

## TERMINATION CONDITIONS FOR APPROXIMATING LINEAR PROBLEMS WITH NOISY INFORMATION

B. Z. KACEWICZ AND L. PLASKOTA

**ABSTRACT.** We study the diameter termination criterion for approximating linear continuous problems. It is assumed that only nonexact information about the problem is available. We evaluate the quality of the diameter termination criterion by comparing it with the theoretically best stopping condition. The comparison is made with respect to the cost of computing an  $\varepsilon$ -approximation. Although the diameter termination criterion is independent of a particular problem, it turns out to be essentially equivalent to the theoretical condition. Optimal information and the best way of constructing an  $\varepsilon$ -approximation are exhibited.

### 1. INTRODUCTION

Many problems in numerical analysis can be formulated in terms of approximating a linear continuous operator. Denoting such an operator by  $S$ , the problem is to approximate  $S(f)$  for elements  $f$  belonging to a certain ball  $K$ . To find an approximation, we gather *information* about the problem by successively computing some numbers  $z_1, z_2, \dots$  dependent on  $f$ . We assume that each  $z_n$  is a noisy evaluation of a linear continuous functional at  $f$ . For instance, for the integration problem, where  $f$  is a function and  $S(f)$  its integral, the numbers  $z_n$  may be given as perturbed values of  $f$  at some points. The noise in information may result, e.g., from measurement, representation or computational errors. There is a growing literature on noisy information, as such problems attract attention of statisticians, engineers, and numerical analysts, see the references.

Given  $\varepsilon > 0$ , we want to produce an approximation to  $S(f)$  for all  $f \in K$ , with error at most  $\varepsilon$ . Information values are gathered until we collect  $n = n(f)$  numbers  $z_1, z_2, \dots, z_n$  from which an  $\varepsilon$ -approximation can be computed. We therefore need a criterion which allows us to decide when to terminate the computations.

The stopping condition based directly on the requirement that the error does not exceed  $\varepsilon$  (although theoretically the best) is impossible to implement, since the error of approximation is usually unknown. Moreover, such a criterion depends on a particular unknown element  $f$ .

We study in this paper a termination criterion based on the use of the diameter of information, a quantity studied by many authors; see for instance

---

Received by the editor May 16, 1991.

1991 *Mathematics Subject Classification*. Primary 65D99, 65J05.

©1992 American Mathematical Society  
0025-5718/92 \$1.00 + \$.25 per page

Babenko [1], Micchelli and Rivlin [5], Marchuk and Osipenko [4], Kacewicz and Plaskota [3], and Traub et al. [6]. This criterion (referred to as the *diameter termination criterion*) does not depend on the unknown error of approximation and  $f$ , but only on information used and the class  $K$ . In spite of this advantage, the diameter termination criterion may seem too strong when compared to the theoretical stopping condition, since it is based on a stronger inequality (see the relation (2.7)). One may ask if the number of steps that we have to perform before terminating (i.e., the number of evaluations that we have to compute) is not unnecessarily large, at least for some elements  $f$ .

The main purpose of this paper is to determine the quality of the diameter termination criterion. As the basis for the evaluation we use the *cost (information cost)* of computing an  $\varepsilon$ -approximation. For the diameter termination criterion, this cost has been studied for some problems; see for instance Kacewicz and Plaskota [2] and the Example in §4. We compare the cost yielded by the diameter termination criterion with that for the theoretical stopping condition. Moreover, we aim to choose the best information and the best way of combining it, in order to obtain an  $\varepsilon$ -approximation with the cost growing as slowly as possible as  $\varepsilon \rightarrow 0^+$ .

We now outline the contents and results of the paper. In §2 we give basic definitions and introduce the theoretical stopping condition and the diameter termination criterion. In §3 we define some special noisy information and the way of using it, and give an upper bound on the cost of obtaining an  $\varepsilon$ -approximation. This bound is expressed in terms of the minimal cost for the diameter termination criterion (Theorem 3.1). In §4 we show that the upper bound derived in §3 cannot be improved. Namely, it turns out that elements  $f$ , for which the diameter termination criterion yields a cost much greater than necessary, are exceptional in the sense that they form a boundary set (Theorems 4.1–4.3). Finally, we give an example showing what the minimal cost is for the problem of approximating multivariate functions.

We conclude that the diameter termination criterion can successfully replace the theoretical stopping condition, as it yields essentially the same cost of computing an  $\varepsilon$ -approximation.

In this paper, we consider only nonadaptive information which uses functionals chosen in advance. We feel that similar results can also be obtained when successive information functionals are adaptively chosen based on previously computed information. The results on this subject are still in preparation.

## 2. PRELIMINARIES

Let  $S$ ,  $S \neq 0$ , be a linear continuous operator acting from a Banach space  $F$  to a linear normed space  $G$ . Let  $K = \{f \in F: \|f\| \leq R\}$ , where  $0 < R < +\infty$ . We wish to approximate  $S(f)$  for all  $f \in K$ , based on certain (noisy) information about  $f$ . For  $f \in K$ , *information*  $N(f)$  is gathered by a successive calculation (or observation) of certain numbers,

$$(2.1) \quad N(f) = [L_1(f), L_2(f), \dots],$$

where  $L_i: F \rightarrow \mathbb{R}$  are linear continuous functionals with  $\|L_i\| \leq 1$  belonging to some class  $\Lambda$ ,  $i \geq 1$ . Collecting information is continued until some stopping condition is fulfilled. The operator  $N: F \rightarrow \mathbb{R}^\infty$  given by (2.1) will

also be called *information*. Since the functionals  $L_i$  are given in advance, the information  $N$  is nonadaptive.

We assume that instead of the exact values  $L_i(f)$  we can only evaluate (or observe) perturbed values  $z_i$  such that

$$(2.2) \quad |z_i - L_i(f)| \leq \Delta_i,$$

where  $\Delta_i \geq 0, \quad i \geq 1$ . The sequence  $\bar{\Delta} = [\Delta_1, \Delta_2, \dots] \in \mathbb{R}^\infty$  is called a *precision sequence*.

The  $n$ th approximation  $g_n$  to  $S(f)$  is obtained as  $g_n = \phi_n(z_1, \dots, z_n)$ , where  $\phi_n$  is a mapping from  $\mathbb{R}^n$  to  $G$ . The sequence  $\phi = \{\phi_n\}_{n=0}^\infty$  is called an (idealized) *algorithm*. (By  $\phi_0$  we mean here a fixed element of  $G$ .) The  $n$ th error of  $\phi$  at  $f$  is defined as

$$(2.3) \quad e_n(\phi, N, \bar{\Delta}, f) = \sup\{ \|S(f) - \phi_n(z_1, \dots, z_n)\| : |z_i - L_i(f)| \leq \Delta_i, 1 \leq i \leq n \}.$$

We now specify the stopping condition in computing information. Given  $\varepsilon > 0$ , we compute the values  $z_1, z_2, \dots$  until the error does not exceed  $\varepsilon$ . Once such an accuracy is achieved, we want to be sure that it will not be lost if for some reason calculations are continued. The minimal number of steps after which we can terminate is thus equal to

$$(2.4) \quad n(\phi, N, \bar{\Delta}, f)(\varepsilon) = \min\{ n \geq 0 : e_j(\phi, N, \bar{\Delta}, f) \leq \varepsilon, \forall j \geq n \}$$

(with the convention  $\min \emptyset = +\infty$ ). In the sequel, we shall define a termination criterion which allows one to check whether the condition (2.4) is satisfied. For the discussion of the case when the condition ‘for all  $j \geq n$ ’ is skipped in (2.4), see the Remark after Theorem 4.3.

In the model described above, the algorithm  $\phi$ , information  $N$  and the precision sequence  $\bar{\Delta}$  do not depend on  $\varepsilon$ . Consequently, the change in accuracy from  $\varepsilon$  to  $\varepsilon_1$  ( $\varepsilon_1 < \varepsilon$ ) does not make it necessary to start computations from the beginning. To achieve  $\varepsilon_1$  we only need to compute additionally some new information values  $z_i$ .

We assume that collecting information is connected with some cost, i.e., we are charged for each noisy evaluation (observation) of a functional. The cost of obtaining a value  $z$  such that  $|z - L(f)| \leq \Delta$  is equal to  $c(\Delta)$ , where  $c : [0, +\infty) \rightarrow [0, +\infty]$  is a given nonincreasing function, independent of  $L, f$ , and  $z$ . We assume that  $c(\Delta) > 0$  for sufficiently small  $\Delta > 0$ .

The *information cost* (or *cost*) of obtaining an  $\varepsilon$ -approximation using the algorithm  $\phi$  with information  $N$ , the precision sequence  $\bar{\Delta}$ , and the termination criterion (2.4) is defined by

$$(2.5) \quad \mathbb{C}(\phi, N, \bar{\Delta}, f)(\varepsilon) = \sum_{i=1}^m c(\Delta_i)$$

for  $m < +\infty$ , and  $\mathbb{C}(\phi, N, \bar{\Delta}, f)(\varepsilon) = +\infty$  for  $m = +\infty$ , where  $m = n(\phi, N, \bar{\Delta}, f)(\varepsilon)$ . (The convention  $\sum_{i=1}^0 = 0$  is used.)

The actual cost of constructing an approximation is equal to the sum of the information cost and combinatory cost of calculating  $\phi_n(z_1, \dots, z_n)$ . The combinatory cost is here neglected. It turns out that for many important linear

problems there always exists a ‘good’ algorithm with the combinatory cost small compared to the information cost; for examples, see Traub et al. [6].

One of our purposes will be to analyze the behavior of  $\mathbb{C}(\phi, N, \bar{\Delta}, f)(\varepsilon)$  as  $\varepsilon \rightarrow 0^+$ .

Clearly, the stopping condition given by (2.4) is theoretical and cannot serve as a practical termination criterion. We shall now give a criterion which allows us to decide whether gaining new information is necessary, or an  $\varepsilon$ -approximation can already be computed using a special algorithm. To this end, we recall the concept of the *n*th diameter of information. It is given by

$$(2.6) \quad d_n(N, \bar{\Delta}) = 2 \cdot \sup\{\|S(h)\| : h \in F, \|h\| \leq 1, |L_i(h)| \leq \Delta_i, 1 \leq i \leq n\}$$

(see, e.g., Micchelli and Rivlin [5], Traub et al. [6]). Furthermore, we define a spline ( $\rho$ -spline) algorithm  $\phi^* = \{\phi_n^*\}_{n \geq 0}$  (see Trojan [7] and Kacewicz and Plaskota [3]) as follows. For  $\rho > 1$ ,  $n \geq 1$ , and  $[z_1, z_2, \dots]$  being perturbed information for some  $f \in K$ , we choose an element  $\sigma_n = \sigma_n(z_1, \dots, z_n) \in F$  such that

$$|L_i(\sigma_n) - z_i| \leq \Delta_i, \quad 1 \leq i \leq n,$$

and

$$\|\sigma_n\| \leq \rho \inf\{\|f\| : |L_i(f) - z_i| \leq \Delta_i, 1 \leq i \leq n\}.$$

A spline algorithm is defined by

$$\phi_0^* = 0, \quad \phi_n^*(z_1, \dots, z_n) = S(\sigma_n), \quad n \geq 1.$$

It is known that

$$(2.7) \quad e_n(\phi^*, N, \bar{\Delta}, f) \leq D(f) d_n(N, \bar{\Delta}) \quad \text{for all } f \in K,$$

where  $D(f) = \max\{1, \frac{1+\rho}{2}\|f\|\}$ , see Kacewicz and Plaskota [3, Theorem 4.1]. Hence, if the algorithm  $\phi^*$  is applied, it is enough to compute  $n^d(N, \bar{\Delta})(\varepsilon)$  pieces of information to obtain an  $\varepsilon$ -approximation, where

$$(2.8) \quad n^d(N, \bar{\Delta})(\varepsilon) = \min \left\{ n \geq 0 : d_n(N, \bar{\Delta}) \leq \frac{\varepsilon}{D_R} \right\}$$

and  $D_R = \max\{1, \frac{1+\rho}{2}R\}$ .

Note that the termination criterion (2.8) does not depend on a particular  $f$ , but only on the class  $K$ . Since for many problems the behavior of  $d_n(N, \bar{\Delta})$  is known (see, e.g., Babenko [1], Marchuk and Osipenko [4], and Traub et al. [6]), the number  $n^d(N, \bar{\Delta})(\varepsilon)$  may often be computed, in contrast with  $n(\phi, N, \bar{\Delta}, f)(\varepsilon)$ . However, the criterion (2.8) is only useful provided that  $n^d(N, \bar{\Delta})(\varepsilon)$  is not much greater than  $n(\phi, N, \bar{\Delta}, f)(\varepsilon)$ . As we shall see, this is indeed the case.

The information cost of obtaining an  $\varepsilon$ -approximation using  $N$ ,  $\bar{\Delta}$  and  $\phi^*$  with the stopping criterion (2.8) is independent of  $f$  and equal to

$$(2.9) \quad \mathbb{C}^d(N, \bar{\Delta})(\varepsilon) = \sum_{i=1}^m c(\Delta_i)$$

for  $m < +\infty$ , and  $\mathbb{C}^d(N, \bar{\Delta})(\varepsilon) = +\infty$  for  $m = +\infty$ , where  $m = n^d(N, \bar{\Delta})(\varepsilon)$ . We call  $\mathbb{C}^d(N, \bar{\Delta})(\varepsilon)$  the *diameter criterion cost*. Obviously, (2.7) yields that, for any  $\varepsilon > 0$  and  $f \in K$ , one has

$$(2.10) \quad \mathbb{C}(\phi^*, N, \bar{\Delta}, f)(\varepsilon) \leq \mathbb{C}^d(N, \bar{\Delta})(\varepsilon).$$

A relation between the costs (2.5) and (2.9) will be discussed in the next sections. We shall show that the upper bound (2.10) is essentially sharp, i.e., the criterion (2.8) is *not* pessimistic. Special attention will be given to the choice of  $\phi$ ,  $N$  and  $\bar{\Delta}$  such that the cost  $C(\phi, N, \bar{\Delta}, f)(\varepsilon)$  grows as slowly as possible as  $\varepsilon \rightarrow 0^+$ , for all  $f \in K$ . In particular, we shall study a relation between  $C(\phi, N, \bar{\Delta}, f)(\varepsilon)$  and the *minimal diameter criterion cost* defined as

$$(2.11) \quad \text{MC}^d(\varepsilon) = \inf_{N, \bar{\Delta}} C^d(N, \bar{\Delta})(\varepsilon),$$

the infimum being taken with respect to information  $N$  consisting of functionals from  $\Lambda$ . For illustration, the cost  $\text{MC}^d(\varepsilon)$  for the approximation problem is given in the Example of §4.

In the next section we shall construct information  $N^*$  and a precision sequence  $\bar{\Delta}^*$  which supply an (almost)  $\varepsilon$ -approximation with cost no greater than  $\text{MC}^d(\varepsilon)$ , for all sufficiently small  $\varepsilon$ . The problem that we have to face is that  $N$  and  $\bar{\Delta}$  minimizing (2.11) depend on  $\varepsilon$ , so that they cannot serve immediately as  $N^*$  and  $\bar{\Delta}^*$ , which must be  $\varepsilon$ -independent.

### 3. THE CONSTRUCTION OF $N^*$ AND $\bar{\Delta}^*$ AND AN UPPER BOUND

We first observe that  $\text{MC}^d(\varepsilon) > 0$  for sufficiently small  $\varepsilon > 0$ . Indeed, let  $0 < \Delta \leq 1$  be such that  $c(\Delta) > 0$ . If  $\text{MC}^d(\varepsilon) = 0, \forall \varepsilon > 0$ , then there exist  $N, \bar{\Delta}$ , and  $n$  for which  $d_n(N, \bar{\Delta}) < 2\Delta\|S\|$  and  $\sum_{i=1}^n c(\Delta_i) < c(\Delta)$ . This, however, is impossible, since then  $\Delta_i > \Delta, 1 \leq i \leq n$ , and

$$d_n(N, \bar{\Delta}) \geq 2 \cdot \sup\{\|S(h)\| : \|h\| \leq 1, |L_i(h)| \leq \Delta, 1 \leq i \leq n\} \geq 2\Delta\|S\|.$$

Without loss of generality, we restrict ourselves to problems solvable with respect to the criterion (2.8), i.e., such that  $\text{MC}^d(\varepsilon) < +\infty, \forall \varepsilon > 0$ . The case  $\text{MC}^d(\varepsilon) = +\infty$  (for small  $\varepsilon$ ) is considered in Theorem 4.3 (ii), which states that the problem is then practically not solvable, even with respect to the theoretical criterion (2.4).

We assume in this section that the problem is ‘hard’ in the following sense:

(A) *There exist  $0 < p < 1$  and  $\alpha > 1$  such that*

$$\text{MC}^d(\alpha \cdot \varepsilon) \leq p \cdot \text{MC}^d(\varepsilon)$$

for all sufficiently small  $\varepsilon > 0$ .  $\square$

Note that (A) always holds with  $p = 1$ . For  $p < 1$  it states that the minimal diameter criterion cost tends to infinity sufficiently fast as  $\varepsilon$  decreases, see the Example.

We now define  $N^*$  and  $\bar{\Delta}^*$ . Let  $\omega > 1$  and  $i \geq 0$ . Choose information  $N^i = [L_1^i, L_2^i, \dots]$  consisting of functionals from  $\Lambda$ , a precision sequence  $\bar{\Delta}^i = [\Delta_1^i, \Delta_2^i, \dots]$  and an integer  $n_i \geq 0$  such that

$$(3.1) \quad d_{n_i}(N^i, \bar{\Delta}^i) \leq \frac{1}{\alpha^i D_R}$$

and

$$(3.2) \quad \sum_{j=1}^{n_i} c(\Delta_j^i) \leq \omega \cdot \text{MC}^d\left(\frac{1}{\alpha^i}\right).$$

This selection is possible for sufficiently large  $i$ ,  $i \geq l$ , where  $l \geq 0$ . Denoting by  $N_{n^i}^i$  and  $\bar{\Delta}_{n^i}^i$  the first  $n^i$  components of  $N^i$  and  $\bar{\Delta}^i$ , respectively, we now define

$$N^* = [N_{n^l}^l, N_{n^{l+1}}^{l+1}, N_{n^{l+2}}^{l+2}, \dots]$$

and

$$\bar{\Delta}^* = [\bar{\Delta}_{n^l}^l, \bar{\Delta}_{n^{l+1}}^{l+1}, \bar{\Delta}_{n^{l+2}}^{l+2}, \dots].$$

In the following theorem we show that the spline algorithm  $\phi^*$  using information  $N^*$  and the precision sequence  $\bar{\Delta}^*$  produces an (almost)  $\varepsilon$ -approximation with the cost proportional at most to  $\text{MC}^d(\varepsilon)$ , even if the criterion (2.8) is applied.

**Theorem 3.1.** *Let  $\text{MC}^d(\varepsilon) < +\infty$  for all  $\varepsilon > 0$ , and let the condition (A) hold. Then, for all  $f \in K$  and all sufficiently small  $\varepsilon > 0$ , we have that*

$$\mathbb{C}(\phi^*, N^*, \bar{\Delta}^*, f)(\alpha \cdot \varepsilon) \leq \mathbb{C}^d(N^*, \bar{\Delta}^*)(\alpha \cdot \varepsilon) \leq \frac{\omega}{1-p} \text{MC}^d(\varepsilon).$$

*Proof.* From (3.1), (3.2), and from the condition (A), we have for sufficiently large  $k$  that

$$\begin{aligned} \mathbb{C}^d(N^*, \bar{\Delta}^*) \left( \frac{1}{\alpha^k} \right) &\leq \sum_{i=l}^k \sum_{j=1}^{n^i} c(\Delta_j^i) \leq \omega \sum_{i=l}^k \text{MC}^d \left( \frac{1}{\alpha^i} \right) \\ &\leq \omega \text{MC}^d \left( \frac{1}{\alpha^k} \right) \sum_{i=0}^{k-l} p^i \leq \frac{\omega}{1-p} \text{MC}^d \left( \frac{1}{\alpha^k} \right). \end{aligned}$$

Now let  $k = k(\varepsilon) \in \mathbb{N}$  be the minimal number such that  $\alpha^{-k} \leq \varepsilon$ . Then, for sufficiently small  $\varepsilon > 0$ ,

$$\begin{aligned} \mathbb{C}^d(N^*, \bar{\Delta}^*)(\varepsilon) &\leq \mathbb{C}^d(N^*, \bar{\Delta}^*) \left( \frac{1}{\alpha^k} \right) \\ &\leq \frac{\omega}{1-p} \text{MC}^d \left( \frac{1}{\alpha^k} \right) \leq \frac{\omega}{1-p} \text{MC}^d \left( \frac{1}{\alpha} \varepsilon \right). \end{aligned}$$

The inequality (2.10) finally yields that

$$\mathbb{C}(\phi^*, N^*, \bar{\Delta}^*, f)(\alpha \cdot \varepsilon) \leq \mathbb{C}^d(N^*, \bar{\Delta}^*)(\alpha \varepsilon) \leq \frac{\omega}{1-p} \text{MC}^d(\varepsilon). \quad \square$$

Hence, information  $N^*$  and the precision sequence  $\bar{\Delta}^*$  with the termination criterion (2.8) (and, obviously, also with the stopping condition (2.4)) give an almost  $\varepsilon$ -approximation with a cost at most (up to a constant)  $\text{MC}^d(\varepsilon)$ . In the next section we shall show that  $N^*$  and  $\bar{\Delta}^*$  are almost optimal, in the sense that the cost of obtaining an  $\varepsilon$ -approximation using arbitrary  $N$  and  $\bar{\Delta}$  cannot be much smaller than  $\text{MC}^d(\varepsilon)$ , even if the theoretical condition (2.4) is used.

#### 4. LOWER BOUNDS

In this section we provide lower bounds on the cost  $\mathbb{C}(\phi, N, \bar{\Delta}, f)(\varepsilon)$ , for elements  $f$  belonging to certain dense subsets of  $K$ . The main result states, roughly speaking, that  $\mathbb{C}(\phi, N, \bar{\Delta}, f)(\varepsilon)$  is bounded from below by  $\text{MC}^d(\varepsilon)$ . This will determine the sharpness of the upper bound derived in the previous section.

Let us first consider fixed  $N$  and  $\bar{\Delta}$ , and start with the case  $\mathbb{C}^d(N, \bar{\Delta})(\varepsilon) < +\infty$ , for all  $\varepsilon > 0$ .

**Theorem 4.1.** Let  $\mathbb{C}^d(N, \bar{\Delta})(\varepsilon) < +\infty$ ,  $\forall \varepsilon > 0$ , and let  $\phi$  be an arbitrary algorithm.

(i) If  $d_n(N, \bar{\Delta}) > 0$ ,  $\forall n \geq 0$ , then for any function  $h: (0, +\infty) \rightarrow (0, +\infty)$  with  $\lim_{\varepsilon \rightarrow 0^+} h(\varepsilon) = 0$ , the set

$$A_1 = \{f \in K : \exists C = C(f) \geq 0 \exists \varepsilon_0 = \varepsilon_0(f) > 0 \text{ such that for all } 0 < \varepsilon \leq \varepsilon_0 \\ \mathbb{C}(\phi, N, \bar{\Delta}, f)(C \cdot h(\varepsilon) \cdot \varepsilon) \leq \mathbb{C}^d(N, \bar{\Delta})(\varepsilon)\}$$

is a boundary set in  $K$ .

(ii) If  $d_n(N, \bar{\Delta}) = 0$  for some  $n$ , then the set

$$A_2 = \{f \in K : \exists C = C(f) \geq 0 \exists \varepsilon_0 = \varepsilon_0(f) > 0 \text{ such that for all } 0 < \varepsilon \leq \varepsilon_0 \\ \mathbb{C}(\phi, N, \bar{\Delta}, f)(C \cdot \varepsilon) < \mathbb{C}^d(N, \bar{\Delta})(\varepsilon)\}$$

is a boundary set in  $K$ .

*Proof.* (i) Observe that there exists a subsequence  $\{\Delta_{n_k}\}$  such that  $\lim_{k \rightarrow +\infty} \Delta_{n_k} = 0$ . Indeed, otherwise  $\Delta_i \geq \Delta$ ,  $\forall i > n$ , for some  $n \in \mathbb{N}$  and  $0 < \Delta \leq 1$ . Since  $d_n(N, \bar{\Delta}) > 0$ , there is an element  $h$  such that  $0 < \|h\| \leq \Delta$ ,  $|L_i(h)| \leq \Delta_i$ ,  $1 \leq i \leq n$ , and  $\|S(h)\| > 0$ . Hence,  $|L_i(h)| \leq \Delta_i$ ,  $\forall i \geq 1$ , and consequently  $d_i(N, \bar{\Delta}) \geq 2\|S(h)\| > 0$ ,  $\forall i \geq 0$ . On the other hand,  $\mathbb{C}^d(N, \bar{\Delta})(\varepsilon) < +\infty$ ,  $\forall \varepsilon > 0$ , implies that  $\lim_{i \rightarrow +\infty} d_i(N, \bar{\Delta}) = 0$ , which is a contradiction.

Hence, we can select a subsequence  $\{n_k\}$  in such a way that

$$(4.1) \quad c(\Delta_{n_k+1}) > 0, \quad \forall k \geq 1.$$

Let  $\varepsilon_k = d_{n_k}(N, \bar{\Delta}) \cdot D_R$  and  $\delta_k = h(\varepsilon_k)$ , for  $k \geq 1$ . A slight modification of Theorem 4.2 from Kacewicz and Plaskota [3], which consists in replacing  $n$  by  $n_k$ , yields that the set

$$\tilde{A}_1 = \left\{ f \in K : \limsup_{k \rightarrow +\infty} \frac{e_{n_k}(\phi, N, \bar{\Delta}, f)}{\delta_k \varepsilon_k} < +\infty \right\}$$

is a boundary set in  $K$ . We shall show that  $A_1 \subset \tilde{A}_1$ . Indeed, suppose that  $f \in K$  is not in  $\tilde{A}_1$ . Then, for any  $C > 0$  there exists an increasing sequence  $\{k_l\}$  such that

$$e_{n_{k_l}}(\phi, N, \bar{\Delta}, f) > C \delta_{k_l} \varepsilon_{k_l} = Ch(\varepsilon_{k_l})\varepsilon_{k_l}$$

for  $l \geq 1$ . This and (4.1) yield that

$$\mathbb{C}(\phi, N, \bar{\Delta}, f)(Ch(\varepsilon_{k_l})\varepsilon_{k_l}) > \sum_{i=1}^{n_{k_l}} c(\Delta_i).$$

Since, on the other hand,

$$\mathbb{C}^d(N, \bar{\Delta})(\varepsilon_{k_l}) \leq \sum_{i=1}^{n_{k_l}} c(\Delta_i),$$

we have

$$\mathbb{C}(\phi, N, \bar{\Delta}, f)(Ch(\varepsilon_{k_l})\varepsilon_{k_l}) > \mathbb{C}^d(N, \bar{\Delta})(\varepsilon_{k_l}),$$

for  $l \geq 1$ , which implies that  $f$  is not an element of  $A_1$ . Hence,  $A_1 \subset \tilde{A}_1$ , and  $A_1$  is a boundary set in  $K$ .

(ii) Let  $K(f, \alpha)$  be a closed ball with center  $f \in A_2$  and radius  $\alpha > 0$ , contained in  $K$ . We shall show that  $K(f, \alpha)$  contains an element which is not in  $A_2$ . Let  $m = \min\{n \geq 1 : d_n(N, \bar{\Delta}) = 0\}$ . Choose  $h \in F$  such that  $\|h\| \leq \alpha$ ,  $|L_i(h)| \leq \Delta_i$ ,  $1 \leq i < m$ , and  $\|S(h)\| > 0$ . Letting  $f_1 = f + h$  and  $z_i = L_i(f)$ ,  $\forall i$ , we have that  $f_1 \in K(f, \alpha)$  and  $|z_i - L_i(f_1)| \leq \Delta_i$ ,  $1 \leq i < m$ . For sufficiently small  $\varepsilon$ , since  $f \in A_2$  and  $n^d(N, \bar{\Delta})(\varepsilon) = m$ , we get that

$$\begin{aligned} \|S(f_1) - \phi_{m-1}(z_1, \dots, z_{m-1})\| &\geq \|S(h)\| - \|S(f) - \phi_{m-1}(z_1, \dots, z_{m-1})\| \\ &\geq \|S(h)\| - C(f) \cdot \varepsilon \geq \frac{1}{2} \|S(h)\|. \end{aligned}$$

Hence, for any  $D \geq 0$  and all sufficiently small  $\varepsilon > 0$ , we have that

$$e_{m-1}(\phi, N, \bar{\Delta}, f_1) > D\varepsilon,$$

which yields

$$\mathbb{C}(\phi, N, \bar{\Delta}, f_1)(D\varepsilon) \geq \sum_{i=1}^m c(\Delta_i) = \mathbb{C}^d(N, \bar{\Delta})(\varepsilon).$$

Thus,  $f_1 \notin A_2$  and the set  $A_2$  is a boundary set in  $K$ . The proof is completed.  $\square$

Theorem 4.1 provides a lower bound on  $\mathbb{C}(\phi, N, \bar{\Delta}, f)(\varepsilon)$  on a dense set of elements  $f$ , for any algorithm  $\phi$ . In the case (i), it shows that the upper bound (2.10) is sharp with respect to the choice of an algorithm. More specifically, an inequality of the type (2.10) cannot hold (except for a boundary set of  $f$ 's) no matter what  $\phi$  is, if  $\varepsilon$  in the left-hand side is replaced by  $h(\varepsilon) \cdot \varepsilon$ . Here, the function  $h(\varepsilon)$  may tend to 0 arbitrarily slowly with  $\varepsilon$ . The inequality (2.10) also shows that the function  $h(\varepsilon)$  cannot be omitted in the formulation of Theorem 4.1. In the case (ii), the theorem states that the weak inequality (2.10) cannot be replaced by a sharp one. For given  $N$  and  $\bar{\Delta}$ , the spline algorithm  $\phi^*$  is thus almost optimal.

In terms of the termination criteria, the above result is somewhat surprising. It says that the theoretical stopping condition (2.4) yields a cost larger than the criterion (2.8), if the accuracy required in (2.4) is only slightly smaller than that in (2.8).

Consider now the case when  $\mathbb{C}^d(N, \bar{\Delta})(\varepsilon) = +\infty$  for sufficiently small  $\varepsilon > 0$ , i.e., when the problem cannot be solved with respect to the criterion (2.8).

**Theorem 4.2.** *Let  $\mathbb{C}^d(N, \bar{\Delta})(\varepsilon) = +\infty$  for sufficiently small  $\varepsilon > 0$ , and let  $\phi$  be an arbitrary algorithm.*

(i) *If  $\lim_{n \rightarrow +\infty} d_n(N, \bar{\Delta}) > 0$  and  $\sum_{i=1}^{\infty} c(\Delta_i) = +\infty$ , then for any function  $H : (0, +\infty) \rightarrow [0, +\infty)$ , the set*

$$A_3 = \{f \in K : \exists C = C(f) \geq 0 \exists \varepsilon_0 = \varepsilon_0(f) > 0 \text{ such that for all } 0 < \varepsilon \leq \varepsilon_0 \\ \mathbb{C}(\phi, N, \bar{\Delta}, f)(C \cdot \varepsilon) \leq H(\varepsilon)\}$$

*is a boundary set in  $K$ .*

(ii) *If*

$$\lim_{n \rightarrow +\infty} d_n(N, \bar{\Delta}) > 0 \quad \text{and} \quad \sum_{i=1}^{\infty} c(\Delta_i) < +\infty,$$



or if  $\lim_{n \rightarrow +\infty} d_n(N, \bar{\Delta}) = 0$ , then the set

$$A_4 = \{f \in K : \mathbb{C}(\phi, N, \bar{\Delta}, f)(\varepsilon) < +\infty \quad \forall \varepsilon > 0\}$$

is a boundary set in  $K$ .

*Proof.* (i) Choose a positive, nonincreasing sequence  $\{\varepsilon_n\}$  such that  $\lim_{n \rightarrow +\infty} \varepsilon_n = 0$  and

$$\sum_{i=1}^n c(\Delta_i) > H(\varepsilon_n)$$

for sufficiently large  $n$ . By a result of Kacewicz and Plaskota [3], the set

$$\tilde{A}_3 = \left\{ f \in K : \limsup_{n \rightarrow +\infty} \frac{e_n(\phi, N, \bar{\Delta}, f)}{\varepsilon_n} < +\infty \right\}$$

is a boundary set in  $K$ . We show that  $A_3 \subset \tilde{A}_3$ . Let  $f \in A_3$ . Then for some  $C \geq 0$  and large  $n$  one has

$$\mathbb{C}(\phi, N, \bar{\Delta}, f)(C\varepsilon_n) \leq H(\varepsilon_n) < \sum_{i=1}^n c(\Delta_i).$$

This yields that  $e_n(\phi, N, \bar{\Delta}, f) \leq C\varepsilon_n$ , so that  $f \in \tilde{A}_3$ , as claimed. Hence,  $A_3$  is a boundary set.

(ii) Suppose first that  $\lim_{n \rightarrow +\infty} d_n(N, \bar{\Delta}) > 0$  and  $\sum_{i=1}^{\infty} c(\Delta_i) < +\infty$ . Let  $K(f, \alpha)$  be a closed ball with center  $f \in A_4$  and radius  $\alpha > 0$ , contained in  $K$ . Since  $\sum_{i=1}^{\infty} c(\Delta_i) < +\infty$ , there are  $m \in \mathbb{N}$  and  $\Delta$ ,  $0 < \Delta \leq \alpha$ , such that  $\Delta_i \geq \Delta$  for all  $i \geq m$ . Similarly as in the proof of Theorem 4.1 (ii), we choose  $h \in F$  satisfying  $\|h\| \leq \Delta$ ,  $|L_i(h)| \leq \Delta_i$ ,  $1 \leq i < m$ , and  $\|S(h)\| > 0$ , and we set  $f_1 = f + h$ . Then  $f_1 \in K(f, \alpha)$ , and for  $z_i = L_i(f)$  there holds  $|z_i - L_i(f_1)| \leq \Delta_i$ ,  $\forall i \geq 1$ . Since  $\lim_{i \rightarrow +\infty} e_i(\phi, N, \bar{\Delta}, f) = 0$ , we have for all sufficiently large  $i$  that

$$\begin{aligned} e_i(\phi, N, \bar{\Delta}, f_1) &\geq \|S(f_1) - \phi_i(z_1, \dots, z_i)\| \\ &\geq \|S(h)\| - \|S(f) - \phi_i(z_1, \dots, z_i)\| \geq \frac{1}{2} \|S(h)\|, \end{aligned}$$

which implies that  $\mathbb{C}(\phi, N, \bar{\Delta}, f_1)(\varepsilon) = +\infty$ , for sufficiently small  $\varepsilon$ . Hence,  $f_1 \notin A_4$ , which yields that  $A_4$  is a boundary set.

Now let  $\lim_{n \rightarrow +\infty} d_n(N, \bar{\Delta}) = 0$ . Since  $\mathbb{C}^d(N, \bar{\Delta})(\varepsilon) = +\infty$  for sufficiently small  $\varepsilon > 0$ , there must be a number  $m \in \mathbb{N}$  such that  $c(\Delta_m) = +\infty$ . We choose  $m$  to be the minimal number with this property. Let  $f \in A_4$ . Observe that  $d_{m-1}(N, \bar{\Delta}) > 0$ , and take  $\alpha$ ,  $h$ ,  $f_1$ , and  $z_i$ ,  $1 \leq i < m$ , as in the proof of Theorem 4.1 (ii). Since  $\phi_{m-1}(z_1, \dots, z_{m-1}) = S(f)$ , we get that

$$e_{m-1}(\phi, N, \bar{\Delta}, f_1) \geq \|S(f_1) - \phi_{m-1}(z_1, \dots, z_{m-1})\| = \|S(h)\| > 0.$$

This yields that for sufficiently small  $\varepsilon$  we must compute  $L_m(f_1)$  if we want to have an  $\varepsilon$ -approximation to  $S(f_1)$ , which implies that  $\mathbb{C}(\phi, N, \bar{\Delta}, f_1)(\varepsilon) = +\infty$ . Thus,  $f_1 \notin A_4$ , and the set  $A_4$  is a boundary set in  $K$ . The proof of Theorem 4.2 is completed.  $\square$

Hence, if the problem cannot be approximated with finite cost using the termination criterion (2.8) then, practically, it also cannot be approximated,

even if the idealized criterion (2.4) is applied. For any algorithm  $\phi$ , the cost is arbitrarily large (in the case (i)), or infinite (in the case (ii)), on a dense set of elements  $f$ .

Theorems 4.1, 4.2, and the obvious inequality  $\text{MC}^d(\varepsilon) \leq \mathbb{C}^d(N, \bar{\Delta})(\varepsilon)$ , for all  $N, \bar{\Delta}, \varepsilon$ , yield the following theorem.

**Theorem 4.3.** *Let  $N, \bar{\Delta}$ , and  $\phi$  be arbitrary information, precision sequence, and algorithm, respectively. We have:*

(i) *If  $\text{MC}^d(\varepsilon) < +\infty, \forall \varepsilon > 0$ , then for any function  $h: (0, +\infty) \rightarrow (0, +\infty)$  with  $\lim_{\varepsilon \rightarrow 0^+} h(\varepsilon) = 0$ , the set*

$$B_1 = \{ f \in K : \exists C = C(f) \geq 0 \exists \varepsilon_0 = \varepsilon_0(f) > 0 \text{ such that for all } 0 < \varepsilon \leq \varepsilon_0 \\ \mathbb{C}(\phi, N, \bar{\Delta}, f)(C \cdot h(\varepsilon) \cdot \varepsilon) < \text{MC}^d(\varepsilon) \}$$

*is a boundary set in  $K$ .*

(ii) *If  $\text{MC}^d(\varepsilon) = +\infty$  for sufficiently small  $\varepsilon > 0$ , then for any function  $H: (0, +\infty) \rightarrow [0, +\infty)$ , the set*

$$B_2 = \{ f \in K : \exists C = C(f) \geq 0 \exists \varepsilon_0 = \varepsilon_0(f) > 0 \text{ such that for all } 0 < \varepsilon \leq \varepsilon_0 \\ \mathbb{C}(\phi, N, \bar{\Delta}, f)(C \cdot \varepsilon) \leq H(\varepsilon) \}$$

*is a boundary set in  $K$ .  $\square$*

We now comment on the results of this paper. In the case  $\text{MC}^d(\varepsilon) < +\infty$  the cost  $\mathbb{C}(\phi, N, \bar{\Delta}, f)(\varepsilon)$  may grow more slowly than  $\text{MC}^d(\varepsilon)$ , as  $\varepsilon \rightarrow 0^+$ , only on a boundary set of elements  $f$ . If the problem satisfies the assumption (A), then the information  $N^*$ , the precision sequence  $\bar{\Delta}^*$ , and the spline algorithm  $\phi^*$  are almost optimal, i.e.,  $\mathbb{C}(\phi^*, N^*, \bar{\Delta}^*, f)(\varepsilon)$  essentially behaves like  $\text{MC}^d(\varepsilon)$ , for all  $f \in K$ . This holds up to a (not significant in practice) function  $h(\varepsilon)$ . In the case when  $\text{MC}^d(\varepsilon) = +\infty$ , the cost  $\mathbb{C}(\phi, N, \bar{\Delta}, f)(\varepsilon)$  grows arbitrarily fast as  $\varepsilon \rightarrow 0^+$  for any  $\phi, N$ , and  $\bar{\Delta}$ , on a dense set of  $f$ .

We conclude that the problem of finding the optimal  $N, \bar{\Delta}$ , and  $\phi$  for the theoretical stopping condition (2.4) can be essentially reduced to the similar problem with the criterion (2.8). In both cases, the minimal cost essentially behaves like  $\text{MC}^d(\varepsilon)$ . Therefore, the diameter termination criterion can be recommended when approximating linear problems.

We end with an example showing how large the costs related to the conditions (2.4) and (2.8) are for an approximation problem.

**Example.** For  $r \geq 1$  and  $s \geq 1$ , let  $F$  be the space of  $r$  times continuously differentiable functions  $f: [0, 1]^s \rightarrow \mathbb{R}$  with the norm

$$\|f\| = \max_{0 \leq k_1 + \dots + k_s = k \leq r} \left\| \frac{\partial^k f(x_1, \dots, x_s)}{\partial x_1^{k_1} \dots \partial x_s^{k_s}} \right\|_\infty,$$

where  $x_i \in [0, 1], 1 \leq i \leq s$ . We want to uniformly approximate functions  $f$  such that  $\|f\| \leq 1$ , based on their noisy evaluations at some points. That is,  $G = \mathbb{C}([0, 1]^s)$  and  $S: F \rightarrow G$  is the embedding operator,  $S(f) = f$ . Information is given by

$$N(f) = [f(t_1), f(t_2), \dots],$$

where  $t_i \in [0, 1]^s$ ,  $i \geq 1$ . We define the cost function by

$$c(\Delta) = \max\{0, \log_2(1/\Delta)\},$$

which corresponds to the number of binary bits required for representing a value  $f(t)$  with precision  $\Delta$ .

From Kacewicz and Plaskota [2] we have for this problem that

$$\text{MC}^d(\varepsilon) = \Theta\left(\varepsilon^{-s/r} \cdot \log_2\left(\frac{1}{\varepsilon}\right)\right), \quad \text{as } \varepsilon \rightarrow 0^+.$$

Note that the condition (A) is now obviously satisfied. The results of this paper show that the cost  $\mathbb{C}(\phi, N, \bar{\Delta}, f)(\varepsilon)$  increases at least as fast as (essentially)  $\varepsilon^{-s/r} \cdot \log_2(\frac{1}{\varepsilon})$ , as  $\varepsilon \rightarrow 0^+$ , except for a boundary set of elements  $f$ . The slowest possible growth is achieved by the algorithm  $\phi^*$ , information  $N^*$  and precision sequence  $\bar{\Delta}^*$  defined in §§2 and 3.  $\square$

*Remark.* Theorem 4.3 does not hold if we drop the requirement that the inequality in (2.4) hold for all  $j \geq n$ . Let (2.4) be replaced by

$$(2.4a) \quad \tilde{n}(\phi, N, \bar{\Delta}, f)(\varepsilon) = \min\{n \geq 0 : e_n(\phi, N, \bar{\Delta}, f) \leq \varepsilon\}.$$

Assume that  $F$  is a separable Banach space, and let  $\{f_n\}_{n=0}^\infty$  be a dense subset in  $F$ . Take a problem such that  $\text{MC}^d(\varepsilon) > 0$  for small  $\varepsilon$ , with the cost function satisfying  $c(1) = 0$  (see for instance the Example). Let  $N$  be arbitrary information,  $\Delta_n = 1$  for all  $n$  and the algorithm  $\phi = \{\phi_n\}_{n=0}^\infty$  be given by

$$\phi_n(z_1, \dots, z_n) = S(f_n)$$

for all  $z_1, \dots, z_n$  and  $n \geq 0$ . Then, for all  $f$  and any  $\varepsilon > 0$ , we have

$$e_n(\phi, N, \bar{\Delta}, f) = \|S(f) - S(f_n)\| \leq \varepsilon$$

for some  $n$ , which yields that  $\tilde{n}(\phi, N, \bar{\Delta}, f)(\varepsilon) < +\infty$  and  $\mathbb{C}(\phi, N, \bar{\Delta}, f)(\varepsilon) = 0$ . Hence, the sets  $B_1$  and  $B_2$  in Theorem 4.3 are both equal to  $K$ , and the assertion of the theorem does not hold.  $\square$

#### BIBLIOGRAPHY

1. K. I. Babenko, *Theoretical background and constructing computational algorithms for mathematical-physical problems*, Nauka, Moscow, 1979. (Russian)
2. B. Z. Kacewicz and L. Plaskota, *On the minimal cost of approximating linear problems based on information with deterministic noise*, Numer. Funct. Anal. Optim., Nos. 5 and 6 (1990), 511–529.
3. ———, *Noisy information for linear problems in the asymptotic setting*, J. Complexity 7 (1991), 35–57.
4. A. A. Marchuk and K. Y. Osipenko, *Best approximation of functions specified with an error at a finite number of points*, Math. Notes 17 (1975), 207–212.
5. C. A. Micchelli and T. J. Rivlin, *A survey of optimal recovery*, Estimation in Approximation Theory, Plenum Press, New York, 1977, pp. 1–54.
6. J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski, *Information-based complexity*, Academic Press, New York, 1988.
7. G. M. Trojan, *Asymptotic setting for linear problems*, unpublished, see Traub et al. [6, pp. 383–395].

INSTITUTE OF APPLIED MATHEMATICS, UNIVERSITY OF WARSAW, UL. BANACHA 2, 02–097 WARSAW, POLAND

E-mail address, B. Z. Kacewicz: bkacew@mimuw.edu.pl  
 bkacew@plearn.bitnet