_j's but are independent of f. For any f, we view $A_{n,\cdot}(f)$ as a random process, and $A_{n,\vec{t}}(f)$ as its specific realization.

We stress that algorithms of the form (6) belong to a restricted class of all randomized algorithms, which are called randomized linear algorithms. Indeed, we assume that n is not randomized, and for a fixed $\vec{t}$ we consider only linear algorithms in $f(t_j)$. In this paper, we also consider algorithms of the form (6). However, in [8, 14, 16] t_j's are assumed to be independent, while in this paper we only assume that $\vec{t}$ is distributed according to some probability measure, and do not assume that t_j's are independent.

The information complexity can be studied using either the Absolute error criterion (ABS) or the Normalized error criterion (NOR). For $\Diamond \in \{\text{wor}, \text{ran}\}$, $\star \in \{\text{ABS}, \text{NOR}\}$, and $\Lambda \in \{\Lambda^{\text{all}}, \Lambda^{\text{std}}\}$, we define the information complexity

$n^{\diamond,\star}(\varepsilon, d; \Lambda)$ as

$$n^{\diamond,\star}(\varepsilon,d;\Lambda) = \inf\{n \,|\, e^{\diamond}(n,d;\Lambda) \le \varepsilon \mathrm{CRI}_d\}, \tag{7}$$

where

$$\mathrm{CRI}_d = \begin{cases} 1, & \text{for } \star = \mathrm{ABS}, \\ e^{\diamond}(0,d,\Lambda), & \text{for } \star = \mathrm{NOR} \end{cases} = \begin{cases} 1, & \text{for } \star = \mathrm{ABS}, \\ (\lambda_{1,\mathrm{d}})^{1/2}, & \text{for } \star = \mathrm{NOR}. \end{cases}$$

We remark that

$$e^{\mathrm{wor}}(0,d,\Lambda^{\mathrm{all}}) = e^{\mathrm{ran}}(0,d,\Lambda^{\mathrm{all}}) = e^{\mathrm{ran}}(0,d,\Lambda^{\mathrm{std}}) = (\lambda_{1,d})^{1/2}.$$

Since $\Lambda^{\mathrm{std}} \subset \Lambda^{\mathrm{all}}$, we get

$$e^{\mathrm{ran}}(n,d;\Lambda^{\mathrm{all}}) \le e^{\mathrm{ran}}(n,d;\Lambda^{\mathrm{std}}).$$

It follows that for $\star \in \{\mathrm{ABS}, \mathrm{NOR}\}$,

$$n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{all}}) \le n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{std}}). \tag{8}$$

2.3 *Notions of tractability*

In this subsection, we briefly recall the various tractability notions. First, we introduce all notions of algebraic tractability. Let $\mathrm{APP} = \{\mathrm{APP}_d\}_{d\in\mathbb{N}}$, $\diamond \in \{\mathrm{wor}, \mathrm{ran}\}$, $\star \in \{\mathrm{ABS}, \mathrm{NOR}\}$, and $\Lambda \in \{\Lambda^{\mathrm{all}}, \Lambda^{\mathrm{std}}\}$. In the $\diamond$ setting for the class Λ, and for error criterion $\star$, we say that APP is

• Algebraically Strongly Polynomially Tractable (ALG-SPT) if there exist $C > 0$ and a non-negative number p such that

$$n^{\diamond,\star}(\varepsilon,d;\Lambda) \le C\varepsilon^{-p}, \text{ for all } \varepsilon \in (0,1). \tag{9}$$

The exponent ALG-$p^{\diamond,\star}(\Lambda)$ of ALG-SPT is defined as the infimum of p for which (9) holds;

• Algebraically Polynomially Tractable (ALG-PT) if there exist $C > 0$ and nonnegative numbers p, q such that

$$n^{\diamond,\star}(\varepsilon,d;\Lambda) \le Cd^{q}\varepsilon^{-p}, \text{ for all } d \in \mathbb{N}, \varepsilon \in (0,1);$$

• Algebraically Quasi-Polynomially Tractable (ALG-QPT) if there exist $C > 0$ and a non-negative number t such that

$$n^{\diamond,\star}(\varepsilon,d;\Lambda) \le C\exp(t(1+\ln d)(1+\ln \varepsilon^{-1})), \text{ for all } d \in \mathbb{N}, \ \varepsilon \in (0,1). \tag{10}$$

The exponent ALG-$t^{\diamond,\star}(\Lambda)$ of ALG-QPT is defined as the infimum of t for which (10) holds;

• Algebraically Uniformly Weakly Tractable (ALG-UWT) if

$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln n^{\diamond,\star}(\varepsilon,d;\Lambda)}{\varepsilon^{-\alpha}+d^{\beta}} = 0, \text{ for all } \alpha,\beta > 0;$$

• Algebraically Weakly Tractable (ALG-WT) if

$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln n^{\Diamond,\star}(\varepsilon,d;\Lambda)}{\varepsilon^{-1}+d} = 0;$$

• Algebraically (s, t)-Weakly Tractable (ALG-(s, t)-WT) for fixed s, $t > 0$ if

$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln n^{\Diamond,\star}(\varepsilon,d;\Lambda)}{\varepsilon^{-s}+d^{t}} = 0.$$

Clearly, ALG-(1, 1)-WT is the same as ALG-WT. If APP is not ALG-WT, then APP is called intractable.

If the nth minimal error is exponentially convergent, then we should study tractability with ε^{-1} being replaced by $1 + \ln\varepsilon^{-1}$, which is called exponential tractability. Recently, there have been many papers studying exponential tractability (see [1, 3-5, 9, 34]).

In the definitions of ALG-SPT, ALG-PT, ALG-QPT, ALG-UWT, ALG-WT, and ALG-(s, t)-WT, if we replace ε^{-1} by $1 + \ln\varepsilon^{-1}$, we get the definitions of Exponential Strong Polynomial Tractability (EXP-SPT), Exponential Polynomial Tractability (EXP-PT), Exponential Quasi-Polynomial Tractability (EXP-QPT), Exponential Uniform Weak Tractability (EXP-UWT), Exponential Weak Tractability (EXP-WT), and Exponential (s, t)-Weak Tractability (EXP-(s, t)-WT), respectively. We now give the above notions of exponential tractability in detail.

Let APP = $\{\mathrm{APP}_d\}_{d\in\mathbb{N}}$, $\Diamond \in \{\mathrm{wor}, \mathrm{ran}\}$, $\star \in \{\mathrm{ABS}, \mathrm{NOR}\}$, and $\Lambda \in \{\Lambda^{\mathrm{all}}, \Lambda^{\mathrm{std}}\}$. In the $\Diamond$ setting for the class Λ, and for error criterion $\star$, we say that APP is

• Exponentially Strongly Polynomially Tractable (EXP-SPT) if there exist $C > 0$ and a non-negative number p such that

$$n^{\Diamond,\star}(\varepsilon,d;\Lambda) \le C(\ln\varepsilon^{-1}+1)^{p}, \text{ for all } \varepsilon\in(0,1). \tag{11}$$

The exponent EXP-$p^{\Diamond,\star}(\Lambda)$ of EXP-SPT is defined as the infimum of p for which (11) holds;

• Exponentially Polynomially Tractable (EXP-PT) if there exist $C > 0$ and nonnegative numbers p, q such that

$$n^{\Diamond,\star}(\varepsilon,d;\Lambda) \le Cd^{q}(\ln\varepsilon^{-1}+1)^{p}, \text{ for all } d\in\mathbb{N}, \varepsilon\in(0,1);$$

• Exponentially Quasi-Polynomially Tractable (EXP-QPT) if there exist $C > 0$ and a non-negative number t such that

$$n^{\Diamond,\star}(\varepsilon,d;\Lambda) \le C\exp\Big(t(1+\ln d)\big(1+\ln(\ln\varepsilon^{-1}+1)\big)\Big), \text{ for all } d\in\mathbb{N},\ \varepsilon\in(0,1). \tag{12}$$

The exponent EXP-$t^{\Diamond,\star}(\Lambda)$ of EXP-QPT is defined as the infimum of t for which (12) holds;

• Exponentially Uniformly Weakly Tractable (EXP-UWT) if

$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln n^{\Diamond,\star}(\varepsilon,d;\Lambda)}{(1+\ln\varepsilon^{-1})^{\alpha}+d^{\beta}} = 0, \text{ for all } \alpha,\beta>0;$$

• Exponentially Weakly Tractable (EXP-WT) if

$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln n^{\Diamond,\star}(\varepsilon,d;\Lambda)}{1+\ln\varepsilon^{-1}+d} = 0;$$

• Exponentially (s, t)-Weakly Tractable (EXP-(s, t)-WT) for fixed s, $t > 0$ if

$$\lim_{\varepsilon^{-1}+d\to\infty}\frac{\ln n^{\Diamond,\star}(\varepsilon,d;\Lambda)}{(1+\ln\varepsilon^{-1})^{s}+d^{t}}=0.$$

2.4 *Equivalences of tractability for Λ^{all} in the worst case and randomized settings*

In this subsection, we introduce the equivalences of tractability for Λ^{all} in the worst case and randomized settings. It follows from [6] that

$$\frac{1}{2}e^{\mathrm{wor}}(4n-1,d;\Lambda^{\mathrm{all}})\le e^{\mathrm{ran}}(n,d;\Lambda^{\mathrm{all}})\le e^{\mathrm{wor}}(n,d;\Lambda^{\mathrm{all}}),$$

which means that for $\star\in\{\mathrm{ABS},\mathrm{NOR}\}$ and $n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{all}})\ge 1$,

$$\frac{1}{4}\left(n^{\mathrm{wor},\star}(2\varepsilon,d;\Lambda^{\mathrm{all}})+1\right)\le n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{all}})\le n^{\mathrm{wor},\star}(\varepsilon,d;\Lambda^{\mathrm{all}}).\tag{13}$$

From (13) we get the equivalences of tractability for Λ^{all} in the worst case and randomized settings. Indeed, for the absolute or normalized error criterion, Corollaries 22.1 and 22.2 in [8] show the equivalences of ALG-SPT, ALG-PT, ALG-QPT, ALG-WT for Λ^{all} in the worst case and randomized settings, and that the exponents of ALG-SPT and ALG-QPT in the worst case and randomized settings are also the same.

Using (13) and the same method as in the proofs of Corollaries 22.1 and 22.2 in [8], for the absolute or normalized error criterion we obtain the equivalences of ALG-UWT, ALG-(s, t)-WT, EXP-SPT, EXP-PT, EXP-QPT, EXP-WT, EXP-UWT, EXP-(s, t)-WT, for Λ^{all} in the worst case and randomized settings, and that the exponents of EXP-SPT and EXP-QPT in the worst case and randomized settings are also the same.

We remark that in showing EXP-$t^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}})$ = EXP-$t^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}})$, we use the following inequalities (see page 43 in [8]): for $\delta\in(0,1)$ and $n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{all}})\ge 1$,

$$\delta^{2}\left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{1-\delta},d;\Lambda^{\mathrm{all}}\right)+1\right)\le n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{all}})\le n^{\mathrm{wor},\star}(\varepsilon,d;\Lambda^{\mathrm{all}}),$$

instead of (13). See the proof of Theorem 5.4.

We summarize these properties in the next corollary.

Corollary 2.1 Consider the approximation problem $\mathrm{APP}=\{\mathrm{APP}_d\}_{d\in\mathbb{N}}$ for the absolute or normalized error criterion in the randomized and worst case settings for Λ^{all}. Then

• ALG-SPT, ALG-PT, ALG-QPT, ALG-UWT, ALG-WT, ALG-(s, t)-WT in the randomized setting is equivalent to ALG-SPT, ALG-PT, ALG-QPT, ALG-UWT, ALG-WT, ALG-(s, t)-WT in the worst case setting;

• EXP-SPT, EXP-PT, EXP-QPT, EXP-UWT, EXP-WT, EXP-(s, t)-WT in the randomized setting is equivalent to EXP-SPT, EXP-PT, EXP-QPT, EXP-UWT, EXP-WT, EXP-(s, t)-WT in the worst case setting;

• the exponents of SPT and QPT are the same in the two settings, i.e., for $\star\in\{\mathrm{ABS},\mathrm{NOR}\}$,

$$\mathrm{ALG}\text{-}p^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}})=\mathrm{ALG}\text{-}p^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}}),$$

$$\mathrm{ALG}\text{-}t^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}})=\mathrm{ALG}\text{-}t^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}}),$$

$$\mathrm{EXP}\text{-}p^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}})=\mathrm{EXP}\text{-}p^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}}),$$

$$\mathrm{EXP}\text{-}t^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}})=\mathrm{EXP}\text{-}t^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}}).$$

2.5 *Main results*

We shall give the main results of this paper in this subsection. The first important progress about the relation between $e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}})$ and $e^{\mathrm{wor}}(n, d; \Lambda^{\mathrm{all}})$ was obtained by Wasilkowski and Woźniakowski in [16] by constructing iterated Monte Carlo methods. They showed that the powers of $e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}})$ and $e^{\mathrm{wor}}(n, d; \Lambda^{\mathrm{all}})$ are the same, and obtained the equivalences of ALG-SPT and ALG-PT for Λ^{all} and Λ^{std} for the normalized error criterion in the randomized setting. Novak and Woźniakowski in [8] and Krieg in [14] refined the above randomized algorithms and showed that $e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}})$ is asymptotically of the same order as $e^{\mathrm{wor}}(n, d; \Lambda^{\mathrm{all}})$ given that $e^{\mathrm{wor}}(n, d; \Lambda^{\mathrm{all}})$ is regularly decreasing. However, the obtained relations are heavily dependent of the initial error, and are not sharp if $e^{\mathrm{wor}}(n, d; \Lambda^{\mathrm{all}})$ is exponentially convergent.

If nodes $\mathrm{X} = (x^1, \ldots, x^n) \in D_d^n$ are drawn independently and identically distributed according to a probability measure, then the samples on the nodes X is called the random information (see [35-37]). Using random information and the least squares method we can obtain the relation between $e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}})$ and $e^{\mathrm{wor}}(n, d; \Lambda^{\mathrm{all}})$ (see [12, 13]). The authors in [38] used random information satisfying some conditions and the least squares method to obtain an inequality between $e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}})$ and $e^{\mathrm{wor}}(n, d; \Lambda^{\mathrm{all}})$ (see Theorem 6.1 in [38]). They remarked in Remark 6.3 in [38] that using the weighed least squares method can improve the above inequality.

In this paper we use the method proposed in Remark 6.3 in [38], i.e., combining the proof of Theorem 6.1 in [38] with the weighted least squares method used in [13], to get an improved inequality between $e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}})$ and $e^{\mathrm{wor}}(n, d; \Lambda^{\mathrm{all}})$. See the following theorem. Compared with the results in [8, 14], our inequality does not depend on the initial error, and are almost sharp if $e^{\mathrm{wor}}(n, d; \Lambda^{\mathrm{all}})$ is exponentially convergent. However, if $e^{\mathrm{wor}}(n, d; \Lambda^{\mathrm{all}})$ is regularly decreasing, then by our inequality we can only obtain that $e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}})$ is at most asymptotically of the order of $e^{\mathrm{wor}}(m, d; \Lambda^{\mathrm{all}})$, where n is at least of order $m \ln m$.

Theorem 2.2 Let $\delta \in (0, 1)$, $m, n \in \mathbb{N}$ be such that

$$m = \left\lfloor \frac{n}{48(\sqrt{2}\ln(2n) - \ln\delta)} \right\rfloor.$$

Then we have

$$e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}}) \le \left(1 + \frac{4m}{n}\right)^{\frac{1}{2}} \frac{1}{\sqrt{1-\delta}} e^{\mathrm{wor}}(m, d; \Lambda^{\mathrm{all}}), \tag{14}$$

where $\lfloor x \rfloor$ denotes the largest integer not exceeding x.

Based on Theorem 2.2, we obtain two relations between the information complexities $n^{\mathrm{ran},\star}(\varepsilon, d; \Lambda^{\mathrm{std}})$ and $n^{\mathrm{wor},\star}(\varepsilon, d; \Lambda^{\mathrm{all}})$ for $\star \in \{\mathrm{ABS}, \mathrm{NOR}\}$.

Theorem 2.3 For $\star \in \{\mathrm{ABS}, \mathrm{NOR}\}$, we have

$$n^{\mathrm{ran},\star}(\varepsilon, d; \Lambda^{\mathrm{std}}) \le 96\sqrt{2}\left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{4}, d; \Lambda^{\mathrm{all}}\right) + 1\right)\left(\ln\left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{4}, d; \Lambda^{\mathrm{all}}\right) + 1\right) + \ln(192\sqrt{2})\right). \tag{15}$$

Furthermore, for $0 < \delta < e^{-2}$, we have

$$n^{\mathrm{ran},\star}(\varepsilon, d; \Lambda^{\mathrm{std}}) \le 48\left(4\left(\ln 48 + \ln\ln\frac{1}{\delta} + \ln\left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{A_\delta}, d; \Lambda^{\mathrm{all}}\right) + 1\right)\right) + \ln\frac{1}{\delta}\right)$$

$$\left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{A_\delta}, d; \Lambda^{\mathrm{all}}\right) + 1\right), \tag{16}$$

where

$$A_\delta := \left(1+\frac{1}{12\ln\frac{1}{\delta}}\right)^{\frac{1}{2}}\frac{1}{\sqrt{1-\delta}}.$$

It is easy to see that for any $\omega > 0$, $0 < \delta < e^{-2}$,

$$\sup_{x\geq 1}\frac{96\sqrt{2}(\ln x+\ln(192\sqrt{2}))}{x^\omega} = C_\omega < +\infty, \tag{17}$$

and

$$\sup_{x\geq 1}\frac{48\left(4\left(\ln 48+\ln\ln\frac{1}{\delta}+\ln x\right)+\ln\frac{1}{\delta}\right)}{x^\omega} = C_{\omega,\delta} < +\infty. \tag{18}$$

According to (15)-(18), we have the following corollary which gives two useful inequalities between the information complexities $n^{\mathrm{ran},\star}(\varepsilon, d; \Lambda^{\mathrm{std}})$ and $n^{\mathrm{wor},\star}(\varepsilon, d; \Lambda^{\mathrm{all}})$ for $\star \in \{\mathrm{ABS}, \mathrm{NOR}\}$.

Corollary 2.4 For $\star \in \{\mathrm{ABS}, \mathrm{NOR}\}$ and $\omega > 0$, we have

$$n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{std}}) \leq C_\omega\left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{4},d;\Lambda^{\mathrm{all}}\right)+1\right)^{1+\omega}. \tag{19}$$

Similarly, for $\omega > 0$, $0 < \delta < e^{-2}$ and $\star \in \{\mathrm{ABS}, \mathrm{NOR}\}$, we have

$$n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{std}}) \leq C_{\omega,\delta}\left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{A_\delta},d;\Lambda^{\mathrm{all}}\right)+1\right)^{1+\omega}, \tag{20}$$

where

$$A_\delta := \left(1+\frac{1}{12\ln\frac{1}{\delta}}\right)^{\frac{1}{2}}\frac{1}{\sqrt{1-\delta}}.$$

In the randomized setting, for the normalized error criterion, Theorems 22.19, 22.21, and 22.5 in [8] give the equivalences of ALG-PT (ALG-SPT), ALG-QPT, ALG-WT for Λ^{all} and Λ^{std}, and shows that the exponents of ALG-SPT and ALG-QPT for Λ^{all} and Λ^{std} are the same. For the absolute error criterion, Theorems 22.20, 22.22, and 22.6 in [8] give the equivalences of ALG-PT (ALG-SPT), ALG-QPT, ALG-WT for Λ^{all} and Λ^{std} under some conditions on the initial error $\sqrt{\lambda_{1,d}}$. Novak and Woźniakowski posed Open problems 98, 101, 102 in [8] which ask whether the above conditions are necessary.

In this paper, we obtain the equivalences of ALG-SPT, ALG-PT, ALG-QPT, ALG-WT for Λ^{all} and Λ^{std} for the absolute error criterion without any condition, which means the above conditions are unnecessary. This solves Open problems 98, 101, 102 in [8]. See the following theorem.

Theorem 2.5 Consider the problem $\mathrm{APP} = \{\mathrm{APP}_d\}_{d\in\mathbb{N}}$ in the randomized setting for the absolute error criterion. Then

• ALG-SPT, ALG-PT, ALG-QPT, ALG-WT for Λ^{all} is equivalent to ALG-SPT, ALG-PT, ALG-QPT, ALG-WT for Λ^{std};

• The exponents ALG-$p^{\mathrm{ran,ABS}}(\Lambda)$ of ALG-SPT for Λ^{all} and Λ^{std} are the same, and the exponents ALG-$t^{\mathrm{ran,ABS}}(\Lambda)$ of ALG-QPT for Λ^{all} and Λ^{std} are also the same.

For exponential convergence in the randomized and worst case settings, we first give an almost complete solution to Open Problem 100 in [8].

In the randomized setting for the normalized or absolute error criterion, the equivalences of ALG-UWT and ALG-(s, t)-WT, and the various notions of EXP-tractability for Λ^{all} and Λ^{std}, as far as we know, have not been studied. In this paper, we investigate these problems and obtain the following theorem which gives the above equivalences without any condition.

Theorem 2.6 Consider the problem $\mathrm{APP} = \{\mathrm{APP}_d\}_{d\in\mathbb{N}}$ in the randomized setting for the absolute or normalized error criterion. Then for $\star \in \{\mathrm{ABS}, \mathrm{NOR}\}$,

• EXP-SPT, EXP-PT, EXP-QPT, EXP-UWT, EXP-WT, EXP-(s, t)-WT, ALG-UWT, ALG-(s, t)-WT for Λ^{all} is equivalent to EXP-SPT, EXP-PT, EXP-QPT, EXP-UWT, EXP-WT, EXP-(s, t)-WT, ALG-UWT, ALG-(s, t)-WT for Λ^{std};

• The exponents EXP-$p^{\mathrm{ran},\star}(\Lambda)$ of EXP-SPT for Λ^{all} and Λ^{std} are the same, and the exponents EXP-$t^{\mathrm{ran},\star}(\Lambda)$ of EXP-QPT for Λ^{all} and Λ^{std} are also the same.

Combining Corollary 2.1 with Theorems 22.19, 22.21, and 22.5 in [8] and Theorems 2.5 and 2.6, we obtain the following corollary.

Corollary 2.7 Consider the approximation problem $\mathrm{APP} = \{\mathrm{APP}_d\}_{d\in\mathbb{N}}$ for the absolute or normalized error criterion in the randomized and worst case settings. Then

• ALG-SPT, ALG-PT, ALG-QPT, ALG-UWT, ALG-WT, ALG-(s, t)-WT in the worst case setting for Λ^{all} is equivalent to ALG-SPT, ALG-PT, ALG-QPT, ALG-UWT, ALG-WT, ALG-(s, t)-WT in the randomized setting for Λ^{all} or for Λ^{std};

• EXP-SPT, EXP-PT, EXP-QPT, EXP-UWT, EXP-WT, EXP-(s, t)-WT in the worst case setting for Λ^{all} is equivalent to EXP-SPT, EXP-PT, EXP-QPT, EXP-UWT, EXP-WT, EXP-(s, t)-WT in the randomized setting for Λ^{all} or for Λ^{std};

• the exponents of SPT and QPT are the same in the worst case setting for Λ^{all} and in the randomized setting for Λ^{all} and Λ^{std}, i.e., for $\star \in \{\mathrm{ABS}, \mathrm{NOR}\}$,

$$\mathrm{ALG}\text{-}p^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}}) = \mathrm{ALG}\text{-}p^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}}) = \mathrm{ALG}\text{-}p^{\mathrm{ran},\star}(\Lambda^{\mathrm{std}}),$$

$$\mathrm{ALG}\text{-}t^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}}) = \mathrm{ALG}\text{-}t^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}}) = \mathrm{ALG}\text{-}t^{\mathrm{ran},\star}(\Lambda^{\mathrm{std}}),$$

$$\mathrm{EXP}\text{-}p^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}}) = \mathrm{EXP}\text{-}p^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}}) = \mathrm{EXP}\text{-}p^{\mathrm{ran},\star}(\Lambda^{\mathrm{std}}),$$

$$\mathrm{EXP}\text{-}t^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}}) = \mathrm{EXP}\text{-}t^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}}) = \mathrm{EXP}\text{-}t^{\mathrm{ran},\star}(\Lambda^{\mathrm{std}}).$$

3. Proofs of Theorems 2.2 and 2.3

Let us keep the notations of Subsection 2.1. For any $m \in \mathbb{N}$, we define the functions $h_{m,d}$ and $\omega_{m,d}$ on D_d by

$$h_{m,d}(x) := \frac{1}{m}\sum_{j=1}^{m}|\eta_{j,d}(x)|^2, \quad \omega_{m,d}(x) := h_{m,d}(x)\rho_d(x),$$

where $\{\eta_{j,d}\}_{j=1}^{\infty}$ is an orthonormal system in $G_d = L_2(D_d, \rho_d(x)d_x)$. Then $\omega_{m,d}$ is a probability density function on D_d, i.e., $\int_{D_d}\omega_{m,d}(x)d_x = 1$. We define the corresponding probability measure $\mu_{m,d}$ by

$$\mu_{m,d}(A) = \int_A \omega_{m,d}(x)dx,$$

where A is a Borel subset of D_d. We use the convention that $\frac{0}{0} := 0$. Then $\{\tilde{\eta}_{j,d}\}_{j=1}^{m}$ is an orthonormal system in $L_2(D_d, \mu_{m,d})$, where

$$\tilde{\eta}_{j,d} := \frac{\eta_{j,d}}{\sqrt{h_{m,d}}},\ 1 \le j \le m.$$

For $\mathrm{X} = (x^1, ..., x^n) \in D_d^n$, we use the following matrices

$$\tilde{L}_m = \tilde{L}_m(\mathrm{X}) = \begin{pmatrix} \tilde{\eta}_{1,d}(x^1) & \tilde{\eta}_{2,d}(x^1) & \cdots & \tilde{\eta}_{m,d}(x^1) \\ \tilde{\eta}_{1,d}(x^2) & \tilde{\eta}_{2,d}(x^2) & \cdots & \tilde{\eta}_{m,d}(x^2) \\ \vdots & \vdots & & \vdots \\ \tilde{\eta}_{1,d}(x^n) & \tilde{\eta}_{2,d}(x^n) & \cdots & \tilde{\eta}_{m,d}(x^n) \end{pmatrix} \quad \text{and} \quad \widetilde{H}_m = \frac{1}{n}\tilde{L}_m^* \tilde{L}_m, \tag{21}$$

where A^* is the conjugate transpose of a matrix A. Note that

$$\widetilde{N}(m) := \sup_{x \in D_d} \sum_{k=1}^{m} |\tilde{\eta}_{k,d}(x)|^2 = m.$$

It follows from the proof of Proposition 5.1 in [38] that whether Proposition 5.1 in [38] holds or not depends only on the orthonormal property of η_j, $j = 1, ..., m$ in $L_2(D, \rho_d)$. Since $\{\tilde{\eta}_{j,d}\}_{j=1}^{m}$ is an orthonormal system in $L_2(D_d, \mu_{m,d})$, Propositions 5.1 and 3.1 in [38] hold for $\{\tilde{\eta}_{j,d}\}$, $j = 1, ..., m$. According to Propositions 5.1 and 3.1 in [38] we have the following results.

Lemma 3.1 Let $n, m \in \mathbb{N}$. Let $x^1, ..., x^n \in D_d$ be drawn independently and identically distributed at random with respect to the probability measure $\mu_{m,d}$. Then it holds for $0 < t < 1$ that

$$\mathbb{P}\left(\|\widetilde{H}_m - I_m\| > t\right) \le (2n)^{\sqrt{2}} \exp\left(-\frac{nt^2}{12m}\right),$$

where $\tilde{L}_m$, $\widetilde{H}_m$ are given by (21), I_m is the identity matrix of order m, and $\|L\|$ denotes the spectral norm (i.e. the largest singular value) of a matrix L.

Lemma 3.2 Let $n, m \in \mathbb{N}$, and let $\tilde{L}_m$, $\widetilde{H}_m$ be given by (21). If

$$\|\widetilde{H}_m - I_m\| \le 1/2,$$

then

$$\|(\tilde{L}_m^* \tilde{L}_m)^{-1}\| \le \frac{2}{n}. \tag{22}$$

Remark 3.3 From Lemma 3.1 we immediately obtain that the matrix $\widetilde{H}_m \in \mathbb{C}^{m \times m}$ has only eigenvalues larger than $t := 1/2$ and satisfies

$$\|\widetilde{H}_m - I_m\| \le 1/2,$$

with probability at least $1 - \delta$ if

$$\widetilde{N}(m) = m \le \frac{n}{48(\sqrt{2}\ln(2n) - \ln\delta)}.$$

Specifically, if

$$m = \left\lfloor \frac{n}{48(\sqrt{2}\ln(2n) - \ln\delta)} \right\rfloor \ge 1, \tag{23}$$

then the matrix $\widetilde{H}_m$ has only eigenvalues larger than 1/2 and satisfies

$$\|\widetilde{H}_m - I_m\| \le 1/2$$

with probability at least $1 - \delta$, where $\lfloor x \rfloor$ denotes the largest integer not exceeding x. It follows that

$$\mathbb{P}\left(\|\widetilde{H}_m - I_m\| \le 1/2\right) \ge 1 - \delta \tag{24}$$

holds given the condition (23).

Now let n, $m \in \mathbb{N}$ satisfying (23), x^1, ..., x^n be independent and identically distributed sample points from D_d that are distributed according to the probability measure $\mu_{m,d}$, and $\widetilde{L}_m$, $\widetilde{H}_m$ be given by (21). If the sample points $\mathrm{X} = (x^1, ..., x^n)$ satisfy $\|\widetilde{H}_m - I_m\| > 1/2$, then we discard these points and re-sample until the re-sample points satisfy $\|\widetilde{H}_m - I_m\| \le 1/2$. That is, we consider the conditional distribution given the event $\|\widetilde{H}_m - I_m\| \le 1/2$, and the conditional expectation

$$\mathbb{E}\left(X \,\middle|\, \|\widetilde{H}_m - I_m\| \le 1/2\right) = \frac{\int_{\|\widetilde{H}_m - I_m\| \le 1/2} X(x^1, \ldots, x^n) d\mu_{m,d}(x^1) \ldots d\mu_{m,d}(x^n)}{\mathbb{P}\left(\|\widetilde{H}_m - I_m\| \le 1/2\right)}$$

of a random variable X.

If $\|\widetilde{H}_m - I_m\| \le 1/2$ for some $\mathrm{X} = (x^1, ..., x^n) \in D_d^n$ then $\widetilde{L}_m = \widetilde{L}_m(\mathrm{X})$ has the full rank. The algorithm is a weighted least squares estimator

$$S_{\mathrm{X}}^m(f) = \arg\min_{g \in V_m} \sum_{i=1}^n \frac{|f(x^i) - g(x^i)|^2}{h_{m,d}(x^i)}, \tag{25}$$

which has a unique solution, where $V_m := \mathrm{span}\{\eta_{1,d}, ..., \eta_{m,d}\}$. Note that

$$\mu_{m,d}\left(\{x \in D_d \mid h_{m,d}(x) = 0\}\right) = 0.$$

This means the above algorithm is meaningful. It follows that $S_{\mathrm{X}}^m(f) = f$ whenever $f \in V_m$. From [13, 38], we can give the following algorithm.

Algorithm Weighted least squares regression.

Input: $\mathrm{X} = (x^1, ..., x^n) \in D_d^n$ set of distinct sampling nodes,

$\tilde{\mathrm{f}} = \left(\frac{f(x^1)}{\sqrt{h_{m,d}(x^1)}}, \ldots, \frac{f(x^n)}{\sqrt{h_{m,d}(x^n)}}\right)^T$ weighted samples of f evaluted at the nodes from X,

$m \in \mathbb{N}$ $m < n$ such that the matrix $\widetilde{L}_m = \widetilde{L}_m(\mathrm{X})$ from (21) has full (column) rank.

Solve the over-determined linear system

$$\tilde{L}_m(\tilde{c}_1,\cdots,\tilde{c}_m)^T = \tilde{\mathrm{f}}$$

via least squares, i.e., compute

$$(\tilde{c}_1,\cdots,\tilde{c}_m)^T = (\tilde{L}_m^* \tilde{L}_m)^{-1} \tilde{L}_m^* \tilde{\mathrm{f}}.$$

Output: $\tilde{c} = (\tilde{c}_1,\cdots,\tilde{c}_m)^T \in \mathbb{C}^m$ coefficients of the approximant $S_{\mathrm{X}}^m(f) \coloneqq \sum_{k=1}^m \tilde{c}_k \eta_{k,d}$ which is the unique solution of (25).

Proof of Theorem 2.2

We use the above notations. We remark that X = (x^1, ..., x^n) obeys the conditional distribution given the event $\|\widetilde{H}_m - I_m\| \le 1/2$, and S_{X}^m is a randomized linear algorithm with fixed cardinality n. It follows that

$$\left(e^{\mathrm{ran}}(n,d;\Lambda^{\mathrm{std}})\right)^2 \le \sup_{f\in F_d,\|f\|_{F_d}\le 1} \mathbb{E}\left(\|f - S_{\mathrm{X}}^m(f)\|_{G_d}^2 \;\middle|\; \|\widetilde{H}_m - I_m\| \le 1/2\right),$$

where $n, m \in \mathbb{N}$ satisfy (23). We estimate $\|f - S_{\mathrm{X}}^m(f)\|_{G_d}^2$ for $f \in F_d$ with $\|f\|_{F_d} \le 1$. We set

$$H_d = L_2(D_d, \mu_{m,d}).$$

We recall that $\{e_{j,d}\}_{j=1}^\infty$ is an orthonormal basis in F_d, $\{\eta_{j,d}\}_{j=1}^\infty$ is an orthonormal system in $G_d = L_2(D_d, \rho_d(x)dx)$, and $\{\tilde{\eta}_{j,d}\}_{j=1}^m$ is an orthonormal system in $H_d = L_2(D_d, \mu_{m,d})$, where

$$\eta_{j,d} = \lambda_{j,d}^{-1/2} e_{j,d},\ j \ge 1 \ \text{ and } \ \tilde{\eta}_{j,d} \coloneqq \frac{\eta_{j,d}}{\sqrt{h_{m,d}}},\ 1 \le j \le m.$$

For $f \in F_d \subset G_d$ with $\|f\|_{Fd} \le 1$, we have

$$f = \sum_{k=1}^\infty \langle f, e_{k,d}\rangle_{F_d} e_{k,d} = \sum_{k=1}^\infty \langle f, \eta_{k,d}\rangle_{G_d} \eta_{k,d},$$

and

$$\|f\|_{F_d}^2 = \sum_{k=1}^\infty |\langle f, e_{k,d}\rangle_{F_d}|^2 = \sum_{k=1}^\infty \lambda_{k,d}^{-1} |\langle f, \eta_{k,d}\rangle_{G_d}|^2.$$

We note that $f - S_{m,d}^*(f)$ is orthogonal to the space V_m with respect to the inner product $\langle .,. \rangle_{G_d}$, and

$$S_{\mathrm{X}}^m(f) - S_{m,d}^*(f) = S_{\mathrm{X}}^m(f - S_{m,d}^*(f)) \in V_m \coloneqq \mathrm{span}\{\eta_{1,d},\ldots,\eta_{m,d}\},$$

where

$$S_{m,d}^*(f) = \sum_{k=1}^m \langle f, \eta_{k,d}\rangle_{G_d} \eta_{k,d}.$$

It follows that

$$\| f - S_{\mathrm{X}}^{m}(f)\|_{G_d}^2 = \| f - S_{m,d}^{*}(f)\|_{G_d}^2 + \| S_{\mathrm{X}}^{m}(f - S_{m,d}^{*}(f))\|_{G_d}^2$$

$$= \| g\|_{G_d}^2 + \| S_{\mathrm{X}}^{m}(g)\|_{G_d}^2,$$

where $g = f - S_{m,d}^{*}(f)$.

We recall that

$$S_{\mathrm{X}}^{m}(g) = \sum_{k=1}^{m} \tilde{c}_k \eta_{k,d},\ \tilde{c} = (\tilde{c}_1, \ldots, \tilde{c}_m)^T = (\tilde{L}_m^* \tilde{L}_m)^{-1} \tilde{L}_m^* \tilde{\mathrm{g}},$$

where

$$\tilde{\mathrm{g}} = (\tilde{g}(x^1), \cdots, \tilde{g}(x^n))^T,\ \ \tilde{g} = \frac{g}{\sqrt{h_{m,d}}}.$$

Since $\{\eta_{k,d}\}_{k=1}^{\infty}$ is an orthonormal system in G_d, we get

$$\| S_{\mathrm{X}}^{m}(g)\|_{G_d}^2 = \| \tilde{c}\|_2^2 = \| (\tilde{L}_m^* \tilde{L}_m)^{-1} \tilde{L}_m^* \tilde{\mathrm{g}}\|_2^2$$

$$\leq \| (\tilde{L}_m^* \tilde{L}_m)^{-1}\|^2 \cdot \| \tilde{L}_m^* \tilde{\mathrm{g}}\|_2^2$$

$$\leq \frac{4}{n^2} \| \tilde{L}_m^* \tilde{\mathrm{g}}\|_2^2,$$

where $\| \cdot \|_2$ is the Euclidean norm of a vector. We have

$$\| \tilde{L}_m^* \tilde{\mathrm{g}}\|_2^2 = \sum_{k=1}^{m} \Big| \sum_{j=1}^{n} \overline{\tilde{\eta}_{k,d}(x^j)} \cdot \tilde{g}(x^j) \Big|^2$$

$$= \sum_{k=1}^{m} \sum_{j=1}^{n} \sum_{i=1}^{n} \overline{\tilde{\eta}_{k,d}(x^j)} \tilde{g}(x^j) \tilde{\eta}_{k,d}(x^i) \overline{\tilde{g}(x^i)}.$$

It follows that

$$J = \int_{\|\widetilde{H}_m - I_m\| \leq \frac{1}{2}} \| \tilde{L}_m^* \tilde{\mathrm{g}}\|_2^2 \, d\mu_{m,d}(x^1) \ldots d\mu_{m,d}(x^n)$$

$$\leq \int_{D_d^n} \| \tilde{L}_m^* \tilde{\mathrm{g}}\|_2^2 \, d\mu_{m,d}(x^1) \ldots d\mu_{m,d}(x^n)$$

$$= \sum_{k=1}^{m} \sum_{i,j=1}^{n} \int_{D_d^n} \overline{\tilde{\eta}_{k,d}(x^j)} \tilde{g}(x^j) \tilde{\eta}_{k,d}(x^i) \overline{\tilde{g}(x^i)} \, d\mu_{m,d}(x^1) \ldots d\mu_{m,d}(x^n)$$

$$= \sum_{k=1}^{m} \sum_{i,j=1}^{n} J_{k,i,j},$$

where

$$J_{k,i,j} = \int_{D_d^n} \overline{\tilde{\eta}_{k,d}(x^j)}\tilde{g}(x^j)\tilde{\eta}_{k,d}(x^i)\overline{\tilde{g}(x^i)}d\mu_{m,d}(x^1)\ldots d\mu_{m,d}(x^n).$$

If $i \neq j$ and $1 \leq k \leq m$, then

$$J_{k,i,j} = |\langle \tilde{g}, \tilde{\eta}_{k,d}\rangle_{H_d}|^2 = |\langle g, \eta_{k,d}\rangle_{G_d}|^2 = 0;$$

If $i = j$, then

$$J_{k,i,j} = \| \tilde{\eta}_{k,d} \cdot \tilde{g} \|_{H_d}^2.$$

Since $h_{m,d}(x) = \frac{1}{m}\sum_{k=1}^{m}|\eta_{k,d}(x)|^2$, we get

$$J \leq \sum_{k=1}^{m}\sum_{i=j=1}^{n} J_{k,i,j} = n\sum_{k=1}^{m}\| \tilde{\eta}_{k,d} \cdot \tilde{g} \|_{H_d}^2$$

$$= n\sum_{k=1}^{m}\int_{D_d^n} |\tilde{g}(x)\tilde{\eta}_{k,d}(x)|^2 \rho_d(x)h_{m,d}(x)dx$$

$$= n\sum_{k=1}^{m}\int_{D_d^n} \frac{|g(x)\eta_{k,d}(x)|^2}{h_{m,d}(x)}\rho_d(x)dx$$

$$\leq n\int_{D_d^n} m|g(x)|^2 \rho_d(x)dx$$

$$= nm\cdot\| g \|_{G_d}^2.$$

Hence, by (3) we have

$$\sup_{f\in F_d, \|f\|_{F_d}\leq 1}\int_{\|\widetilde{H}_m - I_m\|\leq\frac{1}{2}}\| f - S_X^m(f)\|_{G_d}^2 d\mu_{m,d}(x^1)\ldots d\mu_{m,d}(x^n)$$

$$\leq \| g \|_{G_d}^2 + \frac{4}{n^2}J \leq \left(1+\frac{4m}{n}\right)\| g \|_{G_d}^2 \leq \left(1+\frac{4m}{n}\right)\left(e^{\text{wor}}(m,d;\Lambda^{\text{all}})\right)^2.$$

We conclude that

$$\sup_{f\in F_d, \|f\|_{F_d}\leq 1}\mathbb{E}\left(\| f - S_X^m(f)\|_{G_d}^2 \,\middle|\, \|\widetilde{H}_m - I_m\| \leq 1/2\right)$$

$$= \sup_{f\in F_d, \|f\|_{F_d}\leq 1}\frac{\int_{\|\widetilde{H}_m - I_m\|\leq\frac{1}{2}}\| f - S_X^m(f)\|_{G_d}^2 d\mu_{m,d}(x^1)\ldots d\mu_{m,d}(x^n)}{\mathbb{P}\left(\|\widetilde{H}_m - I_m\| \leq \frac{1}{2}\right)}$$

$$\leq \left(1+\frac{4m}{n}\right)\frac{1}{1-\delta}\left(e^{\text{wor}}(m,d;\Lambda^{\text{all}})\right)^2,$$

where in the last inequality we used (24). This completes the proof of Theorem 2.2.

Proof of Theorem 2.3

Applying Theorem 2.2 with $\delta = \frac{1}{2^{\sqrt{2}}}$, we obtain

$$e^{\text{ran}}(n,d;\Lambda^{\text{std}}) \leq \left(1+\frac{4m}{n}\right)^{\frac{1}{2}}\left(\frac{2^{\sqrt{2}}}{2^{\sqrt{2}}-1}\right)^{\frac{1}{2}} e^{\text{wor}}(m,d;\Lambda^{\text{all}}), \tag{26}$$

where $m, n \in \mathbb{N}$, and

$$m = \left\lfloor \frac{n}{48\sqrt{2}\ln(4n)} \right\rfloor.$$

Since $1 + \frac{4m}{n} \leq 1 + \frac{1}{12\sqrt{2}\ln(4n)} \leq 2$, by (26) we get

$$e^{\text{ran}}(n,d;\Lambda^{\text{std}}) \leq 4e^{\text{wor}}(m,d;\Lambda^{\text{all}}). \tag{27}$$

It follows that

$$\begin{aligned} n^{\text{ran},\star}(\varepsilon,d;\Lambda^{\text{std}}) &= \min\left\{n \middle| e^{\text{ran}}(n,d;\Lambda^{\text{std}}) \leq \varepsilon \text{CRI}_d\right\} \\ &\leq \min\left\{n \middle| 4e^{\text{wor}}(m,d;\Lambda^{\text{all}}) \leq \varepsilon \text{CRI}_d\right\} \\ &= \min\left\{n \middle|\ e^{\text{wor}}(m,d;\Lambda^{\text{all}}) \leq \frac{\varepsilon}{4}\text{CRI}_d\right\}. \end{aligned} \tag{28}$$

We note that

$$m = \left\lfloor \frac{n}{48\sqrt{2}\ln(4n)} \right\rfloor \geq \frac{n}{48\sqrt{2}\ln(4n)} - 1.$$

This inequality is equivalent to

$$4n \leq 192\sqrt{2}(m+1)\ln(4n). \tag{29}$$

Taking logarithm on both sides of (29), and using the inequality $\ln x \leq \frac{1}{2}x$ for $x \geq 1$, we get

$$\ln(4n) \leq \ln(m+1) + \ln(192\sqrt{2}) + \ln\ln(4n),$$

and

$$\frac{1}{2}\ln(4n) \le \ln(4n) - \ln\ln(4n) \le \ln(m+1) + \ln(192\sqrt{2}).$$

It follows from (29) that

$$n \le 96\sqrt{2}(m+1)\Big(\ln(m+1) + \ln(192\sqrt{2})\Big). \tag{30}$$

By (28) and (30) we obtain

$$n^{\mathrm{ran},\star}(\varepsilon, d; \Lambda^{\mathrm{std}}) \le 96\sqrt{2}\left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{4}, d; \Lambda^{\mathrm{all}}\right)+1\right)\left(\ln\left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{4}, d; \Lambda^{\mathrm{all}}\right)+1\right)+\ln(192\sqrt{2})\right),$$

proving (15).

For $0 < \delta < e^{-2}$ and $m, n \in \mathbb{N}$ satisfying

$$m = \left\lfloor \frac{n}{48(\sqrt{2}\ln(2n) - \ln\delta)} \right\rfloor,$$

by Theorem 2.2 we have

$$\begin{aligned} e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}}) &\le \left(1 + \frac{4m}{n}\right)^{\frac{1}{2}} \frac{1}{\sqrt{1-\delta}} e^{\mathrm{wor}}(m, d; \Lambda^{\mathrm{all}}) \\ &\le \left(1 + \frac{1}{12\left(\sqrt{2}\ln(2n) + \ln\frac{1}{\delta}\right)}\right)^{\frac{1}{2}} \frac{1}{\sqrt{1-\delta}} e^{\mathrm{wor}}(m, d; \Lambda^{\mathrm{all}}) \\ &\le \left(1 + \frac{1}{12\ln\frac{1}{\delta}}\right)^{\frac{1}{2}} \frac{1}{\sqrt{1-\delta}} e^{\mathrm{wor}}(m, d; \Lambda^{\mathrm{all}}) \\ &= A_\delta e^{\mathrm{wor}}(m, d; \Lambda^{\mathrm{all}}), \end{aligned}$$

where

$$A_\delta = \left(1 + \frac{1}{12\ln\frac{1}{\delta}}\right)^{\frac{1}{2}} \frac{1}{\sqrt{1-\delta}}.$$

Using the same method used in the proof of (28), we have

$$n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{std}}) \le \min\left\{n \mid\ e^{\mathrm{wor}}(m,d;\Lambda^{\mathrm{all}}) \le \frac{\varepsilon}{A_\Lambda}\mathrm{CRI}_d\right\}.$$

We note that

$$n \le 48\left(\sqrt{2}\ln(2n)+\ln\frac{1}{\delta}\right)(m+1).$$

Taking logarithm on both sides, and using the inequalities $\ln x \le \frac{x}{4}$ for $x \ge 9$ and $a + b \le ab$ for $a, b \ge 2$, we get

$$\ln n \le \ln 48 + \ln\left(\sqrt{2}\ln(2n)+\ln\frac{1}{\delta}\right)+\ln(m+1)$$

$$\le \ln 48 + \ln\left(\sqrt{2}\ln(2n)\right)+\ln\ln\frac{1}{\delta}+\ln(m+1)$$

$$\le \ln 48 + \frac{\sqrt{2}}{4}\ln(2n)+\ln\ln\frac{1}{\delta}+\ln(m+1).$$

Since

$$\frac{\sqrt{2}}{4}\ln(2n) \le \ln n - \frac{\sqrt{2}}{4}\ln(2n) \ \text{ for } \ n \ge 9,$$

we get

$$\sqrt{2}\ln(2n) \le 4\left(\ln 48+\ln\ln\frac{1}{\delta}+\ln(m+1)\right).$$

It follows that

$$n \le 48\left(4\left(\ln 48+\ln\ln\frac{1}{\delta}+\ln(m+1)\right)+\ln\frac{1}{\delta}\right)(m+1).$$

We conclude that for $0 < \delta < e^{-2}$,

$$n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{std}}) \le 48\left(4\left(\ln 48+\ln\ln\frac{1}{\delta}+\ln\left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{A_\delta},d;\Lambda^{\mathrm{all}}\right)+1\right)\right)+\ln\frac{1}{\delta}\right)$$

$$\left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{A_\delta},d;\Lambda^{\mathrm{all}}\right)+1\right), \tag{31}$$

proving (16). Theorem 2.3 is proved.

4. Equivalence results of algebraic tractability

First, we consider the equivalences of ALG-PT and ALG-SPT for Λ^{std} and Λ^{all} in the randomized setting. The equivalence results for the normalized error criterion can be found in Theorem 22.19 in [8]. For the absolute error criterion, Theorem 22.20 in [8] shows the equivalence of ALG-PT under the condition

$$\lambda_{1,d} \le C_\lambda d^{s_\lambda} \quad \text{for all } d \in \mathbb{N}, \text{ some } C_\lambda > 0, \text{ and some } s_\lambda \ge 0, \tag{32}$$

and the equivalence of ALG-SPT under the condition (32) with $s_\lambda = 0$.

We obtain the following equivalence results of ALG-PT and ALG-SPT without any condition. Hence, the condition (32) is unnecessary. This solves Open Problem 101 as posed by Novak and Woźniakowski in [8].

Theorem 4.1 We consider the problem $\text{APP} = \{\text{APP}_d\}_{d\in\mathbb{N}}$ in the randomized setting for the absolute error criterion. Then,

• ALG-PT for Λ^{all} is equivalent to ALG-PT for Λ^{std}.

• ALG-SPT for Λ^{all} is equivalent to ALG-SPT for Λ^{std}. In this case, the exponents of ALG-SPT for Λ^{all} and Λ^{std} are the same.

Proof. It follows from (8) that ALG-PT (ALG-SPT) for Λ^{std} means ALG-PT (ALG-SPT) for Λ^{all} in the randomized setting. Since ALG-PT (ALG-SPT) for Λ^{all} in the worst case setting is equivalent to ALG-PT (ALG-SPT) for Λ^{all} in the randomized setting, it suffices to show that ALG-PT (ALG-SPT) for Λ^{all} in the worst case setting means that ALG-PT (ALG-SPT) for Λ^{std} in the randomized setting.

Suppose that ALG-PT holds for Λ^{all} in the worst case setting. Then there exist $C \ge 1$ and non-negative p, q such that

$$n^{\text{wor,ABS}}(\varepsilon, d; \Lambda^{\text{all}}) \le C d^q \varepsilon^{-p}, \text{ for all } d \in \mathbb{N}, \ \varepsilon \in (0,1). \tag{33}$$

It follows from (19) and (33) that

$$n^{\text{ran,ABS}}(\varepsilon, d; \Lambda^{\text{std}}) \le C_\omega \left(C d^q \left(\frac{\varepsilon}{4}\right)^{-p} + 1 \right)^{1+\omega}$$

$$\le C_\omega (2C4^p)^{1+\omega} d^{q(1+\omega)} \varepsilon^{-p(1+\omega)},$$

which means that ALG-PT holds for Λ^{std} in the randomized setting.

If ALG-SPT holds for Λ^{all} in the worst case setting, then (33) holds with $q = 0$. We obtain

$$n^{\text{ran,ABS}}(\varepsilon, d; \Lambda^{\text{std}}) \le C_\omega (2C4^p)^{1+\omega} \varepsilon^{-p(1+\omega)},$$

which means that ALG-SPT holds for Λ^{std} in the randomized setting. Furthermore, since ω can be arbitrary small, by Corollary 2.1 we have

$$\text{ALG-}p^{\text{ran,ABS}}(\Lambda^{\text{std}}) \le \text{ALG-}p^{\text{wor,ABS}}(\Lambda^{\text{all}})$$

$$= \text{ALG-}p^{\text{ran,ABS}}(\Lambda^{\text{all}}) \le \text{ALG-}p^{\text{ran,ABS}}(\Lambda^{\text{std}}),$$

which means that the exponents of ALG-SPT for Λ^{all} and Λ^{std} are the same. This completes the proof of Theorem 4.1.

Next we consider the equivalence of ALG-QPT for Λ^{std} and Λ^{all} in the randomized setting. The result for the normalized error criterion can be found in Theorem 22.21 in [8]. For the absolute error criterion, Theorem 22.22 in [8]

shows the equivalence of ALG-QPT under the condition

$$\limsup_{d\to\infty} \lambda_{1,d} < \infty. \tag{34}$$

We obtain the following equivalence result of ALG-QPT without any condition. Hence, the condition (34) is unnecessary. This solves Open Problem 102 as posed by Novak and Woźniakowski in [8].

Theorem 4.2 We consider the problem APP = $\{\mathrm{APP}_d\}_{d\in\mathbb{N}}$ in the randomized setting for the absolute error criterion. Then, ALG-QPT for Λ^{all} is equivalent to ALG-QPT for Λ^{std}. In this case, the exponents of ALG-QPT for Λ^{all} and Λ^{std} are the same.

Proof. Similar to the proof of Theorem 4.1, it is enough to prove that ALG-QPT for Λ^{all} in the worst case setting implies ALG-QPT for Λ^{std} in the randomized setting.

Suppose that ALG-QPT holds for Λ^{all} in the worst case setting. Then there exist $C \geq 1$ and a non-negative t such that

$$n^{\mathrm{wor,ABS}}(\varepsilon, d; \Lambda^{\mathrm{all}}) \leq C \exp\left(t(1+\ln d)(1+\ln \varepsilon^{-1})\right), \text{ for all } d \in \mathbb{N},\ \varepsilon \in (0,1). \tag{35}$$

It follows from (19) and (35) that for $\omega > 0$,

$$\begin{aligned} n^{\mathrm{ran,ABS}}(\varepsilon, d; \Lambda^{\mathrm{std}}) &\leq C_\omega \left(n^{\mathrm{wor,ABS}}(\varepsilon/4, d; \Lambda^{\mathrm{all}}) + 1\right)^{1+\omega} \\ &\leq C_\omega \left(C \exp\left(t(1+\ln d)\left(1+\ln\left(\frac{\varepsilon}{4}\right)^{-1}\right)\right) + 1\right)^{1+\omega} \\ &\leq C_\omega (2C)^{1+\omega} \exp\left((1+\omega)t(1+\ln d)(1+\ln 4+\ln \varepsilon^{-1})\right) \\ &\leq C_\omega (2C)^{1+\omega} \exp\left(t^*(1+\ln d)(1+\ln \varepsilon^{-1})\right), \end{aligned}$$

where $t^* = (1+\omega)(1+\ln 4)t$. This implies that ALG-QPT holds for Λ^{std} in the randomized setting.

Next we show that the exponents ALG-$t^{\mathrm{ran,ABS}}(\Lambda^{\mathrm{all}})$ and ALG-$t^{\mathrm{ran,ABS}}(\Lambda^{\mathrm{std}})$ are equal if ALG-QPT holds for Λ^{all} in the worst case setting. We have

$$\mathrm{ALG}\text{-}t^{\mathrm{wor,ABS}}(\Lambda^{\mathrm{all}}) = \mathrm{ALG}\text{-}t^{\mathrm{ran,ABS}}(\Lambda^{\mathrm{all}}) \leq \mathrm{ALG}\text{-}t^{\mathrm{ran,ABS}}(\Lambda^{\mathrm{std}}).$$

It suffices to show that

$$\mathrm{ALG}\text{-}t^{\mathrm{ran,ABS}}(\Lambda^{\mathrm{std}}) \leq \mathrm{ALG}\text{-}t^{\mathrm{wor,ABS}}(\Lambda^{\mathrm{all}}).$$

Note that using (19) we can only obtain that

$$\mathrm{ALG}\text{-}t^{\mathrm{ran,ABS}}(\Lambda^{\mathrm{std}}) \leq (1+\ln 4)\cdot \mathrm{ALG}\text{-}t^{\mathrm{wor,ABS}}(\Lambda^{\mathrm{all}}).$$

Instead we use (20). For sufficiently small $\delta > 0$ and $\omega > 0$, it follows from (20) and (35) that

$$n^{\mathrm{ran,ABS}}(\varepsilon,d;\Lambda^{\mathrm{std}}) \le C_{\omega,\delta}\left(n^{\mathrm{wor,ABS}}\left(\frac{\varepsilon}{A_\delta},d;\Lambda^{\mathrm{all}}\right)+1\right)^{1+\omega}$$

$$\le C_{\omega,\delta}(2C)^{1+\omega}\exp\left((1+\omega)t(1+\ln d)(1+\ln A_\delta+\ln\varepsilon^{-1})\right)$$

$$\le C_{\omega,\delta}(2C)^{1+\omega}\exp\left((1+\omega)t(1+\ln A_\delta)(1+\ln d)(1+\ln\varepsilon^{-1})\right),$$

where

$$A_\delta=\left(1+\frac{1}{12\ln\frac{1}{\delta}}\right)^{\frac{1}{2}}\frac{1}{\sqrt{1-\delta}}.$$

Taking the infimum over t for which (35) holds, and noting that $\lim_{(\omega,\delta)\to(0,0)}(1+\omega)(1+\ln A_\delta)=1$, we get that

$$\mathrm{ALG}\text{-}t^{\mathrm{ran,ABS}}(\Lambda^{\mathrm{std}}) \le \mathrm{ALG}\text{-}t^{\mathrm{wor,ABS}}(\Lambda^{\mathrm{all}}).$$

This completes the proof of Theorem 4.2.

Now we consider the equivalence of ALG-WT for Λ^{std} and Λ^{all} in the randomized setting. The result for the normalized error criterion can be found in Theorem 22.5 in [8]. For the absolute error criterion, Theorem 22.6 in [8] shows the equivalence of ALG-WT under the condition

$$\lim_{d\to\infty}\frac{\ln\max(\lambda_{1,d},1)}{d}=0. \tag{36}$$

We obtain the following equivalence result of ALG-WT without any condition. Hence, the condition (36) is unnecessary. This solves Open Problem 98 as posed by Novak and Wo´zniakowski in [8].

Theorem 4.3 We consider the problem $\mathrm{APP}=\{\mathrm{APP}_d\}_{d\in\mathbb{N}}$ in the randomized setting for the absolute error criterion. Then, ALG-WT for Λ^{all} is equivalent to ALG-WT for Λ^{std}.

Proof. The proof is identical to the proof of Theorem 4.4 with $s=t=1$ for the absolute error criterion. We omit the details.

Finally, we consider the equivalences of ALG-(s, t)-WT and ALG-UWT for Λ^{std} and Λ^{all} in the randomized setting. As far as we know, these equivalences have not been studied yet. We obtain the following equivalence results of ALG-(s, t)-WT and ALG-UWT for the absolute or normalized error criterion without any condition.

Theorem 4.4 We consider the problem $\mathrm{APP}=\{\mathrm{APP}_d\}_{d\in\mathbb{N}}$ in the randomized setting for the absolute or normalized error criterion. Then for fixed $s, t>0$, ALG-(s, t)-WT for Λ^{all} is equivalent to ALG-(s, t)-WT for Λ^{std}.

Proof. Again it is enough to prove that ALG-(s, t)-WT for Λ^{all} in the worst case setting implies ALG-(s, t)-WT for Λ^{std} in the randomized setting.

Suppose that ALG-(s, t)-WT holds for Λ^{all} in the worst case setting. Then we have for $\star\in\{\mathrm{ABS},\mathrm{NOR}\}$,

$$\lim_{\varepsilon^{-1}+d\to\infty}\frac{\ln n^{\mathrm{wor},\star}(\varepsilon,d;\Lambda^{\mathrm{all}})}{\varepsilon^{-s}+d^t}=0. \tag{37}$$

It follows from (19) that for $\omega>0$,

$$\frac{\ln n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{std}})}{\varepsilon^{-s}+d^t} \le \frac{\ln\left(C_\omega\left(n^{\mathrm{ran},\star}(\varepsilon/4,d;\Lambda^{\mathrm{all}})+1\right)^{1+\omega}\right)}{\varepsilon^{-s}+d^t}$$

$$\le \frac{\ln(C_\omega 2^{1+\omega})}{\varepsilon^{-s}+d^t} + \frac{4^s(1+\omega)\ln n^{\mathrm{wor},\star}(\varepsilon/4,d;\Lambda^{\mathrm{all}})}{(\varepsilon/4)^{-s}+d^t}.$$

Since $\varepsilon^{-1}+d\to\infty$ is equivalent to $\varepsilon^{-s}+d^t\to\infty$, by (37) we get that

$$\lim_{\varepsilon^{-1}+d\to\infty}\frac{\ln(C_\omega 2^{1+\omega})}{\varepsilon^{-s}+d^t}=0 \quad\text{and}\quad \lim_{\varepsilon^{-1}+d\to\infty}\frac{\ln n^{\mathrm{wor},\star}(\varepsilon/4,d;\Lambda^{\mathrm{all}})}{(\varepsilon/4)^{-s}+d^t}=0.$$

We obtain

$$\lim_{\varepsilon^{-1}+d\to\infty}\frac{\ln n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{std}})}{\varepsilon^{-s}+d^t}=0,$$

which implies that ALG-(s, t)-WT holds for Λ^{std} in the randomized setting. This completes the proof of Theorem 4.4.

Theorem 4.5 We consider the problem APP = $\{\mathrm{APP}_d\}_{d\in\mathbb{N}}$ in the randomized setting for the absolute or normalized error criterion. Then, ALG-UWT for Λ^{all} is equivalent to ALG-UWT for Λ^{std}.

Proof. By definition we know that APP is ALG-UWT if and only if APP is ALG-(s, t)-WT for all s, $t>0$. Then Theorem 4.5 follows from Theorem 4.4 immediately.

Proof of Theorem 2.5

Theorem 2.5 follows from Theorems 4.1-4.3 immediately.

5. Equivalence results of exponential tractability

First we consider exponential convergence. Assume that there exist two constants $A\ge 1$ and $q\in(0,1)$ such that

$$e^{\mathrm{wor}}(n,d;\Lambda^{\mathrm{all}})\le Aq^{n+1}\sqrt{\lambda_{1,d}}. \tag{38}$$

Novak and Woźniakowski proved in [8] that there exist two constants $C_1\ge 1$ and $q_1\in(q,1)$ independent of d and n such that

$$e^{\mathrm{ran}}(n,d;\Lambda^{\mathrm{std}})\le C_1Aq_1^{\sqrt{n}}\sqrt{\lambda_{1,d}}. \tag{39}$$

If A, q in (38) are independent of d, then

$$n^{\mathrm{wor,NOR}}(\varepsilon,d;\Lambda^{\mathrm{all}})\le C_2(\ln\varepsilon^{-1}+1),$$

and

$$n^{\mathrm{ran,NOR}}(\varepsilon,d;\Lambda^{\mathrm{std}})\le C_3(\ln\varepsilon^{-1}+1)^2.$$

Novak and Woźniakowski posed the following Open Problem 100:

(1) Verify if the upper bound in (39) can be improved.

(2) Find the smallest p for which there holds

$$n^{\mathrm{ran,NOR}}(\varepsilon, d; \Lambda^{\mathrm{std}}) \le C_4 (\ln \varepsilon^{-1} + 1)^p.$$

We know that $p \le 2$, and if (38) is sharp then $p \ge 1$.

The following theorem gives a confirmative solution to Open Problem 100 (1). We improve enormously the upper bound $q_1^{\sqrt{n}}$ in (39) to $q_2^{\frac{n}{\ln(4n)}}$ in (42), where $q_1, q_2 \in (q, 1)$.

Theorem 5.1 Let $m, n \in \mathbb{N}$ and

$$m = \left\lfloor \frac{n}{48\sqrt{2}\ln(4n)} \right\rfloor. \tag{40}$$

Then we have

$$e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}}) \le 4 e^{\mathrm{wor}}(m, d; \Lambda^{\mathrm{all}}). \tag{41}$$

Specifically, if (38) holds, then we have

$$e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}}) \le 4Aq_2^{\frac{n}{\ln(4n)}} \sqrt{\Lambda_{1,d}}, \tag{42}$$

where $q_2 = q^{\frac{1}{48\sqrt{2}}} \in (q, 1)$.

Proof. Inequality (41) is just (27), which has been proved. If (38) holds, then by (40) and (41) we get

$$e^{\mathrm{ran}}(n, d; \Lambda^{\mathrm{std}}) \le 4Aq^{\left\lfloor \frac{n}{48\sqrt{2}\ln(4n)} \right\rfloor + 1} \sqrt{\lambda_{1,d}}$$

$$\le 4Aq^{\frac{n}{48\sqrt{2}\ln(4n)}} \sqrt{\lambda_{1,d}}$$

$$= 4Aq_2^{\frac{n}{\ln(4n)}} \sqrt{\lambda_{1,d}}.$$

This completes the proof of Theorem 5.1.

Now we consider the equivalences of various notions of exponential tractability for Λ^{std} and Λ^{all} in the randomized setting. As far as we know, there is hardly any result for these equivalences.

First, we consider the equivalences of EXP-PT and EXP-SPT for Λ^{std} and Λ^{all} in the randomized setting. We obtain the following equivalence results of ALG-PT and ALG-SPT without any condition.

Theorem 5.2 We consider the problem APP = $\{\mathrm{APP}_d\}_{d \in \mathbb{N}}$ in the randomized setting for the absolute or normalized error criterion. Then,

• EXP-PT for Λ^{all} is equivalent to EXP-PT for Λ^{std}.

• EXP-SPT for Λ^{all} is equivalent to EXP-SPT for Λ^{std}. In this case, the exponents of EXP-SPT for Λ^{all} and Λ^{std} are the same.

Proof. Again, it is enough to prove that EXP-PT for Λ^{all} in the worst case setting implies EXP-PT for Λ^{std} in the randomized setting.

Suppose that EXP-PT holds for Λ^{all} in the worst case setting. Then there exist $C \ge 1$ and non-negative p, q, for $\star \in \{\mathrm{ABS}, \mathrm{NOR}\}$ such that

$$n^{\mathrm{wor},\star}(\varepsilon,d;\Lambda^{\mathrm{all}}) \le Cd^q(\ln\varepsilon^{-1}+1)^p, \text{ for all } d\in\mathbb{N},\ \varepsilon\in(0,1). \tag{43}$$

It follows from (19) and (43) that

$$n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{std}}) \le C_\omega\left(Cd^q\left(\ln\left(\frac{\varepsilon}{4}\right)^{-1}+1\right)^p+1\right)^{1+\omega}$$

$$\le C_\omega(2C)^{1+\omega}(1+\ln 4)^{p(1+\omega)}d^{q(1+\omega)}(\ln\varepsilon^{-1}+1)^{p(1+\omega)},$$

which means that EXP-PT holds for Λ^{std} in the randomized setting.

If EXP-SPT holds for Λ^{all} in the worst case setting, then (43) holds with $q = 0$. We obtain

$$n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{std}}) \le C_\omega(2C)^{1+\omega}(1+\ln 4)^{p(1+\omega)}(\ln\varepsilon^{-1}+1)^{p(1+\omega)},$$

which means that EXP-SPT holds for Λ^{std} in the randomized setting. Furthermore, in this case we have

$$\mathrm{EXP}\text{-}p^{\mathrm{ran},\star}(\Lambda^{\mathrm{std}}) \le \mathrm{EXP}\text{-}p^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}})$$

$$= \mathrm{EXP}\text{-}p^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}}) \le \mathrm{EXP}\text{-}p^{\mathrm{ran},\star}(\Lambda^{\mathrm{std}}),$$

which means that the exponents of EXP-SPT for Λ^{all} and Λ^{std} are the same. This completes the proof of Theorem 5.2.

Remark 5.3 We remark that if (38) holds with A, q independent of d, then the problem APP is EXP-SPT for Λ^{all} in the randomized setting for the normalized error criterion, and the exponent $\mathrm{EXP}\text{-}p^{\mathrm{wor,NOR}}(\Lambda^{\mathrm{all}}) \le 1$. If (38) is sharp, then $\mathrm{EXP}\text{-}p^{\mathrm{wor,NOR}}(\Lambda^{\mathrm{all}}) = 1$.

By Theorem 5.2 we obtain that if (38) holds, then $\mathrm{EXP}\text{-}p^{\mathrm{ran,NOR}}(\Lambda^{\mathrm{std}}) \le 1$, and if (38) is sharp, then $\mathrm{EXP}\text{-}p^{\mathrm{ran,NOR}}(\Lambda^{\mathrm{std}}) = 1$. It follows that for any $p > 1$, we have

$$n^{\mathrm{ran,NOR}}(\varepsilon,d;\Lambda^{\mathrm{std}}) \le C_p(\ln\varepsilon^{-1}+1)^p.$$

However, we do not know whether or not the above inequality holds for $p = 1$.

This almost solves Open Problem 100 (2) as posed by Novak and Woźniakowski in [8].

Next, we consider the equivalence of EXP-QPT for Λ^{std} and Λ^{all} in the randomized setting. We obtain the following equivalence result of EXP-QPT without any condition.

Theorem 5.4 We consider the problem $\mathrm{APP} = \{\mathrm{APP}_d\}_{d\in\mathbb{N}}$ in the randomized setting for the absolute or normalized error criterion. Then, EXP-QPT for Λ^{all} is equivalent to EXP-QPT for Λ^{std}. In this case, the exponents of EXP-QPT for Λ^{all} and Λ^{std} are the same.

Proof. Again, it is enough to prove that EXP-QPT for Λ^{all} in the worst case setting implies EXP-QPT for Λ^{std} in the randomized setting.

Suppose that EXP-QPT holds for Λ^{all} in the worst case setting. Then there exist $C \ge 1$ and a non-negative t such that

$$n^{\mathrm{wor},\star}(\varepsilon,d;\Lambda) \le C\exp\Big(t(1+\ln d)\big(1+\ln(\ln\varepsilon^{-1}+1)\big)\Big),\ \text{for all } d\in\mathbb{N},\ \varepsilon\in(0,1). \tag{44}$$

It follows from (19) and (44) that for $\omega > 0$,

$$
\begin{aligned}
& n^{\mathrm{ran},\star}(\varepsilon, d; \Lambda^{\mathrm{std}}) \\
& \le C_\omega \left(n^{\mathrm{wor},\star}(\varepsilon/4, d; \Lambda^{\mathrm{all}}) + 1 \right)^{1+\omega} \\
& \le C_\omega \left(C \exp\left(t(1+\ln d)\left(1+\ln(\ln \varepsilon^{-1} + \ln 4 + 1)\right)\right) + 1 \right)^{1+\omega} \\
& \le C_\omega (2C)^{1+\omega} \exp\left((1+\omega) t (1+\ln d)\left(1+\ln(\ln 4 + 1) + \ln(\ln \varepsilon^{-1} + 1)\right)\right) \\
& \le C_\omega (2C)^{1+\omega} \exp\left(t^* (1+\ln d)\left(1+\ln(\ln \varepsilon^{-1} + 1)\right)\right),
\end{aligned}
$$

where $t^* = (1+\omega)(1+\ln(\ln 4 + 1))t$, in the third inequality we used the fact

$$
\ln(1+a+b) \le \ln(1+a) + \ln(1+b), \quad a, b \ge 0.
$$

This implies that EXP-QPT holds for Λ^{std} in the randomized setting.

Next we show that the exponents EXP-$t^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}})$ and EXP-$t^{\mathrm{ran},\star}(\Lambda^{\mathrm{std}})$ are equal if EXP-QPT holds for Λ^{all} in the worst case setting. We have

$$
\mathrm{EXP}\text{-}t^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}}) = \mathrm{EXP}\text{-}t^{\mathrm{ran},\star}(\Lambda^{\mathrm{all}}) \le \mathrm{EXP}\text{-}t^{\mathrm{ran},\star}(\Lambda^{\mathrm{std}}).
$$

It suffices to show that

$$
\mathrm{EXP}\text{-}t^{\mathrm{ran},\star}(\Lambda^{\mathrm{std}}) \le \mathrm{EXP}\text{-}t^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}}).
$$

Note that using (19) we can only obtain that

$$
\mathrm{EXP}\text{-}t^{\mathrm{ran},\star}(\Lambda^{\mathrm{std}}) \le \left(1+\ln(1+\ln 4)\right) \cdot \mathrm{EXP}\text{-}t^{\mathrm{wor},\star}(\Lambda^{\mathrm{all}}).
$$

Instead we use (20). For sufficiently small $\delta > 0$ and $\omega > 0$, it follows from (20) and (44) that

$$
\begin{aligned}
& n^{\mathrm{ran},\star}(\varepsilon, d; \Lambda^{\mathrm{std}}) \\
& \le C_{\omega,\delta} \left(n^{\mathrm{wor},\star}\left(\frac{\varepsilon}{A_\delta}, d; \Lambda^{\mathrm{all}} \right) + 1 \right)^{1+\omega} \\
& \le C_{\omega,\delta} (2C)^{1+\omega} \exp\left((1+\omega) t (1+\ln d)\left(1+\ln(\ln A_\delta + 1) + \ln(\ln \varepsilon^{-1} + 1)\right)\right) \\
& \le C_{\omega,\delta} (2C)^{1+\omega} \exp\left(t^* (1+\ln d)\left(1+\ln(\ln \varepsilon^{-1} + 1)\right)\right),
\end{aligned}
$$

where

$$t^* = (1+\omega)t\left(1+\ln(\ln A_\delta + 1)\right),\ A_\delta = \left(1+\frac{1}{12\ln\frac{1}{\delta}}\right)^{\frac{1}{2}}\frac{1}{\sqrt{1-\delta}}.$$

Taking the infimum over t for which (44) holds, and noting that $\lim_{(\omega,\delta)\to(0,0)}(1+\omega)\left(1+\ln(\ln A_\delta+1)\right)=1$, we get that

$$\text{EXP-}t^{\text{ran},\star}(\Lambda^{\text{std}}) \le \text{EXP-}t^{\text{wor},\star}(\Lambda^{\text{all}}).$$

This completes the proof of Theorem 5.4.

Finally, we consider the equivalences of EXP-(s, t)-WT (including EXP-WT) and EXP-UWT for Λ^{std} and Λ^{all} in the randomized setting. We obtain the following equivalence results of EXP-(s, t)-WT (including EXP-WT) and EXP-UWT for the absolute or normalized error criterion without any condition.

Theorem 5.5 We consider the problem APP = $\{\text{APP}_d\}_{d\in\mathbb{N}}$ in the randomized setting for the absolute or normalized error criterion. Then for fixed s, $t > 0$, EXP-(s, t)-WT for Λ^{all} is equivalent to EXP-(s, t)-WT for Λ^{std}. Specifically, EXP-WT for Λ^{all} is equivalent to EXP-WT for Λ^{std}.

Proof. Again, it is enough to prove that EXP-(s, t)-WT for Λ^{all} in the worst case setting implies EXP-(s, t)-WT for Λ^{std} in the randomized setting.

Suppose that EXP-(s, t)-WT holds for Λ^{all} in the worst case setting. Then we have for $\star \in \{\text{ABS, NOR}\}$,

$$\lim_{\varepsilon^{-1}+d\to\infty}\frac{\ln n^{\text{wor},\star}(\varepsilon,d;\Lambda^{\text{all}})}{(1+\ln\varepsilon^{-1})^s+d^t}=0. \tag{45}$$

It follows from (19) that for $\omega > 0$,

$$\frac{\ln n^{\text{ran},\star}(\varepsilon,d;\Lambda^{\text{std}})}{(1+\ln\varepsilon^{-1})^s+d^t} \le \frac{\ln\left(C_\omega\left(n^{\text{ran},\star}(\varepsilon/4,d;\Lambda^{\text{all}})+1\right)^{1+\omega}\right)}{(1+\ln\varepsilon^{-1})^s+d^t}$$

$$\le \frac{\ln(C_\omega 2^{1+\omega})}{(1+\ln\varepsilon^{-1})^s+d^t}+\frac{(1+\ln 4)^s(1+\omega)\ln n^{\text{wor},\star}(\varepsilon/4,d;\Lambda^{\text{all}})}{\left(1+\ln(\varepsilon/4)^{-1}\right)^s+d^t}.$$

Since $\varepsilon^{-1}+d\to\infty$ is equivalent to $(1+\ln\varepsilon^{-1})^s+d^t\to\infty$, by (45) we get that

$$\lim_{\varepsilon^{-1}+d\to\infty}\frac{\ln(C_\omega 2^{1+\omega})}{(1+\ln\varepsilon^{-1})^s+d^t}=0 \quad\text{and}\quad \lim_{\varepsilon^{-1}+d\to\infty}\frac{\ln n^{\text{wor},\star}(\varepsilon/4,d;\Lambda^{\text{all}})}{\left(1+\ln(\varepsilon/4)^{-1}\right)^s+d^t}=0.$$

We obtain

$$\lim_{\varepsilon^{-1}+d\to\infty}\frac{\ln n^{\text{ran},\star}(\varepsilon,d;\Lambda^{\text{std}})}{(\ln\varepsilon^{-1})^s+d^t}=0,$$

which implies that EXP-(s, t)-WT holds for Λ^{std} in the randomized setting.

Specifically, EXP-WT is just EXP-(s, t)-WT with $s = t = 1$.

This completes the proof of Theorem 5.5.

Theorem 5.6 We consider the problem APP = $\{\mathrm{APP}_d\}_{d\in\mathbb{N}}$ in the randomized setting for the absolute or normalized error criterion. Then, EXP-UWT for Λ^{all} is equivalent to EXP-UWT for Λ^{std}.

Proof. By definition, we know that APP is EXP-UWT if and only if APP is EXP-(s, t)-WT for all s, $t > 0$. Then Theorem 5.6 follows from Theorem 5.5 immediately.

Proof of Theorem 2.6

Theorem 2.6 follows from Theorems 4.4, 4.5, 5.2, and 5.4-5.6 immediately.

Remark 5.7 While this manuscript was submitted for possible publication, Dolbeault and Cohen in [39] used a weighted least squares method and the weaver subsampling technique (see [23]) to show that

$$e^{\mathrm{ran}}(n,d;\Lambda^{\mathrm{std}}) \le C_1 e^{\mathrm{wor}}(C_2 n,d;\Lambda^{\mathrm{all}}), \tag{46}$$

where C_1, C_2 are two absolute positive constants. It follows from (46) immediately that

$$n^{\mathrm{ran},\star}(\varepsilon,d;\Lambda^{\mathrm{std}}) \le C_2 n^{\mathrm{wor},\star}(\varepsilon / C_1,d;\Lambda^{\mathrm{all}}). \tag{47}$$

By (47) we easily obtain Theorems 2.5 and 2.6 except that the exponents of ALG-QPT and EXP-QPT are equal, and solve Open Problem 100 as posed by E.Novak and H.Woźniakowski [8]. However, we cannot obtain by (47) that the exponents of ALG-QPT and EXP-QPT are equal.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Project no. 11671271).

Conflict of interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] Dick J, Kritzer P, Pillichshammer F, Woźniakowski H. Approximation of analytic functions in Korobov spaces. *Journal of Complexity*. 2014; 30(2): 2-28.

[2] Gnewuch M, Woźniakowski H. Quasi-polynomial tractability. *Journal of Complexity*. 2011; 27(3-4): 312-330.

[3] Irrgeher C, Kritzer P, Pillichshammer F, Woźniakowski H. Tractability of multivariate approximation defined over Hilbert spaces with exponential weights. *Journal of Approximation Theory*. 2016; 207: 301-338.

[4] Kritzer P, Woźniakowski H. Simple characterizations of exponential tractability for linear multivariate problems. *Journal of Complexity*. 2019; 51: 110-128.

[5] Liu YP, Xu GQ. Average case tractability of a multivariate approximation problem. *Journal of Complexity*. 2017; 43: 76-102.

[6] Novak E, Woźniakowski H. Volume I: Linear Information. *Tractablity of Multivariate Problems*. EMS, Zürich; 2008.

[7] Novak E, Woźniakowski H. Volume II: Standard Information for Functionals. *Tractablity of Multivariate Problems*. EMS, Zürich; 2010.

[8] Novak E, Woźniakowski H. Volume III: Standard Information for Operators. *Tractablity of Multivariate Problems*. EMS, Zürich; 2012.

[9] Papageorgiou A, Petras I, Xu GQ, Dong YQ. EC-(s, t)-weak tractability of multivariate linear problems in the average case setting. *Journal of Complexity*. 2019; 55: 26.

[10] Siedlecki P. Uniform weak tractability. *Journal of Complexity*. 2013; 29(6): 438-453.

[11] Siedlecki P, Weimar M. Notes on (s, t)-weak tractability: a refined classification of problems with (sub)exponential

information complexity. *Journal of Approximation Theory*. 2015; 200: 227-258.
[12] Cohen A, Davenport MA, Leviatan D. On the stability and accuracy of least squares approximations. *Foundations of Computational Mathematics*. 2013; 13: 819-834. (Correction in *Foundations of Computational Mathematics*. 2019; 19(1): 239.)
[13] Cohen A, Migliorati G. Optimal weighted least-squares methods. *SMAI Journal of Computational Mathematics*. 2017; 3: 181-203.
[14] Krieg D. Optimal Monte Carlo Methods for L_2-approximation. *Constructive Approximation*. 2019; 49(2): 385-403.
[15] Kuo FY, Wasilkowski GW, Woźniakowski H. On the power of standard information for L_∞ approximation in the randomized setting. *BIT Numererical Mathematics*. 2019; 49(3): 543-564.
[16] Wasilkowski GW, Woźniakowski H. The power of standard information for multivariate approximation in the randomized setting. *Mathematics of Computation*. 2007; 76(258): 965-988.
[17] Hickernell FJ, Wasilkowski GW, Woźniakowski H. Tractability of linear multivariate problems in the average case setting. In: Keller A, Heinrich S, Niederreiter H (Eds.). *Monte Carlo and Quasi-Monte Carlo Methods*. Berlin: Springer; 2008. p. 461-494.
[18] Lifshits M, Zani M. Approximation of additive random fields based on standard information: Average case and probabilistic settings. *Journal of Complexity*. 2015; 31(5): 659-674.
[19] Xu GQ. On the power of standard information for L_2-approximation in the average case setting. *Journal of Complexity*. 2020; 59: 101482.
[20] Hinrichs A, Novak E, Vybiral J. Linear information versus function evaluations for L_2-approximation. *Journal of Approximation Theory*. 2008; 153(1): 97-107.
[21] Krieg D, Ullrich M. Function values are enough for L_2-approximation. *Foundations of Computational Mathematics*. 2021; 21(4): 1141-1151.
[22] Kuo FY, Wasilkowski GW, Woźniakowski H. On the power of standard information for multivariate approximation in the worst case setting. *Journal of Approximation Theory*. 2019; 158(5): 97-125.
[23] Nagel N, Schäfer M, Ullrich T. A new upper bound for sampling numbers. *Foundations of Computational Mathematics*. 2021. Available from: doi: 10.1007/s10208-021-09504-0.
[24] Novak E, Woźniakowski H. Tractablity of Multivariate Problems for standard and linear information in the worst case setting: Part I. *Journal of Approximation Theory*. 2016; 207: 177-192.
[25] Novak E, Woźniakowski H. Tractablity of Multivariate Problems for standard and linear information in the worst case setting: Part II. *Contemporary computational mathematics-a celebration of the 80th birthday of Ian Sloan*. 2018; 1(2): 963-977.
[26] Wasilkowski GW, Woźniakowski H. On the power of standard information for weighted approximation. *Foundations of Computational Mathematics*. 2001; 1: 417-434.
[27] Hinrichs A, Novak E, Woźniakowski H. Discontinuous information in the worst case and randomized settings. *Mathematische Nachrichten*. 2013; 286(7): 679-690.
[28] Byrenheid G, Kunsch RJ, Nguyen VK. Monte Carlo methods for the L_∞-approximation on periodic Sobolev spaces with bounded mixed derivative. *Journal of Complexity*. 2018; 46: 90-102.
[29] Heinrich S. Lower bounds for the complexity of Monte Carlo function approximation. *Journal of Complexity*. 1992; 8(3): 277-300.
[30] Kunsch RJ. Breaking the curse for uniform approximation in Hilbert spaces via Monte Carlo methods. *Journal of Complexity*. 2018; 48: 15-35.
[31] Mathé P. Random approximation of sobolev embeddings. *Journal of Complexity*. 1991; 7: 261-281.
[32] Novak E. *Deterministic and Stochastic Error Bound in Numerical Analysis*. Berlin: Springer; 1988.
[33] Traub JF, Wasilkowski GW, Woźniakowski H. *Information-Based Complexity*. New York: Academic Press; 1988.
[34] Xu GQ. Exponential convergence-tractability of general linear problems in the average case setting. *Journal of Complexity*. 2015; 31: 617-636.
[35] Hinrichs A, Krieg D, Novak E, Prochno J, Ullrich M. On the power of random information. In: Hickernell FJ, Kritzer P (Eds.). *Multivariate Algorithms and Information-Based Complexity*. Berlin/Boston: De Gruyter; 2020; p. 43-64.
[36] Hinrichs A, Krieg D, Novak E, Prochno J, Ullrich M. Random sections of ellipsoids and the power of random information. *Transactions of the American Mathematical Society*. Available from: doi: arxiv.org/abs/1901.06639.
[37] Krieg D, Sonnleitner M. Random points are optimal for the approximation of Sobolev functions. 2020. Available from: doi: arxiv.org/abs/2009.11275.

[38] Kämmerer L, Ullrich T, Volkmer T. Worst-case recovery guarantees for least squares approximation using random samples. *Constructive Approximation*. 2021; 54: 295-352.
[39] Dolbeault M, Cohen A. Optimal pointwise sampling for L_2 approximation. *Journal of Complexity*. 2022; 68: 101602. Available from: doi: 10.1016/j.jco.2021.101602.