

## ON DECOMPOSITIONS OF MULTIVARIATE FUNCTIONS

F. Y. KUO, I. H. SLOAN, G. W. WASILKOWSKI, AND H. WOŹNIAKOWSKI

*This is a more elaborated version of our paper to appear in *Mathematics of Computation*.  
Replacement text and additional text are marked by ~~strikeout~~ and **highlight**.*

ABSTRACT. We present formulas that allow us to decompose a function  $f$  of  $d$  variables into a sum of  $2^d$  terms  $f_{\mathbf{u}}$  indexed by subsets  $\mathbf{u}$  of  $\{1, \dots, d\}$ , where each term  $f_{\mathbf{u}}$  depends only on the variables with indices in  $\mathbf{u}$ . The decomposition depends on the choice of  $d$  commuting projections  $\{P_j\}_{j=1}^d$ , where  $P_j(f)$  does not depend on the variable  $x_j$ . We present an explicit formula for  $f_{\mathbf{u}}$ , which is new even for the ANOVA and anchored decompositions; both are special cases of the general decomposition. We show that the decomposition is minimal in the following sense: if  $f$  is expressible as a sum in which there is no term that depends on all of the variables indexed by the subset  $\mathbf{z}$  then, for every choice of  $\{P_j\}_{j=1}^d$ , the terms  $f_{\mathbf{u}} = 0$  for all subsets  $\mathbf{u}$  containing  $\mathbf{z}$ . Furthermore, in a reproducing kernel Hilbert space setting, we give sufficient conditions for the terms  $f_{\mathbf{u}}$  to be mutually orthogonal.

### 1. INTRODUCTION

Functions with very large numbers of variables arise nowadays in many contexts, ranging from molecular chemistry to probability and statistics and the pricing of financial options. Often there is a need to approximate such functions by a small number of terms, either concretely (as in the case of “model reduction” in molecular chemistry), or notionally (as in modern approaches to multivariate integration, which seek to exploit the existence of low-dimensional approximations without actually constructing them).

To add to the existing toolkit for the study of low-dimensional approximations, we introduce in this paper general decomposition formulas of functions of  $d$  variables into sums of  $2^d$  terms, with each term depending on a group of variables indexed by a particular subset of  $\{1, \dots, d\}$ . That is, the decomposition is of the form

$$f = \sum_{\mathbf{u} \subseteq \{1, \dots, d\}} f_{\mathbf{u}},$$

where the summation is over all subsets of  $\{1, \dots, d\}$ , and  $f_{\mathbf{u}}$  depends only on the subset of variables  $\{x_j : j \in \mathbf{u}\}$ .

One important example of such a decomposition formula is the ANOVA decomposition, see e.g., [2, 5, 14, 18, 21]. Another is the *anchored* decomposition, see e.g., [5, 7, 8, 15, 19]. Both were already discussed in Sobol’s 1969 book, see [17]. They are both included here as special cases. The results obtained here for the general

---

Received by the editor July 2009.  
2000 *Mathematics Subject Classification*. 41A63, 41A99.

decomposition include not only those that are well known in special cases, such as the recursive form of the ANOVA decomposition, but also results that are new. These include an explicit expression for the general term in the decomposition, and a certain “minimal” property that applies to all decompositions in the class. This new explicit formula may turn out to be very useful for both theoretical analysis and practical computation. In its simplest form the minimal property states that if there exists a (possibly unknown) decomposition of the function in which there is no term involving one particular variable, then the decomposition (whether ANOVA or anchored or any other in the class) never unnecessarily introduces that variable. (As a simple example with  $d = 2$ , the function  $f(x_1, x_2) := 2x_1$  can obviously be written in many non-minimal ways, such as  $f(x_1, x_2) = x_1 - x_2 + (x_1 + x_2)$ . Our decompositions never commit a crime of this kind.) More generally, if there is a decomposition in which no term contains all members of some subset of the variables (say  $x_1, x_2, x_4$ ), then in our decomposition every term that involves all variables in that subset vanishes.

The general decomposition depends on a set of commuting projections  $\{P_j\}_{j=1}^d$ , where  $P_j$  acting on a multivariate function has the effect of eliminating the variable  $x_j$  in some way, while leaving all other variables unchanged. (For example, in the anchored case  $P_j$  freezes  $x_j$  at some specified value  $c_j$ . In the ANOVA case  $P_j$  integrates with respect to  $x_j$ .) We show that, for a given set of projections  $\{P_j\}_{j=1}^d$ , our decomposition is the unique decomposition of the form  $f = \sum_{\mathbf{u} \subseteq \{1, \dots, d\}} f_{\mathbf{u}}$  with the property  $P_j(f_{\mathbf{u}}) = 0$  whenever  $j \in \mathbf{u}$ . The projections provide a useful tool for the proof of various properties, such as the minimal property mentioned above.

We also show that decompositions defined by the commuting projections are orthogonal decompositions in certain reproducing kernel Hilbert spaces. The orthogonality property is a key to the successful theoretical and practical development of integration and approximation schemes in high-dimensional reproducing kernel Hilbert spaces. Since the orthogonality is with respect to the inner product of a Hilbert space, it is clear that orthogonality can only be achieved if there is a correct association between the set of commuting projections and a particular reproducing kernel Hilbert space.

The general decomposition makes an appearance in the field of computational chemistry under the heading *high-dimensional model representation (HDMR)*. See e.g., the papers [11, 12, 13] and the references therein. In particular, their *cut-HDMR* is essentially our anchored decomposition. In the recent survey article [5] there is also a section discussing ANOVA-like decompositions. These decompositions are for product measures with densities, and for the Dirac measures they become anchored decompositions.

The paper is organized as follows. The class of decompositions and their properties are presented in Section 2. Their minimal property is shown in Section 3, and orthogonality in reproducing kernel Hilbert spaces is considered in Section 4. As examples we consider two Sobolev spaces – the anchored and unanchored Sobolev spaces – both with arbitrary smoothness, and show that the corresponding orthogonal decompositions are precisely the anchored and the ANOVA decompositions, respectively. Section 5 focuses on the computational cost of the anchored decomposition. The final section contains a short discussion about potential applications.

A more elaborated version of this paper, including a derivation of (3.2) and of the unanchored Sobolev space, can be found on [www.maths.unsw.edu.au/~fkuo/pubs/](http://www.maths.unsw.edu.au/~fkuo/pubs/). A derivation of the unanchored Sobolev space<sup>1</sup> can be found in the Appendix.

## 2. A GENERAL DECOMPOSITION FORMULA

Let  $\mathcal{F}$  be a linear space of real functions  $f(\mathbf{x}) = f(x_1, \dots, x_d)$  defined on a domain  $D \subseteq \mathbb{R}^d$ . We introduce  $[1..d] := \{1, \dots, d\}$  as a shorthand notation for the set of indices from 1 to  $d$ , and say that a function  $f \in \mathcal{F}$  *does not depend on*  $x_j$  for some  $j \in [1..d]$  if

$$\mathbf{x}_{[1..d]\setminus\{j\}} = \mathbf{y}_{[1..d]\setminus\{j\}} \implies f(\mathbf{x}) = f(\mathbf{y}) \quad \text{for all } \mathbf{x}, \mathbf{y} \in D.$$

Here and elsewhere for a given  $\mathbf{x} = (x_1, \dots, x_d)$  we use the notation  $\mathbf{x}_{\mathbf{u}} = (x_j)_{j \in \mathbf{u}}$ .

Let  $\{P_j\}_{j=1}^d$  be a set of commuting projections on  $\mathcal{F}$  with the following property:

$$(2.1) \quad P_j(f) = f \text{ if } f \text{ does not depend on } x_j, \text{ and } P_j(f) \text{ does not depend on } x_j.$$

Intuitively, we can think of  $P_j$  as acting only on the  $j$ th variable  $x_j$ , to produce a result that does not depend on  $x_j$ ; and if  $P_j$  acts on a function that is independent of  $x_j$  it leaves the function unchanged.

For  $\mathbf{u} \subseteq [1..d]$ , we define  $P_{\mathbf{u}} := \prod_{j \in \mathbf{u}} P_j$ , with  $P_{\emptyset} := I$ , the identity operator. Since the commuting projections in  $P_{\mathbf{u}}$  can be applied in any order and since  $P_j P_j = P_j$ , we have  $P_{\mathbf{u}} P_{\mathbf{v}} = P_{\mathbf{u} \cup \mathbf{v}}$  for all  $\mathbf{u}, \mathbf{v} \subseteq [1..d]$ .

Formally, for each  $\mathbf{u} \subseteq [1..d]$  we can define  $\mathcal{F}_{\mathbf{u}} := P_{[1..d]\setminus\mathbf{u}}(\mathcal{F})$ . Then  $\mathcal{F}_{\mathbf{u}}$  is a subspace of  $\mathcal{F}$  and consists of functions which depend only on the variables listed in  $\mathbf{u}$ , i.e., on  $\mathbf{x}_{\mathbf{u}} = (x_j)_{j \in \mathbf{u}}$ , and which are constant with respect to the variables outside  $\mathbf{u}$ . For an arbitrary  $g_{\mathbf{u}} \in \mathcal{F}_{\mathbf{u}}$ , it is sometimes useful to ignore the uninteresting variables, by writing  $g_{\mathbf{u}}(\mathbf{x}) = g_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}})$ . We have  $\mathcal{F}_{[1..d]} = \mathcal{F}$ , and  $\mathcal{F}_{\emptyset}$  contains only constant functions. Moreover,  $\mathcal{F}_{\mathbf{v}} \subseteq \mathcal{F}_{\mathbf{u}}$  if  $\mathbf{v} \subseteq \mathbf{u}$ .

The commuting projections  $\{P_j\}_{j=1}^d$  can be very general as long as each  $P_j$  satisfies (2.1). We have in mind two obvious examples:

- Freezing  $x_j$  at  $c_j$ :

$$(2.2) \quad P_j(f)(\mathbf{x}) = f(x_1, \dots, x_{j-1}, c_j, x_{j+1}, \dots, x_d) \quad \text{for } \mathbf{x} \in D.$$

- With  $D = [0, 1]^d$ , integrating out  $x_j$ :

$$(2.3) \quad P_j(f)(\mathbf{x}) = \int_0^1 f(x_1, \dots, x_{j-1}, t, x_{j+1}, \dots, x_d) dt \quad \text{for } \mathbf{x} \in D.$$

Depending on the choice of the projections  $P_j$ , there are some implicit assumptions on the domain  $D$  and the linear space  $\mathcal{F}$ . In the first example we need to assume that for all  $j \in [1..d]$  and all  $\mathbf{x} \in D$  we have  $(x_1, \dots, x_{j-1}, c_j, x_{j+1}, \dots, x_d) \in D$ . In the second example where  $D = [0, 1]^d$ , we require  $f$  to be integrable on  $[0, 1]^d$ . Other obvious examples include a weighted sum  $P_j(f)(\mathbf{x}) = \sum_{m=1}^{M_j} w_{j,m} f(x_1, \dots, x_{j-1}, c_{j,m}, x_{j+1}, \dots, x_d)$ , with  $\sum_{m=1}^{M_j} w_{j,m} = 1$ , and a weighted integral  $P_j(f)(\mathbf{x}) = \int_0^1 f(x_1, \dots, x_{j-1}, t, x_{j+1}, \dots, x_d) \rho_j(t) dt$ , with  $\int_0^1 \rho_j(t) dt = 1$ . There is no requirement that the  $d$  projections have to be of the same form.

We are now ready to introduce the general decomposition formula. There are multiple ways to present the same theorem. Here we take the most elegant approach

<sup>1</sup>This space is derived under the condition that its periodic subspace coincides with the weighted Korobov space of the same smoothness.

which was brought to our attention by Griebel, see also [5, 11]. Let  $a_j$  and  $b_j$  be commuting linear operators from  $\mathcal{F}$  to  $\mathcal{F}$ . Using the simple property

$$(2.4) \quad \prod_{j \in \mathbf{u}} (a_j + b_j) = \sum_{\mathbf{v} \subseteq \mathbf{u}} \left( \prod_{j \in \mathbf{v}} a_j \right) \left( \prod_{j \in \mathbf{u} \setminus \mathbf{v}} b_j \right),$$

we can express the identity operator  $I$  as

$$(2.5) \quad I = \prod_{j=1}^d [(I - P_j) + P_j] = \sum_{\mathbf{u} \subseteq [1..d]} \left( \prod_{j \in \mathbf{u}} (I - P_j) \right) \left( \prod_{j \in [1..d] \setminus \mathbf{u}} P_j \right).$$

**Theorem 2.1.** *Let  $\{P_j\}_{j=1}^d$  be commuting projections on  $\mathcal{F}$  satisfying (2.1). For every function  $f \in \mathcal{F}$  and each subset  $\mathbf{u} \subseteq [1..d]$ , we define*

$$(2.6) \quad f_{\mathbf{u}} := \left( \prod_{j \in \mathbf{u}} (I - P_j) \right) P_{[1..d] \setminus \mathbf{u}}(f).$$

This leads to a decomposition of  $f$  given by

$$(2.7) \quad f = \sum_{\mathbf{u} \subseteq [1..d]} f_{\mathbf{u}},$$

where  $f_{\mathbf{u}}$  depends only on variables with indices in  $\mathbf{u}$ .

- (a) Define the functions  $g_{\mathbf{u}}$  recursively with respect to the cardinality of  $\mathbf{u} \subseteq [1..d]$  by

$$g_{\emptyset} := P_{[1..d]}(f), \quad g_{\mathbf{u}} := P_{[1..d] \setminus \mathbf{u}}(f) - \sum_{\mathbf{v} \subsetneq \mathbf{u}} g_{\mathbf{v}}.$$

Then  $f_{\mathbf{u}} = g_{\mathbf{u}}$  for all  $\mathbf{u} \subseteq [1..d]$ .

- (b) Define the functions  $h_{\mathbf{u}}$  for  $\mathbf{u} \subseteq [1..d]$  by

$$h_{\mathbf{u}} := \sum_{\mathbf{v} \subseteq \mathbf{u}} (-1)^{|\mathbf{u}| - |\mathbf{v}|} P_{[1..d] \setminus \mathbf{v}}(f).$$

Then  $f_{\mathbf{u}} = h_{\mathbf{u}}$  for all  $\mathbf{u} \subseteq [1..d]$ .

- (c) Assume that  $f = \sum_{\mathbf{u} \subseteq [1..d]} w_{\mathbf{u}}$ , where  $w_{\mathbf{u}}$  depends only on the variables with indices in the set  $\mathbf{u}$  and satisfies the ‘‘annihilating’’ property

$$(2.8) \quad P_j(w_{\mathbf{u}}) = 0 \quad \text{for all } j \in \mathbf{u}.$$

Then  $f_{\mathbf{u}} = w_{\mathbf{u}}$  for all  $\mathbf{u} \subseteq [1..d]$ . In other words, (2.7) is the unique decomposition of this form with the annihilating property.

*Proof.* The fact that the functions  $f_{\mathbf{u}}$  defined by (2.6) yield (2.7) follows immediately from (2.5). We now proceed to prove parts (a), (b), and (c).

It is clear from (2.6) that  $f_{\emptyset} = P_{[1..d]}(f) = g_{\emptyset}$ . For  $\mathbf{u} \neq \emptyset$  we can rewrite (2.6) as

$$\begin{aligned} f_{\mathbf{u}} &= P_{[1..d] \setminus \mathbf{u}}(f) - \left( I - \prod_{j \in \mathbf{u}} (I - P_j) \right) P_{[1..d] \setminus \mathbf{u}}(f) \\ &= P_{[1..d] \setminus \mathbf{u}}(f) - \left( \prod_{j \in \mathbf{u}} [(I - P_j) + P_j] - \prod_{j \in \mathbf{u}} (I - P_j) \right) P_{[1..d] \setminus \mathbf{u}}(f). \end{aligned}$$

We use (2.4) for the first product and obtain

$$\begin{aligned} f_{\mathbf{u}} &= P_{[1..d]\setminus\mathbf{u}}(f) - \left( \sum_{\mathbf{v}\subsetneq\mathbf{u}} \left( \prod_{j\in\mathbf{v}} (I - P_j) \right) P_{\mathbf{u}\setminus\mathbf{v}} \right) P_{[1..d]\setminus\mathbf{u}}(f) \\ &= P_{[1..d]\setminus\mathbf{u}}(f) - \sum_{\mathbf{v}\subsetneq\mathbf{u}} \left( \prod_{j\in\mathbf{v}} (I - P_j) \right) P_{[1..d]\setminus\mathbf{v}}(f) = P_{[1..d]\setminus\mathbf{u}}(f) - \sum_{\mathbf{v}\subsetneq\mathbf{u}} f_{\mathbf{v}}, \end{aligned}$$

where we used the definition (2.6) with  $\mathbf{u}$  replaced by  $\mathbf{v}$ . This proves that  $f_{\mathbf{u}}$  satisfies the same recurrence as  $g_{\mathbf{u}}$ . Since  $f_{\emptyset} = g_{\emptyset}$ , we conclude that  $f_{\mathbf{u}} = g_{\mathbf{u}}$  for all  $\mathbf{u} \subseteq [1..d]$ .

To prove that  $f_{\mathbf{u}} = h_{\mathbf{u}}$  for all  $\mathbf{u} \subseteq [1..d]$ , we apply (2.4) to the first product in (2.6):

$$\begin{aligned} f_{\mathbf{u}} &= \left( \sum_{\mathbf{v}\subsetneq\mathbf{u}} \left( \prod_{j\in\mathbf{v}} I \right) \left( \prod_{j\in\mathbf{u}\setminus\mathbf{v}} (-P_j) \right) \right) P_{[1..d]\setminus\mathbf{u}}(f) \\ &= \sum_{\mathbf{v}\subsetneq\mathbf{u}} (-1)^{|\mathbf{u}|-|\mathbf{v}|} P_{[1..d]\setminus\mathbf{v}}(f) = h_{\mathbf{u}}. \end{aligned}$$

Finally, for  $f = \sum_{\mathbf{v}\subseteq[1..d]} w_{\mathbf{v}}$  as given in (c), we have from (2.6) that

$$f_{\mathbf{u}} = \sum_{\mathbf{v}\subseteq[1..d]} \left( \prod_{j\in\mathbf{u}} (I - P_j) \right) P_{[1..d]\setminus\mathbf{u}}(w_{\mathbf{v}}).$$

If  $\mathbf{v} \neq \mathbf{u}$ , then either there exists an index  $k \in \mathbf{v}$  and  $k \notin \mathbf{u}$ , or there exists an index  $k \in \mathbf{u}$  and  $k \notin \mathbf{v}$ . In the former case we have from (2.8) that  $P_k(w_{\mathbf{v}}) = 0$ ; in the latter case we have  $(I - P_k)(w_{\mathbf{v}}) = 0$ . Hence only the  $\mathbf{v} = \mathbf{u}$  term survives, and we have

$$f_{\mathbf{u}} = \left( \prod_{j\in\mathbf{u}} (I - P_j) \right) P_{[1..d]\setminus\mathbf{u}}(w_{\mathbf{u}}) = w_{\mathbf{u}}.$$

This completes the proof.  $\square$

It is clear from the theorem that, instead of defining the decomposition through (2.6), we could use any of the parts (a), (b), or (c) as the definition for  $f_{\mathbf{u}}$ .

We now present two specific examples of the decomposition formula: the first requires integrals of  $f$ , while the second requires only function evaluations of  $f$ .

**Example 2.2** (ANOVA Decomposition). Taking  $D = [0, 1]^d$  and  $\mathcal{F}$  as the space of square integrable functions, and choosing each  $P_j$  according to (2.3), we recover the well-known ANOVA *decomposition* (analysis of variance), see e.g., [2, 5, 14, 18, 21]. This has the form

$$(2.9) \quad f = \sum_{\mathbf{u}\subseteq[1..d]} f_{\mathbf{u},*},$$

with  $f_{\mathbf{u},*}$  depending only on the variables listed in  $\mathbf{u}$  and satisfying

$$\begin{aligned} (a) \quad f_{\emptyset,*} &= \int_{[0,1]^d} f(\mathbf{x}) \, d\mathbf{x}, \quad f_{\mathbf{u},*}(\mathbf{x}_{\mathbf{u}}) = \int_{[0,1]^{d-|\mathbf{u}|}} f(\mathbf{x}) \, d\mathbf{x}_{[1..d]\setminus\mathbf{u}} - \sum_{\mathbf{v}\subsetneq\mathbf{u}} f_{\mathbf{v},*}(\mathbf{x}_{\mathbf{v}}), \\ (b) \quad f_{\mathbf{u},*}(\mathbf{x}_{\mathbf{u}}) &= \sum_{\mathbf{v}\subseteq\mathbf{u}} (-1)^{|\mathbf{u}|-|\mathbf{v}|} \int_{[0,1]^{d-|\mathbf{v}|}} f(\mathbf{x}) \, d\mathbf{x}_{[1..d]\setminus\mathbf{v}}, \\ (c) \quad \int_0^1 f_{\mathbf{u},*}(\mathbf{x}_{\mathbf{u}}) \, dx_j &= 0 \quad \text{for all } j \in \mathbf{u}. \end{aligned}$$

Here  $d\mathbf{x}_{[1..d]\setminus\mathbf{u}}$  denotes  $\prod_{j \in [1..d]\setminus\mathbf{u}} dx_j$ . For  $\mathbf{u} = [1..d]$ , the integral over the empty set is replaced by  $f(\mathbf{x})$ . In the literature, the recursive form (a) of the ANOVA terms is typically used as the definition of ANOVA decomposition. However, Sobol' used the annihilating property (c) as the definition of the ANOVA terms, see e.g., [18]. The explicit form (b) obtained here is new. The ANOVA decomposition might be difficult to use in practice because it requires integrals of  $f$  which are typically unavailable. **Having said that, it does provide a natural variance decomposition of  $f$ . Indeed, it is well-known and easy to show that the ANOVA terms are orthogonal in  $\mathcal{L}_2([0, 1]^d)$ , i.e.,**

$$\int_{[0,1]^d} f_{\mathbf{u},*}(\mathbf{x}_{\mathbf{u}}) f_{\mathbf{v},*}(\mathbf{x}_{\mathbf{v}}) d\mathbf{x} = 0 \quad \text{for all } \mathbf{u} \neq \mathbf{v}.$$

Since  $f$  is square integrable (and hence so too are all the ANOVA terms), it follows that we have the variance decomposition

$$\sigma^2(f) = \sum_{\mathbf{u} \subseteq [1..d]} \sigma^2(f_{\mathbf{u},*}),$$

where  $\sigma^2(f) = \int_{[0,1]^d} f^2(\mathbf{x}) d\mathbf{x} - (\int_{[0,1]^d} f(\mathbf{x}) d\mathbf{x})^2$ ,  $\sigma^2(f_{\emptyset,*}) = 0$ , and for  $\mathbf{u} \neq \emptyset$  we have  $\sigma^2(f_{\mathbf{u},*}) = \int_{[0,1]^{|\mathbf{u}|}} f_{\mathbf{u},*}^2(\mathbf{x}_{\mathbf{u}}) d\mathbf{x}_{\mathbf{u}}$ .

**Example 2.3** (Anchored Decomposition). Let  $\mathbf{c} \in D$ , and let  $(\mathbf{x}; \mathbf{c})_{\mathbf{u}}$  denote the  $d$ -dimensional vector whose  $j$ th component is  $x_j$  if  $j \in \mathbf{u}$  and  $c_j$  if  $j \notin \mathbf{u}$ . Assuming that  $(\mathbf{x}; \mathbf{c})_{\mathbf{u}} \in D$  whenever  $\mathbf{x} \in D$ , and choosing each  $P_j$  according to (2.2), we obtain the *anchored decomposition with respect to the anchor  $\mathbf{c}$* . This has the form

$$(2.10) \quad f = \sum_{\mathbf{u} \subseteq [1..d]} f_{\mathbf{u},\mathbf{c}},$$

with  $f_{\mathbf{u},\mathbf{c}}$  depending only on the variables listed in  $\mathbf{u}$  and satisfying

- (a)  $f_{\emptyset,\mathbf{c}} = f(\mathbf{c})$ ,  $f_{\mathbf{u},\mathbf{c}}(\mathbf{x}_{\mathbf{u}}) = f((\mathbf{x}; \mathbf{c})_{\mathbf{u}}) - \sum_{\mathbf{v} \subsetneq \mathbf{u}} f_{\mathbf{v},\mathbf{c}}(\mathbf{x}_{\mathbf{v}})$ ,
- (b)  $f_{\mathbf{u},\mathbf{c}}(\mathbf{x}_{\mathbf{u}}) = \sum_{\mathbf{v} \subseteq \mathbf{u}} (-1)^{|\mathbf{u}|-|\mathbf{v}|} f((\mathbf{x}; \mathbf{c})_{\mathbf{v}})$ ,
- (c)  $f_{\mathbf{u},\mathbf{c}}(\mathbf{x}_{\mathbf{u}}) = 0$  whenever  $x_j = c_j$  for all  $j \in \mathbf{u}$ .

The recursive form (a) of the anchored decomposition already appeared in earlier literature, see e.g., [7, 8]. The explicit form (b) is new. This is attractive because it involves only function evaluations of  $f$  with respect to the anchor  $\mathbf{c}$ ; see Section 5 for a discussion on computational cost. We add that, in some specific setting, the terms  $f_{\mathbf{u},\mathbf{c}}$  from the anchored decomposition can also be expressed as integrals of the mixed first partial derivatives of  $f$ , see [7, Proposition 1]. This is useful for some theoretical error analysis.

As a simple illustration, we consider  $d = 2$  and  $f(x_1, x_2) = 1 + x_1 + x_2 + x_1x_2$ . We have the following decompositions:

	$f_{\emptyset}$	$f_{\{1\}}(x_1)$	$f_{\{2\}}(x_2)$	$f_{\{1,2\}}(x_1, x_2)$
Anchored at $\mathbf{c} = (0, 0)$	1	$x_1$	$x_2$	$x_1x_2$
Anchored at $\mathbf{c} = (1, 1)$	4	$2x_1 - 2$	$2x_2 - 2$	$x_1x_2 - x_1 - x_2 + 1$
ANOVA	$\frac{9}{4}$	$\frac{3x_1 - 3}{2}$	$\frac{3x_2 - 3}{2}$	$x_1x_2 - \frac{x_1}{2} - \frac{x_2}{2} + \frac{1}{4}$

(For this function, the ANOVA decomposition happens to be the same as the anchored decomposition with  $\mathbf{c} = (\frac{1}{2}, \frac{1}{2})$ .) Clearly the choice of the anchor plays an important role.

### 3. THE DECOMPOSITION IS MINIMAL

In this section, we prove that the decomposition from Theorem 2.1 with respect to any choice of commuting projections  $\{P_j\}_{j=1}^d$  is minimal in the sense of *never introducing unnecessary terms*.

In general, there are infinitely many ways to express a function  $f$  of  $d$  variables  $\mathbf{x} = (x_1, \dots, x_d)$  in the form

$$(3.1) \quad f = \sum_{\mathbf{u} \subseteq [1..d]} t_{\mathbf{u}},$$

with  $t_{\mathbf{u}}$  depending only on variables listed in  $\mathbf{u}$ , i.e., on  $\mathbf{x}_{\mathbf{u}} = (x_j)_{j \in \mathbf{u}}$ . With respect to a particular decomposition (3.1), it is common to refer to the terms  $\{t_{\mathbf{u}} : |\mathbf{u}| = \ell\}$  collectively as the “order- $\ell$  terms”. Similarly, in the discussion below we shall refer to the terms  $\{t_{\mathbf{u}} : \mathbf{z} \subseteq \mathbf{u} \subseteq [1..d]\}$  for a given  $\mathbf{z} \subseteq [1..d]$  as the “super- $\mathbf{z}$  terms”. In other words, the super- $\mathbf{z}$  terms are all the terms  $t_{\mathbf{u}}$  for which  $\mathbf{u}$  contains  $\mathbf{z}$ .

The minimal property of the decomposition from Theorem 2.1 is stated below.

**Theorem 3.1.** *Let  $f$  be a function from  $\mathcal{F}$  and let  $\mathbf{z}$  be a given subset of  $[1..d]$ . Assume that there exists a decomposition (3.1) of  $f$  in which all super- $\mathbf{z}$  terms are zero, i.e.,*

$$t_{\mathbf{u}} = 0 \text{ for all } \mathbf{u} \text{ containing } \mathbf{z}.$$

*Then the decomposition from Theorem 2.1 also has all super- $\mathbf{z}$  terms equal to zero, i.e.,*

$$f_{\mathbf{u}} = 0 \text{ for all } \mathbf{u} \text{ containing } \mathbf{z}.$$

*Proof.* Suppose that  $f = \sum_{\mathbf{w} \subseteq [1..d]} t_{\mathbf{w}}$  with  $t_{\mathbf{w}}$  depending only on the variables listed in  $\mathbf{w}$ , and with  $t_{\mathbf{w}} = 0$  for all  $\mathbf{w}$  containing  $\mathbf{z}$ . For each  $\mathbf{u}$  satisfying  $\mathbf{z} \subseteq \mathbf{u} \subseteq [1..d]$ , we want to prove that  $f_{\mathbf{u}} = 0$  in the decomposition in Theorem 2.1. We see from (2.6) that

$$f_{\mathbf{u}} = \left( \prod_{j \in \mathbf{u}} (I - P_j) \right) P_{[1..d] \setminus \mathbf{u}}(f) = \sum_{\mathbf{w} \subseteq [1..d]} \left( \prod_{j \in \mathbf{u}} (I - P_j) \right) P_{[1..d] \setminus \mathbf{u}}(t_{\mathbf{w}}).$$

Note that for  $\mathbf{u} \subseteq \mathbf{w}$  we have  $\mathbf{z} \subseteq \mathbf{w}$ , and it follows from the assumption on  $t_{\mathbf{w}}$  that  $t_{\mathbf{w}} = 0$ . For  $\mathbf{u} \not\subseteq \mathbf{w}$ , there exists an index  $k \in \mathbf{u}$  and  $k \notin \mathbf{w}$ , and for this index  $k$  we have  $(I - P_k)(t_{\mathbf{w}}) = 0$  since  $t_{\mathbf{w}}$  does not depend on  $x_k$ . Hence we conclude that  $f_{\mathbf{u}} = 0$ .  $\square$

From Theorem 3.1 we see that if for different sets  $\mathbf{z}$  there exist different decompositions of  $f$  in which all super- $\mathbf{z}$  terms are zero, then the decomposition from Theorem 2.1 has all super- $\mathbf{z}$  terms equal to zero for *all* of these different sets  $\mathbf{z}$ . In other words, the decomposition from Theorem 2.1 uses lower order terms as much as possible, and does not introduce unnecessary higher order terms. Furthermore, this holds for all possible choices of commuting projections  $\{P_j\}_{j=1}^d$  satisfying (2.1). Indeed, if the decomposition with respect to one choice of operators  $P_j$  has all super- $\mathbf{z}$  terms equal to zero, then the decompositions with respect to *all* choices of operators  $P_j$  have all super- $\mathbf{z}$  terms equal to zero.

This minimal property has a strong link with the following notions of dimension:

- We say that a function of  $d$  variables has *cutoff dimension*  $k$  iff it depends only on the first  $k$  variables, and  $k$  is the smallest number for which this holds.
- We say that a function of  $d$  variables is *of order*  $q$  iff it can be written as a sum of functions each depending on at most  $q$  of the  $d$  variables, and  $q$  is the smallest number for which this holds.

These two notions correspond respectively to the more familiar concepts of *truncation dimension* and *superposition dimension*, see e.g., [2, 14, 21], except that the latter two allow small contributions from other groups of variables. Obviously the cutoff dimension of  $f$  is the smallest  $k$  for which a decomposition  $f = \sum_{\mathbf{u} \subseteq [1..k]} t_{\mathbf{u}}$  exists. Similarly, the order of  $f$  is the smallest  $q$  for which a decomposition  $f = \sum_{\mathbf{u} \subseteq [1..d], |\mathbf{u}| \leq q} t_{\mathbf{u}}$  holds. Note that the cutoff dimension and the order of a given function are both uniquely defined, and that the definition is not related to any particular decomposition method. **It is easy to see that**

- $f$  having cutoff dimension  $k$  implies the existence of a decomposition in which all super- $\{j\}$  terms are zero for all  $j \in \{k+1, k+2, \dots, d\}$ , and
- $f$  being of order  $q$  implies the existence of a decomposition in which all super- $\mathbf{z}$  terms are zero for all  $\mathbf{z}$  with  $|\mathbf{z}| > q$ .

**With these two observations in mind,** Theorem 3.1 implies that the decomposition from Theorem 2.1 is minimal with respect to both cutoff dimension and order:

- A function  $f$  of  $d$  variables has cutoff dimension  $k < d$  iff  $k$  is the smallest number such that  $f_{\mathbf{u}} = 0$  for all subsets  $\mathbf{u} \subseteq [1..d]$  for which  $\mathbf{u} \not\subseteq [1..k]$ .
- A function  $f$  of  $d$  variables is of order  $q < d$  iff  $q$  is the smallest number such that  $f_{\mathbf{u}} = 0$  for all subsets  $\mathbf{u} \subseteq [1..d]$  with  $|\mathbf{u}| > q$ .

The order of  $f$  may be defined as the smallest value of  $q$  for which  $f = F_q$ , with the quantity  $F_q$  defined by

$$F_q := \sum_{\substack{\mathbf{u} \subseteq [1..d] \\ |\mathbf{u}| \leq q}} f_{\mathbf{u}}, \quad 0 \leq q \leq d.$$

Using simple algebraic manipulation and induction, it can be proved that  $F_d = f$  and

$$(3.2) \quad F_q = \sum_{j=0}^q (-1)^{q-j} \binom{d-1-j}{q-j} \sum_{\substack{\mathbf{u} \subseteq [1..d] \\ |\mathbf{u}|=j}} P_{[1..d] \setminus \mathbf{u}}(f) \quad \text{for } q < d.$$



We now derive (3.2). It follows from the explicit expression for  $f_{\mathbf{u}}$  in Theorem 2.1(b) that for  $0 \leq i \leq d$

$$\begin{aligned} \sum_{\substack{\mathbf{u} \subseteq [1..d] \\ |\mathbf{u}|=i}} f_{\mathbf{u}} &= \sum_{\substack{\mathbf{u} \subseteq [1..d] \\ |\mathbf{u}|=i}} \sum_{\mathbf{v} \subseteq \mathbf{u}} (-1)^{i-|\mathbf{v}|} P_{[1..d] \setminus \mathbf{v}}(f) \\ &= \sum_{\substack{\mathbf{v} \subseteq [1..d] \\ |\mathbf{v}| \leq i}} \sum_{\substack{\mathbf{u} \subseteq [1..d] \\ |\mathbf{u}|=i}} (-1)^{i-|\mathbf{v}|} P_{[1..d] \setminus \mathbf{v}}(f) \\ &= \sum_{\substack{\mathbf{v} \subseteq [1..d] \\ |\mathbf{v}| \leq i}} (-1)^{i-|\mathbf{v}|} \binom{d-|\mathbf{v}|}{i-|\mathbf{v}|} P_{[1..d] \setminus \mathbf{v}}(f), \end{aligned}$$

and this leads to

$$\begin{aligned} F_q &= \sum_{\substack{\mathbf{u} \subseteq [1..d] \\ |\mathbf{u}| \leq q}} f_{\mathbf{u}} = \sum_{i=0}^q \sum_{\substack{\mathbf{u} \subseteq [1..d] \\ |\mathbf{u}|=i}} f_{\mathbf{u}} = \sum_{i=0}^q \sum_{\substack{\mathbf{v} \subseteq [1..d] \\ |\mathbf{v}| \leq i}} (-1)^{i-|\mathbf{v}|} \binom{d-|\mathbf{v}|}{i-|\mathbf{v}|} P_{[1..d] \setminus \mathbf{v}}(f) \\ &= \sum_{\substack{\mathbf{v} \subseteq [1..d] \\ |\mathbf{v}| \leq q}} \left[ \sum_{i=|\mathbf{v}|}^q (-1)^{i-|\mathbf{v}|} \binom{d-|\mathbf{v}|}{i-|\mathbf{v}|} \right] P_{[1..d] \setminus \mathbf{v}}(f). \end{aligned}$$

Using the property  $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ , we can show by induction on  $q \in [|\mathbf{v}|, d]$

$$\sum_{i=|\mathbf{v}|}^q (-1)^{i-|\mathbf{v}|} \binom{d-|\mathbf{v}|}{i-|\mathbf{v}|} = (-1)^{q-|\mathbf{v}|} \binom{d-1-|\mathbf{v}|}{q-|\mathbf{v}|},$$

which leads to the desired formula.

#### 4. ORTHOGONALITY

In this section, we identify a *reproducing kernel Hilbert space* setting and conditions for which the decomposition is precisely an orthogonal decomposition. It turns out that the setting is very general and covers many interesting function spaces, including certain well-known Sobolev spaces. For the theory of reproducing kernels, we refer the readers to [1], and in particular, to the section on “sum of reproducing kernels”. We mention here only that a reproducing kernel  $K : D \times D \rightarrow \mathbb{R}$  corresponding to the reproducing kernel Hilbert space  $\mathcal{H}(K)$  satisfies  $K(\mathbf{x}, \cdot) \in \mathcal{H}(K)$  for all  $\mathbf{x} \in D$  and  $K(\mathbf{x}, \mathbf{y}) = K(\mathbf{y}, \mathbf{x})$  for all  $\mathbf{x}, \mathbf{y} \in D$ , and has the reproducing property

$$\langle f, K(\mathbf{x}, \cdot) \rangle_{\mathcal{H}(K)} = f(\mathbf{x}) \quad \text{for all } f \in \mathcal{H}(K) \text{ and } \mathbf{x} \in D,$$

where  $\langle \cdot, \cdot \rangle_{\mathcal{H}(K)}$  denotes the inner product in  $\mathcal{H}(K)$ . Additionally,  $\mathcal{H}(K)$  is the completion of  $\text{span}\{K(\cdot, \mathbf{y}) : \mathbf{y} \in D\}$ . Therefore, the space of all finite linear combinations of functions  $K(\cdot, \mathbf{y})$  is dense in  $\mathcal{H}(K)$ .

Following [24], we consider a general reproducing kernel Hilbert space  $\mathcal{H}(K_d)$  of real functions defined on  $D \subseteq \mathbb{R}^d$  with a reproducing kernel of the form

$$(4.1) \quad K_d(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{u} \subseteq [1..d]} K_{d,\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}}), \quad K_{d,\emptyset} = 1.$$

For each  $\mathbf{u} \subseteq [1..d]$ ,  $K_{d,\mathbf{u}}$  is the reproducing kernel for a Hilbert space  $\mathcal{H}(K_{d,\mathbf{u}})$  of real functions defined on  $D_{\mathbf{u}} := \{\mathbf{x}_{\mathbf{u}} : \mathbf{x} \in D\}$ . Functions in  $\mathcal{H}(K_{d,\mathbf{u}})$  depend only on variables listed in  $\mathbf{u}$ . In particular, the space  $\mathcal{H}(K_{d,\emptyset}) = \text{span}\{1\}$  contains just constant functions. We allow for more generality than [24] by *not* assuming that  $K_{d,\mathbf{u}}$  is of a product form.

Functions from  $\mathcal{H}(K_d)$  are sums of functions from  $\mathcal{H}(K_{d,\mathbf{u}})$ ,

$$(4.2) \quad f = \sum_{\mathbf{u} \subseteq [1..d]} f_{\mathbf{u}} \quad \text{for some} \quad f_{\mathbf{u}} \in \mathcal{H}(K_{d,\mathbf{u}}).$$

The representation (4.2) is generally *not* unique, because some nonzero functions may belong to  $\mathcal{H}(K_{d,\mathbf{u}})$  for different subsets  $\mathbf{u}$ . If  $\|\cdot\|_{\mathcal{H}(K_d)}$  denotes the norm in  $\mathcal{H}(K_d)$  and  $\|\cdot\|_{\mathcal{H}(K_{d,\mathbf{u}})}$  denotes the norm in  $\mathcal{H}(K_{d,\mathbf{u}})$  for each  $\mathbf{u} \subseteq [1..d]$ , then we have

$$(4.3) \quad \|f\|_{\mathcal{H}(K_d)}^2 = \inf \left\{ \sum_{\mathbf{u} \subseteq [1..d]} \|f_{\mathbf{u}}\|_{\mathcal{H}(K_{d,\mathbf{u}})}^2 : f_{\mathbf{u}} \in \mathcal{H}(K_{d,\mathbf{u}}) \text{ such that } f = \sum_{\mathbf{u} \subseteq [1..d]} f_{\mathbf{u}} \right\},$$

see [1]. The representation (4.2) is unique iff

$$(4.4) \quad \mathcal{H}(K_{d,\mathbf{u}}) \cap \mathcal{H}(K_{d,\mathbf{v}}) = \{0\} \quad \text{for all subsets } \mathbf{u} \neq \mathbf{v}.$$

When (4.4) holds, the space  $\mathcal{H}(K_d)$  is the direct and orthogonal sum of the spaces  $\mathcal{H}(K_{d,\mathbf{u}})$ , and we have that is, for  $f = \sum_{\mathbf{u} \subseteq [1..d]} f_{\mathbf{u}}$  and  $g = \sum_{\mathbf{u} \subseteq [1..d]} g_{\mathbf{u}}$  with  $f_{\mathbf{u}}, g_{\mathbf{u}} \in \mathcal{H}(K_{d,\mathbf{u}})$  we then have

$$\langle f, g \rangle_{\mathcal{H}(K_d)} = \sum_{\mathbf{u} \subseteq [1..d]} \langle f_{\mathbf{u}}, g_{\mathbf{u}} \rangle_{\mathcal{H}(K_{d,\mathbf{u}})},$$

with the norms satisfying

$$\|f\|_{\mathcal{H}(K_d)}^2 = \sum_{\mathbf{u} \subseteq [1..d]} \|f_{\mathbf{u}}\|_{\mathcal{H}(K_{d,\mathbf{u}})}^2.$$

Now we relate reproducing kernel Hilbert spaces constructed in this manner to the decompositions introduced in Section 2.

**Theorem 4.1.** *Let  $\mathcal{H}(K_d)$  be a reproducing kernel Hilbert space of real functions defined on  $D \subseteq \mathbb{R}^d$  which is constructed from reproducing kernel Hilbert spaces  $\mathcal{H}(K_{d,\mathbf{u}})$  for  $\mathbf{u} \subseteq [1..d]$  as in (4.1) and (4.3). Let  $\{P_j\}_{j=1}^d$  be a set of commuting projections on  $\mathcal{H}(K_d)$  satisfying (2.1) and such that for all nonempty  $\mathbf{u} \subseteq [1..d]$  and all  $\mathbf{y}_{\mathbf{u}} \in D_{\mathbf{u}}$  we have*

$$(4.5) \quad P_j(K_{d,\mathbf{u}}(\cdot, \mathbf{y}_{\mathbf{u}})) = 0 \quad \text{for all } j \in \mathbf{u}.$$

*Then the representation (4.2) is unique and is an orthogonal decomposition that coincides with the decomposition from Theorem 2.1.*

*Proof.* Take a representation (4.2) of  $f$ . Since  $\text{span}\{K_{d,\mathbf{u}}(\cdot, \mathbf{y}_{\mathbf{u}}) : \mathbf{y}_{\mathbf{u}} \in D_{\mathbf{u}}\}$  is dense in  $\mathcal{H}(K_{d,\mathbf{u}})$  and the projections  $P_j$  are continuous, (4.5) immediately implies that  $P_j(f_{\mathbf{u}}) = 0$  if  $j \in \mathbf{u}$ . Hence Theorem 2.1(e) implies that (4.2) coincides with (2.7), and because of uniqueness, (4.4) then gives orthogonality of the subspaces  $\mathcal{H}(K_{d,\mathbf{u}})$  in  $\mathcal{H}(K_d)$ . Take a representation (4.2) of  $f = \sum_{\mathbf{u} \subseteq [1..d]} f_{\mathbf{u}}$ , with  $f_{\mathbf{u}} \in \mathcal{H}(K_{d,\mathbf{u}})$ . It is enough to show that for all nonempty  $\mathbf{u} \subseteq [1..d]$  we have  $P_j(f_{\mathbf{u}}) = 0$  for all  $j \in \mathbf{u}$ . Indeed if this holds then Theorem 2.1 implies uniqueness of the decomposition, from which it follows that (4.2) must coincide with (2.7).

Recall that  $\text{span}\{K_{d,\mathbf{u}}(\cdot, \mathbf{y}_{\mathbf{u}}) : \mathbf{y}_{\mathbf{u}} \in D_{\mathbf{u}}\}$  is dense in  $\mathcal{H}(K_{d,\mathbf{u}})$ . Thus for arbitrary  $\varepsilon > 0$  there exist  $N(\varepsilon) > 0$ ,  $\mathbf{y}_1, \dots, \mathbf{y}_{N(\varepsilon)} \in D_{\mathbf{u}}$  and  $\alpha_1, \dots, \alpha_{N(\varepsilon)} \in \mathbb{R}$  such that

$$\|\theta_{\mathbf{u}}\|_{\mathcal{H}(K_{d,\mathbf{u}})} \leq \varepsilon, \quad \text{with} \quad \theta_{\mathbf{u}} := f_{\mathbf{u}} - \sum_{i=1}^{N(\varepsilon)} \alpha_i K_{d,\mathbf{u}}(\cdot, \mathbf{y}_i).$$

Now consider the projections  $\{P_j\}_{j=1}^d$ . Since  $P_j$  is continuous over  $\mathcal{H}(K_d)$ , it is continuous over  $\mathcal{H}(K_{d,\mathbf{u}})$ . From (4.5) and the definition of  $\theta_{\mathbf{u}}$  we have, for  $j \in \mathbf{u}$ ,

$$\|P_j(f_{\mathbf{u}})\|_{\mathcal{H}(K_{d,\mathbf{u}})} = \|P_j(\theta_{\mathbf{u}})\|_{\mathcal{H}(K_{d,\mathbf{u}})} \leq \|P_j\| \|\theta_{\mathbf{u}}\|_{\mathcal{H}(K_{d,\mathbf{u}})} \leq \|P_j\| \varepsilon.$$

Since  $\varepsilon$  is arbitrary, it follows that  $P_j(f_{\mathbf{u}}) = 0$ . Theorem 2.1 then tells us that the representation (4.2) coincides (2.7), and because of uniqueness, (4.4) then gives orthogonality of the subspaces  $\mathcal{H}(K_{d,\mathbf{u}})$  in  $\mathcal{H}(K_d)$ .  $\square$

For the anchored decomposition, see Example 2.3, the condition (4.5) requires

$$(4.6) \quad K_{d,\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \cdot) = 0 \quad \text{whenever} \quad x_j = c_j \\ \text{for all } j \in \mathbf{u} \text{ and for all nonempty } \mathbf{u} \subseteq [1..d].$$

In the case of the ANOVA decomposition, see Example 2.2, the condition (4.5) requires

$$(4.7) \quad \int_0^1 K_{d,\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \cdot) dx_j = 0 \quad \text{for all } j \in \mathbf{u} \text{ and for all nonempty } \mathbf{u} \subseteq [1..d].$$

We now give examples of reproducing kernel Hilbert spaces satisfying (4.6) and (4.7).

**Example 4.2** (Anchored Sobolev spaces). Let  $D = [0, 1]^d$ . The weighted Sobolev space anchored at  $\mathbf{c} \in D$  with smoothness parameter  $r \geq 1$  has a reproducing kernel of the form (4.1), with  $K_{d,\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}}) = \gamma_{d,\mathbf{u}} \prod_{j \in \mathbf{u}} \eta_j(x_j, y_j)$ , where

$$\eta_j(x, y) = \begin{cases} \int_{c_j}^{\min(x,y)} \frac{(x-t)^{r-1} (y-t)^{r-1}}{((r-1)!)^2} dt + \sum_{i=1}^{r-1} \frac{(x-c_j)^i (y-c_j)^i}{(i!)^2} & \text{if } x, y > c_j, \\ \int_{\max(x,y)}^{c_j} \frac{(t-x)^{r-1} (t-y)^{r-1}}{((r-1)!)^2} dt + \sum_{i=1}^{r-1} \frac{(c_j-x)^i (c_j-y)^i}{(i!)^2} & \text{if } x, y < c_j, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly the kernel is continuous, and  $\eta_j(c_j, \cdot) = 0$  for all  $j = 1, 2, \dots, d$ . Thus the condition (4.6) holds, and the anchored decomposition (2.10) is precisely the

orthogonal decomposition in the anchored Sobolev spaces for all  $r \geq 1$ . The numbers  $\gamma_{d,\mathbf{u}} > 0$  are referred to as *weights*. They model the relative importance between various groups of variables by weighting the norms. For  $r = 1$ , the norm of  $f = \sum_{\mathbf{u} \subseteq [1..d]} f_{\mathbf{u},\mathbf{c}}$  satisfies  $\|f\|_{\mathcal{H}(K_d)}^2 = \sum_{\mathbf{u} \subseteq [1..d]} \|f_{\mathbf{u},\mathbf{c}}\|_{\mathcal{H}(K_{d,\mathbf{u}})}^2$ , with

$$\|f_{\emptyset,\mathbf{c}}\|_{\mathcal{H}(K_{d,\emptyset})}^2 = f^2(\mathbf{c}),$$

$$\|f_{\mathbf{u},\mathbf{c}}\|_{\mathcal{H}(K_{d,\mathbf{u}})}^2 = \frac{1}{\gamma_{d,\mathbf{u}}} \int_{[0,1]^{|\mathbf{u}|}} \left( \frac{\partial^{|\mathbf{u}|} f_{\mathbf{u},\mathbf{c}}(\mathbf{x}_{\mathbf{u}})}{\partial \mathbf{x}_{\mathbf{u}}} \right)^2 d\mathbf{x}_{\mathbf{u}}, \quad \mathbf{u} \neq \emptyset,$$

where  $\partial^{|\mathbf{u}|}/\partial \mathbf{x}_{\mathbf{u}} = \prod_{j \in \mathbf{u}} \partial/\partial x_j$  means to differentiate once with respect to each variable  $x_j$  with  $j \in \mathbf{u}$ . Equivalently, for  $r = 1$  we can express the norm solely in terms of  $f$  as

$$\|f\|_{\mathcal{H}(K_d)}^2 = f^2(\mathbf{c}) + \sum_{\emptyset \neq \mathbf{u} \subseteq [1..d]} \frac{1}{\gamma_{d,\mathbf{u}}} \int_{[0,1]^{|\mathbf{u}|}} \left( \frac{\partial^{|\mathbf{u}|} f((\mathbf{x}; \mathbf{c})_{\mathbf{u}})}{\partial \mathbf{x}_{\mathbf{u}}} \right)^2 d\mathbf{x}_{\mathbf{u}}.$$

The norm for higher smoothness  $r > 1$  is difficult to express explicitly when  $d > 1$ . For  $d = 1$  and  $r \geq 1$ , it takes the following form

$$\|f\|_{\mathcal{H}(K_1)}^2 = f^2(c) + \frac{1}{\gamma} \left( \sum_{i=1}^{r-1} [f^{(i)}(c)]^2 + \int_0^1 [f^{(r)}(x)]^2 dx \right).$$

Special cases of anchored Sobolev spaces have been considered in many papers (specializing to  $\mathbf{c} = \mathbf{0}$  or  $\mathbf{c} = \mathbf{1}$ , and/or  $r = 1$ , and/or  $\gamma_{d,\mathbf{u}} = 1$ ), see e.g., [16] and the references therein, and [20, Chapter 1.2]. See also [23] for extensions. Weighted Sobolev spaces defined over the unit cube  $[0, 1]^d$ , with anchor  $\mathbf{c}$  (sometimes specializing to  $\mathbf{c} = \mathbf{0}$  or  $\mathbf{c} = \mathbf{1}$ ) and with smoothness parameter  $r = 1$ , have been studied in many recent papers, see e.g., [16] and the references therein. The anchored Sobolev spaces with a general smoothness parameter  $r \geq 1$  were analyzed in the book [20, Chapter 1.2], which has the anchor  $\mathbf{c} = \mathbf{0}$  and the weights  $\gamma_{d,\mathbf{u}} = 1$ . The paper [23] considered weighted Sobolev spaces with higher smoothness  $r$  for a more general domain together with a weight function  $\psi$ , but assumed that the functions have mixed derivatives of order  $< r$  vanishing at  $\mathbf{c}$ .

**Example 4.3** (Unanchored Sobolev spaces). Let  $D = [0, 1]^d$ . The weighted and unanchored Sobolev space with smoothness parameter  $r \geq 1$  has a reproducing kernel of the form (4.1), with  $K_{d,\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}}) = \gamma_{d,\mathbf{u}} \prod_{j \in \mathbf{u}} \eta(x_j, y_j)$ , where

$$\eta(x, y) = \frac{B_{2r}(|x - y|)}{(-1)^{r+1}(2r)!} + \sum_{i=1}^r \frac{B_i(x) B_i(y)}{(i!)^2}.$$

Here  $B_i$  denotes the Bernoulli polynomial of degree  $i$ , which has the property  $\int_0^1 B_i(x) dx = 0$ . It is easy to see that  $\int_0^1 \eta(x, \cdot) dx = 0$ . Thus the condition (4.7) holds, and the ANOVA decomposition (2.9) is precisely the orthogonal decomposition in this unanchored Sobolev space. It is a striking and sometimes useful consequence that the terms in the ANOVA decomposition are orthogonal not only in  $\mathcal{L}_2([0, 1]^d)$ , but also in the unanchored Sobolev spaces for every value of  $r$ . For

$r = 1$ , we have

$$\|f_{\emptyset,*}\|_{\mathcal{H}(K_{d,\emptyset})}^2 = \left( \int_{[0,1]^d} f(\mathbf{x}) \, d\mathbf{x} \right)^2,$$

$$\|f_{\mathbf{u},*}\|_{\mathcal{H}(K_{d,\mathbf{u}})}^2 = \frac{1}{\gamma_{d,\mathbf{u}}} \int_{[0,1]^{|\mathbf{u}|}} \left( \frac{\partial^{|\mathbf{u}|} f_{\mathbf{u},*}(\mathbf{x}_{\mathbf{u}})}{\partial \mathbf{x}_{\mathbf{u}}} \right)^2 d\mathbf{x}_{\mathbf{u}}, \quad \mathbf{u} \neq \emptyset,$$

or equivalently, we have

$$\|f\|_{\mathcal{H}(K_d)}^2 = \left( \int_{[0,1]^d} f(\mathbf{x}) \, d\mathbf{x} \right)^2 + \sum_{\emptyset \neq \mathbf{u} \subseteq [1..d]} \frac{1}{\gamma_{d,\mathbf{u}}} \int_{[0,1]^{|\mathbf{u}|}} \left( \int_{[0,1]^{d-|\mathbf{u}|}} \frac{\partial^{|\mathbf{u}|} f(\mathbf{x})}{\partial \mathbf{x}_{\mathbf{u}}} d\mathbf{x}_{[1..d] \setminus \mathbf{u}} \right)^2 d\mathbf{x}_{\mathbf{u}}.$$

For  $d = 1$  and general  $r \geq 1$ , we have

$$\|f\|_{\mathcal{H}(K_1)}^2 = \left( \int_0^1 f(x) \, dx \right)^2 + \frac{1}{\gamma} \left[ \sum_{i=1}^{r-1} \left( \int_0^1 f^{(i)}(x) \, dx \right)^2 + \int_0^1 [f^{(r)}(x)]^2 \, dx \right].$$

Special cases of unanchored Sobolev spaces have been considered in many papers (specializing to  $r = 1$  and/or  $\gamma_{d,\mathbf{u}} = 1$ ), see e.g., [4] and the references therein, as well as [3] and [20, Chapter 10.2]. In particular, it was already shown in [4] that the ANOVA and the orthogonal decompositions for  $r = 1$  coincide. The unanchored Sobolev spaces with smoothness parameter  $r = 1$  has been analyzed in many papers, see e.g., [4] and the references therein. It was already shown in [4] that the ANOVA decomposition coincides with the orthogonal decomposition for  $r = 1$ . The spaces with general smoothness parameter  $r$  appeared in [3] and in [20, Chapter 10.2]. We include the derivation in the Appendix for the benefit of interested readers: this space is derived under the condition that *its periodic subspace coincides with the weighted Korobov space of the same smoothness*.

**Example 4.4** (General tensor product spaces). We now show how to construct operators  $P_j$  when the space  $\mathcal{H}(K_d)$  is given as the  $d$ -fold tensor product of a general reproducing kernel Hilbert space  $\mathcal{H}(K_1)$ . We only assume that  $K_1 : D \times D \rightarrow \mathbb{R}$ , with  $D \subseteq \mathbb{R}$ , is of the form  $K_1(x, y) = 1 + \eta(x, y)$ , where  $\eta : D \times D \rightarrow \mathbb{R}$  is a reproducing kernel such that  $\mathcal{H}(\eta)$  is separable and  $1 \notin \mathcal{H}(\eta)$ . Thus we have  $K_d(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^d (1 + \eta(x_i, y_i)) = \sum_{\mathbf{u} \subseteq [1..d]} \prod_{i \in \mathbf{u}} \eta(x_i, y_i)$ . Let  $\xi_0 \equiv 1$ , and let  $\{\xi_k\}_{k=1}^{\infty}$  be a complete orthonormal system in  $\mathcal{H}(\eta)$ . Then  $\{\xi_k\}_{k=0}^{\infty}$  is a complete orthonormal system in  $\mathcal{H}(K_1)$ . For  $d \geq 1$  and  $\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{N}^d$  (where  $\mathbb{N} = \{0, 1, \dots\}$ ), let  $\xi_{d,\mathbf{k}}(\mathbf{x}) = \prod_{i=1}^d \xi_{k_i}(x_i)$  for  $\mathbf{x} \in D^d$ . Then  $\{\xi_{d,\mathbf{k}}\}_{\mathbf{k} \in \mathbb{N}^d}$  is a complete orthonormal system in  $\mathcal{H}(K_d)$ . Note that if  $k_i = 0$  then  $\xi_{d,\mathbf{k}}$  does not depend on the  $i$ th variable. Therefore, we set  $P_j$  to be orthogonal projections onto  $\text{span}\{\xi_{d,\mathbf{k}} : \mathbf{k} \in \mathbb{N}^d \text{ with } k_j = 0\}$ , since they clearly satisfy (2.1) and (4.5). For  $f \in \mathcal{H}(K_d)$ , we can write

$$f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{N}^d} \langle f, \xi_{d,\mathbf{k}} \rangle_{\mathcal{H}(K_d)} \xi_{d,\mathbf{k}}(\mathbf{x}).$$

We now define  $P_j$  by

$$(*) \quad P_j(\xi_{d,\mathbf{k}}) = \begin{cases} \xi_{d,\mathbf{k}} & \text{if } k_j = 0, \\ 0 & \text{if } k_j \geq 1, \end{cases}$$

for all  $\mathbf{k} \in \mathbb{N}^d$ . Then we have

$$P_j(f)(\mathbf{x}) = \sum_{\substack{\mathbf{k} \in \mathbb{N}^d \\ k_j = 0}} \langle f, \xi_{d,\mathbf{k}} \rangle_{\mathcal{H}(K_d)} \xi_{d,\mathbf{k}}(\mathbf{x}).$$

Clearly each operator  $P_j$  is the orthogonal projection of the space  $\mathcal{H}(K_d)$  into a subspace  $\mathcal{H}(\kappa_j)$ , where

$$\kappa_j(\mathbf{x}, \mathbf{y}) := \prod_{\substack{i=1 \\ i \neq j}}^d \eta(x_i, y_i) \quad \text{for } \mathbf{x}, \mathbf{y} \in D^d.$$

It is easy to check that  $P_j$  defined above indeed satisfies (2.1). The condition (4.5) is also satisfied since

$$K_{d,\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}}) = \prod_{i \in \mathbf{u}} \eta(x_i, y_i) = \sum_{\substack{\mathbf{k} \in \mathbb{N}^d \\ k_i = 0 \forall i \notin \mathbf{u} \\ k_i \geq 1 \forall i \in \mathbf{u}}} \xi_{d,\mathbf{k}}(\mathbf{x}) \xi_{d,\mathbf{k}}(\mathbf{y}) \quad \text{for all } \mathbf{x}, \mathbf{y} \in D^d.$$

Suppose now that the kernel  $\eta$  vanishes at a point  $a \in D$ , i.e.,  $\eta(a, a) = 0$ . It is easy to see that then the projections of the form (2.2) with  $c_j = a$  also satisfy (\*). Hence, they coincide with the projections  $P_j$  defined by (\*), and the orthogonal decomposition discussed here is just the anchored decomposition with the anchor  $\mathbf{c} = (a, \dots, a)$ .

For simplicity, we have restricted ourselves to unweighted tensor product spaces in this example. The same argument can be generalized to any reproducing kernel of the form (4.1), where  $K_{d,\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}}) = \gamma_{d,\mathbf{u}} \prod_{i \in \mathbf{u}} \eta_i(x_i, y_i)$ , with  $\eta_i : D \times D \rightarrow \mathbb{R}$  being any reproducing kernel such that  $\mathcal{H}(\eta_i)$  is separable and  $1 \notin \mathcal{H}(\eta_i)$  for all  $i = 1, 2, \dots, d$ .

Although the above examples have reproducing kernel of the product form, the product form is not a necessary property for the existence of a suitable set of projections  $\{P_j\}_{j=1}^d$ . To see this, let  $\tilde{K}_{d,\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}})$  be an arbitrary (non-product) reproducing kernel on  $D_{\mathbf{u}}$ . Then  $K_{d,\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}}) := (\prod_{j \in \mathbf{u}} x_j y_j) \tilde{K}_{d,\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}})$  is the reproducing kernel of some Hilbert space  $\mathcal{H}(K_{d,\mathbf{u}})$  which is anchored at  $\mathbf{0}$ , since every element of the space will vanish if any component of  $\mathbf{x}_{\mathbf{u}}$  has the value 0. Thus Theorem 4.1 holds for  $\mathcal{H}(K_{d,\mathbf{u}})$  with  $P_j$  defined by (2.2) and  $c_j = 0$  for all  $j = 1, 2, \dots, d$ . Although the present theory has wider scope than just product kernels  $K_{d,\mathbf{u}}$ , we do *not* claim that a suitable set of projections  $\{P_j\}_{j=1}^d$  exists for every choice of reproducing kernel Hilbert spaces  $\mathcal{H}(K_{d,\mathbf{u}})$ .

5. COMPUTATIONAL COST

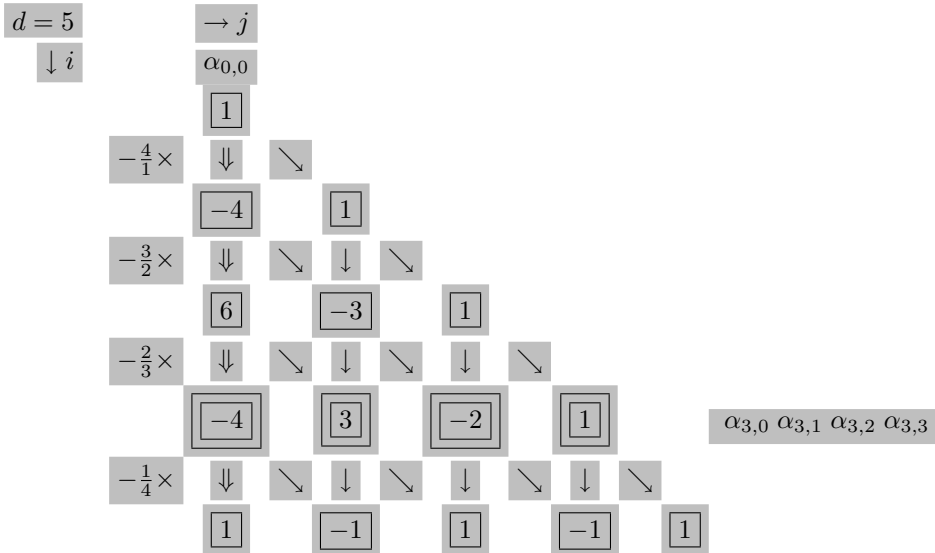
The anchored decomposition is arguably more useful in practice than the ANOVA decomposition because it requires only function evaluations of  $f$  with respect to a chosen anchor  $\mathbf{c}$ . Let  $\mathbf{x} \in D$  be given and fixed. It is clear from the explicit formula in Example 2.3(b) that for each  $\mathbf{u}$ ,  $f_{\mathbf{u},\mathbf{c}}(\mathbf{x}_{\mathbf{u}})$  can be computed using  $2^{|\mathbf{u}|} - 1$  additions and subtractions and  $2^{|\mathbf{u}|}$  samples of  $f$ .

Suppose we want to compute  $f_{\mathbf{u},\mathbf{c}}(\mathbf{x}_{\mathbf{u}})$  for all  $\mathbf{u}$  with  $|\mathbf{u}| \leq q$ . We have  $A_{d,q} := \sum_{j=0}^q \binom{d}{j}$  such subsets  $\mathbf{u}$ . It is easy to check, see also [25], that  $A_{d,q} \leq 2d^q$  and  $A_{d,q} = (d^q/q!)(1 + o(1))$  as  $d \rightarrow \infty$ . We need to evaluate  $f$  at the points  $(\mathbf{x}; \mathbf{c})_{\mathbf{u}}$  for all  $\mathbf{u}$  with  $|\mathbf{u}| \leq q$ , that is, we use at most  $q$  components of  $\mathbf{x}$  and the rest are replaced by the corresponding components of  $\mathbf{c}$ . The total number of such points is again  $A_{d,q}$ . Hence, all  $f_{\mathbf{u},\mathbf{c}}(\mathbf{x}_{\mathbf{u}})$  with  $|\mathbf{u}| \leq q$  can be computed using at most  $A_{d,q}$  samples of  $f$  and at most  $B_{d,q} := \sum_{j=0}^q 2^j \binom{d}{j}$  arithmetic operations, where  $B_{d,q} \leq 2^{q+1}d^q$  and  $B_{d,q} = (2^q d^q/q!)(1 + o(1))$  as  $d \rightarrow \infty$ .

If instead we just want to compute  $F_q(\mathbf{x})$ , that is, the combined contribution from all terms  $f_{\mathbf{u},\mathbf{c}}(\mathbf{x}_{\mathbf{u}})$  with  $|\mathbf{u}| \leq q$ , then we have from (3.2) that for  $q < d$ ,

$$F_q(\mathbf{x}) = \sum_{j=0}^q \alpha_{q,j} \sum_{\substack{\mathbf{u} \subseteq [1..d] \\ |\mathbf{u}|=j}} f((\mathbf{x}; \mathbf{c})_{\mathbf{u}}), \quad \alpha_{q,j} = (-1)^{q-j} \binom{d-1-j}{q-j}.$$

This would again require at most  $A_{d,q}$  samples of  $f$  at the points  $(\mathbf{x}; \mathbf{c})_{\mathbf{u}}$  for all  $\mathbf{u}$  with  $|\mathbf{u}| \leq q$ . However, the total number of arithmetic operations will be significantly reduced. Starting with  $\alpha_{0,0} = 1$ , the numbers  $\alpha_{i,j}$  for  $i = 1, \dots, q$  and  $j = 1, \dots, i$  can be computed using the general recurrence relations  $\alpha_{i,0} = -\alpha_{i-1,0}(d-i)/i$ ,  $\alpha_{i,j} = \alpha_{i-1,j} + \alpha_{i-1,j-1}$ , and  $\alpha_{i,i} = 1$ . **Since the numbers  $\alpha_{i,j}$  are essentially signed binomial coefficients, they can be computed using a sideways Pascal's triangle, see the illustration below for case  $d = 5$  (conceptually, the triangle can be extended indefinitely toward the left).**



Once the numbers  $\alpha_{q,j}$  are computed, we need to perform  $A_{d,q} - 1$  additions and  $q + 1$  multiplications. For large  $q$ , the total number of arithmetic operations is then roughly a factor of  $2^q$  smaller than  $B_{d,q}$ .

## 6. DISCUSSION

In recent times, the success of quasi-Monte Carlo (QMC) methods for many practical problems of high dimension has been attributed to the fact that these problems have *low effective dimension*. Two notions of effective dimension have been defined through the variance allocation of the ANOVA terms: *truncation dimension* and *superposition dimension*, see e.g., [2] for their precise definitions and [21] for known techniques to approximate them.

Although the explanation in terms of effective dimension is undoubtedly broadly true, the existing framework is somewhat unsatisfactory since, unlike the Monte Carlo (MC) error, the QMC error is *not* expressible in terms of the variance. Rather, QMC errors are typically bounded in terms of some discrepancy times the variation of the function, or more recently, in terms of the worst case error in a reproducing kernel Hilbert space times the norm of the function. Moreover, the ANOVA decomposition may not be orthogonal in this function space. Once the underlying function space is decided, a more suitable approach is to identify a set of projections that lead to orthogonal decomposition in the function space as discussed in our paper, and then capture the notion of low dimensionality through the allocation of the square norm.

Many recent papers have considered *weighted* reproducing kernel Hilbert spaces, with the weights intended to characterize the relative importance of various groups of variables. The choice of weights can have practical as well as theoretical importance, since modern constructions of lattice rules (for a survey see [9]) make explicit use of the weights. The question of how best to choose the weights is a topic of active current research. In considering quadrature errors for QMC in the unanchored space of Example 4.3 with  $r = 1$ , [22] showed that the quadrature error bound of the form worst case error times the norm of  $f$  achieves its minimum value for a given function  $f$  if the weights are given by  $\gamma_{d,\mathbf{u}} = 6^{|\mathbf{u}|/2} \beta_{\mathbf{u}}(f)$ , where

$$\beta_{\mathbf{u}}(f) := \left[ \int_{[0,1]^{|\mathbf{u}|}} \left( \int_{[0,1]^{d-|\mathbf{u}|}} \frac{\partial^{|\mathbf{u}|} f}{\partial \mathbf{x}_{\mathbf{u}}} d\mathbf{x}_{[1..d]\setminus\mathbf{u}} \right)^2 d\mathbf{x}_{\mathbf{u}} \right]^{1/2} = \|f_{\mathbf{u},*}\|_{\mathcal{H}(K_{d,\mathbf{u}})},$$

where in the last step we used two easily verified properties of the ANOVA decomposition (2.9), namely that  $\beta_{\mathbf{u}}(f_{\mathbf{v},*}) = 0$  if  $\mathbf{v} \neq \mathbf{u}$  and  $\beta_{\mathbf{u}}(f_{\mathbf{v},*}) = \|f_{\mathbf{u},*}\|_{\mathcal{H}(K_{d,\mathbf{u}})}$  if  $\mathbf{v} = \mathbf{u}$ . This opens a path to a good choice of the weights if we can estimate  $\|f_{\mathbf{u},*}\|_{\mathcal{H}(K_{d,\mathbf{u}})}$ , a task that should be made easier by a better understanding of the structure of the ANOVA decomposition. Similar arguments for choosing weights can be used for the anchored Sobolev space with the anchored decomposition.

The explicit decomposition formula can be useful for numerical integration and approximation of functions in high dimensions, since there might be a merit in approximating the solution corresponding to lower order part (see e.g., [10]). We may also use the explicit decomposition formula in a theoretical analysis where no actual computation is required, noting that it may be easier to observe or identify certain properties of one particular decomposition term when it is expressed explicitly rather than recursively (see e.g., [6]).



## APPENDIX

Here we derive the reproducing kernel of the *unanchored Sobolev space* under the condition that *its periodic subspace coincides with the weighted Korobov space of the same smoothness*. We begin by reviewing the weighted Korobov space.

For  $d = 1$ , the weighted Korobov space with arbitrary smoothness parameter  $\alpha > 1/2$  consists of 1-periodic functions  $f$  that have finite norms given by

$$\|f\|_{\text{per}}^2 = \sum_{h=-\infty}^{\infty} |\hat{f}(h)|^2 \rho_{\alpha}(\gamma, h),$$

with

$$\hat{f}(h) = \int_0^1 f(x) e^{-2\pi i h x} dx \quad \text{and} \quad \rho_{\alpha}(\gamma, h) = \begin{cases} 1 & \text{if } h = 0, \\ \gamma^{-1} (2\pi)^{2\alpha} |h|^{2\alpha} & \text{if } h \neq 0. \end{cases}$$

Weighted Korobov spaces have been studied in many papers. However, here we have adopted an unconventional notation: the  $\alpha$  and  $\gamma$  in the traditional notation are replaced by  $2\alpha$  and  $\gamma/(2\pi)^{2\alpha}$ , respectively. We did this to simplify the expression for the norm in terms of derivatives. Indeed, if  $\alpha = r \geq 1$  is an integer then the norm can be written as

$$(6.1) \quad \|f\|_{\text{per}}^2 = \left( \int_0^1 f(t) dt \right)^2 + \frac{1}{\gamma} \int_0^1 [f^{(r)}(t)]^2 dt,$$

and thus  $r$  is precisely the number of derivatives available. Moreover, the periodicity combined with the smoothness means that

$$f^{(i)}(0) = f^{(i)}(1) \quad \text{for all } i = 0, 1, \dots, r-1.$$

The weighted Korobov space has reproducing kernel

$$K_{\text{per}}(x, y) = \sum_{h=-\infty}^{\infty} \frac{e^{2\pi i h(x-y)}}{\rho_{\alpha}(\gamma, h)} = 1 + \frac{2}{(2\pi)^{2\alpha}} \sum_{h=1}^{\infty} \frac{\cos(2\pi h(x-y))}{h^{2\alpha}}.$$

When  $\alpha = r$  is an integer, the kernel can be written as

$$(6.2) \quad K_{\text{per}}(x, y) = 1 + \gamma \frac{B_{2r}(|x-y|)}{(-1)^{r+1} (2r)!},$$

where  $B_{2r}$  denotes the Bernoulli polynomial of degree  $2r$ .

We are now ready to derive the reproducing kernel of the unanchored Sobolev space, which consists of real functions  $f$  defined on  $[0, 1]$  such that  $f^{(r-1)}$  is absolutely continuous and  $f^{(r)}$  is square integrable, with a positive integer  $r$ .

Let  $f$  denote an arbitrary function in our Sobolev space. For  $j = 0, 1, \dots, r-1$ , we define the functionals

$$T_j(f) := f^{(j)}(1) - f^{(j)}(0) = \int_0^1 f^{(j+1)}(t) dt.$$

For  $i = 0, 1, \dots, r-1$ , let  $b_i$  be a polynomial of degree  $\leq i+1$  such that

$$T_j(b_i) = \delta_{i,j} \quad \text{for } j = 0, 1, \dots, r-1, \quad \text{and} \quad \int_0^1 b_i(t) dt = 0.$$

Moreover, we define

$$p(x) := f(x) - \sum_{j=0}^{r-1} T_j(f) b_j(x),$$

that is,  $p$  is obtained from  $f$  by a rank- $r$  modification. It then follows that

$$T_k(p) = T_k(f) - \sum_{j=0}^{r-1} T_j(f) T_k(b_j) = 0 \quad \text{for all } k = 0, 1, \dots, r-1.$$

Thus  $p$  is a periodic function satisfying

$$p^{(j)}(1) = p^{(j)}(0) \quad \text{for all } j = 0, 1, \dots, r-1.$$

Hence, every function  $f$  in our one-dimensional Sobolev space can be written as a sum of a periodic function and a correction term

$$(6.3) \quad f(x) = p(x) + \sum_{j=0}^{r-1} T_j(f) b_j(x).$$

We define the norm of the periodic subspace of our Sobolev space to be the same as the norm in the corresponding Korobov space, see (6.1),

$$\|p\|^2 := \|p\|_{\text{per}}^2 = \left( \int_0^1 p(t) dt \right)^2 + \frac{1}{\gamma} \int_0^1 [p^{(r)}(t)]^2 dt.$$

We then define the norm of  $f$  by

$$(6.4) \quad \begin{aligned} \|f\|^2 &:= \left( \int_0^1 p(t) dt \right)^2 + \frac{1}{\gamma} \left( \int_0^1 [p^{(r)}(t)]^2 dt + \sum_{j=0}^{r-1} [T_j(f)]^2 \right) \\ &= \left( \int_0^1 f(t) dt \right)^2 + \frac{1}{\gamma} \left( \int_0^1 [f^{(r)}(t) - T_{r-1}(f) b_{r-1}^{(r)}(t)]^2 dt + \sum_{j=0}^{r-1} [T_j(f)]^2 \right), \end{aligned}$$

where the second equality follows from properties of the polynomials  $b_i$ . Observe that we do not need to know the decomposition (6.3) to be able to evaluate the norm of  $f$ .

Let  $B_n(x)$  denote the Bernoulli polynomial of degree  $n$ . We now verify that

$$b_i(x) := \frac{B_{i+1}(x)}{(i+1)!}$$

satisfies our requirements above. Using  $\int_0^1 B_n(x) dx = 0$  and  $B'_n(x) = n B_{n-1}(x)$  for all  $n$ , we have  $\int_0^1 b_i(x) dx = 0$  and

$$T_j(b_i) = \frac{1}{(i+1)!} \int_0^1 B_{i+1}^{(j+1)}(x) dx = \begin{cases} 0 & \text{if } j > i, \\ \frac{(i+1)!}{(i+1)!} = 1 & \text{if } j = i, \\ \frac{1}{(i+1)!} \frac{(i+1)!}{(i-j)!} \int_0^1 B_{i-j}(x) dx = 0 & \text{if } j < i, \end{cases}$$

confirming the required properties for  $b_i$ . Moreover, we have  $b_{r-1}^{(r)}(t) = B_0(t) = 1$  in (6.4). This leads to a simplified expression for the norm

$$\|f\|^2 = \left( \int_0^1 f(t) dt \right)^2 + \frac{1}{\gamma} \left[ \sum_{j=1}^{r-1} \left( \int_0^1 f^{(j)}(t) dt \right)^2 + \int_0^1 [f^{(r)}(t)]^2 dt \right],$$

from which we see that the inner product should be defined as

$$\begin{aligned} \langle f, g \rangle &= \left( \int_0^1 f(t) dt \right) \left( \int_0^1 g(t) dt \right) \\ &+ \frac{1}{\gamma} \left[ \sum_{j=1}^{r-1} \left( \int_0^1 f^{(j)}(t) dt \right) \left( \int_0^1 g^{(j)}(t) dt \right) + \int_0^1 f^{(r)}(t) g^{(r)}(t) dt \right]. \end{aligned}$$

The reproducing kernel for the periodic subspace of our Sobolev space is precisely  $K_{\text{per}}(x, y)$ , see (6.2). The reproducing kernel for our Sobolev space is given by

$$K(x, y) = K_{\text{per}}(x, y) + \gamma \sum_{k=0}^{r-1} b_k(x) b_k(y).$$

Indeed, it is not hard to verify that

$$\langle b_j, b_k \rangle = \begin{cases} \gamma^{-1} & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Moreover, we have  $\langle p, b_j \rangle = 0$  for all periodic functions  $p$ , including the function  $K_{\text{per}}(x, \cdot)$ . Thus for any  $f$  with decomposition (6.3) we have

$$\begin{aligned} \langle f, K(x, \cdot) \rangle &= \left\langle p(\cdot) + \sum_{j=0}^{r-1} T_j(f) b_j(\cdot), K_{\text{per}}(x, \cdot) + \gamma \sum_{k=0}^{r-1} b_k(x) b_k(\cdot) \right\rangle \\ &= p(x) + \sum_{j=0}^{r-1} T_j(f) b_j(x) \\ &= f(x), \end{aligned}$$

which verifies the reproducing property of  $K(x, y)$ . Combining the formulas for  $K_{\text{per}}(x, y)$  and  $b_k(x)$ , we obtain the explicit formula

$$K(x, y) = 1 + \gamma \left( \frac{B_{2r}(|x-y|)}{(-1)^{r+1}(2r)!} + \sum_{i=1}^r \frac{B_i(x) B_i(y)}{(i!)^2} \right).$$

This kernel has already been used in [3].

## ACKNOWLEDGMENTS

The authors thank Rob Womersley, Michael Griebel, and the anonymous associate editor for valuable suggestions. The support of the Australian Research Council under its Centres of Excellence Program is gratefully acknowledged. The first author is supported by an Australian Research Council Queen Elizabeth II Research Fellowship. The last two authors were partially supported by the National Science Foundation under Grants DMS-0609703 and DMS-0608727, respectively.

## REFERENCES

1. N. Aronszajn, *Theory of reproducing kernels*, Trans. AMS, **68**, 337–404 (1950).
2. R. E. Caflisch, W. Morokoff, and A. Owen, *Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension*, J. Comput. Finance **1**, 27–46 (1997).
3. J. Dick and F. Pillichshammer, *Strong tractability of multivariate integration of arbitrary high order using digitally shifted polynomial lattice rules*, J. Complexity **23**, 436–453 (2007).
4. J. Dick, I. H. Sloan, X. Wang, and H. Woźniakowski, *Liberating the weights*, J. Complexity **20**, 593–623 (2004).
5. M. Griebel, *Sparse grids and related approximation schemes for higher dimensional problems*, Foundations of Computational Mathematics, Santander 2005 (L. M. Pardo, A. Pinkus, E. Süli, M. J. Todd, eds.), Cambridge University Press, London, 2006, pp. 106–161.
6. M. Griebel, F. Y. Kuo and I. H. Sloan *The smoothing effect of the ANOVA decomposition*, in preparation.
7. F. J. Hickernell, I. H. Sloan, and G. W. Wasilkowski, *On tractability of weighted integration over bounded and unbounded regions in  $\mathbb{R}^s$* , Math. Comp., **73**, 1885–1901 (2004).
8. F. J. Hickernell, I. H. Sloan, and G. W. Wasilkowski, *The strong tractability of multivariate integration using lattice rules*, Monte Carlo and Quasi-Monte Carlo Methods 2002 (H. Niederreiter, ed.), Springer-Verlag, Berlin, 2004, pp. 259–273.
9. F. Y. Kuo and I. H. Sloan, *Lifting the curse of dimensionality*, Notices Amer. Math. Soc. **52**, 1320–1329 (2005).
10. F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and H. Woźniakowski, *Liberating the dimension*, submitted.
11. G. Li, J. Schoendorf, T.-S. Ho, and H. Rabitz, *Multicut-HDMR with an application to an ionospheric model*, J. Comput. Chem. **25**, 1149–1156 (2004).
12. G. Li, J. Hu, S.-W. Wang, P. G. Georgopoulos, J. Schoendorf, and H. Rabitz, *Random sampling-high dimensional model representation (RS-HDMR) and orthogonality of its different order component functions*, J. Phys. Chem. A **110**, 2474–2485 (2006).
13. G. Li and H. Rabitz, *Ratio control variate method for efficiently determining high-dimensional model representations*, J. Comput. Chem. **27**, 1112–1118 (2006).
14. R. Liu and A. Owen, *Estimating mean dimensionality of analysis of variance decompositions*, J. Amer. Statist. Assoc. **101**, 712–721 (2006).
15. H. Rabitz, O. F. Alis, J. Shorter and K. Shim, *Efficient input-output model representation*, Comput. Phys. Commun. **117**, 11–20 (1999).
16. I. H. Sloan, X. Wang, and H. Woźniakowski, *Finite-order weights imply tractability of multivariate integration*, J. Complexity **20**, 46–74 (2004).
17. I. M. Sobol', *Multidimensional Quadrature Formulas and Haar Functions*, Nauka, Moscow, 1969 (in Russian).
18. I. M. Sobol', *Sensitivity estimates for nonlinear mathematical models* Matematicheskoe Modelirovanie, 1990, V. 2, N 1, 112–118 (in Russian), English translation in: Mathematical Modeling and Computational Experiment, 407–414 (1993).
19. I. M. Sobol', *Theorems and examples on high dimensional model representation*, Reliability Engrg. and System Safety **79**, 187–193 (2003).
20. G. Wahba, *Spline Models for Observational Data*, SIAM-NSF Regional Conference Series in Appl. Math., Vol. 59, SIAM, Philadelphia, 1990.
21. X. Wang and K. -T. Fang, *Effective dimension and quasi-Monte Carlo integration*, J. Complexity **19**, 101–124 (2003).

22. X. Wang and I. H. Sloan, *Brownian bridge and principal component analysis*, IMA J. Numer. Anal. **27**, 631–654 (2007).
23. G. W. Wasilkowski and H. Woźniakowski, *Tractability of approximation and integration for weighted tensor product problems over unbounded domains*, Monte Carlo and Quasi-Monte Carlo Methods 2000, (K.-T. Fang, F. J. Hickernell, H. Niederreiter, eds.), Springer, Berlin, 2002, pp. 497–522.
24. G. W. Wasilkowski and H. Woźniakowski, *Finite-order weights imply tractability of linear multivariate problems*, J. Approx. Theory **130**, 57–77 (2004).
25. G. W. Wasilkowski and H. Woźniakowski, *Polynomial-Time Algorithms for Multivariate Linear Problems with Finite-Order Weights: Worst Case Setting*, Found. Comput. Math., **5**, 451–491 (2005).

SCHOOL OF MATHEMATICS AND STATISTICS, UNIVERSITY OF NEW SOUTH WALES, SYDNEY NSW  
2052, AUSTRALIA  
*E-mail address:* `f.kuo@unsw.edu.au`

SCHOOL OF MATHEMATICS AND STATISTICS, UNIVERSITY OF NEW SOUTH WALES, SYDNEY NSW  
2052, AUSTRALIA  
*E-mail address:* `i.sloan@unsw.edu.au`

DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF KENTUCKY, LEXINGTON, KY 40506,  
USA  
*E-mail address:* `greg@cs.uky.edu`

DEPARTMENT OF COMPUTER SCIENCE, COLUMBIA UNIVERSITY, NEW YORK, NY 10027, USA,  
AND INSTITUTE OF APPLIED MATHEMATICS, UNIVERSITY OF WARSAW, UL. BANACHA 2, 02-097  
WARSZAWA, POLAND  
*E-mail address:* `henryk@cs.columbia.edu`