

TRACTABILITY OF LINEAR ILL-POSED PROBLEMS IN HILBERT SPACE

PETER MATHÉ AND BERND HOFMANN

ABSTRACT. We introduce a notion of tractability for ill-posed operator equations in Hilbert space. For such operator equations the asymptotics of the best possible rate of reconstruction in terms of the underlying noise level is known in many cases. However, the relevant question is, which level of discretization, again driven by the noise level, is required in order to achieve this best possible accuracy. The proposed concept adapts the one from Information-based Complexity. Several examples indicate the relevance of this concept in the light of the curse of dimensionality.

1. INTRODUCTION

We shall introduce the concept of *tractability* for linear *ill-posed problems*, which are modeled in the form of *operator equations*

$$(1) \quad Ax = y,$$

where some injective bounded linear operator $A: X \rightarrow Y$ is acting between real infinite dimensional Hilbert spaces X and Y . We consider the noise model

$$(2) \quad y^\delta := Ax + \delta\xi,$$

where the unknown noise element ξ is norm bounded by one, such that $\|Ax - y^\delta\|_Y \leq \delta$. There is vast literature available referring to the analysis of such ill-posed problems, and the standard monograph is [3]. The optimal rates of reconstruction, under appropriate smoothness assumptions, are known in many cases, and often these are given in terms of the noise level $\delta > 0$.

As usual for ill-posed operator equations (1), we consider smoothness *relative* to the given operator A . Hence we assume that solution smoothness is given in terms of source sets as follows:

Date: April 18, 2024.

Key words and phrases. curse of dimensionality, tractability, multivariate problems.

Definition 1 (source set). There is an index function¹ φ such that the unknown solution obeys

$$x \in H_\varphi := \{x, x = \varphi(A^*A)v, \|v\|_X \leq 1\}.$$

Often the ill-posed operator equation (1) refers to Hilbert spaces of functions on a d -dimensional domain, see e.g. the examples given in the monograph [19, Chapt. 5]. In many cases the reconstruction rate deteriorates when the spatial dimension d grows. This is best seen in the decay rate of the singular values of embeddings between Sobolev spaces in Example 3 below, which exhibits a behavior of the form $n^{-a/d}$, where n is the cardinality of data, and a encodes smoothness information. This phenomenon is often called ‘curse of dimensionality’², and it indicates that the amount of computational effort that is required to find a suitable reconstruction increases exponentially with the dimension d . However, this may be foiled when the leading constant in the asymptotics decays rapidly with increasing dimension. In such a case, the curse of dimensionality does not necessarily reflect the difficulties in solving the problem at hand.

In contrast, the associated decay rate may be dimension independent ‘up to some logarithmic factor’, which depends on a power of the dimension d . This phenomenon is also well known, specifically for Sobolev embeddings when the underlying spaces are anisotropic. Here we highlight the inverse problem of *reconstruction of copula densities*, representing the correlation structure of a family of assets, where the *dimensionality* is given by the number d of assets (cf. [2]). The recent note [4], for example, outlines the mathematical model, where the forward operator is a *multivariate integration* operator. Details for this approach will be given below in Section 5. This model has also been investigated in [21] with respect to the stable numerical solution of the inverse problem of copula density identification, which is severely hampered by limited computational resources. Regarding the reconstruction rate, as $\delta \rightarrow 0$, there seems to be no impact arising from the number of assets. However, as will be made precise below, the *discretization level* (amount on linear information) may depend on d , even *exponentially*. Consequently, for large d it may be *infeasible in practice* to reach the region with the optimal rate of reconstruction.

¹We call a function $\varphi: (0, \infty) \rightarrow [0, \infty)$ an index function, if it is continuous, non-decreasing and satisfies the limit condition $\lim_{t \searrow 0} \varphi(t) = 0$.

²The meaning of this notion differs in varying places. Here we just mean that the rate as a function of the amount of information, deteriorates exponentially in the dimension.

Therefore, in order to understand the difficulty to solve a numerical problem to a given accuracy, a more precise understanding is required. This purpose is achieved by studying the tractability of such problems. Within the theory of *Information-based Complexity* there is a long discussion in this direction, and we refer to the three volumes [13, 14, 15]. As far as we are aware of, in the literature there is by now no discussion of the tractability of ill-posed problems, which takes into account the impact of both occurring facets: the dimension d and the reconstruction rate in terms of δ . Our goal here is to adapt this approach to the theory of ill-posed operator equations. Recently, the authors in [16] consider tractability of problems when the information is noisy. Still this concerns direct problems. In *ibid.* the focus is on the comparison of tractability with and without noisy information. For ill-posed operator equations the presence of noise is constitutive for the error analysis. Reconstruction rates in the presense of exact data (often called bias decay) differ significantly from the rates under noise.

Thus, the outline is as follows. We shall formulate the problem, including some examples in Section 2, and we give the formal definition of *tractability* in Section 3. The main observation here is presented as Theorem 1, showing that our notion of tractability of a family of inverse problems is consistent with some appropriate family of direct ones.

We discuss examples for operators with power-type decay of singular values in Section 4. Finally, we discuss the family of multivariate integration problems in Section 5. The tractability of this problem family is stated as Theorem 2.

2. INFORMATION COMPLEXITY, PROBLEM FORMULATION

Our focus is on operator equations (1) with *compact forward operator* $A: X \rightarrow Y$. These operators allow for a singular value decomposition (SVD) $\{s_j, u_j, v_j\}_{j=1}^{\infty}$ with a non-increasing infinite sequence $\{s_j\}_{j=1}^{\infty}$ of *singular values*, tending to zero as j tends to infinity, and sequences $u_j, v_j, j = 1, 2, \dots$, of corresponding *eigenfunctions* for the self-adjoint operators A^*A and AA^* , respectively.

The standard noise model (2) is not realistic, as elements in infinite dimensional Hilbert spaces do not fit numerical computations, and discretization is required to do so. We follow the standard approach as presented in [19, Chapt. 3].

2.1. Discretization. Here we restrict ourselves to *(non-adaptive) linear information*, i.e., we choose linear functionals $L_1, \dots, L_n \in Y^*$ and

build the n -vector

$$(3) \quad N_n(y^\delta) := (L_1(y^\delta), \dots, L_n(y^\delta)), \quad y^\delta \in Y,$$

which we shall call information N_n . For suitable reconstructions R to the unknown elements $x \in X$ we then may use some arbitrary mapping ψ into the space X , acting on $N_n(y^\delta) \in \mathbb{R}^n$. Thus the admissible reconstructions are given as

$$(4) \quad x_n^\delta := R(y^\delta) = \psi(N_n(y^\delta)), \quad y^\delta \in Y.$$

Such approach to ill-posed problems which mimics Information-based Complexity was first used in [11], and we follow this pathway.

We consider the worst-case error for any such reconstruction R which uses information of cardinality at most n , uniformly on source sets H_φ , given as in Definition 1. At any instance $x \in H_\varphi$ the error is given as

$$e(R, x, \delta) := \sup_{\|\xi\| \leq 1} \|x - R(Ax + \delta\xi)\|_X,$$

and the error uniformly for $x \in H_\varphi$ is given as

$$e(R, H_\varphi, \delta) := \sup_{x \in H_\varphi} e(R, x, \delta).$$

The minimal error $e_n(H_\varphi, \delta)$ is then the minimum over all reconstructions R using information of cardinality at most n .

2.2. Problem formulation. A fundamental observation from [11] is here rephrased as follows.

Proposition 1. *Suppose that the operator A has an SVD with singular values $\{s_j\}_{j=1}^\infty$, and that smoothness is given as in Definition 1 with an index function φ . For $j \in \mathbb{N}$ we have that*

$$e_j(H_\varphi, \delta) \geq \varphi(s_{j+1}^2).$$

Remark 1. We mention that the assertion of Proposition 1 is stated in [11, Theorem 1] for the *Gelfand widths*, but by virtue of [11, Lemma 1] these coincide with the singular values.

Remark 2. In the noiseless case ($\delta = 0$) the right hand side above is attained by the error of spectral cut-off

$$R(y^\delta) := \sum_{i=1}^j \frac{1}{s_i} \langle y^\delta, v_i \rangle u_i.$$

uniformly on the smoothness class H_φ , taking into account the j largest singular values. In this case the information N_j is based on using the first eigenfunctions v_1, \dots, v_j of the operator AA^* .

The above lower bound from Proposition 1 is contrasted to the well-known upper bound from the Melkman-Micchelli construction. We denote by $e(H_\varphi, \delta)$ the best possible accuracy of any reconstruction making use of the full data y^δ . Referring to [5] we have

$$(5) \quad e(H_\varphi, \delta) \leq \varphi(\Theta^{-1}(\delta)),$$

where $\Theta(t) := \sqrt{t}\varphi(t)$ is the companion index function to φ .

Consequently, the rate $\delta \mapsto \varphi(\Theta^{-1}(\delta))$ is the best possible reconstruction rate. It represents the asymptotic regime, and hence the question is how far discretization needs to go, in order to reach this regime.

Therefore it is interesting to discuss the index

$$(6) \quad k_*(\delta) := \begin{cases} 1, & \text{if } \Theta(s_1^2) \leq \delta, \\ \max \{k \in \mathbb{N} : \Theta(s_k^2) > \delta\}, & \text{otherwise,} \end{cases}$$

which describes the minimal level of discretization required to achieve order optimal regularization.

Proposition 2. *For the cardinality k_* from (6) we have that*

$$\varphi(s_{k_*+1}^2) \leq \varphi(\Theta^{-1}(\delta)).$$

Proof. By construction we see that $\Theta(s_{k_*+1}^2) \leq \delta$. Hence

$$\varphi(s_{k_*+1}^2) \leq \varphi(\Theta^{-1}(\delta)),$$

by the monotonicity of the index function φ . \square

Thus, from k_* on we see the asymptotic regime given in (5) with the rate $\varphi(\Theta^{-1}(\delta))$, and the size of k_* is a measure of the computational difficulty of the problem at hand. We highlight this by the following first example.

Example 1 (moderately ill-posed operator). Here we focus on the error behavior as $\delta \rightarrow 0$ for a problem with operator whose singular values tend to zero polynomially. Precisely, fix A with $s_j(A) \sim j^{-a}$ ($j \in \mathbb{N}, a > 0$). Let us assume power type smoothness as in Definition 1 for a function $\varphi(t) = t^p$, $t > 0$, for some exponent $p > 0$. In this case we see that $k_*(\delta) \asymp \left(\frac{1}{\delta}\right)^{\frac{1}{2a(p+1/2)}}$ as $\delta \rightarrow 0$, such that the discretization level k_* increases polynomially in $1/\delta$, which is assumed to be feasible. The asymptotically optimal rate will be seen as $\delta \rightarrow \delta^{p/(p+1/2)}$, regardless of the exponent a .

We point out the following. When the solution smoothness is measured in terms of source conditions, as this is done in Definition 1, the obtained error bounds are given in terms of the corresponding index

functions, and hence these are independent of the decay rates of the singular values of the governing operator A . In contrast, the cardinality k_* very well depends on the singular values, as well as the noise level δ . The question that we are to address is the following: If we need a discretization level k_* in order to enter the asymptotic regime, will this level be feasible for a problem at hand? For further motivation we present the next example, which is a slight variation of the first one.

Example 2 (mildly ill-posed operator). Here we consider a similar problem as in Example 1, and we assume power type smoothness as for a function $\varphi(t) = t^p$, $t > 0$, for some exponent $p > 0$. However, the singular values of the operator A tend to zero slowly. Precisely, fix A with $s_j(A) \sim 1/\log(j)$ ($j \in \mathbb{N}$). We need to describe k_* based on formula (6) and using the companion function $\Theta(t) = t^{p+1/2}$ to φ . Again, the order optimal rate is given as $\delta \rightarrow \delta^{p/(p+1/2)}$, but here we have

$$k_*(\delta) \sim \exp \left(\left(\frac{1}{\delta} \right)^{1/(2p+1)} \right) \quad \text{as } \delta \rightarrow 0.$$

Thus k_* is exponential in $1/\delta$. For a noise level $\delta := 10^{-4}$ and with an exponent $p = 1/2$ in the index function φ , which characterizes the solution smoothness as $x \in \mathcal{R}(A^*)$, we have that $k_* \sim e^{100}$. This is certainly not feasible.

We conclude that it is difficult (intractable) to discretize ill-posed problems with mildly ill-posed operator to enter the asymptotic regime, even if the noise level δ is moderate. \square

For d -dimensional (d -variate) situations, singular values and hence the index k_* depend on both the dimension $d > 1$ and the noise level $\delta > 0$. We present in this context another illustrative example.

Example 3 (Sobolev embeddings). Here we shall consider an ill-posed problem when the forward operator acts along a given scale of Sobolev spaces $W_2^\mu(\Omega)$ on some bounded C^∞ -domain $\Omega \subset \mathbb{R}^d$ (boundary conditions may be imposed). This means that d is the spatial dimension of the domain Ω .

The spaces $W_2^\mu(\Omega)$ form, for a given interval $[-a, a]$ ($a > 0$), a scale of Hilbert spaces (see, e.g., [12]) with two-sided estimates generated by some unbounded self-adjoint operator, say L . For $0 \leq \nu \leq a$ we have that $x \in W_2^\nu(\Omega)$ exactly if $x \in \mathcal{D}(L^\nu)$ belongs to the domain of L^ν . Negative smoothness is given by duality.

The above phrase “along a scale” means that A acts from $W^{-a}(\Omega)$ to Y , and there are constants $0 < m \leq M < \infty$, for which

$$(7) \quad m \|x\|_{-a} \leq \|Ax\|_Y \leq M \|x\|_{-a}, \quad x \in W_2^{-a}(\Omega).$$

Also, we assume that the given solution smoothness is $x \in W_2^p(\Omega) = \mathcal{R}(L^{-p})$, and hence there is some source element v with $x = L^{-p}v$. For details we also refer to the monograph [20, Chapt. 4.9].

The crucial tool for understanding this situation is given by Douglas' range inclusion theorem, see its formulation in [9]. Thus we see that condition (7) is equivalent to $\mathcal{R}((A^*A)^{1/2}) = \mathcal{R}(L^{-a})$ (with corresponding norm bounds). Using Heinz' inequality ([3, Prop. 8.21]) this yields for $0 < p \leq a$ that $\mathcal{R}((A^*A)^{p/(2a)}) = \mathcal{R}(L^{-p})$. The link condition also has a consequence given by Weyl's monotonicity theorem, see [1, Chapt. III.2.3], and this reads as $ms_j(L^{-a}) \leq s_j(A) \leq Ms_j(L^{-a})$ for $j \in \mathbb{N}$, where $s_j := s_j(A)$ and $s_j(L^{-a})$ denote the singular values of A and L^{-a} , respectively. It will be important to know that asymptotically $s_j(L^{-a}) \asymp j^{-a/d}$ as $j \rightarrow \infty$. Hence, the dimensionality of the domain enters the decay rate of the singular values.

For convenience let us consider the index function $\varphi(t) := t^{p/(2a)}$ for describing the solution smoothness and its companion function $\Theta(t) := \sqrt{t}\varphi(t) = t^{(a+p)/(2a)}$. Moreover, for simplicity, we assume that the singular value decomposition of A is available. Then we can use this to build reconstructions

$$(8) \quad x_n^\delta := \sum_{j=1}^n \frac{1}{s_j} \langle y^\delta, v_j \rangle u_j \quad (n \in \mathbb{N})$$

from the singular value decomposition. Straightforward calculations yield the order optimal error bound

$$(9) \quad \|x - x_n^\delta\|_X \leq \varphi(s_{n+1}^2) + \frac{\delta}{s_n}, \quad (n \in \mathbb{N}).$$

For the cardinality k_* from (6) this gives the error bound

$$\|x - x_{k_*}^\delta\|_X \leq 2\varphi(s_{k_*}^2).$$

What will be the size of k_* in terms of the dimension d and of the noise level δ ? Taking into account Weyl's monotonicity theorem and the present structure of the function Θ , we see that, for small $0 < \delta < 1$,

$$(10) \quad k_*(\delta, d) \asymp \left(\frac{1}{\delta}\right)^{\frac{d}{a+p}} \quad \text{as } \delta \rightarrow 0,$$

with an implied rate $\delta \rightarrow \delta^{p/(a+p)}$ as $\delta \rightarrow 0$. This rate is order optimal under the present conditions, but we stress that the discretization level k_* depends on $1/\delta$ with a power having the dimension d in its numerator. Assuming that the leading constant, say $C(d) > 0$ in (10) is bounded away from zero, say $C(d) \geq 1$ for simplicity, and for $a = p = 1/2$ we find that $k_* \asymp 100^d$ in the case of a moderate

noise level $\delta = 10^{-2}$. Such values k_* , however, are not feasible for large dimensions. Hence the problem is difficult (intractable), even for moderate values of $\delta > 0$, when the dimension d is large enough.

3. TRACTABILITY

Following the monographs [13]–[15], we will now suggest a definition that tries to formalize the notion of ‘difficulty’ that we have highlighted in the Examples 2 and 3.

3.1. Tractability of families of ill-posed equations. We want to capture both, difficulties for small level δ , as well as difficulties for large spatial dimensions d . In order to capture the dimensionality d of the problems, the concept of tractability is based on the following construction: For $d \in \mathbb{N}$ we introduce a family of linear (multivariate) problems by considering an associated family of operator equations (1) with compact linear operators $A := A_d$. Having such a family A_d ($d \in \mathbb{N}$), with corresponding singular values $s_j(A_d)$ fixed, the following definition seems to be appropriate, for the index $k_* = k_*(\delta, d)$ from (6).

Definition 2 (weak tractability). We call the family of operator equations (1) with compact linear operators $A := A_d$ ($d \in \mathbb{N}$) *weakly tractable* if we have

$$(11) \quad \lim_{d+1/\delta \rightarrow \infty} Q(\delta, d) = 0 \quad \text{for} \quad Q(\delta, d) := \frac{\log(k_*(\delta, d))}{d + 1/\delta}.$$

Otherwise we call it *intractable*.

Remark 3. The above limit $d + 1/\delta \rightarrow \infty$ means that for each subsequence $\{(\delta_k, d_k)\}_{k=1}^\infty$ with $d_k + 1/\delta_k \rightarrow \infty$ the quotient $Q(\delta_k, d_k)$ tends to zero as $k \rightarrow \infty$. If there are subsequences such that the quotient $Q(\delta_k, d_k)$ is bounded away from zero, then intractability is seen.

In particular, intractability (non-vanishing limit Q) may occur if either

- (I) $d \in \mathbb{N}$ is fixed and $\delta \rightarrow 0$ (intractability in δ), or
- (II) $\delta > 0$ is fixed and $d \rightarrow \infty$ (intractability in d).

By definition, a problem is intractable if k_* is exponential in d or $1/\delta$. Of course, often a sequence of problems may be intractable in both factors, simultaneously.

Remark 4. Formally, we just have a family of problems with the corresponding sequence of compact operators A_d ($d \in \mathbb{N}$), and these do not need to have anything in common. In order to interpret the tractability vs. intractability, “the reader must be convinced that the definition

of A_d ($d \in \mathbb{N}$) properly models the same problem for varying dimensions" (see [22, p. 101]). This is the case for the problem studied here.

With Definition 2 at hand, we may rephrase that the problem family of Example 2 is intractable in δ , whereas the problem family in Example 3 is intractable in d .

Remark 5. We add that there is a zoo of modifications of the notion of tractability, often emphasizing power-type behavior, or other special features, see [13, Chapt. 8]. Here we constrain to the weakest notion of tractability.

3.2. Relation to tractability of direct problems. We now relate the tractability for ill-posed equations to the one as studied in Information-based complexity. As in [13, Chapt. 4] this is based on the notion of a mapping $S: H \rightarrow G$, called *solution operator*. Here $H \subset F$ is the set of problem elements, and the goal is to approximate $S(f)$, $f \in F$ with error measured in G . Within the present context we thus need to find Hilbert spaces, say H and G , and a solution operator acting between these Hilbert spaces H, G , where the unit ball in H represents the set of problem elements from F . In our case the set $F := H_\varphi \subset X$ of problem elements is given in Definition 1. This is seen to be the unit ball in the Hilbert space X_φ , given as

$$(12) \quad X_\varphi := \{x \in \ker^\perp(A), \quad x = \varphi(A^*A)v, \quad \|v\|_X < \infty\},$$

endowed with the natural scalar product

$$(13) \quad \langle x, y \rangle_{X_\varphi} := \langle v, w \rangle_X,$$

where the elements v, w are the unique source elements, because $X_\varphi \subset \ker^\perp(A)$ is restricted to the orthogonal complement to the kernel of A . Actually, the spaces X_φ with φ being index functions generate scales of Hilbert spaces, and we refer to the study [10].

Hence, the numerical problem (in the sense of Information-based complexity) that we consider is $A: X_\varphi \rightarrow Y$. Given $x \in H_\varphi$, the goal is to approximately compute $Ax \in Y$. The error for any algorithm, say $R_k = \psi(N_k(x)): X \rightarrow Y$, using at most information of cardinality k , is given as

$$(14) \quad e(A, R_k) := \sup_{x \in H_\varphi} \|Ax - R_k(x)\|_Y.$$

Next, having a family A_d ($d \in \mathbb{N}$) at hand, for this absolute error criterion the (weak) tractability is derived from the information complexity $n(\varepsilon, d)$, given as

$$(15) \quad n(\varepsilon, d) := \max \left\{ k, \quad \inf_{R_k} e(A_d, R_k) > \varepsilon \right\}.$$

(We tentatively assume that $\varepsilon > 0$ is small enough, such that $\varepsilon < \|A_d\|_{X_\varphi \rightarrow Y}$.) According to [13, § 4.4.2] the problem is (weakly) tractable if

$$(16) \quad \lim_{d+1/\varepsilon \rightarrow \infty} \frac{\log(n(\varepsilon, d))}{d+1/\varepsilon} = 0,$$

otherwise it is intractable. For numerical problems in Hilbert space the information complexity $n(\varepsilon, d)$ is characterized by the singular values $s_{d,k}$, $k \in \mathbb{N}$ of the maps $A_d: X_\varphi(d) \rightarrow Y(d)$ via

$$(17) \quad n(\varepsilon, d) = \max \{k, \quad s_{d,k} > \varepsilon\}.$$

The following holds true.

Theorem 1. *Suppose that we have a problem family A_d ($d \in \mathbb{N}$) with $\text{rank}(A_d) = \infty$ and SVD $(s_{d,k}, u_{d,k}, v_{d,k})$, $k \in \mathbb{N}$, with smoothness given as in Definition 1. Then we have for $k_*(\delta, d)$ from (6) the identity*

$$k_*(\delta, d) = n(\delta, d).$$

Thus, the ill-posed problem family A_d ($d \in \mathbb{N}$) is tractable if and only if the related family of direct problems is tractable.

Proof. We need to find the SVDs of the maps $A_d: X_\varphi(d) \rightarrow Y(d)$. Knowing the SVD of $A_d: X(d) \rightarrow Y(d)$ we argue as follows. For every $d \in \mathbb{N}$, and corresponding SVD $(\tilde{s}_{d,j}, u_{d,j}, v_{d,j})$, $j = 1, 2, \dots$ of the map A_d , we see that

$$(18) \quad A_d \varphi(A_d^* A_d) u_{d,j} = \tilde{s}_{d,j} \varphi(s_{d,j}^2) v_{d,j}, \quad j \in \mathbb{N}.$$

Consequently, the map $A_d: X_\varphi(d) \rightarrow Y(d)$ has the following SVD, namely, in terms of the companion function Θ to φ ,

$$(\Theta(\tilde{s}_{d,j}^2), \varphi(A^* A) u_{d,j}, v_{d,j}), \quad j \in \mathbb{N}.$$

Notice, that by construction, for every $d \in \mathbb{N}$ the elements $\varphi(A_d^* A_d) u_{d,j}$ form an orthonormal basis in X_φ . Thus, we see that

$$(19) \quad n(\delta, d) = \max \{k, \quad \Theta(\tilde{s}_{d,k}^2) > \delta\} = k_*(\delta, d).$$

The proof is complete. \square

Looking into the proof of the theorem we obtain the following

Corollary 1. *Suppose that we have a problem family A_d ($d \in \mathbb{N}$). The following holds true.*

- (1) *If the family of direct problems is tractable, then this holds true for the corresponding family of inverse problems, regardless of the smoothness given in Definition 1.*
- (2) *If there is smoothness with index function φ , for which the family of inverse problems is intractable, then the family of direct problems is intractable.*

Sketch of the proof. Let us temporarily abbreviate $k_*^\varphi(\delta, d) = k_*(\delta, d)$ to highlight the dependency on the smoothness, given in terms of the index function φ . Furthermore, we confine the the case when $k_*^\varphi(\delta, d) \rightarrow \infty$ as $d + 1/\delta \rightarrow \infty$. Hence, if $d + 1/\delta$ is large enough, we may assume that $\varphi(s_{d,k_*}^2) \leq 1$. In this case we learn from (19) that

$$k_*^\varphi(\delta, d) = \max \{k, \Theta(\tilde{s}_{d,k}^2) > \delta\} \leq \max \{k, \tilde{s}_{d,k}^2 > \delta\} = n(\delta, d).$$

The assertions of the corollary are a direct consequence of this relation. \square

Remark 6. A less technical argument is as follows. After rescaling we may assume that $H_\varphi \subseteq B_X$, where B_X is the unit ball in $X(d)$. Thus, if the family is tractable on B_X then it will also be tractable on H_φ , which corresponds to the family of inverse problems. Similarly, if for some φ the family will be intractable, then the family of direct problems (on B_X) will also be intractable.

We close this section with the following observation. Often it may be hard to get grip on the exact value for k_* . Clearly, tractability/intractability is an asymptotic property. If $k_*(\delta, d)$ is uniformly bounded then tractability is clearly seen. So, the interesting case is when $k_*(\delta, d) \rightarrow \infty$ as $d + 1/\delta \rightarrow \infty$. Then the following auxiliary result can be used.

Proposition 3. *Let $a > 0$ and $b \in \mathbb{R}$. The following holds true, whenever $ak_* + b$ is positive.*

- (i) *The problem is tractable if and only if*

$$\lim_{d+1/\delta \rightarrow \infty} \frac{\log(ak_*(\delta, d) + b)}{d + 1/\delta} = 0.$$

- (ii) *If along a subsequence $\{(\delta_k, d_k)\}_{k=1}^\infty$ we have that*

$$Q(\delta_k, d_k) \geq \underline{c} > 0,$$

then $\frac{\log(ak_(\delta, d) + b)}{d + 1/\delta} \geq \underline{c}/2 > 0$ for $d_k + 1/\delta_k$ large enough. Hence the problem family is intractable.*

Sketch of a proof. The analysis is simple, and we just hint that

$$\frac{\log(ak_* + b)}{d + 1/\delta} = \frac{\log\left(ak_*\left(1 + \frac{b}{ak_*}\right)\right)}{d + 1/\delta} = \frac{\log(k_*)}{d + 1/\delta} + \frac{\log(a)}{d + 1/\delta} + \frac{\log\left(1 + \frac{b}{ak_*}\right)}{d + 1/\delta}$$

The last two terms on the right tend to zero as $d + 1/\delta \rightarrow \infty$, and hence have no impact on the asymptotics. \square

4. POWER-TYPE DECAY OF THE SINGULAR VALUES

Here we discuss in more detail another example, showing that the joint limit $d + 1/\delta \rightarrow \infty$ may be crucial. We assume a power-type decay of the singular values of the operator family A_d ($d \in \mathbb{N}$). Specifically we assume that for some power $a > 0$, and a leading term³ $c\left(\frac{1}{d}\right)$ there are constants $0 < m \leq M < \infty$ such that we have

$$(20) \quad m c\left(\frac{1}{d}\right) j^{-a/d} \leq s_j(A_d) \leq M c\left(\frac{1}{d}\right) j^{-a/d} \quad (j \in \mathbb{N}).$$

For the tractability analysis the following observation is useful: Looking at the definition of k_* in (6), and taking into account the bounds from (20), we see the following:

- (1) If $j < \left(\frac{m^2 c\left(\frac{1}{d}\right)^2}{\Theta^{-1}(\delta)}\right)^{d/2a}$ then $k_*(\delta, d) \geq j$.
- (2) If $j \geq \left(\frac{M^2 c\left(\frac{1}{d}\right)^2}{\Theta^{-1}(\delta)}\right)^{d/2a}$ then $k_*(\delta, d) \leq j$.

By taking largest and smallest such values j , respectively, we end up with the following two-sided bound.

$$(21) \quad \left(\frac{m^2 c\left(\frac{1}{d}\right)^2}{\Theta^{-1}(\delta)}\right)^{d/2a} - 1 \leq k_*(\delta, d) \leq \left(\frac{M^2 c\left(\frac{1}{d}\right)^2}{\Theta^{-1}(\delta)}\right)^{d/2a} + 1.$$

Thus, in order to see *tractability* for a given situation, we need to consider the right hand side, whereas the left hand side will be used to show *intractability*. In this context, the assertion of Proposition 3 may be used.

We discuss the impact of the behavior of the leading constant $c\left(\frac{1}{d}\right)$ on tractability/intractability. We will distinguish three benchmark situations. If $c\left(\frac{1}{d}\right)$ is bounded away from zero as $d \rightarrow \infty$, intractability occurs. If $c\left(\frac{1}{d}\right)$ goes to zero quickly, at least linear in $1/d$, then tractability is seen. In the intermediate cases, specifically if $c\left(\frac{1}{d}\right)$ is

³We chose the parametrization in terms of $1/d$, because then the constant $c\left(\frac{1}{d}\right)$ may be decreasing with increasing dimension d .

sublinear in $1/d$, then intractability can be seen for low smoothness. We give details for these cases, next.

a) $c(\frac{1}{d})$ is bounded from below: In this case the analysis is particularly simple.

Proposition 4. *If $c(\frac{1}{d}) \geq \underline{c} > 0$ then the problem family A_d ($d \in \mathbb{N}$) is intractable.*

Proof. If $c(\frac{1}{d})$ is bounded away from zero as $d \rightarrow \infty$, then the lower bound in (21) yields for $\delta_0 < \Theta(m^2 \underline{c}^2)$ an exponential increase in d , and the problem is *intractable* in d . \square

b) $c(1/d)$ is at least linear in $1/d$: In this case the following is seen.

Proposition 5. *If the function $t \mapsto c(t)$ is at least linear, i.e., $c(t) \leq \bar{c}t$ for some constant \bar{c} , then the problem family A_d ($d \in \mathbb{N}$) is tractable.*

For the proof we need the following simple but important observation, which takes into account the special form of the function Θ .

Lemma 1. *For each constant $C_0 > 0$ there is $\delta_0 > 0$ such that*

$$C_0^2 \delta^2 \leq \Theta^{-1}(\delta) \quad \text{whenever } 0 < \delta < \delta_0.$$

Proof. Clearly, for the given smoothness function φ we find $\delta_0 > 0$ such that

$$C_0 \delta \varphi(C_0^2 \delta^2) \leq \delta, \quad \text{for } 0 < \delta \leq \delta_0.$$

By the definition of Θ this yields $\Theta(C_0^2 \delta^2) \leq \delta$ for $0 < \delta \leq \delta_0$, and the proof can easily be completed. \square

Proof of Proposition 5. We distinguish two cases, relevant for the analysis. First let us assume that $\delta d \geq \underline{c} > 0$. Then we apply Lemma 1 with $C_0 := \frac{M\bar{c}}{\underline{c}}$. This gives

$$M^2 \frac{\bar{c}^2}{d^2} \leq C_0^2 \delta^2 \leq \Theta^{-1}(\delta) \quad \text{for } \underline{c}/d \leq \delta \leq \delta_0,$$

which by virtue of (21) implies $k_* \leq 2$ in this case.

Next, let us assume that $\delta d \rightarrow 0$, in particular $\delta \rightarrow 0$. Specifically we may assume $\delta d \leq \bar{c}$, and $\delta \leq \delta_0$. By using the right hand side in (21), we can bound from above as

$$k_* - 1 \leq \left(\frac{M^2 c(\frac{1}{d})^2}{\Theta^{-1}(\delta)} \right)^{d/2a} \leq \left(\frac{M^2 \bar{c}^2}{d^2 \Theta^{-1}(\delta)} \right)^{d/2a}.$$

Lemma 1 with $C_0 := M\bar{c}$ yields $\delta_0 > 0$ such that

$$M^2 \bar{c}^2 / \Theta^{-1}(\delta) \leq 1/\delta^2 \quad (\delta \leq \delta_0),$$

and we can estimate as

$$\frac{\log(k_*(\delta, d) - 1)}{d + 1/\delta} \leq \frac{\delta d}{2a} \log \left(\frac{M^2 \bar{c}^2}{d^2 \Theta^{-1}(\delta)} \right) \leq \frac{\delta d}{2a} \log \left(\frac{1}{(\delta d)^2} \right),$$

which implies that $\frac{\log(k_*(\delta, d) - 1)}{d + 1/\delta}$ tends to zero as $\delta d \rightarrow 0$. By virtue of Proposition 3 the proof is complete, and we have tractability. \square

c) $c\left(\frac{1}{d}\right)$ is sublinear: This case exhibits an interesting feature. To be precise we assume here that the function $1/d \rightarrow c\left(\frac{1}{d}\right)$ is the restriction of a strictly increasing continuous sublinear⁴ index function $c: (0, \infty) \rightarrow (0, \infty)$.

Proposition 6. *Suppose that the function $t \rightarrow c(t)$ is strictly increasing and sublinear. For a constant $C > 1/m^2$, with m from (20), let φ be any index function which satisfies*

$$(22) \quad \varphi(t) \geq \frac{c^{-1}(\sqrt{Ct})}{\sqrt{t}} \quad (0 < t \leq 1/C).$$

Then the problem family is intractable for this smoothness φ .

Proof. First notice, that the right hand side constitutes an index function, since with $s := c^{-1}(\sqrt{Ct})$ we see that

$$\frac{c^{-1}(\sqrt{Ct})}{\sqrt{t}} = \sqrt{C} \frac{s}{c(s)},$$

and the sublinearity of c applies.

If the function φ obeys (22), then with letting $t := \frac{c\left(\frac{1}{d}\right)^2}{C}$ we find that

$$\Theta \left(\frac{c\left(\frac{1}{d}\right)^2}{C} \right) \geq \frac{1}{d},$$

and hence that

$$c\left(\frac{1}{d}\right)^2 \geq C \Theta^{-1}\left(\frac{1}{d}\right),$$

which implies

$$(23) \quad \frac{m^2 c\left(\frac{1}{d}\right)^2}{\Theta^{-1}\left(\frac{1}{d}\right)} \geq C m^2 > 1 \quad (d \in \mathbb{N}).$$

Consequently, for the sequence $(\delta(d), d)$ with $\delta(d) = 1/d$ we can see that $k_*(1/d, d) + 1$ is growing exponentially with d , because the associated constant $C m^2$ is greater than one. By using Proposition 3, this shows intractability under the given smoothness φ from (22). \square

⁴An index function f is called sublinear if the function $t \mapsto t/f(t)$ is an index function.

Example 4. Let us consider the case when $c(t) = t^q$ for some $0 < q < 1$, to maintain sublinearity. Then the benchmark smoothness on the right in (22) is seen to be $t \rightarrow t^{\frac{1-q}{2q}}$. This highlights that we need to have low smoothness if q is close to one, whereas smoothness can be arbitrarily large when q tends to zero.

5. MULTIVARIATE INTEGRATION OPERATOR

The current study was inspired by the investigations in [4] for the operator of d -variate mixed integration $A_d: L^2(0, 1)^d \rightarrow L^2(0, 1)^d$ ($d \in \mathbb{N}$), which is defined for $0 < s_1, \dots, s_d < 1$ as

$$(24) \quad (A_d x)(s_1, \dots, s_d) := \int_0^{s_d} \cdots \int_0^{s_1} x(t_1, \dots, t_d) dt_1 \cdots dt_d.$$

It was shown in [4], and it was evaluated in a different context also in [7], that the singular values behave as

$$(25) \quad s_j(A_d) \asymp C(d) \frac{\log^{d-1}(j)}{j} \quad \text{as } j \rightarrow \infty.$$

Remark 7. Formulas of the form (25) possessing the associated asymptotics as $j \rightarrow \infty$ occur in different applications. For example, we have such formulas in [17] for characterizing the j -th entropy numbers of an embedding operator from Sobolev spaces of dominating mixed smoothness to $L^2(0, 1)^d$, and in [18] for the Kolmogorov j -th-width in case of periodic functions over the d -dimensional torus \mathbb{T}^d . Within the present context, the study [8] is most important. These authors analyze approximation numbers (coinciding with singular values) of Sobolev embeddings over the d -dimensional torus \mathbb{T}^d . So, different behavior of $C(d)$ under the auspices of the formula (25) can be seen in the literature.

The major point in [4] was to stress that the degree of ill-posedness of the operators A_d from (24) with the asymptotics (25) for the singular values proves to be *one*, regardless of the spatial dimension d . Here, the degree of ill-posedness measures the power-type decay of the singular numbers of the operators A_d , and we have that

$$\liminf_{j \rightarrow \infty} \frac{-\log(s_j(A_d))}{\log(j)} = \limsup_{j \rightarrow \infty} \frac{-\log(s_j(A_d))}{\log(j)} = \lim_{j \rightarrow \infty} \frac{-\log(s_j(A_d))}{\log(j)} = 1,$$

see [4, Def. 1.1].

However, such a family A_d ($d \in \mathbb{N}$) of problems may be intractable. Then the degree of ill-posedness does not necessarily reflect the difficulty for solving the related ill-posed problem, and the decision between

tractability and intractability is influenced by the behavior of the leading constant $C(d)$.

If $C(d) \geq \underline{c} > 0$, then the problem is obviously intractable in d , because the function $f(t) = \frac{\log^{d-1}(t)}{t}$ ($t \geq 1$) is growing for $t \in [1, e^{d-1}]$ and thus $\log(k_*(\delta, d)) \geq d - 1$ for sufficiently small $\delta > 0$.

The situation may change when $C(d)$ tends to zero as $d \rightarrow \infty$. To this end, let us assume smoothness given by any index function φ with related companion Θ , as described in Definition 1. The following result is relevant.

Proposition 7. *Suppose that, for a family A_d ($d \in \mathbb{N}$) of compact operators obeying a singular value behavior of the form (25), there is a constant $c_0 > 0$ such that $C(d) \geq c_0 \left(\frac{e}{d-1}\right)^{d-1}$ ($d = 2, 3, \dots$), and hence there is some constant $\underline{c} > 0$, for which the singular values $s_j(A_d)$ are bounded from below by*

$$(26) \quad s_j(A_d) \geq \underline{c} \left(\frac{e}{d-1}\right)^{d-1} \frac{\log^{d-1}(j)}{j} \quad (j, d = 2, 3, \dots).$$

Then the family A_d ($d \in \mathbb{N}$) is intractable in d .

Proof. By the definition of k_* in (6) we can argue as follows: If for an index l we see that $s_l^2(A_d) > \Theta^{-1}(\delta)$ then $k_*(\delta, d) \geq l$. We let $l := \lceil e^{d-1} \rceil$, the smallest integer larger than e^{d-1} , and we fix some $\delta_0 \leq 1/2$ such that $4\Theta^{-1}(\delta_0) < \underline{c}$. Then, for $d \geq 2$, we can bound

$$\begin{aligned} s_l(A_d) &\geq \underline{c} \left(\frac{e}{d-1}\right)^{d-1} \frac{\log^{d-1}(l)}{l} \\ &\geq \underline{c} \left(\frac{e}{d-1}\right)^{d-1} \frac{\log^{d-1}(e^{d-1})}{e^{d-1} + 1} \\ &\geq \frac{\underline{c}}{2} \left(\frac{e}{d-1}\right)^{d-1} \left(\frac{d-1}{e}\right)^{d-1} = \frac{\underline{c}}{2} > \sqrt{\Theta^{-1}(\delta_0)}. \end{aligned}$$

Therefore $s_l^2(A_d) > \Theta^{-1}(\delta_0)$, and hence $k_*(d, \delta_0) \geq e^{d-1}$. But then, for $d \geq 1/\delta_0 \geq 2$ we see that

$$\frac{\log(k_*(d, \delta_0))}{d + 1/\delta_0} \geq \frac{d-1}{2d} \geq 1/4,$$

such that the family A_d ($d \in \mathbb{N}$) of operators is intractable in d . \square

Now we return to the family of multivariate integration operators from (24), and we will show that this class of ill-posed problems is weakly tractable.

Theorem 2. *For the family of compact multivariate integration operators $A_d : L^2(0, 1)^d \rightarrow L^2(0, 1)^d$ defined in (24) we have for the constant $C(d)$ in (25) that*

(27)

$$C(d) \sim \frac{1}{(d-1)! \pi^d} \left(\asymp \frac{1}{\sqrt{d-1}} \left(\frac{e}{\pi(d-1)} \right)^{d-1} \right) \quad \text{as } d \rightarrow \infty,$$

and hence this class of problems is weakly tractable in d .

Proof. Formula (27) is known from [6, Thm. 2]. Indeed, this author bounds the singular values of the multivariate problem from known bounds for the underlying univariate ones by noting that the singular numbers $s_j(A_d)$ are the nonincreasing rearrangement of the d th tensor power of the univariate numbers $s_j(A_1)$. Solving this problem is not a trivial task, and the proof [6, Thm. 2] is based on [8, Thm. 4.3]. The singular values for the univariate integration operator A_1 are known to behave like $s_j(A_1) \sim \frac{1}{\pi^j}$, as $j \rightarrow \infty$. Therefore, we can apply [6, Thm. 2] by setting $s = 1$ and $c = \pi^{-1}$.

The weak tractability in d for the multivariate integration operator can be seen by comparison with the decay rate of $C(d)$ with

$$(28) \quad \tilde{C}(d) \asymp \frac{2^d}{(d-1)!} \quad \text{as } d \rightarrow \infty.$$

The latter has been established in [8, Theorem 4.3] (case $s = 1$) for approximation numbers s_j of embedding operators that follow a rule analog to (25). By virtue of [8, Corollary 5.3] the situation (28) represents (quasi-polynomial) tractability, and hence weak tractability. Since the decay rate of $C(d)$ from (27) is higher than the rate of $\tilde{C}(d)$ from (28), this implies that also the direct problem with multivariate integration operators is weakly tractable in d . By virtue of Corollary 1 this also implies the tractability of the inverse and ill-posed problem. \square

ACKNOWLEDGMENTS

The authors are very grateful to Thomas Kühn (Leipzig) for pointing out the asymptotics in formula (27) for the singular values of the multivariate integration operator. The first named author gratefully acknowledges the support of the Leibniz Center for Informatics, where several discussions about this research were held during the Dagstuhl Seminar “Algorithms and Complexity for Continuous Problems” (Seminar ID 23351). The second named author has been supported by the German Science Foundation (DFG) under grant HO 1454/13-1 (Project No. 453804957).

REFERENCES

- [1] R. Bhatia. *Matrix Analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997.
- [2] A. Charpentier, J.-D. Fermanian, and O. Scaillet. The estimation of copulas: theory and practice. In J. Rank, editor, *Copulas: From Theory to Application in Finance*, pages 35–64. Risk Books, London, 2006.
- [3] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.
- [4] B. Hofmann and H.-J. Fischer. A note on the degree of ill-posedness for mixed differentiation on the d -dimensional unit cube. *J. Inverse Ill-Posed Probl.*, 31(6):949–957, 2023.
- [5] B. Hofmann, P. Mathé, and M. Schieck. Modulus of continuity for conditionally stable ill-posed problems in Hilbert space. *J. Inverse Ill-Posed Probl.*, 16(6):567–585, 2008.
- [6] D. Krieg. Tensor power sequences and the approximation of tensor product operators. *J. Complexity*, 44:30–51, 2018.
- [7] T. Kühn and W. Linde. Optimal series representation of fractional Brownian sheets. *Bernoulli*, 8(5):669–696, 2002.
- [8] T. Kühn, W. Sickel, and T. Ullrich. Approximation of mixed order Sobolev functions on the d -torus: asymptotics, preasymptotics, and d -dependence. *Constr. Approx.*, 42(3):353–398, 2015.
- [9] P. Mathé. Bayesian inverse problems with non-commuting operators. *Math. Comp.*, 88(320):2897–2912, 2019.
- [10] P. Mathé and S. V. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19(3):789–803, 2003.
- [11] P. Mathé and S. V. Pereverzev. Complexity of linear ill-posed problems in Hilbert space. *J. Complexity*, 38:50–67, 2017.
- [12] A. Neubauer. When do Sobolev spaces form a Hilbert scale? *Proc. Amer. Math. Soc.*, 103(2):557–562, 1988.
- [13] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Vol. 1: Linear information*, volume 6 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2008.
- [14] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Vol. II: Standard information for functionals*, volume 12 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2010.
- [15] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Vol. III: Standard information for operators*, volume 18 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2012.
- [16] L. Plaskota and P. Siedlecki. Worst case tractability of linear problems in the presence of noise: linear information. *J. Complexity*, 79:Paper No. 101782, 20, 2023.
- [17] H.-J. Schmeisser. Recent developments in the theory of function spaces with dominating mixed smoothness. In *NAFSA 8—Nonlinear Analysis, Function Spaces and Applications. Vol. 8*, pages 144–204. Czech. Acad. Sci., Prague, 2007.

- [18] V. Temlyakov. *Multivariate Approximation*, volume 32 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2018.
- [19] J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski. *Information-based complexity*. Computer Science and Scientific Computing. Academic Press Inc., Boston, MA, 1988. With contributions by A. G. Werschulz and T. Boulton.
- [20] H. Triebel. *Interpolation theory, function spaces, differential operators*. Johann Ambrosius Barth, Heidelberg, second edition, 1995.
- [21] D. Uhlig and R. Unger. Nonparametric copula density estimation using a Petrov-Galerkin projection. In M. Scherer K. Glau and R. Zagst, editors, *Innovations in Quantitative Risk Management: TU München, September 2013*, pages 423–438. Springer open, Cham/Heidelberg/New York, 2015. <https://link.springer.com/book/10.1007/978-3-319-09114-3>.
- [22] H. Woźniakowski. Tractability and strong tractability of linear multivariate problems. *J. Complexity*, 10(1):96–128, 1994.

WEIERSTRASS INSTITUTE FOR APPLIED ANALYSIS AND STOCHASTICS, MOHREN-
STRASSE 39, 10117 BERLIN, GERMANY

FACULTY OF MATHEMATICS, CHEMNITZ UNIVERSITY OF TECHNOLOGY, 09107
CHEMNITZ, GERMANY