Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

Exponential tractability of linear weighted tensor product problems in the worst-case setting for arbitrary linear functionals

P. Kritzer, F. Pillichshammer, H. Wozniakowski

RICAM-Report 2020-06

Exponential tractability of linear weighted tensor product problems in the worst-case setting for arbitrary linear functionals

Peter Kritzer*, Friedrich Pillichshammer†, Henryk Woźniakowski‡

January 31, 2020

Abstract

We study the approximation of compact linear operators defined over certain weighted tensor product Hilbert spaces. The information complexity is defined as the minimal number of arbitrary linear functionals which is needed to obtain an ε-approximation for the d-variate problem. It is fully determined in terms of the weights and univariate singular values. Exponential tractability means that the information complexity is bounded by a certain function which depends polynomially on d and logarithmically on ε^{-1}. The corresponding un-weighted problem was studied in [4] with many negative results for exponential tractability. The product weights studied in the present paper change the situation. Depending on the form of polynomial dependence on d and logarithmic dependence on ε^{-1}, we study exponential strong polynomial, exponential polynomial, exponential quasi-polynomial, and exponential (s,t)-weak tractability with $\max(s,t) \geq 1$. For all these notions of exponential tractability, we establish necessary and sufficient conditions on weights and univariate singular values for which it is indeed possible to achieve the corresponding notion of exponential tractability. The case of exponential (s,t)-weak tractability with $\max(s,t) < 1$ is left for future study. The paper uses some general results obtained in [4] and [12].

Keywords: exponential tractability, weighted linear tensor product problem, approximation of compact linear operators, worst-case setting
MSC 2010: 65Y20, 65D15

*P. Kritzer is supported by the Austrian Science Fund (FWF), Project F5506-N26, which is part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications".

†F. Pillichshammer is supported by the Austrian Science Fund (FWF), Project F5509-N26, which is part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications".

‡H. Woźniakowski is supported in part by the National Science Center, Poland, based on the decision DEC-2017/25/B/ST1/00945.

The authors thank the Johann Radon Institute for Computational and Applied Mathematics in Linz, where most research for this paper has been done during the Special Semester on Multivariate Algorithms and their Foundations in Number Theory, October 15 - December 21, 2018

1 Introduction

This paper deals with tractability of linear multivariate problems, which has been studied in a large number of papers, and is at the core of the research field of Information-Based Complexity (IBC). For introductions to IBC, we refer to the books [29] and [30]. For a recent and comprehensive overview of results on tractability, we refer the interested reader to the trilogy [17, 18, 19].

In the present paper we study the information complexity of a compact linear operator S_d from a separable Hilbert space H_d into another Hilbert space G_d. The information complexity $n(\varepsilon, S_d)$ is defined as the minimal number of linear functionals needed by an algorithm which approximates S_d to within an error threshold of $\varepsilon > 0$. As shown in [5], without loss of generality we may consider only continuous linear functionals from the class $\Lambda_d^{\rm all} = H_d^*$.

We use different notions of tractability to describe how the information complexity of a given problem depends on ε^{-1} and d as $\max(\varepsilon^{-1}, d)$ tends to infinity in an arbitrary way. A problem is called intractable if its information complexity depends exponentially on ε^{-1} or d. If a problem is tractable, we describe sub-exponential dependence on some powers of ε^{-1} and d by using the classification into various notions of tractability, which can be summarized by the algebraic (abbreviated ALG) and exponential (abbreviated EXP) cases. For the algebraic case, we need to verify that $n(\varepsilon, S_d)$ is bounded by certain functions of d and ε^{-1} which are, in particular, not exponential in some powers of d and ε^{-1}. For the exponential case, we replace ε^{-1} by $1+\log \varepsilon^{-1}$, and consider the same notions of tractability as for the algebraic case.

Most papers on tractability have dealt with the notions of algebraic tractability, which can be said to be the "standard" case of tractability. For an overview of results we again refer to [17, 18, 19] and the references therein, as well as [20] and [32].

However, there is also a recent stream of work on exponential tractability, and this is what we are going to study here. For results on exponential tractability for general linear problems (without necessarily assuming tensor product structure), we refer the reader to the recent paper [12], for exponential tractability for linear problems on un-weighted tensor product spaces, we refer the reader to [4, 21, 22]. Further results on exponential tractability can, e.g., be found in the papers [1, 2, 6, 7, 8, 10, 11, 13, 14, 15, 16, 23, 27, 28, 31, 33].

For un-weighted tensor product problems, such as studied in [4], we have many negative results for exponential tractability. Our point of departure is to verify if these negative results can be changed if we switch to weighted tensor products with product weights. Indeed, this is the case. We now illustrate a sample of our results leaving the general results to Section 3.

For weighted tensor product problems, the information complexity depends on two non-increasing sequences, $\{\lambda_j\}_{j\in\mathbb{N}}$ and $\{\gamma_j\}_{j\in\mathbb{N}}$. Here, the λ_j's are the squares of the ordered singular values of the univariate operator, and the γ_j's are product weights which moderate the importance of the $j^{\rm th}$ univariate problem in the definition of the tensor product for the multivariate case. Without loss of generality we may assume that $\lambda_1 = 1$ and, to omit the trivial case, that $\lambda_2 > 0$. To make the presentation of our results easier we also assume that all weights γ_j are positive. The un-weighted case is obtained if we take $\gamma_j = 1$ for all $j \in \mathbb{N}$.

The concept of exponential strong polynomial tractability (EXP-SPT) is defined when there are two non-negative numbers C and p such that the information complexity is bounded by

$$C(1+\log\,\varepsilon^{-1})^p \qquad \text{for all } \varepsilon\in(0,1] \text{ and } d\in\mathbb{N}.$$

Similarly, the concept of exponential polynomial tractability (EXP-PT) is defined when there are three non-negative numbers C, q, and p such that the information complexity is bounded by

$$C\,d^q(1+\log\,\varepsilon^{-1})^p \quad \text{for all } \varepsilon\in(0,1] \text{ and } d\in\mathbb{N}.$$

Note that especially EXP-SPT is a quite demanding property since the upper bound on the information complexity must be independent of d and at most polynomially dependent on $1+\log\,\varepsilon^{-1}$. For EXP-PT we allow that the information complexity may depend polynomially on d and $1+\log\,\varepsilon^{-1}$. One might suspect that EXP-SPT and also EXP-PT hold only for extremely small λ_j's and γ_j's. As we shall see, this is indeed the case.

We prove that

- EXP-SPT and EXP-PT are equivalent,
- EXP-SPT holds if and only if

$$\lim_{j\to\infty}\lambda_j=\lim_{j\to\infty}\gamma_j=0 \quad \text{and} \quad B_{\mathrm{EXP-SPT}}:=\limsup_{\varepsilon\to 0}\frac{d(\varepsilon)\,\log\,j(\varepsilon)}{\log\,\log\,\varepsilon^{-1}}<\infty,$$

  where

$$\begin{aligned} d(\varepsilon) &= \max\{\,d\in\mathbb{N}\,:\,\gamma_d\,>\,\varepsilon^2\,\},\\ j(\varepsilon) &= \max\{\,j\in\mathbb{N}\,:\,\lambda_j\,>\,\varepsilon^2\,\}. \end{aligned}$$

  Furthermore, the exponent of EXP-SPT, defined as the infimum of those p for which the estimate on the information complexity for EXP-SPT holds, is equal to $B_{\mathrm{EXP-SPT}}$.

As we now see, the relaxation from EXP-SPT to EXP-PT is not essential. For the un-weighted case, $\gamma_j=1$, both EXP-SPT and EXP-PT do not hold.

Let us check for which λ_j's and γ_j's we have EXP-SPT (and EXP-PT) for the weighted case. Since $d(\varepsilon)$ goes to infinity as ε tends to zero, we see that $\log\,j(\varepsilon)$ must go to infinity slower than $\log\,\log\,\varepsilon^{-1}$. It is easy to check that for $\lambda_j=\exp(-\alpha\,j)$ for some (maybe very large) $\alpha>0$, we get $B_{\mathrm{EXP-SPT}}=\infty$. Consider thus $\lambda_j=\exp(-\exp(\alpha\,j))$, this time with (maybe very small) $\alpha>0$. Again, it is easy to check that now $\log\,j(\varepsilon)/(\log(\log(\varepsilon^{-1})))$ goes to zero. Obviously we cannot yet claim EXP-SPT since it also depends on γ_j's. By the same token we conclude that $d(\varepsilon)$ must go to infinity slower than $\log\,\log\,\varepsilon^{-1}$. So $\gamma_j=\exp(-\exp(\alpha\,j))$ is not enough. For positive α_1,α_2, consider then

$$\lambda_j=\exp(-\exp(\alpha_1\,j)) \quad \text{and} \quad \gamma_j=\exp(-\exp(\exp(\alpha_2\,j))).$$

We now have

$$j(\varepsilon)=\frac{1+o(1)}{\alpha_1}\,\log\,\log\,\varepsilon^{-1} \quad \text{and} \quad d(\varepsilon)=\frac{1+o(1)}{\alpha_2}\,\log\,\log\,\log\,\varepsilon^{-1}.$$

Then $B_{\mathrm{EXP-SPT}}=0$, so that EXP-SPT now holds with the zero exponent. Further examples of different notions of exponential tractability will be presented in Section 3.

Of course, we may say that EXP-SPT (or EXP-PT) is a too strong notion of exponential tractability. We now present a result for a much weaker notion, namely for exponential weak tractability (EXP-WT) which holds if the logarithm of the information complexity divided by $d + \log \varepsilon^{-1}$ goes to zero if $\max(d, \varepsilon^{-1})$ goes to infinity. We prove that EXP-WT holds if and only if

$$\lim_{j\to\infty} \gamma_j = 0 \quad \text{and} \quad \lim_{j\to\infty} \frac{\log \lambda_j^{-1}}{\log j} = \infty.$$

The conditions on the λ_j's and the γ_j's are now much more lenient but still do not hold for the un-weighted case. For the weighted case, we obtain EXP-WT if for $\beta > 1$ we have

$$\lambda_j = \mathcal{O}\left(\exp\left(-(\log j)^{\beta}\right)\right)$$

and γ_j goes to zero arbitrarily slowly.

We also consider other notions of exponential tractability such as EXP-QPT, exponential quasi-polynomial tractability, and EXP-(s,t)-WT, exponential (s,t)-weak tractability for $\max(s,t) \geq 1$. The corresponding necessary and sufficient conditions on these notions of exponential tractability are presented in Theorem 1. The case of exponential (s,t)-weak tractability with $\max(s,t) < 1$ as well as EXP-UWT, exponential uniform weak tractability, are left for future research.

We end the introduction by presenting a couple of other open problems.

- In this paper, we assume the class Λ_d^{all} of all continuous linear functionals as information evaluations. It is of a practical interest to consider the class Λ_d^{std} of only function values. In this case we assume that H_d is a reproducing kernel Hilbert space so that function values are continuous linear functionals. The open problem is to find necessary and sufficient conditions for various notions of exponential tractability for the class Λ_d^{std}, and to compare them to necessary and sufficient conditions for the class Λ_d^{all}. We believe that the recent paper [9] may be very helpful for the solution of this problem.

- As we already mentioned, we consider in this paper only product weights. It would be of interest to study more general weights and to see how the conditions for product weights can be changed.

We summarize the contents of the rest of this paper. In Section 2 we define the problem we study here. In Section 3 we present the results, and in Section 4 the proofs.

2 Problem Setting

We outline the formal setting considered in this paper. Let H_1 be a separable infinite-dimensional Hilbert space with inner product denoted by $\langle \cdot, \cdot \rangle_{H_1}$. Let G_1 be an arbitrary Hilbert space, and let $S_1 : H_1 \to G_1$ be a compact linear operator. We stress that compactness of S_1 is a necessary condition to get a finite information complexity and any type of algebraic or exponential tractability. Then

$$W_1 = S_1^* S_1 \;:\; H_1 \to H_1$$

is also a compact and self-adjoint non-negative operator. Let $(\lambda_j, e_j)_{j\in\mathbb{N}}$ denote its j^{th} eigenpair,

$$W_1 e_j = \lambda_j e_j,$$

where the e_j's are orthonormal and the λ_j's non-increasing. Without loss of generality we assume that $\lambda_1 = 1$ and, to omit the trivial problem, that $\lambda_2 > 0$. Due to compactness of S_1 we have $\lim_{j\to\infty} \lambda_j = 0$.

Let $d \in \mathbb{N}$. Define

$$H_d := \underbrace{H_1 \otimes H_1 \otimes \cdots \otimes H_1}_{d \text{ times}}$$

to be the d-fold tensor product of H_1. The inner product is denoted by $\langle \cdot, \cdot \rangle_{H_d}$. Similarly, define

$$G_d := \underbrace{G_1 \otimes G_1 \otimes \cdots \otimes G_1}_{d \text{ times}}$$

and the d-fold tensor product operator

$$S_d := \underbrace{S_1 \otimes S_1 \otimes \cdots \otimes S_1}_{d \text{ times}} : H_d \to G_d.$$

Obviously, S_d is compact. Then

$$W_d = S_d^* S_d \;:\; H_d \to H_d$$

is also a compact and self-adjoint non-negative operator. Let $\boldsymbol{j} = (j_1, j_2, \ldots, j_d) \in \mathbb{N}^d$. The eigenpairs of W_d are $(\lambda_{d,\boldsymbol{j}}, e_{d,\boldsymbol{j}})_{\boldsymbol{j}\in\mathbb{N}^d}$ with

$$\lambda_{d,\boldsymbol{j}} = \prod_{k=1}^{d} \lambda_{j_k} \quad \text{and} \quad e_{d,\boldsymbol{j}} = e_{j_1} \otimes e_{j_2} \otimes \cdots \otimes e_{j_d}.$$

Hence, we have at least 2^d positive eigenvalues of W_d. The square roots of the $\lambda_{d,\boldsymbol{j}}$ are the singular values of S_d.

In the following we write $[d] := \{1, 2, \ldots, d\}$. Subsets of $[d]$ will be denoted by $\mathfrak{u}$. From [17, Sec. 5.3.1], elements $f \in H_d$ can be decomposed as a sum of mutually orthogonal elements $f_{\mathfrak{u}}$, $\mathfrak{u} \subseteq [d]$, each of which belongs to $\bigotimes_{j\in\mathfrak{u}} H_1$, in the form

$$f = \sum_{\mathfrak{u}\subseteq[d]} f_{\mathfrak{u}}.$$

Furthermore, for $f, g \in H_d$ the inner product is

$$\langle f, g \rangle_{H_d} = \sum_{\mathfrak{u}\subseteq[d]} \langle f_{\mathfrak{u}}, g_{\mathfrak{u}} \rangle_{H_d}$$

and

$$\|f\|_{H_d}^2 = \sum_{\mathfrak{u}\subseteq[d]} \|f_{\mathfrak{u}}\|_{H_d}^2. \tag{1}$$

Further information on this orthogonal decomposition can be found in [17, Sec. 5.3.1].

Eq. (1) shows that the contribution of each $f_{\mathfrak{u}}$ is the same, which suggests that any group of $f_{\mathfrak{u}}$'s is equally important in their contribution to the norm of f. However, in this paper we are interested in the weighted setting which is motivated by the assumption that some groups of $f_{\mathfrak{u}}$'s are more important than others, or that an element f does not depend on some groups of variables at all. Such a behavior can be modeled with the help of so-called weights. Here we restrict ourselves to *product weights* as in the first paper on weighted spaces [26].

Let $\boldsymbol{\gamma} = \{\gamma_j\}_{j\in\mathbb{N}}$ be a sequence of non-increasing positive reals, which are called product weights. The case when some product weights γ_j's are zero is considered in Remark 4. For simplicity we assume that $\gamma_j \in (0,1]$. For $\mathfrak{u} \subseteq [d]$ put

$$\gamma_{\mathfrak{u}} = \prod_{j\in\mathfrak{u}} \gamma_j,$$

where the empty product is considered to be one, i.e., $\gamma_\emptyset = 1$.

Now we define the weighted Hilbert space $H_{d,\boldsymbol{\gamma}}$ as a separable Hilbert space that is algebraically the same as the space H_d but whose inner product for $f, g \in H_d$ is given by

$$\langle f, g\rangle_{H_{d,\boldsymbol{\gamma}}} := \sum_{\mathfrak{u}\subseteq[d]} \gamma_{\mathfrak{u}}^{-1} \langle f_{\mathfrak{u}}, g_{\mathfrak{u}}\rangle_{H_d}. \tag{2}$$

For product weights, the space $H_{d,\boldsymbol{\gamma}}$ can also be described as a tensor product space, namely,

$$H_{1,\gamma_1} \otimes H_{1,\gamma_2} \otimes \cdots \otimes H_{1,\gamma_d}.$$

In particular, if $\gamma_j = 1$ for all $j \in \mathbb{N}$, then we have $H_{d,\boldsymbol{\gamma}} = H_d$.

We study the sequence $\mathcal{S}_{\boldsymbol{\gamma}} = \{S_{d,\boldsymbol{\gamma}}\}_{d\in\mathbb{N}}$ of operators given by

$$S_{d,\boldsymbol{\gamma}} : H_{d,\boldsymbol{\gamma}} \to G_d, \quad S_{d,\boldsymbol{\gamma}}(f) = S_d(f)$$

and consider the problem of approximating $S_{d,\boldsymbol{\gamma}}(f)$ in the norm of G_d for elements f from the unit ball of $H_{d,\boldsymbol{\gamma}}$. We shall show in the next section how the weights affect the singular values of $S_{d,\boldsymbol{\gamma}}$.

The elements $S_{d,\boldsymbol{\gamma}}(f)$ are approximated by algorithms $A_{d,n}(f)$ which use at most n information evaluations from the class $\Lambda_d^{\mathrm{all}} = H_{d,\boldsymbol{\gamma}}^*$ which consists of all continuous linear functionals defined on $H_{d,\boldsymbol{\gamma}}$. The general form of an algorithm $A_{d,n}$ is

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

where $L_j \in \Lambda_d^{\mathrm{all}}$ and $\phi_{d,n} : \mathbb{R}^n \to G_d$ is any mapping. The choice of L_j can be adaptive, i.e., it may depend on the previously computed information $L_1(f), L_2(f), \ldots, L_{j-1}(f)$.

The error is studied in the worst-case setting and is defined as

$$e(A_{d,n}) = \sup_{\substack{f\in H_{d,\boldsymbol{\gamma}} \\ \|f\|_{H_{d,\boldsymbol{\gamma}}}\le 1}} \|S_{d,\boldsymbol{\gamma}}(f) - A_{d,n}(f)\|_{G_d}.$$

Denote by e_0 the initial error, i.e.,

$$e_0 = \sup_{\substack{f\in H_{d,\boldsymbol{\gamma}} \\ \|f\|_{H_{d,\boldsymbol{\gamma}}}\le 1}} \|S_{d,\boldsymbol{\gamma}}(f)\|_{G_d},$$

which is just the operator norm of $S_{d,\boldsymbol{\gamma}}$.

We are interested in the minimal number n of information evaluations from the class $\Lambda_d^{\rm all}$ in order to reduce the initial error by a factor of $\varepsilon \in (0,1]$. To this end let

$$e(n, S_{d,\boldsymbol{\gamma}}) = \inf_{A_{d,n}} e(A_{d,n})$$

be the $n^{\rm th}$ minimal error, where the infimum is extended over all algorithms $A_{d,n}$ which use at most n information evaluations from the class $\Lambda_d^{\rm all}$. Then we study the information complexity for the normalized error criterion, which is defined by

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) = \min\{n \,:\, e(n, S_{d,\boldsymbol{\gamma}}) \le \varepsilon\, e_0\}.$$

It is known that for the class $\Lambda_d^{\rm all}$ the information complexity is fully characterized in terms of the singular values of $S_{d,\boldsymbol{\gamma}}$, or equivalently, in terms of the eigenvalues of $W_{d,\boldsymbol{\gamma}} = S_{d,\boldsymbol{\gamma}}^* S_{d,\boldsymbol{\gamma}}$.

We are interested in the behavior of $n(\varepsilon, S_{d,\boldsymbol{\gamma}})$ when $\max(d, \varepsilon^{-1})$ goes to infinity in an arbitrary way. This is the subject of tractability, see [17, 18, 19]. The notions of tractability classify the order of growth of the information complexity. Standard tractability is studied in the *algebraic setting* (ALG), as it is called nowadays. In this case, one describes the dependence of $n(\varepsilon, S_{d,\boldsymbol{\gamma}})$ on the dimension and the error threshold, i.e., with respect to the pair (d, ε). Recently also the *exponential setting* (EXP) gained much attention, and this setting is the central topic of the present paper. In the exponential setting, we study how $n(\varepsilon, S_{d,\boldsymbol{\gamma}})$ behaves with respect to the pair $(d, 1 + \log \varepsilon^{-1})$. We are ready to define various notions of EXP tractabilities.

The problem $S_{\boldsymbol{\gamma}} = \{S_{d,\boldsymbol{\gamma}}\}$ is said to be:

- *Exponentially strongly polynomially tractable (EXP-SPT)* if there are $C, p \ge 0$ such that
$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le C(1 + \log \varepsilon^{-1})^p \quad \forall d \in \mathbb{N},\ \forall \varepsilon \in (0,1]. \tag{3}$$
The infimum over all exponents $p \ge 0$ such that (3) holds for some $C \ge 0$ is called the exponent of EXP-SPT and is denoted by p^*.

- *Exponentially polynomially tractable (EXP-PT)* if there are $C, p, q \ge 0$ such that
$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le C d^q (1 + \log\, \varepsilon^{-1})^p \quad \forall d \in \mathbb{N},\ \forall \varepsilon \in (0,1].$$

- *Exponentially quasi-polynomially tractable (EXP-QPT)* if there are $C, t \ge 0$ such that
$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le C \exp(t(1 + \log\, d)(1 + \log(1 + \log\, \varepsilon^{-1}))) \quad \forall d \in \mathbb{N},\ \forall \varepsilon \in (0,1]. \tag{4}$$
The infimum over all exponents $t \ge 0$ such that (4) holds for some $C \ge 0$ is called the exponent of EXP-QPT and is denoted by t^*.

- *Exponentially (s,t)-weakly tractable (EXP-(s,t)-WT)* for positive s and t if
$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\log \max(1, n(\varepsilon, S_{d,\boldsymbol{\gamma}}))}{d^t + (1 + \log \varepsilon^{-1})^s} = 0.$$
If $s = t = 1$, we speak of *exponential weak tractability (EXP-WT)*.

- *Exponentially uniformly weakly tractable (EXP-UWT)* if EXP-(s,t)-WT holds for all positive s and t.

To shorten the notation, we often say that the problem $S_{\boldsymbol{\gamma}}$ is EXP-SPT, EXP-PT, etc., by saying that EXP-SPT, EXP-PT, etc., holds. As already mentioned, we do not consider EXP-(s,t)-WT with $\max(s,t)<1$ and exponential uniform tractability in this paper.

Remark 1. In some papers, for example in [22], the notion of EXP-(s,t)-WT is called $(s,\ln^{\kappa})$-weak tractability, where s corresponds to t and κ corresponds to s in our notation.

The notions of quasi-polynomial, (s,t)-weak and uniform weak tractabilities in the algebraic case were for the first time defined correspondingly in [3, 24, 25]. Here we adopt these concepts for exponential tractability by replacing ε^{-1} by $1+\log\varepsilon^{-1}$.

The main result of this work, a characterization of weighted linear tensor product problems with respect to exponential tractability, will be stated in the next section as Theorem 1. The proofs will be presented in Section 4.

3 The results

To begin with we introduce another, for our purpose more convenient, representation of the information complexity. It is known from [29], see also [17], how the information complexity depends on the singular values of $S_{d,\boldsymbol{\gamma}}$, which are the same as the square-roots of the eigenvalues of the compact self-adjoint and positive definite linear operator $W_{d,\boldsymbol{\gamma}} = S_{d,\boldsymbol{\gamma}}^* S_{d,\boldsymbol{\gamma}} : H_{d,\boldsymbol{\gamma}} \to H_{d,\boldsymbol{\gamma}}$.

Let $\{\lambda_j\}_{j\in\mathbb{N}}$ and $\{\gamma_j\}_{j\in\mathbb{N}}$ be as in the previous section. For $\boldsymbol{j}=(j_1,\dots,j_d)\in\mathbb{N}^d$ define

$$\mathfrak{u}(\boldsymbol{j}) := \{k\in[d] \ : \ j_k\ge 2\} \quad \text{and} \quad \lambda_{d,\boldsymbol{j}} := \lambda_{j_1}\lambda_{j_2}\cdots\lambda_{j_d}.$$

From [17, Section 5.3]) we know that the eigenvalues of $W_{d,\boldsymbol{\gamma}}$ are

$$\lambda_{d,\boldsymbol{\gamma},\boldsymbol{j}} := \gamma_{\mathfrak{u}(\boldsymbol{j})}\lambda_{d,\boldsymbol{j}} = \left(\prod_{\substack{k=1\\ j_k\ge 2}}^{d}\gamma_k\right)\lambda_{j_1}\cdots\lambda_{j_d} = \left(\prod_{k\in\mathfrak{u}(\boldsymbol{j})}\gamma_k\right)\lambda_{d,\boldsymbol{j}}.$$

Clearly, $\lambda_{d,\boldsymbol{\gamma},\boldsymbol{j}}$ is maximized for $\boldsymbol{j}=(1,1\dots,1)$ and then it is equal to 1. Hence, $\|W_{d,\boldsymbol{\gamma}}\| = \|S_{d,\boldsymbol{\gamma}}\| = 1$ and the initial error e_0 is also one. This means that the problem is well normalized for all $d\in\mathbb{N}$ and all product weights $\boldsymbol{\gamma}$.

The information complexity is now

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) = \left|\left\{\boldsymbol{j}\in\mathbb{N}^d \ : \ \lambda_{d,\boldsymbol{\gamma},\boldsymbol{j}} > \varepsilon^2\right\}\right|. \tag{5}$$

Define

$$\lambda_{k,j} = \begin{cases} 1 & \text{if } j=1, \\ \gamma_k\lambda_j & \text{if } j\ge 2. \end{cases}$$

Then

$$\lambda_{d,\boldsymbol{\gamma},\boldsymbol{j}} = \prod_{k=1}^{d}\lambda_{k,j_k}$$

and hence

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) = |A_{\varepsilon,d}|, \quad \text{where } A_{\varepsilon,d} = \{(n_1,\dots,n_d)\in\mathbb{N}^d \ : \ \lambda_{1,n_1}\cdots\lambda_{d,n_d} > \varepsilon^2\}. \tag{6}$$

Clearly,

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le n(\varepsilon_1, S_{d_1,\boldsymbol{\gamma}}) \quad \text{for all } \varepsilon_1 \le \varepsilon \text{ and } d_1 \ge d.$$

Hence, for decreasing ε and increasing d, the information complexity is non-increasing.

In the sequel we will work with the representation of the information complexity in (6). We show how weighted tensor product problems can be classified with respect to different notions of EXP tractability by means of the eigenvalues $\{\lambda_j\}_{j\in\mathbb{N}}$ of the operator $W_1 := S^*S : H_1 \to H_1$ and of the weights $\{\gamma_j\}_{j\in\mathbb{N}}$. We remind the reader of what we assume about the λ_j's and γ_j's. We have

$$1 = \lambda_1 \ge \lambda_2 \ge \lambda_3 \ge \ldots \ge 0, \quad \text{with } \lambda_2 > 0. \tag{7}$$

Note that for $\lambda_2 = 0$ the problem becomes trivial since $n(\varepsilon, S_{d,\boldsymbol{\gamma}}) = 1$ for all $\varepsilon \in [0,1)$ and $d \in \mathbb{N}$. On the other hand, if $\lambda_1 \ne 1$ then the problem is not well normalized. In this case, the initial error e_0 is λ_1^d and for the normalized error criterion we may work with $\lambda_{k,n_j}/\lambda_1$ instead of λ_{k,n_j}. By assuming that $\lambda_1 = 1$ we simplify the notation. Note also that $\lim_{j\to\infty}\lambda_j = 0$ iff S_1 (as well as $S_{d,\boldsymbol{\gamma}}$) is compact. This assumption implies that the information complexity $n(\varepsilon, S_{d,\boldsymbol{\gamma}})$ is finite for all $\varepsilon > 0$ and all $d \in \mathbb{N}$.

For $\varepsilon \in (0,1)$ and $\lim_{j\to\infty}\lambda_j = 0$, define

$$j(\varepsilon) = \max\{j \in \mathbb{N} \ : \ \lambda_j > \varepsilon^2\}. \tag{8}$$

Then $j(\varepsilon)$ is well defined and always finite. Since $\lambda_1 = 1$, we have $j(\varepsilon) \ge 1$. Furthermore, $j(\varepsilon)$ goes to infinity if and only if all λ_j's are positive.

We also assume that the weights satisfy

$$1 \ge \gamma_1 \ge \gamma_2 \ge \gamma_3 \ge \ldots > 0. \tag{9}$$

The ordering of the γ_j's tells us that the successive subproblems are less and less important. The assumption that the weights are at most one is made for simplicity to guarantee that $\lambda_{k,j} \le \lambda_{k,1} = 1$. The case of more general λ_j's and γ_j's is considered for algebraic tractability in [17, Section 5.3].

For $\varepsilon \in (0,1)$ and $\lim_{j\to\infty}\gamma_j = 0$, define

$$d(\varepsilon) = \max\{d \in \mathbb{N} \ : \ \gamma_d > \varepsilon^2\}.$$

Then $d(\varepsilon)$ is well defined. We put $d(\varepsilon) = 0$ for $\gamma_1 \le \varepsilon^2$, and note that $d(\varepsilon) \ge 1$ for $\gamma_1 > \varepsilon^2$. Both $j(\varepsilon)$ and $d(\varepsilon)$ are non-decreasing, and $\lim_{\varepsilon\to 0} d(\varepsilon) = \infty$.

Now we are able to state our main result. To shorten the notation we write "iff" instead of "if and only if".

Theorem 1. *We have*

1. *EXP-SPT holds iff*

$$\lim_{j\to\infty}\lambda_j = \lim_{j\to\infty}\gamma_j = 0 \quad \textit{and} \quad B_{\mathrm{EXP-SPT}} := \limsup_{\varepsilon\to 0} \frac{d(\varepsilon)\log j(\varepsilon)}{\log\log\frac{1}{\varepsilon}} < \infty.$$

If this holds then the exponent of EXP-SPT is $p^* = B_{\mathrm{EXP-SPT}}$.

2. *EXP-SPT and EXP-PT are equivalent.*

3. *EXP-QPT holds iff*

$$\lim_{j\to\infty} \lambda_j = \lim_{j\to\infty} \gamma_j = 0 \quad \text{and} \quad B_{\mathrm{EXP-QPT}} := \limsup_{\varepsilon\to 0} \frac{d(\varepsilon)\log j(\varepsilon)}{[\log d(\varepsilon)]\log\log\frac{1}{\varepsilon}} < \infty.$$

If this holds then the exponent of EXP-QPT is $t^* = B_{\mathrm{EXP-QPT}}$.

4. *Let* $s = t = 1$.

EXP-WT holds iff

$$\lim_{j\to\infty} \gamma_j = 0 \quad \text{and} \quad \lim_{j\to\infty} \frac{\log\frac{1}{\lambda_j}}{\log j} = \infty.$$

5. *Let* $s = 1$ *and* $t < 1$.

EXP-$(1,t)$*-WT holds iff*

$$\lim_{j\to\infty} \frac{\log\frac{1}{\gamma_j}}{\log j} = \infty \quad \text{and} \quad \lim_{j\to\infty} \frac{\log\frac{1}{\lambda_j}}{\log j} = \infty.$$

6. *Let* $s = 1$ *and* $t > 1$.

EXP-(s,t)*-WT holds iff*

$$\gamma_j\text{'s are arbitrary} \quad \text{and} \quad \lim_{j\to\infty} \frac{\log\frac{1}{\lambda_j}}{\log j} = \infty.$$

7. *Let* $s > 1$, $t \le 1$ *and* $\lambda_2 < 1$.

EXP-(s,t)*-WT holds iff*

$$\gamma_j\text{'s are arbitrary} \quad \text{and} \quad \lim_{j\to\infty} \frac{(\log\frac{1}{\lambda_j})^s}{\log j} = \infty.$$

8. *Let* $s > 1$, $t \le 1$ *and* $\lambda_2 = 1$.

EXP-(s,t)*-WT holds iff*

$$\exists\, p \in \mathbb{N} \ \text{with} \ \gamma_p < 1 \quad \text{and} \quad \lim_{j\to\infty} \frac{(\log\frac{1}{\lambda_j})^s}{\log j} = \infty.$$

9. *Let* $s > 1$ *and* $t > 1$.

EXP-(s,t)*-WT holds iff*

$$\gamma_j\text{'s are arbitrary} \quad \text{and} \quad \lim_{j\to\infty} \frac{(\log\frac{1}{\lambda_j})^s}{\log j} = \infty.$$

10. *Let* $s < 1$ *and* $t > 1$. *EXP-*(s,t)*-WT holds for arbitrary* γ_j*'s iff*

$$\lim_{j\to\infty} \frac{\left(\log\frac{1}{\lambda_j}\right)^\eta}{\log j} = \infty \quad \text{with} \quad \eta = \frac{s(t-1)}{t-s}.$$

11. Let $s < 1$ and $t = 1$.

EXP-$(s,1)$-WT holds iff for arbitrary integers d, k, j with $j \geq 2$ and $k \leq d$ it is true that

$$\lim_{d+\gamma_k^{-d}\lambda_j^{-d}\to\infty} \frac{\left(\log \frac{1}{\gamma_k}\right)^s + \left(\log \frac{1}{\lambda_j}\right)^s}{d^{1-s}\log j} = \infty. \tag{10}$$

Before we present the proof of Theorem 1 we illustrate some of the results and discuss their meaning.

Example 1. Let $\lambda_j = \exp(-\exp(j^\alpha))$ and $\gamma_j = \exp(-\exp(j^\beta))$ for positive α and β. Then we have

$$j(\varepsilon) = \left\lceil \left(\log\log\frac{1}{\varepsilon^2}\right)^{1/\alpha}\right\rceil - 1 \quad \text{and} \quad d(\varepsilon) = \left\lceil \left(\log\log\frac{1}{\varepsilon^2}\right)^{1/\beta}\right\rceil - 1$$

and hence

- $\alpha > 0$ and $\beta > 1$ imply EXP-SPT with $p^* = 0$;
- $\alpha > 0$ and $\beta = 1$ imply EXP-QPT with $t^* = 1$, but EXP-SPT does not hold.

Note that Items *4.-11.* of Theorem 1 give a full characterization of EXP-(s,t)-WT for all $(s,t) \in (0,\infty)^2 \setminus (0,1)^2$. The following remarks are in order.

Remark 2. The condition

$$\lim_{j\to\infty} \frac{\log\frac{1}{\lambda_j}}{\log j} = \infty$$

is satisfied if and only if λ_j is of the form

$$\lambda_j = \frac{1}{j^{h(j)}}$$

where $h : \mathbb{N} \to \mathbb{R}$ satisfies $\lim_{j\to\infty} h(j) = \infty$. So, for example, we have EXP-WT if $\lim_{j\to\infty}\gamma_j = 0$ and $\lambda_1 = 1$, $\lambda_2 = 0.95$ and $\lambda_j = j^{-\log\log j}$ for $j \geq 3$.

Remark 3. Consider $s > 1$, $t \leq 1$ and $\lambda_2 \leq 1$ described in Items *7.* and *8.* of Theorem 1: For $\lambda_2 < 1$, we have a single largest eigenvalue and EXP-(s,t)-WT holds for arbitrary γ_j's as long as $\lim_{j\to\infty}(\log \lambda_j^{-1})^s/\log j = \infty$. In particular, this holds for the un-weighted case, $\gamma_j = 1$ for all $j \in \mathbb{N}$.

For $\lambda_2 = 1$, we have a multiple largest eigenvalue and EXP-(s,t)-WT holds under the same conditions on the λ_j's but now we need to assume that not all γ_j's are one. In particular, this holds for

$$1 = \gamma_1 = \ldots = \gamma_{p-1} > \gamma_p = \gamma_{p+1} = \ldots > 0.$$

Remark 4. Consider $s \geq 1$ and $t > 1$ described in Items *6.* and *9.* of Theorem 1. Then EXP-(s,t)-WT holds for arbitrary γ_j's, i.e., even for the un-weighted case $\gamma_j = 1$, and for λ_j satisfying the same condition as before. This case was proved in [4].

Remark 5. We briefly note what happens if some weights in (9) are zero, say $\gamma_{j^*} = 0$ for some $j^* \in \mathbb{N}$. Obviously, monotonicity of the γ_j's implies that $\gamma_j = 0$ for all $j \geq j^*$. Then $\gamma_{\mathfrak{u}} = 0$ for all $\mathfrak{u}$ containing one or more indices at least equal to j^*. For such $\mathfrak{u}$, we must assume in (2) that $f_{\mathfrak{u}} = 0$ and adopt the convention that $0/0 = 0$. In this case, $H_{d,\boldsymbol{\gamma}}$ is algebraically a proper subset of H_d.

Assume first that $j^* = 1$. Then the only non-zero eigenvalues are $\lambda_{k,1} = 1$. This means that the problem is trivial since $n(\varepsilon, S_{d,\boldsymbol{\gamma}}) = 1$ for all $\varepsilon \in [0,1)$ and $d \in \mathbb{N}$.

Assume then that $j^* \geq 2$. It is easy to check that we now have

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \leq n(\varepsilon, S_{j^*-1,\boldsymbol{\gamma}}) \quad \text{for all } \varepsilon \in (0,1) \text{ and } d \in \mathbb{N}.$$

Hence, $d(\varepsilon) \leq j^* - 1$ and $d(\varepsilon) = j^* - 1$ for $\varepsilon < \gamma_{j^*-1}^{1/2}$. The factors $d(\varepsilon)$ and $d(\varepsilon)/\log d(\varepsilon)$ cannot change the fact that $B_{\text{EXP}-\text{SPT}}$ or $B_{\text{EXP}-\text{QPT}}$ are finite, and then EXP-SPT, EXP-PT and EXP-QPT are equivalent.

4 The proofs

We first show how the information complexity can be bounded in terms of $j(\varepsilon)$ and $d(\varepsilon)$.

Lemma 1. *If* $\lim_{j\to\infty} \lambda_j = \lim_{j\to\infty} \gamma_j = 0$ *then for* $\varepsilon \in (0,1)$ *we have*

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \leq j(\varepsilon)^{\min(d,d(\varepsilon))} \leq n(\varepsilon^{2d(\varepsilon)}, S_{d(\varepsilon),\boldsymbol{\gamma}}),$$

and for $d \geq d(\varepsilon)$

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) = n(\varepsilon, S_{d(\varepsilon),\boldsymbol{\gamma}}).$$

Proof. We use (6). Consider the eigenvalue $\lambda_{1,n_1}\lambda_{2,n_2}\cdots\lambda_{d,n_d}$.

- If $n_k \geq j(\varepsilon) + 1$ (in particular $n_k \geq 2$) for some $k \leq d$, then we have

  $$\lambda_{1,n_1}\lambda_{2,n_2}\cdots\lambda_{d,n_d} \leq \lambda_{k,n_k} = \gamma_k\lambda_{n_k} \leq \lambda_{n_k} \leq \varepsilon^2$$

  so that $(n_1, \dots, n_d) \notin A_{\varepsilon,d}$.

- If $d \geq d(\varepsilon) + 1$ and $n_d \geq 2$, then

  $$\lambda_{1,n_1}\lambda_{2,n_2}\cdots\lambda_{d,n_d} \leq \gamma_d\lambda_2 \leq \gamma_d \leq \varepsilon^2$$

  and again $(n_1, \dots, n_d) \notin A_{\varepsilon,d}$.

Hence, only

$$(n_1, n_2, \dots, n_{\min(d,d(\varepsilon))}, 1, 1, \dots, 1) \in \mathbb{N}^d$$

for $n_j \in \{1, 2, \dots, j(\varepsilon)\}$ with $j = 1, 2, \dots, \min(d, d(\varepsilon))$ may belong to $A_{\varepsilon,d}$, and therefore

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) = |A_{\varepsilon,d}| \leq j(\varepsilon)^{\min(d,d(\varepsilon))}.$$

Furthermore for $d \geq d(\varepsilon)$, we have

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) = |A_{\varepsilon,d(\varepsilon)}| = n(\varepsilon, S_{d(\varepsilon),\boldsymbol{\gamma}}),$$

as claimed.

In order to show the remaining inequality we consider the eigenvalues

$$\lambda_{1,n_1}\lambda_{2,n_2}\cdots\lambda_{d(\varepsilon),n_{d(\varepsilon)}} \text{ for } n_j \in \{1,\dots,j(\varepsilon)\}.$$

For these eigenvalues we have

$$\lambda_{1,n_1}\lambda_{2,n_2}\cdots\lambda_{d(\varepsilon),n_{d(\varepsilon)}} \ge \left(\gamma_{d(\varepsilon)}\lambda_{j(\varepsilon)}\right)^{d(\varepsilon)} > \varepsilon^{4d(\varepsilon)}.$$

This implies that we have at least $j(\varepsilon)^{d(\varepsilon)}$ eigenvalues no less than $\varepsilon^{4d(\varepsilon)}$. Hence

$$j(\varepsilon)^{\min(d,d(\varepsilon))} \le j(\varepsilon)^{d(\varepsilon)} \le n(\varepsilon^{2d(\varepsilon)}, S_{d(\varepsilon),\boldsymbol{\gamma}}).$$

This completes the proof. □

The next technical lemma will help to state the conditions for various notions of exponential tractability in a concise form.

Lemma 2. *Let $\{a_j\}_{j\in\mathbb{N}}$ be a non-increasing sequence of positive reals. Then we have*

$$M_c := \sum_{j=1}^{\infty} a_j^c < \infty \qquad \textit{for all } c > 0 \tag{11}$$

if and only if

$$\lim_{j\to\infty} \frac{\log \frac{1}{a_j}}{\log j} = \infty. \tag{12}$$

Proof. Assume that (11) holds. We have

$$n a_n^c \le a_1^c + \cdots + a_n^c \le M_c$$

and hence

$$\frac{1}{a_n} \ge \frac{n^{1/c}}{M_c^{1/c}}.$$

Taking the logarithm we obtain

$$\log\frac{1}{a_n} \ge \frac{1}{c}\log n - \frac{1}{c}\log M_c$$

and therefore

$$\liminf_{n\to\infty} \frac{\log\frac{1}{a_n}}{\log n} \ge \frac{1}{c}.$$

Now (12) follows by letting $c \to 0$.

If (12) holds then for every $c > 0$ there exists a number $j_c > 0$ such that

$$\frac{\log\frac{1}{a_j}}{\log j} \ge \frac{2}{c} \quad \text{for all } j \ge j_c.$$

This implies

$$a_j^c \le \frac{1}{j^2} \quad \text{for all } j \ge j_c.$$

Hence (11) holds. □

Lemma 3. *For $s, a_1, a_2, \dots, a_m \geq 0$ and $m \in \mathbb{N}$, we have*

$$(a_1 + \cdots + a_m)^s = \alpha_{s,m}\,(a_1^s + \cdots + a_m^s) \quad \text{with} \quad 0^0 = 1,$$

where $\alpha_{s,m}$ also depends on $a_1, \dots, a_m$ but is uniformly bounded in the a_j's,

$$\alpha_{s,m} \in [1, m^{s-1}] \ \text{ for } s \geq 1 \quad \text{and} \quad \alpha_{s,m} \in [m^{s-1}, 1] \ \text{ for } s < 1.$$

Proof. It is well known that for $s > 1$ we have

$$m^{1-s}(a_1 + \cdots + a_m)^s \leq a_1^s + \cdots + a_m^s \leq (a_1 + \cdots + a_m)^s,$$

whereas for $s < 1$ we have

$$(a_1 + \cdots + a_m)^s \leq a_1^s + \cdots + a_m^s \leq m^{1-s}(a_1 + \cdots + a_m)^s.$$

This can be rewritten as

$$(a_1 + \cdots + a_m)^s = \alpha_{s,m}(a_1^s + \cdots + a_m^s)$$

with $\alpha_{s,m}$ satisfying the bounds in Lemma 3. □

We need a necessary condition on EXP-(s,t)-WT.

Lemma 4. *For any positive s, t and integers $k_1, k_2, \dots, k_d$ with $k_j \geq 2$, EXP-(s,t)-WT implies that*

$$\lim_{d+\max_{j\in[d]} k_j \to \infty} \frac{d^t + \left(\sum_{j=1}^d \log \frac{1}{\gamma_j}\right)^s + \left(\sum_{j=1}^d \log \frac{1}{\lambda_{k_j}}\right)^s}{\sum_{j=1}^d \log\, k_j} = \infty.$$

Proof. EXP-(s,t)-WT implies that

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\log n(\varepsilon, S_{d,\boldsymbol{\gamma}})}{d^t + (\log \varepsilon^{-1})^s} = 0.$$

Take $\varepsilon^2 = \gamma_1 \lambda_{k_1} \cdots \gamma_d \lambda_{k_d} \alpha$ with $\alpha < 1$. We take k_d large enough so that $\varepsilon < 1$. Then $\lambda_{1,j_1} \cdots \lambda_{d,j_d} > \varepsilon^2$ for all $j_1 = 1, 2, \dots, k_1,\ j_2 = 1, 2, \dots, k_2$ and $j_d = 1, 2, \dots, k_d$. Hence $n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \geq \prod_{j=1}^d k_j$ and

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\sum_{j=1}^d \log k_j}{d^t + \frac{1}{2^s}\left(\sum_{j=1}^d \log \frac{1}{\gamma_j} + \sum_{j=1}^d \log \frac{1}{\lambda_{k_j}} + \log \frac{1}{\alpha}\right)^s} = 0.$$

Applying now Lemma 3 with $m = 3$ we obtain

$$\left(\sum_{j=1}^d \log \frac{1}{\gamma_j} + \sum_{j=1}^d \log \frac{1}{\lambda_{k_j}} + \log \frac{1}{\alpha}\right)^s$$
$$= \left(\left(\sum_{j=1}^d \log \frac{1}{\gamma_j}\right)^s + \left(\sum_{j=1}^d \log \frac{1}{\lambda_{k_j}}\right)^s + \left(\log \frac{1}{\alpha}\right)^s\right) \alpha_{s,3}.$$

Note that $d+\varepsilon^{-1}\to\infty$ is equivalent to $d+\max_{j\in[d]}k_j\to\infty$. Taking the reciprocal this yields

$$\lim_{d+\max_{j\in[d]}k_j\to\infty}\frac{d^t+\frac{\alpha_{s,3}}{2^s}\left(\left(\sum_{j=1}^d\log\frac{1}{\gamma_j}\right)^s+\left(\sum_{j=1}^d\log\frac{1}{\lambda_{k_j}}\right)^s+\left(\log\frac{1}{\alpha}\right)^s\right)}{\sum_{j=1}^d\log\,k_j}=\infty.$$

Since $(\log\frac{1}{\alpha})^s/(\sum_{j=1}^d\log k_j)$ tends to zero, and since we may increase the numerator of the last expression by multiplying d^t by $\max(1,\alpha_{s,3}/2^s)$, we obtain Lemma 4. □

We are ready to turn to the proof of the main result of the paper.

Proof of Theorem 1.

1. Assume first that $\lim_{j\to\infty}\lambda_j=\lim_{j\to\infty}\gamma_j=0$ and $B:=B_{\text{EXP-SPT}}<\infty$. Let $\delta\in(0,\infty)$. Then from the definition of B we have that there exists an $\varepsilon_\delta\in(0,\gamma_1^{1/2}/\mathrm{e})$, with $\mathrm{e}=\exp(1)$, such that

$$\log j(\varepsilon)^{d(\varepsilon)}\le\log\left(\log\frac{1}{\varepsilon}\right)^{B+\delta}\qquad\text{for all }\varepsilon\in(0,\varepsilon_\delta].$$

Hence, according to Lemma 1 we have

$$n(\varepsilon,S_{d,\boldsymbol{\gamma}})\le\left(\log\frac{1}{\varepsilon}\right)^{B+\delta}\le\left(1+\log\frac{1}{\varepsilon}\right)^{B+\delta}\qquad\text{for all }\varepsilon\in(0,\varepsilon_\delta].$$

Consider now $\varepsilon\in[\varepsilon_\delta,1]$ and $d\in\mathbb{N}$. Let

$$C_\delta=\left(1+\log\frac{1}{\varepsilon_\delta}\right)^{B+\delta}.$$

Since

$$\left(1+\log\frac{1}{\varepsilon_\delta}\right)^{B+\delta}\le C_\delta\left(1+\log\frac{1}{\varepsilon}\right)^{B+\delta},$$

we have

$$n(\varepsilon,S_{d,\boldsymbol{\gamma}})\le n(\varepsilon_\delta,S_{d,\boldsymbol{\gamma}})\le\left(1+\log\frac{1}{\varepsilon_\delta}\right)^{B+\delta}\le C_\delta\left(1+\log\frac{1}{\varepsilon}\right)^{B+\delta}.$$

Hence,

$$n(\varepsilon,S_{d,\boldsymbol{\gamma}})\le C_\delta\left(1+\log\frac{1}{\varepsilon}\right)^{B+\delta}\qquad\text{for all }\varepsilon\in(0,1)\text{ and }d\in\mathbb{N},$$

which means EXP-SPT with the exponent $p^*\le B$.

On the other hand, assume that we have EXP-SPT with the exponent $p^*<\infty$. Hence, for every $\delta>0$ there exists a number C_δ such that for all $\varepsilon\in(0,1)$ and $d\in\mathbb{N}$

$$n(\varepsilon,S_{d,\boldsymbol{\gamma}})\le C_\delta\left(1+\log\frac{1}{\varepsilon}\right)^{p^*+\delta}.$$

We first show that $\lim_{j\to\infty}\lambda_j = \lim_{j\to\infty}\gamma_j = 0$ which is even known for the algebraic complexity. For completeness we provide a short proof.

The condition $\lim_{j\to\infty}\lambda_j = 0$ easily follows from the compactness of $S_{d,\boldsymbol{\gamma}}$ since otherwise $n(\varepsilon, S_{d,\boldsymbol{\gamma}}) = \infty$ for small positive ε.

The condition $\lim_{j\to\infty}\gamma_j = 0$ is also easy to show since otherwise due to the monotonicity of the γ_j's we have $\lim_{j\to\infty}\gamma_j = \gamma^* > 0$. Then we can take 2^d eigenvalues $\lambda_{1,n_1}\lambda_{2,n_2}\cdots\lambda_{d,n_d}$ with $n_j \in \{1,2\}$. Then each such eigenvalue is at least $(\gamma^*\lambda_2)^d$ and

$$n(\tfrac{1}{2}(\gamma^*\lambda_2)^{d/2}, S_{d,\boldsymbol{\gamma}}) \geq 2^d. \tag{13}$$

This contradicts EXP-SPT.

We now apply EXP-SPT for $n(\varepsilon^{2d(\varepsilon)}, S_{d(\varepsilon)})$ with $d = d(\varepsilon)$. Due to the second inequality in Lemma 1 with $\varepsilon \in (0,1)$ we have

$$j(\varepsilon)^{d(\varepsilon)} \leq n(\varepsilon^{2d(\varepsilon)}, S_{d(\varepsilon),\boldsymbol{\gamma}}) \leq C_\delta \left[1 + \log\left(\frac{1}{\varepsilon}\right)^{2d(\varepsilon)}\right]^{p^*+\delta}.$$

This yields

$$j(\varepsilon)^{d(\varepsilon)} \leq C_\delta \left[(1+2d(\varepsilon))\log\frac{1}{\varepsilon}\right]^{p^*+\delta},$$

and hence

$$\begin{aligned} d(\varepsilon)\log j(\varepsilon) &\leq (p^*+\delta)\left[\log(1+2d(\varepsilon)) + \log\log\frac{1}{\varepsilon}\right] + \log C_\delta \\ &\leq (p^*+\delta)\left[\log d(\varepsilon) + \log\log\frac{1}{\varepsilon}\right] + (p^*+\delta)\log 3 + \log C_\delta. \end{aligned}$$

Now, since $d(\varepsilon)\to\infty$ and $\log\log\frac{1}{\varepsilon}\to\infty$ as $\varepsilon\to 0$ it follows that

$$\limsup_{\varepsilon\to 0} \frac{d(\varepsilon)\log j(\varepsilon)}{\log d(\varepsilon) + \log\log\frac{1}{\varepsilon}} \leq p^* + \delta. \tag{14}$$

For $\varepsilon \to 0$, we have

$$\frac{d(\varepsilon)}{\log d(\varepsilon)} \to \infty$$

and $\log j(\varepsilon)\to\infty$ or $\log j(\varepsilon)\to\log k$, with $k\geq 2$, where the latter case appears when $\lambda_1\geq\lambda_2\geq\ldots\geq\lambda_k>0$ and $\lambda_j = 0$ for $j\geq k+1$. This means that we have

$$p^*+\delta \geq \limsup_{\varepsilon\to 0} \underbrace{\frac{d(\varepsilon)}{\log d(\varepsilon)}}_{\to\infty} \frac{\overbrace{\log j(\varepsilon)}^{\to\,\infty \text{ or } \log k}}{1+\frac{\log\log\frac{1}{\varepsilon}}{\log d(\varepsilon)}}.$$

From this we deduce that

$$\lim_{\varepsilon\to 0}\frac{\log\log\frac{1}{\varepsilon}}{\log d(\varepsilon)} = \infty \quad \text{or, equivalently,} \quad \log d(\varepsilon) = o\left(\log\log\frac{1}{\varepsilon}\right).$$

Hence

$$\log d(\varepsilon) + \log\log\frac{1}{\varepsilon} = (1+o(1))\log\log\frac{1}{\varepsilon}$$

and, inserting this into (14),

$$\limsup_{\varepsilon\to 0}\frac{d(\varepsilon)\log j(\varepsilon)}{\log\log\frac{1}{\varepsilon}} \le p^* + \delta.$$

Since $\delta > 0$ can be arbitrarily small, this means that $B \le p^*$ as needed. Therefore, the proof for EXP-SPT is completed with $p^* = B$.

2. We only need to show that EXP-PT implies EXP-SPT. The conditions $\lim_{j\to\infty}\lambda_j = \lim_{j\to\infty}\gamma_j = 0$ can be shown as before. Under the assumption of EXP-PT and Lemma 1 with $d = d(\varepsilon)$, there exist non-negative numbers C, q, and τ such that

$$j(\varepsilon)^{d(\varepsilon)} \le n(\varepsilon^{2d(\varepsilon)}, S_{d(\varepsilon),\boldsymbol{\gamma}}) \le C d(\varepsilon)^q \left(1 + 2d(\varepsilon)\log\frac{1}{\varepsilon}\right)^{\tau} \le C_1 d(\varepsilon)^{q+\tau}\left(\log\frac{1}{\varepsilon}\right)^{\tau},$$

where in the last estimate we assumed that $\varepsilon \le 1/\mathrm{e}$, and $C_1 := 3^{\tau} C$. Taking the logarithm yields

$$d(\varepsilon)\log j(\varepsilon) \le (q+\tau)\log d(\varepsilon) + \tau\log\log\frac{1}{\varepsilon} + \log C_1. \tag{15}$$

This shows that

$$\limsup_{\varepsilon\to 0}\frac{d(\varepsilon)\log j(\varepsilon)}{\log d(\varepsilon) + \log\log\frac{1}{\varepsilon}} \le q + \tau.$$

Now we argue as in the first part of this proof. We have

$$\limsup_{\varepsilon\to 0}\underbrace{\frac{d(\varepsilon)}{\log d(\varepsilon)}}_{\to\infty}\frac{\overbrace{\log j(\varepsilon)}^{\to\ \infty \text{ or } \log k}}{1 + \frac{\log\log\frac{1}{\varepsilon}}{\log d(\varepsilon)}} \le q + \tau$$

and hence

$$\lim_{\varepsilon\to 0}\frac{\log\log\frac{1}{\varepsilon}}{\log d(\varepsilon)} = \infty \quad \text{or, equivalently,} \quad \log d(\varepsilon) = o\left(\log\log\frac{1}{\varepsilon}\right).$$

Going back to (15) we obtain

$$d(\varepsilon)\log j(\varepsilon) \le \tau\log\log\frac{1}{\varepsilon} + o\left(\log\log\frac{1}{\varepsilon}\right)$$

and hence $B_{\mathrm{EXP-SPT}} \le \tau$. This means that we have EXP-SPT and we are done.

3. Assume first that we have EXP-QPT with the exponent t, i.e., for all $\varepsilon \in (0,1)$ and $d \in \mathbb{N}$ we have

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le C\exp\left(t(1+\log\, d)(1+\log(1+\log\varepsilon^{-1}))\right).$$

Then $\lim_{j\to\infty}\lambda_j = 0$ follows from the fact that $n(\varepsilon, S_{d,\boldsymbol{\gamma}})$ is finite. Using again (13) we conclude that $\lim_{j\to\infty}\gamma_j = 0$. Indeed, if $\lim_{j\to\infty}\gamma_j = \gamma^* > 0$ we have

$$d\ \log 2 \le \log\ n(\tfrac{1}{2}(\gamma^*\lambda_2)^{d/2}, S_{d,\boldsymbol{\gamma}}) = \mathcal{O}(t(\log d)^2),$$

which is a contradiction for large d.

We now apply Lemma 1 for $\varepsilon \in (0, 1/\mathrm{e})$ and $d = d(\varepsilon) \ge 1$. From the definition of EXP-QPT we obtain

$$\begin{aligned} j(\varepsilon)^{d(\varepsilon)} &\le C \exp\left(t(1+\log d(\varepsilon)) \left(1 + \log\left(1 + \log\left(\frac{1}{\varepsilon}\right)^{2d(\varepsilon)} \right)\right)\right) \\ &\le C \exp\left(t(1+\log d(\varepsilon)) \left(1 + 2\log 2 + \log d(\varepsilon) + \log\log\frac{1}{\varepsilon}\right)\right). \end{aligned}$$

Taking the logarithm yields

$$d(\varepsilon)\log j(\varepsilon) \le \log C + t(1+\log d(\varepsilon)) \left(1 + 2\log 2 + \log d(\varepsilon) + \log\log\frac{1}{\varepsilon}\right) \tag{16}$$

and hence

$$\limsup_{\varepsilon\to 0} \frac{d(\varepsilon)\log j(\varepsilon)}{\log(1+d(\varepsilon))[\log\log\frac{1}{\varepsilon} + \log d(\varepsilon)]} \le t.$$

We use, at this stage, the already familiar argument:

$$\limsup_{\varepsilon\to 0} \underbrace{\frac{d(\varepsilon)}{\log d(\varepsilon)\log(1+d(\varepsilon))}}_{\to\infty} \frac{\overbrace{\log j(\varepsilon)}^{\to\ \infty\text{ or }\log k}}{1 + \frac{\log\log\frac{1}{\varepsilon}}{\log d(\varepsilon)}} \le t$$

and hence

$$\log d(\varepsilon) = o\left(\log\log\frac{1}{\varepsilon}\right).$$

Inserting this into (16) we obtain that

$$B_{\mathrm{EXP-QPT}} := \limsup_{\varepsilon\to 0} \frac{d(\varepsilon)\log j(\varepsilon)}{[\log d(\varepsilon)]\log\log\frac{1}{\varepsilon}} \le t, \tag{17}$$

as needed.

Now assume that $\lim_{j\to\infty}\lambda_j = \lim_{j\to\infty}\gamma_j = 0$ and that $B := B_{\mathrm{EXP-QPT}} < \infty$. Then for every $\delta > 0$ there exists an $\varepsilon_\delta > 0$ such that

$$\frac{d(\varepsilon)\log j(\varepsilon)}{[\log d(\varepsilon)]\log\log\frac{1}{\varepsilon}} \le B + \delta \quad \text{for all } \varepsilon \in (0, \varepsilon_\delta). \tag{18}$$

We further assume that ε_δ is small enough such that $d(\varepsilon) \ge 2$ for all $\varepsilon \in (0, \varepsilon_\delta)$.

From this and Lemma 1 we obtain for all $\varepsilon \in (0, \varepsilon_\delta)$,

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le j(\varepsilon)^{d(\varepsilon)} \le \exp\left((B+\delta)[\log d(\varepsilon)]\log\log\frac{1}{\varepsilon}\right).$$

Hence, for all $\varepsilon \in (0, \varepsilon_\delta)$ and all $d \ge d(\varepsilon)$ we have

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le \exp\left((B+\delta)[\log d] \log\log \frac{1}{\varepsilon} \right).$$

If $3 \le d \le d(\varepsilon)$ then we obtain from (18) for all $\varepsilon \in (0, \varepsilon_\delta)$

$$\log j(\varepsilon) \le (B+\delta) \frac{\log d(\varepsilon)}{d(\varepsilon)} \log\log \frac{1}{\varepsilon} \le (B+\delta) \frac{\log d}{d} \log\log \frac{1}{\varepsilon},$$

where in the last estimate we use the fact that $d(\varepsilon) \ge 3$ and that the function $d \mapsto (\log d)/d$ is decreasing for $d \in \{3, \dots, d(\varepsilon)\}$.

For $d = 2$ we obtain again from (18) for all $\varepsilon \in (0, \varepsilon_\delta)$,

$$\log j(\varepsilon) \le (B+\delta) \frac{\log d(\varepsilon)}{d(\varepsilon)} \log\log \frac{1}{\varepsilon} \le (B+\delta) \frac{1+\log 2}{2} \log\log \frac{1}{\varepsilon},$$

and for $d = 1$ we obtain from (18) for all $\varepsilon \in (0, \varepsilon_\delta)$,

$$\log j(\varepsilon) \le (B+\delta) \frac{\log d(\varepsilon)}{d(\varepsilon)} \log\log \frac{1}{\varepsilon} \le (B+\delta) \log\log \frac{1}{\varepsilon}.$$

Hence we obtain

$$\log j(\varepsilon)^d \le (B+\delta)[1+\log d] \log\log \frac{1}{\varepsilon}.$$

The first estimate in Lemma 1 tells us that

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le j(\varepsilon)^{\min(d(\varepsilon), d)}.$$

Hence, for all $\varepsilon \in (0, \varepsilon_\delta)$ and all $d \in \{1, \dots, d(\varepsilon)\}$ we have

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le j(\varepsilon)^d \le \exp\left((B+\delta)[1+\log d] \log\log \frac{1}{\varepsilon} \right).$$

This implies

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le \exp\left((B+\delta)[1+\log d] \log\log \frac{1}{\varepsilon} \right)$$

for all $\varepsilon \in (0, \varepsilon_\delta)$ and for all $d \in \mathbb{N}$.

Finally, for $\varepsilon \in (\varepsilon_\delta, 1)$ we set $C_\delta = n(\varepsilon_\delta, S_{d(\varepsilon_\delta),\boldsymbol{\gamma}})$ and conclude

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le n(\varepsilon_\delta, S_{d(\varepsilon_\delta),\boldsymbol{\gamma}}) \le C_\delta \exp\left((B+\delta)[\log d] \log\log \frac{1}{\varepsilon} \right).$$

Hence there exists a $C_\delta > 0$ such that

$$n(\varepsilon, S_{d,\boldsymbol{\gamma}}) \le C_\delta \exp\left((B+\delta)[1+\log d] \log\log \frac{1}{\varepsilon} \right)$$

for all $\varepsilon \in (0,1)$ and $d \in \mathbb{N}$. This implies EXP-QPT with $t \le B + \delta$. Since $t \ge B$ due to (17) and the positive δ can be arbitrarily small, the infimum of such t is B, as claimed.

4. From [12, Theorem 3] we know that EXP-(s,t)-WT holds for any positive s and t if and only if

$$\sup_{d\in\mathbb{N}} \exp(-cd^t) \sum_{j=1}^{\infty} \mathrm{e}^{-c\left(1+\log\frac{2}{\lambda_{d,\gamma,j}}\right)^s} < \infty \quad \text{for all } c>0, \tag{19}$$

where $\lambda_{d,\gamma,1} \geq \lambda_{d,\gamma,2} \geq \ldots$ denote the eigenvalues of $W_{d,\gamma} = S_{d,\gamma}^* S_{d,\gamma} : H_{d,\gamma} \to H_{d,\gamma}$ ordered in a non-increasing fashion.

Assume first that $s=1$. For the above sum we have

$$\begin{aligned}\sum_{j=1}^{\infty} \mathrm{e}^{-c\left(1+\log\frac{2}{\lambda_{d,\gamma,j}}\right)} &= \frac{1}{\mathrm{e}^{c(1+\log 2)}} \sum_{j_1,\ldots,j_d=1}^{\infty} \mathrm{e}^{c\log(\lambda_{1,j_1}\lambda_{2,j_2}\cdots\lambda_{d,j_d})} \\ &= \frac{1}{(2\mathrm{e})^c} \prod_{k=1}^{d} \left(\sum_{j=1}^{\infty} \lambda_{k,j}^c\right) = \frac{1}{(2\mathrm{e})^c} \prod_{k=1}^{d} \left(1+\gamma_k^c \sum_{j=2}^{\infty} \lambda_j^c\right).\end{aligned}$$

Hence, EXP-$(1,t)$-WT is equivalent to

$$\sup_{d\in\mathbb{N}} \exp(-cd^t) \prod_{k=1}^{d} \left(1+\gamma_k^c \sum_{j=2}^{\infty} \lambda_j^c\right) < \infty \quad \text{for all } c>0.$$

Taking the logarithm, we find that EXP-$(1,t)$-WT is equivalent to

$$\sup_{d\in\mathbb{N}} \left(\sum_{k=1}^{d} \log\left(1+\gamma_k^c \sum_{j=2}^{\infty} \lambda_j^c\right) - cd^t\right) < \infty \quad \text{for all } c>0. \tag{20}$$

Assume first that (20) holds. Then we have $\sum_{j=2}^{\infty} \lambda_j^c < \infty$ for all $c>0$ and hence, by Lemma 2,

$$\lim_{j\to\infty} \frac{\log\frac{1}{\lambda_j}}{\log j} = \infty.$$

Consider the case $t=1$. Assume that $\lim_{j\to\infty} \gamma_j = \gamma_* > 0$. Then (20) implies

$$\sup_{d\in\mathbb{N}} \left(\log\left(1+\gamma_*^c \sum_{j=2}^{\infty} \lambda_j^c\right) - c\right) d < \infty \quad \text{for all } c>0.$$

This, however, yields a contradiction, since $\gamma_*^c \sum_{j=2}^{\infty} \lambda_j^c$ tends to infinity with c approaching zero, and therefore for small $c>0$ we have

$$\log\left(1+\gamma_*^c \sum_{j=2}^{\infty} \lambda_j^c\right) - c > 0.$$

Hence, we must have

$$\lim_{j\to\infty} \gamma_j = 0.$$

Thus we have shown the necessary conditions for EXP-WT.

Now assume that

$$\lim_{k\to\infty} \gamma_k = 0 \quad \text{ and } \lim_{j\to\infty} \frac{\log \frac{1}{\lambda_j}}{\log j} = \infty.$$

From Lemma 2 we obtain that $M_c^* := \sum_{j=2}^{\infty} \lambda_j^c < \infty$ for all $c > 0$. Hence, for every fixed $c > 0$ we have

$$\begin{aligned} \sum_{k=1}^{d} \log\left(1 + \gamma_k^c \sum_{j=2}^{\infty} \lambda_j^c\right) - cd &\le M_c^* \sum_{k=1}^{d} \gamma_k^c - cd \\ &= M_c^* \sum_{k=1}^{d} \left(\gamma_k^c - \frac{c}{M_c^*}\right) \\ &\le M_c^* \sum_{k=1}^{k_c^*} \left(\gamma_k^c - \frac{c}{M_c^*}\right) < \infty, \end{aligned}$$

where k_c^* is the largest $k \in \mathbb{N}$ such that $\gamma_k^c - \frac{c}{M_c^*} > 0$. This number is well defined, since $\lim_{k\to\infty} \gamma_k = 0$. Hence (20) holds for $t = 1$ and this implies EXP-WT.

5. Let $s = 1$ and $t < 1$. Assume we have EXP-$(1,t)$-WT. Then we have EXP-WT and hence

$$\lim_{k\to\infty} \gamma_k = 0 \text{ and } \lim_{j\to\infty} \frac{\log \frac{1}{\lambda_j}}{\log j} = \infty.$$

Now (20) can be rewritten as

$$\sum_{k=1}^{d} \log\left(1 + \gamma_k^c M_c^*\right) - cd^t \le A_c \quad \text{ for all } d \in \mathbb{N},$$

with a positive $A_c < \infty$ for every $c > 0$. We use the inequality

$$\log(1 + x) \ge (\log 2)\, x \quad \text{for all } x \in [0, 1].$$

Let k_c^* be the largest k such that $\gamma_k^c M_c^* > 1$. If such a k_c^* does not exist, we set $k_c^* := 0$. Then we have

$$\begin{aligned} A_c &\ge \sum_{k=1}^{d} \log\left(1 + \gamma_k^c M_c^*\right) - cd^t \\ &\ge \sum_{k=1}^{k_c^*} \log\left(1 + \gamma_k^c M_c^*\right) + (\log 2) M_c^* \sum_{k=k_c^*+1}^{d} \gamma_k^c - cd^t \\ &\ge \sum_{k=1}^{k_c^*} \log\left(1 + \gamma_k^c M_c^*\right) + (\log 2) M_c^* (d - k_c^*) \gamma_d^c - cd^t. \end{aligned}$$

From here it follows by an argument similar to the one used in the proof of Lemma 2 that

$$\liminf_{d\to\infty} \frac{\log \frac{1}{\gamma_d}}{\log d} \ge \frac{1-t}{c} \quad \text{for all } c > 0.$$

Since c can be arbitrarily small and $t < 1$ we have

$$\lim_{d\to\infty} \frac{\log \frac{1}{\gamma_d}}{\log d} = \infty,$$

as desired.

On the other hand for

$$\lim_{j\to\infty} \frac{\log \frac{1}{\gamma_j}}{\log j} = \infty \quad \text{and} \quad \lim_{j\to\infty} \frac{\log \frac{1}{\lambda_j}}{\log j} = \infty,$$

Lemma 2 yields

$$\Gamma_c := \sum_{k=1}^{\infty} \gamma_k^c < \infty \ \text{ and } \ M_c^* := \sum_{j=2}^{\infty} \lambda_j^c < \infty \quad \text{for all } c > 0.$$

Now, for every fixed $c > 0$ we have

$$\sum_{k=1}^{d} \log\left(1 + \gamma_k^c \sum_{j=2}^{\infty} \lambda_j^c\right) - cd^t \ \leq \ M_c^* \Gamma_c - cd^t$$

and hence

$$\sup_{d\in\mathbb{N}} \sum_{k=1}^{d} \log\left(1 + \gamma_k^c \sum_{j=2}^{\infty} \lambda_j^c\right) - cd^t < \infty.$$

Hence (20) holds and this implies EXP-$(1,t)$-WT.

6. Let $s = 1$ and $t > 1$.

   The sufficiency of the condition on the eigenvalues is shown for the un-weighted case (i.e., all γ_j equal 1) in Case (N.1) of [4, Theorem 2]. If some of the weights satisfy $\gamma_j < 1$, then the problem is easier than in the un-weighted case, which means that the sufficient condition holds as well.

   Regarding necessity of the condition, we have shown that EXP-$(1,t)$-WT implies (20) and hence we find as above that the condition

$$\lim_{j\to\infty} \frac{\log \frac{1}{\lambda_j}}{\log j} = \infty$$

   is indeed necessary.

7. Let $s > 1$, $t \leq 1$ and $\lambda_2 < 1$. The sufficiency of the condition on the eigenvalues is shown for the un-weighted case (i.e., all γ_j equal 1) in Case (N.3) of [4, Theorem 2]. If some of the weights satisfy $\gamma_j < 1$, then the problem is easier than in the un-weighted case, which means that the sufficient condition holds as well.

   We show that the condition on the λ_j's is also necessary. We use (19) and Lemma 3 to obtain

$$\sum_{j=1}^{\infty} \mathrm{e}^{-c\left(1+\log \frac{2}{\lambda_{d,\boldsymbol{\gamma},j}}\right)^s}$$

$$
\begin{aligned}
&\geq \mathrm{e}^{-c2^{s-1}(1+\log 2)^s} \sum_{j=1}^{\infty} \mathrm{e}^{-c2^{s-1}\left(\log \frac{1}{\lambda_{d,\gamma,j}}\right)^s} \\
&= \mathrm{e}^{-c2^{s-1}(1+\log 2)^s} \sum_{j_1,\ldots,j_d=1}^{\infty} \mathrm{e}^{-c2^{s-1}\left(\sum_{k=1}^d \log \frac{1}{\lambda_{k,j_k}}\right)^s} \\
&\geq \mathrm{e}^{-c2^{s-1}(1+\log 2)^s} \sum_{j_1,\ldots,j_d=1}^{\infty} \mathrm{e}^{-c(2d)^{s-1}\sum_{k=1}^d \left(\log \frac{1}{\lambda_{k,j_k}}\right)^s} \\
&= \mathrm{e}^{-c2^{s-1}(1+\log 2)^s} \prod_{k=1}^{d} \left(1+\sum_{j=2}^{\infty} \mathrm{e}^{-c(2d)^{s-1}\left(\log \frac{1}{\gamma_k}+\log \frac{1}{\lambda_j}\right)^s}\right) \\
&\geq \mathrm{e}^{-c2^{s-1}(1+\log 2)^s} \prod_{k=1}^{d} \left(1+\mathrm{e}^{-c(4d)^{s-1}\left(\log \frac{1}{\gamma_k}\right)^s} \sum_{j=2}^{\infty} \mathrm{e}^{-c(4d)^{s-1}\left(\log \frac{1}{\lambda_j}\right)^s}\right).
\end{aligned}
$$

Put

$$
\Gamma_k := \mathrm{e}^{-(\log \frac{1}{\gamma_k})^s} \quad \text{and} \quad \Lambda_j := \mathrm{e}^{-(\log \frac{1}{\lambda_j})^s}. \tag{21}
$$

Then we have

$$
\sum_{j=1}^{\infty} \mathrm{e}^{-c\left(1+\log \frac{2}{\lambda_{d,\gamma,j}}\right)^s} \geq \mathrm{e}^{-c2^{s-1}(1+\log 2)^s} \prod_{k=1}^{d} \left(1+\Gamma_k^{c(4d)^{s-1}} \sum_{j=2}^{\infty} \Lambda_j^{c(4d)^{s-1}}\right).
$$

Assume that EXP-(s,t)-WT holds true. Then according to (19) together with the above lower bound (for $d=1$) we obtain

$$
\exp(-c(1+2^{s-1}(1+\log 2)^s)) \left(1+\Gamma_1^{c4^{s-1}} \sum_{j=2}^{\infty} \Lambda_j^{c4^{s-1}}\right) < \infty \quad \text{for all } c>0.
$$

This requires that

$$
\sum_{j=2}^{\infty} \Lambda_j^c < \infty \quad \text{for all } c>0.
$$

According to Lemma 2 this is equivalent to

$$
\lim_{j\to\infty} \frac{\log \frac{1}{\Lambda_j}}{\log j} = \infty,
$$

and this condition holds if and only if

$$
\lim_{j\to\infty} \frac{\left(\log \frac{1}{\lambda_j}\right)^s}{\log j} = \infty.
$$

This finishes the proof for the necessary condition. Note that for this part we did not use that $\lambda_2<1$.

8. Let $s>1$, $t\leq 1$ and $\lambda_2=1$. The necessary condition on the λ_j's follows from the above where we did not use that $\lambda_2<1$. To show that EXP-(s,t)-WT implies

$\gamma_p < 1$ we apply Lemma 4 with all $k_j = 2$ for which $\lambda_2 = 1$. If all $\gamma_j = 1$ then $\lim_{d\to\infty} d^t/(d\log 2)$ is zero for $t < 1$ and $1/\log 2$ for $t = 1$ and it is never infinity. It is infinity only if $\gamma_p < 1$ for some p, as claimed.

In order to show that the conditions on the λ_j's and γ_j's imply EXP-(s,t)-WT, we switch to a possibly harder problem for which the weights are given by

$$1 = \gamma_1 = \ldots = \gamma_{p-1} > \gamma_p = \gamma_{p+1} = \ldots.$$

and the eigenvalues

$$\widetilde{\lambda}_1 = \ldots = \widetilde{\lambda}_{p-1} = 1 \quad \text{and} \quad \widetilde{\lambda}_j = \gamma_p \lambda_j \ \text{ for } j = p, p+1, \ldots.$$

Note that $\widetilde{\lambda}_p < \widetilde{\lambda}_{p-1} = 1$.

For $d \geq p$, we have

$$\begin{aligned}
&\sum_{j=1}^{\infty} \mathrm{e}^{-c\left(1+\log \frac{2}{\lambda_{d,\gamma,j}}\right)^s} \\
&\quad\leq \sum_{j_1,\ldots,j_d=1}^{\infty} \mathrm{e}^{-c\left(\sum_{k=1}^{d} \log \frac{1}{\lambda_{k,j_k}}\right)^s} \\
&\quad\leq \sum_{j_1,\ldots,j_{p-1}=1}^{\infty} \mathrm{e}^{-c\sum_{k=1}^{p-1}\left(\log \frac{1}{\lambda_{k,j_k}}\right)^s} \sum_{j_p,\ldots,j_d=1}^{\infty} \mathrm{e}^{-c\left(\sum_{k=p}^{d} \log \frac{1}{\widetilde{\lambda}_{j_k}}\right)^s} \\
&\quad\leq \left(1 + \sum_{j=2}^{\infty} \Lambda_j^c\right)^{p-1} \sum_{j_p,\ldots,j_d=1}^{\infty} \mathrm{e}^{-c\left(\sum_{k=p}^{d} \log \frac{1}{\widetilde{\lambda}_{j_k}}\right)^s},
\end{aligned}$$

where $\Lambda_j = \mathrm{e}^{-\left(\log \frac{1}{\lambda_j}\right)^s}$.

The series $\sum_{j=2}^{\infty} \Lambda_j^c$ is convergent due to Lemma 2 and the conditions on the λ_j's. Therefore the first factor is of order 1. The second factor

$$\sum_{j_p,\ldots,j_d=1}^{\infty} \mathrm{e}^{-c\left(\sum_{k=p}^{d} \log \frac{1}{\widetilde{\lambda}_{j_k}}\right)^s} = \exp\left(\Theta\left((\log d)^{\max(1,s/(s-1))}\right)\right),$$

as proved in [4, (A.4)]. Therefore

$$\sup_d \mathrm{e}^{-cd^t} \sum_{j=1}^{\infty} \mathrm{e}^{-c\left(1+\log \frac{2}{\lambda_{d,\gamma,j}}\right)^s} = \sup_d \exp\left(-cd^t + \Theta\left((\log d)^{\max(1,s/(s-1))}\right)\right) < \infty,$$

which completes the proof of this item.

9. Let $s > 1$ and $t > 1$. The sufficiency of the condition on the eigenvalues is shown for the un-weighted case (i.e., all γ_j equal 1) in Case (N.1) of [4, Theorem 2]. If some of the weights satisfy $\gamma_j < 1$, then the problem is easier than in the un-weighted case, which means that the sufficient condition holds as well.

   In order to show that the condition on the λ_j's is also necessary one proceeds as above in Item 7.

10. Let $s<1$ and $t>1$. Note that EXP-(s,t)-WT holds for arbitrary γ_j's iff this holds for $\gamma_j=1$ for all $j\in\mathbb{N}$. This case is proved in Case (N.2) of [4, Theorem 2].

11. Let $s<1$ and $t=1$. Suppose first that EXP-$(s,1)$-WT holds. Then EXP-WT also holds and $\lim_{k\to\infty}\gamma_k=\lim_{j\to\infty}\lambda_j=0$. Take integers d,k,j with $j\ge 2$ and $k\le d$, and define

$$\varepsilon^2=\gamma_k^d\lambda_j^d\alpha,$$

where $\alpha\in(0,1)$ such that $\varepsilon<1$. Then $\log n(\varepsilon,S_{d,\boldsymbol{\gamma}})\ge d\log j$ and, proceeding as before, we conclude

$$\begin{aligned}\infty &= \lim_{d+\gamma_k^{-d}\lambda_j^{-d}\to\infty}\frac{d+\left(\log\frac{1}{\varepsilon}\right)^s}{d\log j}\\ &= \lim_{d+\gamma_k^{-d}\lambda_j^{-d}\to\infty}\frac{d+2^{-s}\left[d\left(\log\frac{1}{\gamma_k}\right)+d\left(\log\frac{1}{\lambda_j}\right)+\log\frac{1}{\alpha}\right]^s}{d\log j}\\ &\le \lim_{d+\gamma_k^{-d}\lambda_j^{-d}\to\infty}\frac{d+\left(\log\frac{1}{\alpha}\right)^s+d^s\left(\left(\log\frac{1}{\gamma_k}\right)^s+\left(\log\frac{1}{\lambda_j}\right)^s\right)}{d\log j}.\end{aligned}$$

Since $[d+(\log\frac{1}{\alpha})^s]/(d\log j)$ does not go to infinity, we obtain (10).

Suppose now that (10) holds. We show EXP-$(s,1)$-WT by using (19). From Lemma 3 we get

$$\begin{aligned}\alpha_d &:= \sum_{j=1}^{\infty}\mathrm{e}^{-c\left(1+\log\frac{2}{\lambda_{d,\boldsymbol{\gamma},j}}\right)^s}\\ &\le \sum_{j_1,\dots,j_d=1}^{\infty}\mathrm{e}^{-c\left(\sum_{k=1}^d\log\frac{1}{\lambda_{d,j_k}}\right)^s}\\ &\le \sum_{j_1,\dots,j_d=1}^{\infty}\mathrm{e}^{-cd^{s-1}\sum_{k=1}^d\left(\log\frac{1}{\lambda_{k,j_k}}\right)^s}.\end{aligned}$$

Then, again by Lemma 3, we obtain

$$\alpha_d\le\prod_{k=1}^d\left(1+\sum_{j=2}^{\infty}\mathrm{e}^{-c(2d)^{s-1}\left[\left(\log\frac{1}{\gamma_k}\right)^s+\left(\log\frac{1}{\lambda_j}\right)^s\right]}\right).$$

From (10) we conclude that for any (large) M there exists C_M such that for all $d\ge C_M$ we have

$$\mathrm{e}^{-c(2d)^{s-1}\left[\left(\log\frac{1}{\gamma_k}\right)^s+\left(\log\frac{1}{\lambda_j}\right)^s\right]}\le j^{-c2^{s-1}M}.$$

Note that the exponent $c2^{s-1}M$ can be sufficiently large for large M, and therefore the series $\sum_{j=2}^{\infty}j^{-c2^{s-1}M}$ is convergent and sufficiently small, say it is $o(c)$. Then (19) implies that

$$\begin{aligned}\sup_{d\in\mathbb{N}}\exp(-c\,d)\alpha_d &= \sup_{d\in\mathbb{N}}\exp(-c\,d)(1+o(c))^d\\ &= \sup_{d\in\mathbb{N}}\exp(d(-c+o(c)))<\infty.\end{aligned}$$

Hence, EXP-$(s,1)$-WT holds, and the proof is complete.

□

References

[1] J. Dick, P. Kritzer, F. Pillichshammer, H. Woźniakowski. Approximation of analytic functions in Korobov spaces. J. Complexity 30, 2–28, 2014.

[2] J. Dick, G. Larcher, F. Pillichshammer, H. Woźniakowski. Exponential convergence and tractability of multivariate integration for Korobov spaces, Math. Comp. 80, 905–930, 2011.

[3] M. Gnewuch, H. Woźniakowski. Quasi-polynomial tractability. J. Complexity 27, 312–330, 2011.

[4] F.J. Hickernell, P. Kritzer, H. Woźniakowski. Exponential tractability of linear tensor product problems. To appear in: D.R. Wood, J. DeGier, C. Praeger, T. Tao (eds.), *MATRIX Annals*, Springer, Cham, 2020.

[5] A. Hinrichs, E. Novak, H. Woźniakowski. Discontinuous information in the worst case and randomized settings. Math. Nachr. 286, 679–690, 2013. Special issue dedicated to Professor Hans Triebel's 75's birthday, edited by D. Haroske ans H-J. Schmeisser.

[6] C. Irrgeher, P. Kritzer, G. Leobacher, F. Pillichshammer. Integration in Hermite spaces of analytic functions. J. Complexity 31, 380–404, 2015.

[7] C. Irrgeher, P. Kritzer, F. Pillichshammer, H. Woźniakowski. Tractability of multivariate approximation defined over Hilbert spaces with exponential weights. J. Approx. Theory 207, 301–338, 2016.

[8] C. Irrgeher, P. Kritzer, F. Pillichshammer, H. Woźniakowski. Tractability of multivariate approximation defined over Hilbert spaces with exponential weights. J. Approx. Theory 207, 98–126, 2016.

[9] D. Krieg, M. Ullrich. Function values are enough for L_2-approximation. Submitted, 2020.

[10] P. Kritzer, F. Pillichshammer, H. Woźniakowski. Multivariate integration of infinitely many times differentiable functions in weighted Korobov spaces. Math. Comp. 83, 1189–1206, 2014.

[11] P. Kritzer, F. Pillichshammer, H. Woźniakowski. $\mathbb{L}_\infty$-Approximation in Korobov spaces with exponential weights. J. Complexity 41, 102–125, 2017.

[12] P. Kritzer, H. Woźniakowski. Simple characterizations of exponential tractability for linear multivariate problems. J. Complexity 51, 110–128, 2019.

[13] Y. Liu, G. Xu. A note on tractability of multivariate analytic problems. J. Complexity 34, 42–49, 2016.

[14] Y. Liu, G. Xu. Exponential convergence-tractability of general multivariate linear problems (in Chinese). Numer. Math. J. Chinese Univ. 38, 375–384, 2016.

[15] Y. Liu, G. Xu. Average case tractability of a multivariate approximation problem. J. Complexity 43, 76–102, 2017.

[16] Y. Liu, J. Zhang. EC-tractability of approximation problems in function spaces defined over products of simplices. J. Complexity 55, Article 101411, 2019.

[17] E. Novak, H. Woźniakowski. *Tractability of Multivariate Problems, Volume I: Linear Information*. EMS, Zurich, 2008.

[18] E. Novak, H. Woźniakowski. *Tractability of Multivariate Problems, Volume II: Standard Informations for Functionals*. EMS, Zurich, 2010.

[19] E. Novak, H. Woźniakowski. *Tractability of Multivariate Problems, Volume III: Standard Informations for Operators*. EMS, Zurich, 2012.

[20] A. Papageorgiou and I. Petras. On the tractability of linear tensor product problems in the worst case. J. Complexity 25, 415–419, 2009.

[21] A. Papageorgiou and I. Petras. A new criterion for tractability of multivariate problems. J. Complexity 30, 604–619, 2014.

[22] A. Papageorgiou, I. Petras, and H. Woźniakowski. $(s, \ln^{\kappa})$-weak tractability of linear problems. J. Complexity 40, 1–16, 2017.

[23] A. Papageorgiou, I. Petras, G. Xu, D. Yanqi. EC-(s,t)-weak tractability of multivariate linear problems in the average case setting. J. Complexity 55, Article 101425, 2019.

[24] P. Siedlecki. Uniform weak tractability. J. Complexity 29, 438–453, 2013.

[25] P. Siedlecki, M. Weimar. Notes on (s,t) weak tractability; a refined classification of problems with (sub)exponential information complexity, J. Approx. Theory 200, 227–258, 2015.

[26] I.H. Sloan, H. Woźniakowski. When are quasi-Monte Carlo algorithms efficient for high dimensional integrals? J. Complexity 14, 1–33, 1998.

[27] I.H. Sloan, H. Woźniakowski. Multivariate approximation for analytic functions with Gaussian kernels. J. Complexity 45, 1–21, 2018.

[28] K. Suzuki. Super-polynomial convergence and tractability of multivariate integration for infinitely times differentiable functions. J. Complexity 39, 51–68, 2017.

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

[30] J.F. Traub, A.G. Werschulz. *Complexity and Information*. Cambridge University Press, Cambridge, 1998.

[31] H. Wang. A note about EC-(s,t)-weak tractability of multivariate approximation with analytic Korobov kernels. J. Complexity 55, Article 101412, 2019.

[32] A.G. Werschulz, H. Woźniakowski. A new characterization of (s,t)-weak tractability. J. Complexity 38, 68–79, 2017.

[33] G. Xu. Exponential convergence-tractability of general linear problems in the average case setting. J. Complexity 31, 617–636, 2015.

Authors' addresses:

Peter Kritzer,
Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstr. 69, 4040 Linz, Austria

Friedrich Pillichshammer,
Department of Financial Mathematics and Applied Number Theory, Johannes Kepler University Linz, Altenbergerstr. 69, 4040 Linz, Austria

Henryk Woźniakowski,
Department of Computer Science, Columbia University, New York 10027, USA and Institute of Applied Mathematics, University of Warsaw, ul. Banacha 2, 02-097 Warszawa, Poland

E-mail:
peter.kritzer@oeaw.ac.at
friedrich.pillichshammer@jku.at
henryk@cs.columbia.edu