The complexity of definite elliptic problems with noisy data

Technical Report CUCS-035-96

Arthur G. Werschulz

Department of Computer and Information Sciences, Fordham University
Fordham College at Lincoln Center
New York, NY 10023
and
Department of Computer Science
Columbia University
New York, NY 10023

September 26, 1996

Abstract. We study the complexity of $2m$th order definite elliptic problems $Lu = f$ (with homogeneous Dirichlet boundary conditions) over a d-dimensional domain Ω, error being measured in the $H^m(\Omega)$-norm. The problem elements f belong to the unit ball of $W^{r,p}(\Omega)$, where $p \in [2, \infty]$ and $r > d/p$. Information consists of (possibly-adaptive) noisy evaluations of f or the coefficients of L. The absolute error in each noisy evaluation is at most δ. We find that the nth minimal radius for this problem is proportional to $n^{-r/d} + \delta$, and that a noisy finite element method with quadrature (FEMQ), which uses only function values, and not derivatives, is a minimal error algorithm. This noisy FEMQ can be efficiently implemented using multigrid techniques. Using these results, we find tight bounds on the ε-complexity (minimal cost of calculating an ε-approximation) for this problem, said bounds depending on the cost $c(\delta)$ of calculating a δ-noisy information value. As an example, if the cost of a δ-noisy evaluation is $c(\delta) = \delta^{-s}$ (for $s > 0$), then the complexity is proportional to $(1/\varepsilon)^{d/r+s}$.

1. Introduction

The majority of research (see, e.g., [9]) in information-based complexity has concentrated on problems for which we have partial information that is exact. There has recently been a stream of work (much of which has been done by L. Plaskota, and is described in his monograph [7]) on the complexity of problems with partial information that is contaminated by noise. In this paper, we study the complexity of elliptic partial differential equations $Lu = f$, with noisy partial information.

Most previous work (see, e.g., [10], [11], and [12], as well as the references cited therein) on the complexity of elliptic PDEs has assumed that we have complete information about the coefficients of L, and exact (but partial) information about the right-hand side f. As a typical result, consider the $2m$th order elliptic boundary value problem $Lu = f$ (with homogeneous Dirichlet boundary conditions), defined on a d-dimensional domain Ω. The right-hand sides f belong to the unit ball $\mathscr{B}W^{r,p}(\Omega)$ of the Sobolev space $W^{r,p}(\Omega)$, so that they have r derivatives in the L_p sense. We require that $p \in [2, \infty]$ and $r > d/p$. Error of an approximation is measured in the $H^m(\Omega)$-norm. Information about a problem element f consists of the values of f (or some of its derivatives) at a finite number of points in Ω. Then the minimal error over all algorithms using at most n evaluations is $\Theta(n^{-r/d})$. It then follows that the ε-complexity (i.e., the minimal cost of calculating an ε-approximation) is $\Theta\big((1/\varepsilon)^{d/r}\big)$. Moreover, a finite element method using quadrature (FEMQ), which only uses function values (and no derivatives) is optimal. The details for the special case

This research was supported in part by the National Science Foundation under Grant CCR-95-00850.

$p = 2$ can be found in [10, Section 5.5]; the proof for the general case $p = [2, \infty]$ is not much different than that for this special case.

Of course, it is more realistic to assume that we have only partial information about the coefficients of L. This means that we are studying classes of elliptic Dirichlet problems $L_{\mathbf{a}} u = f$. Here $L_{\mathbf{a}}$ is a linear elliptic operator of order $2m$ with coefficients $\mathbf{a}$, defined on a d-dimensional domain Ω. The right-hand sides f once again belong to $\mathscr{B}W^{r,p}(\Omega)$, and the coefficient vectors a now belong to a class $\mathscr{A}$ of functions.

Note that our problem elements are now of the form $[f; \mathbf{a}]$. Since the solution $u = L_{\mathbf{a}}^{-1} f$ depends nonlinearly on a, we are now dealing with a nonlinear problem. There has been little work on the complexity of nonlinear problems arising in partial differential equations. One such result is the following, from [10, pp. 110–111]:

Assume that we can compute f and the coefficients $\mathbf{a}$ of $L_{\mathbf{a}}$ (or their derivatives) at points in Ω. Then the nth minimal error is $\Theta(n^{-r/d})$, this error being achieved by an FEMQ using n evaluations. Although [10] does not derive the complexity from this minimal error result, it is not too difficult to show that the ε-complexity is still $\Theta\big((1/\varepsilon)^{d/r}\big)$. Indeed, we can use multigrid techniques (see [2], especially Chapter 7) to get a sufficiently-good approximation to the FEMQ, in time proportional to the number of information evaluations used.

However, we can ask that the information be made even more realistic. So far, we have only dealt with the case of exact partial information about problem elements $[f; \mathbf{a}]$. But in practice, these evaluations are contaminated by noise. In this paper, we study the complexity of elliptic problems in which we have noisy information about the coefficients of $L_{\mathbf{a}}$ and the function f. How does this change the problem complexity? What algorithms are optimal?

Note that Plaskota's monograph [7] on complexity and noisy information mainly deals with linear problems. Hence, we cannot directly apply the results of [7]. However, it turns out that we can obtain lower bounds by considering only problem elements $[f; \mathbf{a}]$ with fixed $\mathbf{a}$, and then applying the ideas in [7]; we can get upper bounds by using some perturbation arguments, along with the results in [10, pp. 110–111].

We will slightly restrict the generality of the problem in two respects, mainly to simplify the exposition:

(1) We consider only *definite elliptic problems.* These are self-adjoint problems whose variational formulations involve strongly coercive bilinear forms.
(2) We measure error in the norm $\|\cdot\|_{H^m(\Omega)}$, which is equivalent to the problem's natural energy norm.

Information about any particular $[f; \mathbf{a}]$ consists of a finite number of noisy samples. We can calculate approximate values of (some derivative of) either f or a coefficient of $L_{\mathbf{a}}$ at any point in Ω, the error in each approximate value being at most $\delta \geq 0$. In other words, let ρ be a multi-index (which tells us which derivative, possibly the zeroth, to evaluate) and let x be a point in Ω (at which we will evaluate). Rather than having an exact value of $(D^\rho f)(x)$ or of $(D^\rho a)(x)$, with a some coefficient appearing in $L_{\mathbf{a}}$, we have a value y for which $|y - (D^\rho f)(x)| \leq \delta$ or $|y - (D^\rho a)(x)| \leq \delta$, respectively. We assume that the noise level δ of all evaluations is the same. The extension of the results of this paper to include the case where the noise levels of evaluations vary is an open problem.

Let us outline the contents and results of this paper. In Section 2, we give a precise description of the class of problems to be solved, namely $2m$th order elliptic problems over a d-dimensional domain, with problem elements of smoothness r. Next, we describe noisy information for this problem, said information being possibly-adaptive. We define algorithms using said information and the error of such algorithms. Finally, we describe our model of computation, which allows us to define the cost of an algorithm and the complexity of our problem. Note that since we are using noisy information values, the cost $c(\delta)$ of calculating a noisy sample value will depend on δ, see [7, Section 2.9] for further discussion.

In Section 3, we prove a lower bound of $n^{-r/d}+\delta$ for the nth minimal radius of δ-noisy information for this problem. This means that if we want to be able to calculate ε-approximations for arbitrarily small ε, we need to both increase n and decrease the noise level δ. This means that if we cannot decrease the noise level, then there is a cutoff error value ε_0 such that we can only calculate ε-approximations for $\varepsilon \geq \varepsilon_0$.

Once we know a lower bound on the minimal radius, we want to find an algorithm whose error matches this bound. We describe the noisy finite element method with quadrature (FEMQ) in Section 4. Although we allow the evaluation of derivatives of problem elements, the noisy FEMQ only evaluates function values, and not higher-order derivatives. Furthermore, the FEMQ uses nonadaptive information, even though adaptive information is permissible.

In Section 5, we show that the error of the FEMQ using n noisy samples is proportional to $n^{-r/d}+\delta$ when the parameters defining the noisy FEMQ are properly chosen. Thus the noisy FEMQ is a minimal error algorithm, and adaption is no stronger than nonadaption for our problem.

Note that the n-evaluation noisy FEMQ requires the solution of an $n \times n$ linear system $G_{\mathbf{a}}x = b$, where $G_{\mathbf{a}}$ depends on the coefficients $\mathbf{a}$ of the differential operator and b depends on the right-hand side f. If we were only considering a single fixed operator L, then we could precompute the inverse (or LU-decomposition) of $G_{\mathbf{a}}$, since this is independent of any problem element f. We could then ignore the cost of this precomputation, considering it as a fixed overhead, since it would only be done once. However, for the problems studied in this paper, not only do the right-hand sides f vary, but also the operators $L_{\mathbf{a}}$, since we consider arbitrary $[f;\mathbf{a}] \in F$. This means that the factorization of $G_{\mathbf{a}}$ is no longer independent of the problem element considered, and so we cannot ignore its cost. We discuss the efficient implementation of the noisy FEMQ in Section 6. Using a multigrid technique, we can calculate an approximation to the noisy FEMQ solution. This multigrid approximation uses $\Theta(n)$ noisy evaluations and has error proportional to $n^{-r/d}+\delta$. Moreover, we can calculate this approximation using $\Theta(n)$ arithmetic operations, which is optimal.

Finally, in Section 7, we determine the ε-complexity of our problem. Recall that $c(\delta)$ is the cost of calculating a δ-accurate function value. We find that

$$\operatorname{comp}(\varepsilon) = \Theta\left(\inf_{0<\delta<C^{-1}\varepsilon}\left\{c(\delta)\left(\frac{1}{C^{-1}\varepsilon-\delta}\right)^{d/r}\right\}\right)$$

for some constant C. The noisy FEMQ using n evaluations having noise level δ is an optimal algorithm for our problem, with δ minimizing the expression above and

$$n = \left\lceil\left(\frac{1}{C^{-1}\varepsilon-\delta}\right)^{d/r}\right\rceil.$$

As a specific example, suppose that $c(\delta) = \delta^{-s}$, where $s > 0$. We then find that the optimal δ is proportional to ε, and that the complexity is proportional to $(1/\varepsilon)^{d/r+s}$. (The details are in Section 7.) Let us see how much we lose when we go from exact information to noisy information. For exact information, we assume that one function (or derivative) evaluation has cost c. Then the complexity for exact information is proportional to $c(1/\varepsilon)^{d/r}$. For the sake of comparison, let us write the complexity for noisy information as $(1/\varepsilon)^{d/r'}$, where

$$r' = \frac{d}{d+rs}r.$$

Since $r' < r$, we see that the complexity of our problem using noisy information of smoothness r is the same as the complexity using exact information of lesser smoothness r'.

Remark: We previously mentioned that this paper deals with definite elliptic problems and that we only give results for the norm $\|\cdot\|_{H_0^m(\Omega)}$. One can apply the relevant techniques found in [10, Section 5.5] to see that the error estimates of this paper (both lower and upper bounds) also hold for the lower norms $\|\cdot\|_{H^l(\Omega)}$ for $0 \le l \le m$. We will consider extensions to indefinite problems involving weakly coercive forms in a later paper [13]. $\square$

2. Problem Description

In what follows, we assume that the reader is familiar with the usual terminology and notations arising in the variational study of elliptic boundary value problems, such as multi-indices, Sobolev spaces, and the like. See [10, Chapter 5 and Appendix] for further details, as well as the references cited therein. For any ordered ring $\mathscr{X}$, we let $\mathscr{X}^+$ and $\mathscr{X}^{++}$ respectively denote the nonnegative and strictly positive elements of $\mathscr{X}$, this notation being used when $\mathscr{X} = \mathbb{R}$ or $\mathscr{X} = \mathbb{Z}$. The unit ball of the normed linear space X will be denoted by $\mathscr{B}X$. All O-, Ω-, and Θ-relations will be independent of n, δ, and ε.

We are given $p \in [2, \infty]$ and $m \in \mathbb{Z}^+$, as well as $d \in \mathbb{Z}^{++}$ and $r \in \mathbb{R}$ with $r > d/p$. Let $\Omega \subset \mathbb{R}^d$ be a given bounded, simply-connected region with $\partial\Omega \in C^{2m+r}$. For sufficiently smooth $v \colon \overline{\Omega} \to \mathbb{R}$, we define the partial differential operator

$$(L_{\mathbf{a}} v)(x) = \sum_{|\alpha|,|\beta| \le m} (-1)^{|\alpha|} D^\alpha \big(a_{\alpha,\beta}(x) D^\beta v(x)\big) \qquad \forall\, x \in \Omega,$$

Here, $\mathbf{a} = [a_{\alpha,\beta}]_{|\alpha|,|\beta| \le m}$, where the $a_{\alpha,\beta}$ are real-valued functions for $|\alpha|, |\beta| \le m$. We will assume that $a_{\alpha,\beta} = a_{\beta,\alpha}$ for all multi-indices $\alpha, \beta \in (\mathbb{Z}^+)^d$, i.e., the elliptic operator $L_{\mathbf{a}}$ is formally self-adjoint. Associated with the operator $L_{\mathbf{a}}$ is the bilinear form

$$B_{\mathbf{a}}(v, w) = \sum_{|\alpha|,|\beta| \le m} \int_\Omega a_{\alpha,\beta} D^\alpha v D^\beta w$$

on $H_0^m(\Omega)$.

We will be interested in elliptic Dirichlet problems. The *classical formulation* of such a problem is to find, for $f \colon \Omega \to \mathbb{R}$, a function $u \colon \overline{\Omega} \to \mathbb{R}$ such that

$$\begin{aligned} L_{\mathbf{a}} u &= f \quad \text{in } \Omega, \\ \partial_{\mathbf{n}}^j u &= 0 \quad \text{on } \partial\Omega \qquad (0 \le j \le m-1), \end{aligned} \tag{2.1}$$

with $\partial_{\mathbf{n}}^j$ denoting the jth outward-oriented normal derivative. The *variational formulation* is to find, for $f \in W^{r,p}(\Omega)$, an element $u \in H_0^m(\Omega)$ such that

$$B_{\mathbf{a}}(u, v) = \langle f, v \rangle_{L_2(\Omega)} \quad \forall\, v \in H_0^m(\Omega). \tag{2.2}$$

We will let $\mathscr{A}$ denote a class of coefficient vectors, each giving an elliptic problem. More precisely, for given positive γ_0, M, and γ, we will let $\mathscr{A}$ denote the class of all $\mathbf{a}$ such that the following conditions hold:

(1) The operators $L_{\mathbf{a}}$ are strongly elliptic in Ω, i.e.,

$$(-1)^m \sum_{|\alpha|,|\beta| = m} a_{\alpha,\beta}(x) \xi^{\alpha+\beta} \ge \gamma_0 |\xi|^{2m} \qquad \forall\, x \in \Omega, \forall\, \xi \in \mathbb{R}^d, \forall\, \mathbf{a} \in \mathscr{A}.$$

(2) The coefficients of the operators $L_{\mathbf{a}}$ are bounded in the $W^{r,p}(\Omega)$ senes, i.e.,

$$\|a_{\alpha,\beta}\|_{W^{r,\infty}(\Omega)} \le M \qquad \forall\,|\alpha|,|\beta| \le m, \forall\,\mathbf{a} \in \mathscr{A}.$$

(3) The bilinear forms $B_{\mathbf{a}}$ are *uniformly strongly* $H_0^m(\Omega)$*-coercive*, i.e.,

$$B_{\mathbf{a}}(v,v) \ge \gamma\|v\|_{H_0^m(\Omega)} \qquad \forall\, v \in H_0^m(\Omega), \forall\,\mathbf{a} \in \mathscr{A}. \tag{2.3}$$

Roughly speaking, $\mathbf{a} \in \mathscr{A}$ if (2.1) is a self-adjoint elliptic boundary value problem, the only novelty being that we require a "uniformity condition." Note that for the sake of simplicity, we have assumed that the coefficient vector $\mathbf{a}$ and the right-hand side f all have the same smoothness, i.e., the same number r of derivatives (in the Sobolev sense).

Our class of problem elements will be $F = \mathscr{B}W^{r,p}(\Omega) \times \mathscr{A}$. We define a solution operator $S\colon F \to H_0^m(\Omega)$ by letting $u = S([f;\mathbf{a}])$ iff u satisfies (2.2), i.e., u is the variational solution to the Dirichlet problem (2.1). The operator S is nonlinear. However, $S([f;\mathbf{a}])$ depends nonlinearly only on $\mathbf{a}$, i.e., for any fixed $\mathbf{a}$, the operator $S([\cdot;\mathbf{a}])$ is a linear operator. Hence we may use the generalized Lax-Milgram Lemma ([1, pg. 112], [6, pg. 310]) to see that for any $[f;\mathbf{a}] \in F$, there exists a unique solution $u \in H_0^m(\Omega)$ to (2.2). Hence, the solution operator S is well-defined.

We wish to calculate approximate solutions to this problem, using *noisy standard information.* To be specific, we will be using uniformly sup-norm-bounded noise. Our notation and terminology is that of [7] and [8].

Let $\delta \in [0,1]$ be a *noise level*. For $[f;\mathbf{a}] \in F$, we calculate δ*-noisy information*

$$N_\delta([f;\mathbf{a}]) = y = [y_1, \dots, y_{n(y)}], \tag{2.4}$$

about $[f;\mathbf{a}]$, where for each index $i \in \{1,\dots,n(y)\}$, there exist a multi-index $\rho(i)$ and a point $x_i \in \Omega$ such that either

$$|\rho(i)| < r - \frac{d}{p} \qquad \text{and} \qquad |y_i - (D^{\rho(i)}f)(x_i)| \le \delta$$

or, for some multi-indices α and β of order at most m,

$$|\rho(i)| < r \qquad \text{and} \qquad |y_i - (D^{\rho(i)}a_{\alpha,\beta})(x_i)| \le \delta.$$

(The Sobolev embedding theorem guarantees that these derivatives are well-defined.) Note that for any i,

- ▷ whether to terminate at the ith step,
- ▷ the points x_i,
- ▷ the multi-indices $\rho(i)$,
- ▷ the choice of whether to evaluate (a derivative of) the right-hand side f or a coefficient function $a_{\alpha,\beta}$

may all be determined adaptively, depending on the previously-calculated $y_1, \dots, y_{i-1}$.

Let $\mathbb{N}_\delta([f;\mathbf{a}])$ denote the set of all such y, i.e., the set of all such noisy information about $[f;\mathbf{a}]$, and we let $Y = \bigcup_{[f;\mathbf{a}]\in F} \mathbb{N}_\delta([f;\mathbf{a}])$ denote the set of all possible noisy information values. Then an *algorithm* using the noisy information $\mathbb{N}_\delta$ is a mapping $\phi\colon Y \to H_0^m(\Omega)$.

We want to solve this problem in the worst case setting. This means that the *cardinality* of information $\mathbb{N}_\delta$ is given by

$$\operatorname{card} \mathbb{N}_\delta = \sup_{y\in Y} n(y),$$

and the *error* of an algorithm ϕ using $\mathbb{N}_\delta$ is given by

$$e(\phi, \mathbb{N}_\delta) = \sup_{[f;\mathbf{a}]\in F} \sup_{y\in\mathbb{N}_\delta([f;\mathbf{a}])} \|S([f;\mathbf{a}]) - \phi(y)\|_{H^m(\Omega)}.$$

Next, we describe our model of computation. We will use the model found in [7, Section 2.9]. Here are the most important features of this model:

(1) For any multi-index ρ, any point $x \in \Omega$, and any function v defined on Ω, the cost of calculating a δ-noisy value of $(D^\rho v)(x)$ is $c(\delta)$. Here, the cost function $c\colon \mathbb{R}^+ \to \mathbb{R}^+$ is a nonincreasing function, with $c(\delta) > 0$ for sufficiently small positive δ.
(2) Arithmetic operations and comparisons are done exactly, with unit cost.
(3) We are not charged for Boolean operations.
(4) Linear operations over $H_0^m(\Omega)$ are done exactly, with cost $\mathbf{g}$.

For any noisy information $\mathbb{N}_\delta$ and any algorithm ϕ using $\mathbb{N}_\delta$, we shall let $\operatorname{cost}(\phi, \mathbb{N}_\delta)$ denote the worst case cost of calculating $\phi\big(N_\delta([f;\mathbf{a}])\big)$ over all $[f;\mathbf{a}] \in F$.

Now that we have defined the error and cost of an algorithm, we can finally define the complexity of our problem. We shall say that

$$\operatorname{comp}(\varepsilon) = \inf\{ \operatorname{cost}(\phi, \mathbb{N}_\delta) : \mathbb{N}_\delta \text{ and } \phi \text{ such that } e(\phi, \mathbb{N}_\delta) \le \varepsilon \}$$

is the *ε-complexity* of our problem. An algorithm ϕ using noisy information $\mathbb{N}_\delta$ for which

$$e(\phi, \mathbb{N}_\delta) \le \varepsilon \qquad \text{and} \qquad \operatorname{cost}(\phi, \mathbb{N}_\delta) = \Theta\big(\operatorname{comp}(\varepsilon)\big)$$

is said to be an *optimal algorithm.*

3. A Lower Bound on the Minimal Radius

The most commonly-used idea (see, e.g., [9, Section 4.4]) for determining the problem complexity and optimal algorithms is as follows: we first determine the minimal error possible using a given number of evaluations, and then invert this relationship to determine the minimal number of evaluations necessary to achieve a given error. We will use this idea in this paper.

Let $n \in \mathbb{Z}^+$ and $\delta \in [0, 1]$. If $\mathbb{N}_\delta$ is δ-noisy information of cardinality at most n, then

$$r(\mathbb{N}_\delta) = \inf_{\phi \text{ using } \mathbb{N}_\delta} e(\phi, \mathbb{N}_\delta).$$

is the *radius of information*, i.e., the minimal error among all algorithms using given information $\mathbb{N}_\delta$. The nth *minimal radius*

$$r_n(\delta) = \inf\{ r(\mathbb{N}_\delta) : \operatorname{card} \mathbb{N}_\delta \le n \},$$

is the minimal error among all algorithms using noisy information of cardinality at most n. Noisy information $\mathbb{N}_{n,\delta}$ of cardinality n such that

$$r(\mathbb{N}_{n,\delta}) = \Theta\big(r_n(\delta)\big)$$

is said to be nth *optimal information.* An optimal error algorithm using nth optimal information is said to be an nth *minimal error algorithm.*

In this Section, we show that the nth minimal radius of noisy information is bounded from below by $n^{-r/d} + \delta$, i.e., the sum of the nth minimal radius of exact information and the noise level. In the next Section, we show that the finite element method with quadrature (FEMQ) of degree at least r using n noisy evaluations achieves this error, and hence this FEMQ is a minimal error algorithm. In Section 7, we use these results to find the problem complexity and to determine when the FEMQ is an optimal algorithm.

The main result of this section is a lower bound on the nth minimal radius:

THEOREM 3.1. $r_n(\delta) = \Omega(n^{-r/d} + \delta)$.

PROOF: We first claim that

$$r_n(\delta) = \Omega(\delta). \tag{3.1}$$

Indeed, choose an arbitrary, but fixed, element $\mathbf{a}^*$ of $\mathscr{A}$. Let $\mathbb{N}_\delta$ be (possibly-adaptive) noisy information of cardinality at most n. Define a new solution operator $S_{\mathbf{a}^*} : \mathscr{B}W^{r,p}(\Omega) \to H_0^m(\Omega)$ as

$$S_{\mathbf{a}^*}(f) = S([f; \mathbf{a}^*]) \qquad \forall f \in \mathscr{B}W^{r,p}(\Omega).$$

Define information $\overline{\mathbb{N}}_\delta$ for the problem $\big(S_{\mathbf{a}}, \mathscr{B}W^{r,p}(\Omega)\big)$ as follows. For any $f \in \mathscr{B}W^{r,p}(\Omega)$, write

$$\mathbb{N}_\delta([f; \mathbf{a}^*]) = [y_1, \dots, y_l]$$

for some $l \le n$. Each y_i is a noisy evaluation of (a derivative of) either f or of some coefficient $a^*_{\alpha,\beta}$. Let l' be the number of noisy evaluations of f in $\mathbb{N}_\delta([f; \mathbf{a}^*])$. Without loss of generality, suppose that $y_1, \dots, y_{l'}$ are these noisy f-evaluations, i.e.,

$$|y_j - (D^{\rho(i)} f)(x_j)| \le \delta \qquad (1 \le j \le l')$$

for points $x_1, \dots, x_{l'} \in \Omega$ and multi-indices $\rho(1), \dots, \rho(l')$. Then

$$\overline{\mathbb{N}}_\delta(f) = [y_1, \dots, y_{l'}].$$

Extending our notation for radius of information to include the solution operator and problem element class, it is obvious that $\mathscr{B}W^{r,p}(\Omega) \subset F$ implies that

$$r(\mathbb{N}_\delta; S, F) \ge r\big(\overline{\mathbb{N}}_\delta; S_{\mathbf{a}^*}, \mathscr{B}W^{r,p}(\Omega)\big).$$

Since $\overline{\mathbb{N}}_\delta$ is noisy information for a linear problem $(S_{\mathbf{a}^*}, \mathscr{B}W^{r,p}(\Omega))$, there exists nonadaptive information $\overline{\mathbb{N}}_\delta^{\text{non}}$ such that

$$r\big(\overline{\mathbb{N}}_\delta; S_{\mathbf{a}^*}, \mathscr{B}W^{r,p}(\Omega)\big) \ge \tfrac{1}{2} r\big(\overline{\mathbb{N}}_\delta^{\text{non}}; S_{\mathbf{a}^*}, \mathscr{B}W^{r,p}(\Omega)\big),$$

see [7, Chapter 2.7]. It is easy to see that the hypotheses of [7, Lemma 2.8.2] are satisfied, and so

$$r\big(\overline{\mathbb{N}}_\delta^{\text{non}}; S_{\mathbf{a}^*}, \mathscr{B}W^{r,p}(\Omega)\big) = \Omega(\delta),$$

and the desired result (3.1) follows, as claimed.

We next claim that

$$r_n(\delta) = \Omega(n^{-r/d}). \tag{3.2}$$

Indeed, since $r_n(\delta) \ge r_n(0)$, it suffices to show that $r_n(0) = \Omega(n^{-r/d})$. This latter inequality was proved for the case $p = 2$ in [10, pg. 111], the only dependence on the assumption that $p = 2$ being in its use of [10, Theorem 5.5.1]. It is easy to see that the proof of this latter Theorem easily extends to the case of $p \in [2, \infty]$. Hence the desired result (3.2) holds, as claimed.

Our Theorem now follows immediately from (3.1) and (3.2). □

4. The Noisy FEMQ

In this section, we define the noisy finite element method with quadrature (FEMQ). This is an algorithm using standard information consisting only of function evaluations, i.e., no derivative evaluations are used. Our notation is the standard one found in, e.g., [4] and [10, Chapter 5].

The easiest way to describe the noisy FEMQ is by following three steps. First, we describe the noise-free "pure" finite element method (FEM), which uses non-standard information. Next, we describe the noise-free FEMQ, which uses exact standard information. Finally, we describe the noisy FEMQ.

Before describing each of these FEMs, we first establish some notation. Let $\hat{K}$ be a fixed polyhedron in $\mathbb{R}^d$. We call $\hat{K}$ a *reference element*. We next let K be a (small) *finite element*, i.e., the affine image of $\hat{K}$ under a bijection F_K, where

$$F_K(\hat{x}) = B_K\hat{x} + b_K \qquad \forall\, \hat{x} \in \hat{K}, \tag{4.1}$$

where $B_K \in \mathbb{R}^{d\times d}$ is invertible and $b_K \in \mathbb{R}^d$. Next, we let $\mathscr{T}$ be a triangulation of Ω consisting of finite elements, where each $K \in \mathscr{T}$ is the image of the reference element $\hat{K}$ under the affine bijection F_K. Select a fixed value of $k \in \mathbb{Z}^{++}$, and let $P_k(K)$ denote the space of polynomials having total degree at most k, considered as functions over K. Given this triangulation $\mathscr{T}$, we define a *finite element space*

$$\mathscr{S}(\mathscr{T}) = \left\{\, s \in H_0^m(\Omega) : s\big|_K \in P_k(K)\ \forall\, K \in \mathscr{T} \,\right\}$$

of *degree* k. We will assume that the following conditions hold:

(1) $\{\mathscr{T}_n\}_{n=1}^{\infty}$ is a family of triangulations of Ω such that $\mathscr{S}_n = \mathscr{S}(\mathscr{T}_n)$ is a finite element space of dimension n.

(2) $\{\mathscr{T}_n\}_{n=1}^{\infty}$ is a *quasi-uniform* family of triangulations, i.e.,

$$\limsup_{n\to\infty} \sup_{K\in\mathscr{T}_n} \frac{h_K}{\rho_K} < \infty,$$

where h_K is the diameter of K and ρ_K is the diameter of the largest sphere contained in K.

(3) Let $\|\cdot\|$ denote the ℓ_2 matrix norm on $\mathbb{R}^d$. Then $\|B_K\| \le 1$ for any element $K \in \mathscr{T}_n$ and any triangulation $\mathscr{T}_n$.

We first recall how the noise-free "pure" FEM is defined. Let $n \in \mathbb{Z}^+$, and let $\{s_1, \dots, s_n\}$ be a basis for $\mathscr{S}_n$. For $[f; \mathbf{a}] \in F$, find

$$u_n = \sum_{j=1}^{n} \alpha_j s_j,$$

in $\mathscr{S}_n$ such that

$$B_{\mathbf{a}}(u_n, s_i) = \langle f, s_i\rangle_{L_2(\Omega)} \qquad (1 \le i \le n). \tag{4.2}$$

Note that the coefficient vector

$$a = [\alpha_1, \dots, \alpha_n]^T$$

satisfies

$$Ga = b,$$

where

$$G = [B_{\mathbf{a}}(s_j, s_i)]_{1\le i,j\le n}$$

and

$$b = [\langle f, s_1 \rangle_{L_2(\Omega)}, \dots, \langle f, s_n \rangle_{L_2(\Omega)}]^T.$$

Since the bilinear forms $B_{\mathbf{a}}$ are uniformly strongly coercive, it follows that for any $n \in \mathbb{Z}^{++}$ and any $[f; \mathbf{a}] \in F$, there exists a unique $u_n \in \mathscr{S}_n$ satisfying (4.2). Hence, the pure FEM is well-defined.

Of course, if we want to calculate u_n, we will need to calculate the inner products appearing in the matrix G and the vector b, which means that we have to calculate the various integrals

$$\int_\Omega a_{\alpha,\beta} D^\alpha s_j D^\beta s_i \qquad (1 \le i, j \le n \text{ and } |\alpha|, |\beta| \le m)$$

and

$$\int_\Omega f s_i \qquad (1 \le i \le n).$$

Since only standard information is available to us, we cannot calculate these integrals for arbitrary $[f; \mathbf{a}] \in F$. Instead, we shall use numerical quadrature to approximate these integrals, which gives us the (noise-free) FEMQ.

The quadrature rule used to define the FEMQ is initially defined on the reference element. This reference quadrature rule has the form

$$\hat{I}\hat{v} = \sum_{j=1}^{J} \hat{\omega}_j \hat{v}(\hat{b}_j)$$

for functions $\hat{v}$ defined on $\hat{K}$. This rule is said to be *exact* of degree q if

$$\int_{\hat{K}} v = \hat{I}\hat{v} \qquad \forall\, \hat{v} \in P_q(\hat{K}).$$

We define a local quadrature rule over a particular finite element K as

$$I_K v = \sum_{j=1}^{J} \omega_{j,K}\, v(b_{j,K}),$$

where

$$\omega_{j,K} = \det B_K \cdot \hat{\omega}_j \quad \text{and} \quad b_{j,K} = F_K(\hat{b}_j) \qquad (1 \le j \le J) \tag{4.3}$$

for $K = F_K(\hat{K})$, with F_K given by (4.1). Next, for any $\ell \in \mathbb{Z}^+$, we let

$$\mathscr{N}_\ell = \bigcup_{K \in \mathscr{T}_\ell} \bigcup_{j=1}^{J} \{b_{j,K}\}$$

denote the set of all quadrature nodes in all the elements belonging to $\mathscr{T}_\ell$. This is usually *not* a disjoint union, since a quadrature node on the boundary of one element will be on the boundary of an adjacent element sharing a common face.

We can now define the noise-free FEMQ. Let

$$\kappa = \binom{m+d}{d}^2 \tag{4.4}$$

denote the maximum number of coefficients that can appear in a $2m$th order elliptic operator defined on a d-dimensional domain. Given $n \in \mathbb{Z}^+$, we define

$$\tilde{n} = \max\{ \operatorname{card} \mathscr{N}_\ell : \ell \in \mathbb{Z}^+ \text{ and } (\kappa+1) \operatorname{card} \mathscr{N}_\ell \le n \}. \tag{4.5}$$

Roughly speaking, $\tilde{n} = \lfloor n/(\kappa+1) \rfloor$, allowing for the fact that $\tilde{n}$ must be the cardinality of the set $\mathscr{N}_\ell$ of quadrature nodes for some triangulation $\mathscr{T}_\ell$. Let $\{s_1, \dots, s_{\tilde{n}}\}$ denote a basis for the finite element space $\mathscr{S}_{\tilde{n}}$. For $[f; \mathbf{a}] \in F$, we define a new bilinear form $B_{\mathbf{a},\tilde{n}}$ on $\mathscr{S}_{\tilde{n}}$ by

$$\left.\begin{aligned} B_{\mathbf{a},\tilde{n}}(v,w) &= \sum_{|\alpha|,|\beta|\le m} \sum_{K \in \mathscr{T}_{\tilde{n}}} I_K(a_{\alpha,\beta} D^\alpha v D^\beta w) \\ &= \sum_{|\alpha|,|\beta|\le m} \sum_{K \in \mathscr{T}_{\tilde{n}}} \sum_{j=1}^{J} \omega_{j,K} \cdot a_{\alpha,\beta}(b_{j,K}) \cdot (D^\alpha v)(b_{j,K}) \cdot (D^\beta w)(b_{j,K}) \end{aligned}\right\} \quad \forall\, v, w \in \mathscr{S}_{\tilde{n}}.$$

and a linear functional $f_{\tilde{n}}$ on $\mathscr{S}_{\tilde{n}}$ by

$$f_{\tilde{n}}(v) = \sum_{K \in \mathscr{T}_{\tilde{n}}} I_K(fv) = \sum_{K \in \mathscr{T}_{\tilde{n}}} \sum_{j=1}^{J} \omega_{j,K} \cdot f(b_{j,K}) \cdot v(b_{j,K}) \qquad \forall\, v \in \mathscr{S}_{\tilde{n}}.$$

Then we seek

$$u_{\tilde{n}}^{\mathrm{Q}} = \sum_{j=1}^{\tilde{n}} \alpha_j s_j,$$

such that

$$B_{\mathbf{a},\tilde{n}}(u_{\tilde{n}}^{\mathrm{Q}}, s_i) = f_{\tilde{n}}(s_i) \qquad (1 \le i \le n). \tag{4.6}$$

The new coefficient vector

$$a = [\alpha_1, \dots, \alpha_n]^T$$

satisfies

$$Ga = b,$$

where now

$$G = [B_{\mathbf{a},\tilde{n}}(s_i, s_j)]_{1 \le i,j \le \tilde{n}}$$

and

$$b = [f_{\tilde{n}}(s_1), \dots, f_{\tilde{n}}(s_{\tilde{n}})]^T.$$

Note that since $r > d/p$, the entries in the matrix G and the coefficient vector b are well-defined.

Let

$$\nu = \min\{k+1, r\}.$$

In the remainder of this paper, we shall assume that the following conditions hold:

(1) The smoothness r of the problem elements F satisfies $r \ge 1$ (as well as our previous requirement $r > d/p$).
(2) The degree k of the finite element subspaces $\mathscr{S}_{\tilde{n}}$ satisfies $k > d/p - 1$.
(3) $\hat{I}$ is exact of degree $2k + \nu - 1$ over the reference element $\hat{K}$.

Let us write

$$N_n([f;\mathbf{a}]) = [N_n(f), N_n(\mathbf{a})]$$

where[1]

$$N_n(f) = \{\, f(b_{j,K}) : 1 \le j \le J \text{ and } K \in \mathscr{T}_{\tilde{n}} \,\}.$$

and

$$N_n(\mathbf{a}) = \{\, a_{\alpha,\beta}(b_{j,K}) : 1 \le j \le J \text{ and } K \in \mathscr{T}_{\tilde{n}} \text{ and } |\alpha|, |\beta| \le m \,\}$$

We see that $u_{\tilde{n}}^{\mathrm{Q}}$ depends on $[f;\mathbf{a}]$ only through $N_n([f;\mathbf{a}])$, and so we write $u_{\tilde{n}}^{\mathrm{Q}} = \phi_n(N_n([f;\mathbf{a}]))$, with ϕ_n an algorithm using N_n, which is exact standard information of cardinality at most n.

We are finally ready to define the noisy FEMQ. Given $n \in \mathbb{Z}^+$, we once again choose the largest $\tilde{n} \in \mathbb{Z}^+$ satisfying (4.5), and a basis $\{s_1, \dots, s_{\tilde{n}}\}$ for the finite element space $\mathscr{S}_{\tilde{n}}$. We now calculate a noisy version of $N_n([f;\mathbf{a}])$. That is, for each element $K \in \mathscr{T}_{\tilde{n}}$, each index $j \in \{1, \dots, J\}$, and each pair of multi-indices (α, β) with $|\alpha| \le m$ and $|\beta| \le m$, we obtain real numbers $\tilde{a}_{\alpha,\beta,j,K,\delta}$ and $\tilde{f}_{j,K,\delta}$ satisfying

$$|\tilde{a}_{\alpha,\beta,j,K,\delta} - a_{\alpha,\beta}(b_{j,K})| \le \delta \tag{4.7}$$

and

$$|\tilde{f}_{j,K,\delta} - f(b_{j,K})| \le \delta. \tag{4.8}$$

Let $\tilde{\mathbb{N}}_{n,\delta}$ denote this noisy version of N_n, i.e.,

$$\tilde{\mathbb{N}}_{n,\delta}([f;\mathbf{a}]) = \{\tilde{\mathbb{N}}_{n,\delta}(f), \tilde{\mathbb{N}}_{n,\delta}(\mathbf{a})\}$$

where[2]

$$\tilde{\mathbb{N}}_{n,\delta}(f) = \{\tilde{f}_{j,K,\delta} \text{ satisfying (4.8)} : 1 \le j \le J \text{ and } K \in \mathscr{T}_{\tilde{n}} \,\}.$$

and

$$\tilde{\mathbb{N}}_{n,\delta}(\mathbf{a}) = \{\tilde{a}_{\alpha,\beta,j,K,\delta} \text{ satisfying (4.7)} : 1 \le j \le J \text{ and } K \in \mathscr{T}_{\tilde{n}} \text{ and } |\alpha|, |\beta| \le m \,\}.$$

Clearly, $\tilde{\mathbb{N}}_{n,\delta}$ is noisy information of cardinality at most n. For $[f;\mathbf{a}] \in F$, we define a new bilinear form $\tilde{B}_{\mathbf{a},\tilde{n},\delta}$ on $\mathscr{S}_{\tilde{n}}$ by

$$\tilde{B}_{\mathbf{a},\tilde{n},\delta}(v,w) = \sum_{|\alpha|,|\beta| \le m} \sum_{K \in \mathscr{T}_{\tilde{n}}} \sum_{j=1}^{J} \omega_{j,K} \cdot \tilde{a}_{\alpha,\beta,j,K,\delta} \cdot (D^\alpha v)(b_{j,K}) \cdot (D^\beta w)(b_{j,K}) \qquad \forall\, v, w \in \mathscr{S}_{\tilde{n}}.$$

and a linear functional $\tilde{f}_{\tilde{n},\delta}$ on $\mathscr{S}_{\tilde{n}}$ by

$$\tilde{f}_{\tilde{n},\delta}(v) = \sum_{K \in \mathscr{T}_{\tilde{n}}} \sum_{j=1}^{J} \omega_{j,K} \cdot \tilde{f}_{j,K,\delta} \cdot v(b_{j,K}) \qquad \forall\, v \in \mathscr{S}_{\tilde{n}}.$$

Then we seek

$$\tilde{u}_{\tilde{n}}^{\mathrm{Q}} = \sum_{j=1}^{\tilde{n}} \alpha_j s_j$$

[1]We really should use lists of elements, set out in a specified order, for $N_n(f)$ and $N_n(\mathbf{a})$, so that $N_n([f;\mathbf{a}])$ will be a vector. The reader will indulge this slight abuse of notation, since any precisely-correct alternative would be far more long-winded.

[2]This is also a slight abuse of notation.

such that

$$\tilde{B}_{\mathbf{a},\tilde{n},\delta}(u^{\mathrm{Q}}_{\tilde{n}}, s_i) = \tilde{f}_{\tilde{n},\delta}(s_i) \qquad (1 \le i \le n). \tag{4.9}$$

The new coefficient vector

$$a = [\alpha_1, \dots, \alpha_n]^T$$

satisfies

$$Ga = b,$$

where now

$$G = [\tilde{B}_{\mathbf{a},\tilde{n},\delta}(s_i, s_j)]_{1\le i,j\le \tilde{n}}$$

and

$$b = [\tilde{f}_{\tilde{n},\delta}(s_1), \dots, \tilde{f}_{\tilde{n},\delta}(s_{\tilde{n}})]^T.$$

We see that $\tilde{u}^{\mathrm{Q}}_{\tilde{n}}$ depends on $[f;\mathbf{a}]$ only through $\tilde{\mathbb{N}}_{n,\delta}([f;\mathbf{a}])$, and so we write $\tilde{u}^{\mathrm{Q}}_{\tilde{n}} = \tilde{\phi}_{n,\delta}(\tilde{\mathbb{N}}_{n,\delta}([f;\mathbf{a}]))$, with $\tilde{\phi}_{n,\delta}$ an algorithm using our noisy standard information $\tilde{\mathbb{N}}_{n,\delta}$.

Remark: Recall that we have stated that the solution operator S, the pure FEM, and the noiseless FEMQ are all well-defined. We have not stated such a result for the noisy FEMQ. We will prove that the noisy FEMQ is well-defined in the next section. □

5. The Noisy FEMQ is a Minimal Error Algorithm

In this section, we prove that the noisy FEMQ is well-defined, and that it is a minimal error algorithm. In particular, we give conditions on the degree k of the finite element space which are guarantee that the FEMQ using n noisy evaluations with a noise level of δ has error proportional to $n^{-r/d} + \delta$.

Our starting point is *Strang's Lemma* (see [10, pp. 310–312] for a proof of a version having slightly more restrictive hypotheses). Recall that the bilinear forms $B_{\mathbf{a}}$ are uniformly strongly coercive, with constant γ; see (2.3).

Lemma 5.1. *Suppose that there exists $\delta_0 \in (0,1]$ and $n^* \in \mathbb{Z}^{++}$ such that for any $\delta \in [0,\delta_0]$, any $n \ge n^*$ and any $\mathbf{a} \in \mathscr{A}$, we have*

$$|B_{\mathbf{a}}(v,w) - \tilde{B}_{\mathbf{a},\tilde{n},\delta}(v,w)| \le \tfrac{1}{2}\gamma \|v\|_{H^m(\Omega)} \|w\|_{H^m(\Omega)} \qquad \forall\, v,w \in \mathscr{S}_{\tilde{n}}. \tag{5.1}$$

Then for any $n \ge n^*$, any $\delta \in [0,\delta_0]$, and any $[f;\mathbf{a}] \in F$, there is a unique $\tilde{u}^{\mathrm{Q}}_{\tilde{n}} \in \mathscr{S}_{\tilde{n}}$ such that (4.9) holds. Moreover, there exists a positive constant C, such that if $u = S([f;\mathbf{a}])$ is the solution to (2.2), then

$$\|u - \tilde{u}^{\mathrm{Q}}_{\tilde{n}}\|_{H^m(\Omega)} \le C \inf_{v\in\mathscr{S}_{\tilde{n}}} \left[\|u - v\|_{H^m(\Omega)} + \sup_{w\in\mathscr{S}_{\tilde{n}}} \left(\frac{|B_{\mathbf{a}}(v,w) - \tilde{B}_{\mathbf{a},\tilde{n},\delta}(v,w)|}{\|w\|_{H^m(\Omega)}} + \frac{|f(w) - \tilde{f}_{\tilde{n},\delta}(w)|}{\|w\|_{H^m(\Omega)}} \right) \right],$$

the constant C being independent of n, δ, and $[f;\mathbf{a}]$. □

Before we can use Strang's Lemma, we need to prove some preliminary estimates. In what follows, we use the standard notational technique of letting C denote a generic constant whose value may change from one place to another.

Lemma 5.2. *There exists a positive constant C such that*

$$|B_{\mathbf{a},\tilde{n}}(v,w) - \tilde{B}_{\mathbf{a},\tilde{n},\delta}(v,w)| \le C\delta \|v\|_{H^m(\Omega)} \|w\|_{H^m(\Omega)} \qquad \forall\, v,w \in \mathscr{S}_{\tilde{n}}$$

and

$$|f_{\tilde{n}}(v) - \tilde{f}_{\tilde{n},\delta}(v)| \le C\delta \|v\|_{L_2(\Omega)} \qquad \forall\, v \in \mathscr{S}_{\tilde{n}},$$

for any $[f;\mathbf{a}] \in F$ and any $n \in \mathbb{Z}^+$, with $\tilde{n} = \tilde{n}(n)$ satisfying (4.5).

PROOF: Let $[f;\mathbf{a}] \in F$, and $n \in \mathbb{Z}^+$. We establish the first inequality. For any $v, w \in \mathscr{S}_{\tilde{n}}$, we have

$$\begin{aligned}
&|B_{\mathbf{a},\tilde{n}}(v,w) - \tilde{B}_{\mathbf{a},\tilde{n},\delta}(v,w)| \\
&\qquad = \left| \sum_{K\in\mathscr{T}_{\tilde{n}}} \sum_{|\alpha|,|\beta|\le m} \sum_{j=1}^{J} \omega_{j,K}[a_{\alpha,\beta}(b_{j,K}) - \tilde{a}_{\alpha,\beta,j,K,\delta}](D^\alpha v)(b_{j,K})(D^\beta w)(b_{j,K}) \right| \\
&\qquad \le \delta \sum_{K\in\mathscr{T}_{\tilde{n}}} \sum_{|\alpha|,|\beta|\le m} \sum_{j=1}^{J} |\omega_{j,K}(D^\alpha v)(b_{j,K})(D^\beta w)(b_{j,K})|.
\end{aligned} \tag{5.2}$$

Consider a particular element $K \in \mathscr{T}_{\tilde{n}}$, as well as particular multi-indices α and β. Using (4.3), we have

$$\sum_{j=1}^{J} |\omega_{j,K}(D^\alpha v)(b_{j,K})(D^\beta w)(b_{j,K})| = |\det B_K| \cdot \sum_{j=1}^{J} |\hat{\omega}_j(D^\alpha v)(b_{j,K})(D^\beta w)(b_{j,K})|. \tag{5.3}$$

Let D^l denote the Frechet derivative, where $l = |\alpha|$. As on [4, pg. 118], there exists a subset $\{e_{\alpha_1}, \dots, e_{\alpha_l}\}$ of the standard basis for $\mathbb{R}^d$ such that for any $x \in K$, we have

$$(D^\alpha v)(x) = (D^l v)(x)(e_{\alpha_1}, \dots, e_{\alpha_l}) = (D^l \hat{v})(\hat{x})(B_K^{-1} e_{\alpha_1}, \dots, B_K^{-1} e_{\alpha_l}),$$

with $\hat{x} = F_K^{-1}(x)$. Letting

$$\|(D^l v)(x)\| = \sup\{\, |(D^l v)(x)(\xi_1, \dots, \xi_l)| : \xi_1, \dots, \xi_l \text{ in the Euclidean unit ball of } \mathbb{R}^d \,\}$$

we have

$$\|(D^\alpha v)(x)| \le \|B_K\|^{-l} \|(D^l \hat{v})(\hat{x})\| \le C\|B_K\|^{-l} \sup_{|\alpha|=l} \|(D^\alpha \hat{v})(\hat{x})\|$$

for some constant C, independent of K and v. Since $l \le m$ and $\|B_K\| \le 1$, we see that

$$\|(D^\alpha v)(x)| \le C\|B_K\|^{-m} \sup_{|\alpha|\le m} \|(D^\alpha \hat{v})(\hat{x})\|.$$

Using this inequality, along with the analogous inequality

$$\|(D^\beta w)(x)\| \le C\|B_K\|^{-m} \sup_{|\beta|\le m} \|(D^\beta \hat{w})(\hat{x})\|,$$

in (5.3), and then summing over the multi-indices α and β for which $|\alpha| \le m$ and $|\beta| \le m$, we see that

$$\begin{aligned}
&\sum_{|\alpha|,|\beta|\le m} \sum_{j=1}^{J} |\omega_{j,K}(D^\alpha v)(b_{j,K})(D^\beta w)(b_{j,K})| \\
&\qquad\qquad \le C\|B_K\|^{-2m} |\det B_K| \cdot \sum_{j=1}^{J} |\hat{\omega}_j| \sup_{|\alpha|\le m} |(D^\alpha \hat{v})(\hat{b}_j)| \sup_{|\beta|\le m} |(D^\beta \hat{w})(\hat{b}_j)|.
\end{aligned}$$

Now

$$
\begin{aligned}
\sum_{j=1}^{J} |\hat\omega_j| \sup_{|\alpha|\le m} |(D^\alpha \hat v)(\hat b_j)| \sup_{|\beta|\le m} |(D^\beta \hat w)(\hat b_j)| \\
&\le \left[\sum_{j=1}^{J} |\hat\omega_j| \sup_{|\alpha|\le m} |(D^\alpha \hat v)(\hat b_j)|^2\right]^{1/2} \left[\sum_{j=1}^{J} |\hat\omega_j| \sup_{|\beta|\le m} |(D^\beta \hat w)(\hat b_j)|^2\right]^{1/2} \\
&\le C\|\hat v\|_{H^m(\hat K)} \|\hat w\|_{H^m(\hat K)} \\
&\le C\|B_K\|^{2m} |\det B_K|^{-1} \|v\|_{H^m(K)} \|w\|_{H^m(K)},
\end{aligned}
$$

where we have used [4, Theorem 3.1.2] in the last inequality above. Hence,

$$
\sum_{|\alpha|,|\beta|\le m} \sum_{j=1}^{J} |\omega_{j,K} (D^\alpha v)(b_{j,K}) (D^\beta w)(b_{j,K})| \le C\|v\|_{H^m(K)} \|w\|_{H^m(K)}.
$$

Substituting this inequality into (5.2), we find

$$
\begin{aligned}
|B_{\mathbf{a},\tilde n}(v,w) - \tilde B_{\mathbf{a},\tilde n,\delta}(v,w)| &\le C\delta \sum_{K\in\mathscr{T}_{\tilde n}} \|v\|_{H^m(K)} \|w\|_{H^m(K)} \\
&\le C\delta \left[\sum_{K\in\mathscr{T}_{\tilde n}} \|v\|^2_{H^m(K)}\right]^{1/2} \left[\sum_{K\in\mathscr{T}_{\tilde n}} \|w\|^2_{H^m(K)}\right]^{1/2} \\
&= C\delta \|v\|_{H^m(\Omega)} \|w\|_{H^m(\Omega)},
\end{aligned}
$$

as required.

Next, we establish the second inequality. For any $v \in \mathscr{S}_{\tilde n}$, we have

$$
|f_{\tilde n}(v) - \tilde f_{\tilde n,\delta}(v)| = \left| \sum_{K\in\mathscr{T}_{\tilde n}} \sum_{j=1}^{J} \omega_{j,K} [f(b_{j,K}) - \tilde f_{j,K,\delta}] v(b_{j,K}) \right| \le \delta \sum_{K\in\mathscr{T}_{\tilde n}} \sum_{j=1}^{J} |\omega_{j,K}\, v(b_{j,K})|. \tag{5.4}
$$

Let $K \in \mathscr{T}_{\tilde n}$. Using (4.3), we have

$$
\sum_{j=1}^{J} |\omega_{j,K}\, v(b_{j,K})| = |\det B_K| \sum_{j=1}^{J} |\hat\omega_j \hat v(\hat b_j)|. \tag{5.5}
$$

Now $\hat v \mapsto \sum_{j=1}^{J} |\hat\omega_j \hat v(\hat b_j)|$ is a linear functional on the finite-dimensional space $P_k(\hat K)$, and is thus a bounded linear functional, with respect to any norm on $P_k(\hat K)$. Hence there is a constant C, independent of $\hat v$, such that

$$
\sum_{j=1}^{J} |\hat\omega_j \hat v(\hat b_j)| \le C\|\hat v\|_{L_2(\hat K)} \qquad \forall\, \hat v \in P_k(\hat K).
$$

Applying this result to (5.5), using [4, Theorem 3.1.2] to estimate $\|\hat v\|_{L_2(\hat K)}$ in terms of $\|v\|_{L_2(K)}$, and using the quasi-uniformity of the sequence of triangulations, we see that there exists a constant C, independent of v, K, and n, such that

$$
\sum_{j=1}^{J} |\omega_{j,K}\, v(b_{j,K})| \le C|\det B_K|\, \|\hat v\|_{L_2(\hat K)} \le C|\det B_K|^{1/2} \|v\|_{L_2(K)} \le C\tilde n^{-1/2} \|v\|_{L_2(K)}.
$$

Substituting this inequality into (5.4), we find that there exist constants C such that

$$\begin{aligned}|f_{\tilde{n}}(v) - \tilde{f}_{\tilde{n},\delta}(v)| &\le C\delta\tilde{n}^{-1/2} \sum_{K\in\mathscr{T}_{\tilde{n}}} \|v\|_{L_2(K)} \\ &\le C\delta\tilde{n}^{-1/2}\,\mathrm{card}(\mathscr{T}_{\tilde{n}})^{1/2}\left[\sum_{K\in\mathscr{T}_{\tilde{n}}} \|v\|^2_{L_2(K)}\right]^{1/2} \\ &\le C\delta\|v\|_{L_2(\Omega)},\end{aligned}$$

as required. □

We are now ready to prove the main result of this section.

THEOREM 5.1. *There exist $n^* \in \mathbb{Z}^{++}$ and $\delta_0 > 0$ such that $\tilde{\phi}_{n,\delta}$ is well-defined for all $n \ge n^*$ and all $\delta \in [0, \delta_0]$. Furthermore,*

$$e(\tilde{\phi}_{n,\delta}, \tilde{\mathbb{N}}_{n,\delta}) = O(n^{-\mu/d} + \delta),$$

where

$$\mu = \min\{k, r\}.$$

PROOF: We first show that $\tilde{\phi}_{n,\delta}$ is well-defined. As in [10, pg. 106], we see that there exists a positive constant C such that

$$\begin{aligned}|B_{\mathbf{a}}(v,w) - B_{\mathbf{a},\tilde{n}}(v,w)| &\le C\left[\sum_{|\alpha|,|\beta|\le m} \|a_{\alpha,\beta}\|_{W^{r,p}(\Omega)}\right] n^{-\nu/d}\|v\|_{H^m(\Omega)}\|w\|_{H^m(\Omega)} \\ &\le C\kappa M n^{-\nu/d}\|v\|_{H^m(\Omega)}\|w\|_{H^m(\Omega)}\end{aligned} \qquad \forall\, v,w \in \mathscr{S}_{\tilde{n}}$$

for any $n \in \mathbb{Z}^{++}$. (Recall that κ is given by (4.4) and that M is given by condition (2) defining $\mathscr{A}$.) Using the first inequality in Lemma 5.2, we have

$$|B_{\mathbf{a}}(v,w) - \tilde{B}_{\mathbf{a},\tilde{n},\delta}(v,w)| \le C(\kappa M n^{-\nu/d} + \delta)\|v\|_{H^m(\Omega)}\|w\|_{H^m(\Omega)} \qquad \forall\, v,w \in \mathscr{S}_{\tilde{n}} \tag{5.6}$$

for any $n \in \mathbb{Z}^{++}$ and any $\delta \in [0, 1]$. It now follows that there exists $\delta_0 \in (0, 1]$ and $n^* \in \mathbb{Z}^{++}$ such that (5.1) holds for any $\delta \in [0, \delta_0]$, any $n \ge n^*$ and any $\mathbf{a} \in \mathscr{A}$. Hence, Strang's Lemma implies that if $\delta \in [0, \delta_0]$ and $n \ge n^*$, then for any $[f; \mathbf{a}] \in F$, there is a unique $\tilde{u}^{\mathrm{Q}}_{\tilde{n}} \in \mathscr{S}_{\tilde{n}}$ such that (4.9) holds. Thus the noisy FEMQ $\tilde{\phi}_{n,\delta}$ is well-defined for any such δ and n.

Before we bound the error of the noisy FEMQ, we first note that by the conditions defining $\mathscr{A}$, the so-called "shift theorem" for elliptic problems holds for a constant that is independent of $\mathbf{a} \in \mathscr{A}$. That is, if $f \in H^r(\Omega)$, then for any $\mathbf{a} \in \mathscr{A}$, we have $S([f; \mathbf{a}]) \in H^{r+2}(\Omega)$. Moreover,

$$\sigma^{-1}\|S([f;\mathbf{a}])\|_{H^{r+2}(\Omega)} \le \|f\|_{H^r(\Omega)} \le \sigma\|S([f;\mathbf{a}])\|_{H^{r+2}(\Omega)}, \tag{5.7}$$

where the constant σ is independent of $\mathbf{a} \in \mathscr{A}$, depending only on m, M, and r. See, for instance, the proof in [5], noting that the shift constant depends mainly on the geometry of the region Ω and the size of the coefficients in the partial differential operator $L_{\mathbf{a}}$.

We now turn to the error of the noisy FEMQ. Let $\delta \in [0, \delta_0]$ and $n \ge n^*$. For $[f; \mathbf{a}] \in F$, let $u = S([f; \mathbf{a}])$. From [10, pg. 107], there exists $v \in \mathscr{S}_{\tilde{n}}$ such that

$$\|u - v\|_{H^m(\Omega)} \le Cn^{-\mu/d}\|u\|_{H^{r+2m}\Omega)}. \tag{5.8}$$

Using (5.7), we find that

$$\|u\|_{H^{r+2m}(\Omega)} \le \sigma \|f\|_{H^r(\Omega)}. \tag{5.9}$$

Since $p \ge 2$, there exists a positive constant C, independent of f, such that

$$\|f\|_{H^r(\Omega)} \le C\|f\|_{W^{r,p}(\Omega)} \le C, \tag{5.10}$$

since $f \in \mathscr{B}W^{r,p}(\Omega)$. Combining (5.8)–(5.10), we find that

$$\|u - v\|_{H^m(\Omega)} \le Cn^{-\mu/d}\|f\|_{H^r(\Omega)} \le Cn^{-\mu/d}. \tag{5.11}$$

Now for any $w \in \mathscr{S}_{\tilde{n}}$, we find from [10, pg. 106] that

$$|f(w) - f_{\tilde{n}}(w)| \le Cn^{-\nu/d}\|f\|_{H^r(\Omega)}\|w\|_{H^m(\Omega)} \le Cn^{-\nu/d}\|w\|_{H^m(\Omega)},$$

where we have again used (5.10). Using this inequality and the second inequality in Lemma 5.2, we have

$$|f(x) - \tilde{f}_{\tilde{n},\delta}(w)| \le C(n^{-\nu/d} + \delta)\|w\|_{H^m(\Omega)}. \qquad (5.12).$$

Use (5.6), (5.12), and (5.11) in Strang's Lemma. Since $\mu \le \nu$, we find

$$\|u - \tilde{u}^{\mathrm{Q}}_{\tilde{n}}\|_{H^m(\Omega)} \le C(n^{-\mu/d} + \delta),$$

as required. □

Remark: Theorem 5.1 gives an upper bound on the error of the noisy FEMQ. This upper bound is sharp for the case $p = 2$, i.e.,

$$e(\tilde{\phi}_{n,\delta}, \tilde{\mathbb{N}}_{n,\delta}) = \Theta(n^{-\mu/d} + \delta) \qquad \text{for } p = 2.$$

Indeed, clearly (3.1) implies that

$$e(\tilde{\phi}_{n,\delta}, \tilde{\mathbb{N}}_{n,\delta}) \ge r_n(\delta) = \Omega(\delta).$$

On the other hand, the exact FEMQ is an instance of a noisy FEMQ, and so

$$e(\tilde{\phi}_{n,\delta}, \tilde{\mathbb{N}}_{n,\delta}) \ge e(\phi_n, N_n).$$

But for $p = 2$, we have

$$e(\phi_n, N_n) = \Omega(n^{-\mu/d}),$$

see [10, pg. 106]. Combining these last three inequalities, we get

$$e(\tilde{\phi}_{n,\delta}, \tilde{\mathbb{N}}_{n,\delta}) = \Omega(n^{-\mu/d} + \delta),$$

the desired lower bound matching the upper bound in Theorem 5.1, when $p = 2$. □

Combining Theorems 3.1 and 5.1, we find

Corollary 5.1.

(1) $r_n(\delta) = \Theta(n^{-r/d} + \delta)$.
(2) *The noisy FEMQ, using a quadrature rule that is exact of degree at least $2k + r - 1$, is a minimal error algorithm if $k \ge r$.*
(3) *Adaption is no stronger than non-adaption.* □

6. Multigrid Implementation of the Noisy FEMQ

As we mentioned in the Introduction, both the matrix G and the vector b in the linear system $Ga = b$ characterizing the noisy FEMQ depend on the problem element $[f; \mathbf{a}] \in F$. This means that the standard technique of ignoring the cost of reducing G to a form more suitable for solving linear systems cannot be ignored, as we often do when said matrix does not depend on any particular problem element. Hence we need to find an efficient implementation of the noisy FEMQ.

One idea is to use a multigrid technique. The main ideas underlying multigrid methods are as follows:

(1) We do not need an exact solution of the linear system $Ga = b$, but only one whose error is comparable to the error of the noisy FEMQ.
(2) We can use an iteration for solving the linear system. Moreover:
 (a) A sufficiently-accurate solution corresponding to the coarser grid is a good initial guess for the solution corresponding to the finer grid.
 (b) The iteration on the finer grid has the effect of smoothing, i.e., damping out the oscillatory part of the error, so that this smoothed solution is well approximated on the coarser grid.

Our presentation (and analysis) of the multigrid technique will be based on that in [3, Chapter 6], which covers only the definite problems.

We first establish notation. Recall that $\{\mathscr{T}_n\}_{n=1}^{\infty}$ is a quasi-uniform grid sequence. Let us write

$$h_j = \max_{K \in \mathscr{T}_j} h_K$$

for the meshsize of $\mathscr{T}_j$. Recall (from Theorem 5.1) that the noisy FEMQ $\tilde{\phi}_{n,\delta}$ is well-defined if $n \geq n^*$. Let

$$n_1 = n^* < n_2 < \cdots < n_{l-1} < n_l$$

be a sequence of integers, chosen so that

$$\mathscr{T}_{n_{j-1}} \supset \mathscr{T}_{n_j}, \text{ and thus } \mathscr{S}_{n_{j-1}} \subset \mathscr{S}_{n_j}$$

and

$$h_{n_j} \sim \tfrac{1}{2} h_{n_{j-1}} \qquad (2 \leq j \leq l). \tag{6.1}$$

We let j be fixed, but arbitrary, index in $\{1, \ldots, l\}$. If $p_1, \ldots, p_{n_j}$ are the interior nodes of the triangulation $\mathscr{T}_{n_j}$, then we get the standard finite element basis $\{s_1, \ldots, s_{n_j}\}$ for $\mathscr{S}_{n_j}$ by requiring that $s_i(p_{i'}) = \delta_{i,i'}$ for $1 \leq i, i' \leq n_j$ (see, e.g., the discussion in [10, Sections 5.7 and A.2.3]).

We define a mesh-dependent inner product $\langle \cdot, \cdot \rangle_j$ on $\mathscr{S}_{n_j}$ by

$$\langle v, w \rangle_j = h_{n_j}^d \sum_{i=1}^{n_j} v(p_i) w(p_i) \qquad \forall\, v, w \in \mathscr{S}_{n_j}.$$

Then the operator A_j on $\mathscr{S}_{n_j}$ is defined by

$$\langle A_j v, w \rangle_j = \tilde{B}_{\mathbf{a}, n_j, \delta}(v, w) \qquad \forall\, v, w \in \mathscr{S}_{n_j}.$$

Note that we may follow the proof of [3, Lemma 6.2.8] to find an upper bound

$$\rho(A_j) \leq \Lambda_j = C h_{n_j}^{-2m} \tag{6.2}$$

on the spectral radius of A_j, where the constant C is independent of the index j and the coefficient vector $\mathbf{a}$.

Let us define $f_j \in \mathscr{S}_{n_j}$ by requiring that

$$\langle f_j, s\rangle_j = \tilde{f}_{n_j}(s) \qquad \forall\, s \in \mathscr{S}_{n_j}$$

and let us write $\tilde{u}_j$ for the solution $\tilde{u}_j = \tilde{u}^{\mathrm{Q}}_{n_j}$ of the noisy FEMQ for $\mathscr{S}_{n_j}$, so that

$$A_j \tilde{u}_j = f_j.$$

We then let $I^j_{j-1} \colon \mathscr{S}_{n_{j-1}} \to \mathscr{S}_{n_j}$ be the natural embedding, and let $I^{j-1}_j \colon \mathscr{S}_{n_j} \to \mathscr{S}_{n_{j-1}}$ be its adjoint, i.e.,

$$\langle I^{j-1}_j w, v\rangle_{j-1} = \langle w, I^j_{j-1} v\rangle_j = \langle w, v\rangle_j \qquad \forall\, v \in \mathscr{S}_{n_{j-1}}, w \in \mathscr{S}_{n_j}.$$

Recalling that Λ_j is an upper bound on $\rho(A_j)$, we now define the *jth-level multigrid iteration* recursively, in terms of the multigrid iterations at lower levels:

function $\mathrm{MG}(j\colon \mathbb{Z}^+;\ z_0, g\colon \mathscr{S}_{n_j})\colon \mathscr{S}_{n_j}$;
begin
  if $k = 1$ **then**
    $\mathrm{MG} := A_1^{-1} g$
  else
    begin
      $z_1 := z_0 + \Lambda_j^{-1}(g - A_j z_0)$; { pre-smoothing }
      $\overline{g} := I^{j-1}_j (g - A_j z_1)$; { fine-to-coarse intergrid transfer }
      $q_1 := \mathrm{MG}(j-1, 0, \overline{g})$; { error correcting }
      $z_2 := z_1 + I^j_{j-1} q_1$; { coarse-to-fine intergrid transfer }
      $z_3 := z_2 + \Lambda_j^{-1}(g - A_j z_2)$; { post-smoothing }
    end;
  $\mathrm{MG} := z_3$
end

Then for any index t, the t-fold *full multigrid scheme* produces an approximation $\hat{u}_j$ to $\tilde{u}_j$ as follows:

function $\mathrm{FMG}(j, t\colon \mathbb{Z}^+)\colon \mathscr{S}_{n_j}$;
begin
  if $j = 1$ **then**
    $\hat{u}_j := A_1^{-1} f_1$
  else
    begin
      $u^j_0 := I^j_{j-1} \hat{u}_{j-1}$;
      for $i := 1$ **to** t **do**
        $u^j_i := \mathrm{MG}(j, u^j_{i-1}, f_j)$;
      $\hat{u}_j := u^j_t$
    end;
  $\mathrm{FMG} := \hat{u}_j$
end

Let

$$\overline{\mathbb{N}}_{n,\delta} = [\tilde{\mathbb{N}}_{n_1,\delta}, \tilde{\mathbb{N}}_{n_2,\delta}, \ldots, \tilde{\mathbb{N}}_{n_l,\delta}],$$

with l the maximal index for which $\operatorname{card} \overline{\mathbb{N}}_{n,\delta} \le n$. Then we may write

$$\hat{u}_l = \overline{\phi}_{n,\delta}\big(\overline{\mathbb{N}}_{n,\delta}([f;\mathbf{a}])\big),$$

where $\overline{\phi}_{n,\delta}$ is the *full multigrid algorithm.*

The main result for this section is

THEOREM 6.1.

(1) *The full multigrid algorithm is well-defined.*
(2) *There exists an index t such that the error of the full multigrid algorithm is*

$$e(\overline{\phi}_{n,\delta}, \overline{\mathbb{N}}_{n,\delta}) = O(n^{-\mu/d} + \delta),$$

where (as in Theorem 5.1)

$$\mu = \min\{k, r\}.$$

(3) *The combinatory cost of the full multigrid scheme* $\mathrm{FMG}(l,t)$ *is* $\Theta(n)$.

PROOF: The well-definedness follows from Theorem 5.1. To prove the desired error estimate, let us first consider the jth-level multigrid iteration. Let $\|\cdot\|_{E_j}$ be the energy norm defined by

$$\|v\|_{E_j} = \tilde{B}_{\mathbf{a},n_j,\delta}(v,v)^{1/2},$$

this energy norm being equivalent to the usual $H_0^m(\Omega)$-norm. We claim that there exists a constant C^* such that

$$\|z - \mathrm{MG}(j, z_0, g)\|_{E_j} \le \frac{C^*}{C^*+1}\|z - z_0\|_{E_j}, \tag{6.3}$$

the constant C^* being independent of $g, z, z_0 \in \mathscr{S}_{n_j}$, $j \in \mathbb{Z}^+$, and $[f;\mathbf{a}] \in F$. (There is a "1" in the denominator because we do one pre-smoothing and one post-smoothing step at each level.) Indeed, we only need to (carefully) check that the proof of the analogous result [3, Proposition 6.6.12] applies in our case, once we have made the following changes:

(1) Instead of using [3, Lemma 6.2.8], we use our estimate (6.2) for the spectral radius of A_j.
(2) For any $s \ge 0$, let

$$|||v|||_{s,j} = \langle A_j^s v, v\rangle_j \qquad \forall\, v \in \mathscr{S}_{n_j}.$$

Let $P_j : H_0^m(\Omega) \to \mathscr{S}_{n_j}$ be the orthogonal projection operator with respect to the inner product $\tilde{B}_{\mathbf{a},n_j,\delta}$, i.e., for any $v \in H_0^m(\Omega)$, the element $P_j v \in \mathscr{S}_{n_j}$ satisfies

$$\tilde{B}_{\mathbf{a},n_j,\delta}(P_j v, w) = \tilde{B}_{\mathbf{a},n_j,\delta}(v,w) \qquad \forall\, w \in \mathscr{S}_{n_j}.$$

Then instead of using the approximation property in [3, Corollary 6.4.4], we use the analogous result that there exists a positive constant C such that

$$|||(I - P_{j-1})v|||_{1,j} \le C h_{n_j}^m |||v|||_{2,j} \qquad \forall\, v \in \mathscr{S}_{n_j}.$$

We now consider the error of the full multigrid method, following the proof of [3, Theorem 6.7.1], with a few modifications. Let

$$\theta = \frac{C^*}{C^* + 1}.$$

Choose $[f; \mathbf{a}] \in F$, and let $u = S([f; \mathbf{a}])$. For any j, let $\hat{e}_j = \tilde{u}_j - \hat{u}_j$, noting that $\hat{e}_1 = 0$. Using (6.3) and the definition of FMG, we see that

$$\|\hat{e}_j\|_{E_j} \le \theta^t \|\tilde{u}_j - \hat{u}_{j-1}\|_{E_{j-1}}.$$

Thus, there exist positive constants C such that

$$\|\hat{e}_j\|_{H^m(\Omega)} \le C\theta^t \|\tilde{u}_j - \hat{u}_{j-1}\|_{H^m(\Omega)} \le C\theta^t (\|u - \tilde{u}_{j-1}\|_{H^m(\Omega)} + \|u - \tilde{u}_j\|_{H^m(\Omega)} + \|\hat{e}_{j-1}\|_{H^m(\Omega)}).$$

From Theorem 5.1, we have

$$\begin{aligned} \|u - \tilde{u}_j\|_{H^m(\Omega)} &\le C(h_{n_j}^\mu + \delta) \\ \|u - \tilde{u}_{j-1}\|_{H^m(\Omega)} &\le C(h_{n_{j-1}}^\mu + \delta). \end{aligned}$$

Using (6.1), it follows that

$$\|\hat{e}_j\|_{H^m(\Omega)} \le C\theta^t [(h_{n_j}^\mu + \delta) + \|\hat{e}_{j-1}\|_{H^m(\Omega)}].$$

Solving this inequality, we find that there exist constants C such that

$$\|\hat{e}_j\|_{H^m(\Omega)} \le C \sum_{i=0}^{j-1} (h_{n_{j-i}}^\mu + \delta)(C\theta^t)^i \le C\left[h_{n_j}^\mu \sum_{i=0}^{j-1} (2C\theta^t)^i + \delta \sum_{i=0}^{j-1} (C\theta^t)^i\right],$$

where we have again used (6.1). So if

$$t > \frac{\ln 2C}{\ln 1/\theta},$$

then we find that

$$\|\hat{e}_j\|_{H^m(\Omega)} \le \frac{C\theta^t}{1 - 2C\theta^t} h_{n_j}^\mu + \frac{C\theta^t}{1 - C\theta^t}\delta = O(h_j^\mu + \delta).$$

Hence

$$\begin{aligned} \|S([f; \mathbf{a}]) - \overline{\phi}_{n,\delta}\big(\overline{\mathbb{N}}_{n,\delta}([f; \mathbf{a}])\big)\|_{H^m(\Omega)} &= \|u - \hat{u}_j\|_{H^m(\Omega)} \le \|u - \tilde{u}_j\|_{H^m(\Omega)} + \|\hat{e}_j\|_{H^m(\Omega)} \\ &= O(h_{n_j}^\mu + \delta), \end{aligned}$$

establishing the desired error bound for the full multigrid algorithm.

We now estimate the cost of calculating $\hat{u}_j$, using ideas similar to those in the proof of [3, Proposition 6.7.4]. First, let W_j denote the amount of work in the jth-level scheme. We find

$$W_j \le 2Cn_j + W_{j-1}$$

for some constant C, so that

$$W_j \le 2C(n_j + n_{j-1} + \cdots + n_1).$$

Using (6.1),

$$n_j = \dim \mathscr{S}_{n_j} = \Theta(h_{n_j}^{-d}) = \Theta\big((\tfrac{1}{2}h_{n_{j-1}})^{-d}\big) = \Theta(2^d n_{j-1}),$$

and so

$$W_j = O\bigg(\sum_{i=0}^{j-1} 2^{-di}\bigg) n_j \le C n_j$$

for some constant C. Finally, let $\hat{W}_j$ denote the work done by $\mathrm{FMG}(j,t)$. We find that

$$\hat{W}_j \le \hat{W}_{j-1} + rW_j \le \hat{W}_{j-1} + tCn_j.$$

Hence

$$\hat{W}_j \le tC(n_j + n_{j-1} + \cdots + n_1) \le Cn_j$$

for some constant C. In particular

$$\hat{W}_l = O(n_l) = O(n).$$

Since $\hat{W}_l$ is the combinatory cost of the full multigrid scheme $\mathrm{FMG}(l,t)$, this completes the proof of the Theorem. □

7. Complexity

In this Section, we determine the complexity of the noisy elliptic problem.

It will be useful to explicitly specify some of the order-of-magnitude constants in some of the estimates in the previous sections. Thus, Theorem 3.1 tells us that there exists a positive constant C_1 such that

$$r_n(\delta) \ge C_1(n^{-r/d} + \delta). \tag{7.1}$$

Moreover, let $\tilde{\phi}_{n,\delta}$ be the noisy FEMQ of degree $k \ge r$, using a quadrature rule that is exact of degree at least $2k+r-1$. Then by Theorem 6.1, there exist positive constants C_2 and $C_3 = C_3(\mathbf{g})$ such that

$$e(\overline{\phi}_{n,\delta}, \overline{\mathbb{N}}_{n,\delta}) \le C_2(n^{-r/d} + \delta) \tag{7.2}$$

and

$$\mathrm{cost}(\overline{\phi}_{n,\delta}, \overline{\mathbb{N}}_{n,\delta}) \le C_3 c(\delta) n. \tag{7.3}$$

We now have

Theorem 7.1. *The problem complexity is bounded from below by*

$$\mathrm{comp}(\varepsilon) \ge \inf_{\delta>0} \left\{ c(\delta) \left\lceil \left(\frac{1}{C_1^{-1}\varepsilon - \delta} \right)^{d/r} \right\rceil \right\}, \tag{7.4}$$

and from above by

$$\mathrm{comp}(\varepsilon) \le C_3 \inf_{\delta>0} \left\{ c(\delta) \left\lceil \left(\frac{1}{C_2^{-1}\varepsilon - \delta} \right)^{d/r} \right\rceil \right\}. \tag{7.5}$$

The upper bound is attained by using the noisy FEMQ $\overline{\phi}_{n,\delta}$ described above, with

$$n = \left\lceil \left(\frac{1}{C_2^{-1}\varepsilon - \delta} \right)^{d/r} \right\rceil, \tag{7.6}$$

and with δ chosen minimizing (7.5).

PROOF: To prove (7.4), suppose that ϕ is an algorithm using noisy information $\mathbb{N}_\delta$ such that $e(\phi, \mathbb{N}_\delta) \le \varepsilon$. Then $\operatorname{card} \mathbb{N}_\delta \ge n$, where n must be large enough to make $r_n(\delta) \le \varepsilon$. The lower bound (7.1) immediately tells us that

$$n \ge \left\lceil \left(\frac{1}{C_1^{-1}\varepsilon - \delta} \right)^{d/r} \right\rceil .$$

But the cost of any algorithm using n information evaluations must be at least $n\, c(\delta)$, and so

$$\operatorname{cost}(\phi, \mathbb{N}_\delta) \ge c(\delta) \left\lceil \left(\frac{1}{C_1^{-1}\varepsilon - \delta} \right)^{d/r} \right\rceil .$$

Since ϕ and $\mathbb{N}_\delta$ are an arbitrary algorithm and noisy information such that $e(\phi, \mathbb{N}_\delta) \le \varepsilon$, we find that

$$\operatorname{comp}(\varepsilon) \ge c(\delta) \left\lceil \left(\frac{1}{C_1^{-1}\varepsilon - \delta} \right)^{d/r} \right\rceil .$$

Finally, since $\delta > 0$ is arbitrary, we get the desired lower bound (7.4).

To prove the remainder of this Theorem, let $\delta > 0$. If (7.6) holds, then we may use (7.2) to see that $e(\tilde{\phi}_{n,\delta}, \tilde{\mathbb{N}}_{n,\delta}) \le \varepsilon$. Now using (7.3), we have

$$\operatorname{cost}(\overline{\phi}_{n,\delta}, \overline{\mathbb{N}}_{n,\delta}) \le C_3\, c(\delta) \left\lceil \left(\frac{1}{C_2^{-1}\varepsilon - \delta} \right)^{d/r} \right\rceil .$$

Choosing δ minimizing the right-hand side in this inequality, the desired result follows. $\square$

Comparing the lower and upper bounds in Theorem 7.1, we see that

$$\operatorname{comp}(\varepsilon) = \Theta \left(\inf_{\delta > 0} \left\{ c(\delta) \left(\frac{1}{C^{-1}\varepsilon - \delta} \right)^{d/r} \right\} \right), \tag{7.7}$$

for some constant C, which allows us to determine the complexity for various cost functions $c(\cdot)$. For instance, if c is differentiable, then (7.7) holds if δ satisfies

$$\frac{d/r}{C^{-1}\varepsilon - \delta} = \frac{c'(\delta)}{c(\delta)} .$$

As a specific example, consider the cost function $c(\delta) = c_s(\delta) = \delta^{-s}$, where $s > 0$. After some calculations, we find that for $\varepsilon > 0$, the optimal δ is

$$\delta^* = \frac{rs\varepsilon}{C(rs + d)},$$

so that

$$\operatorname{comp}(\varepsilon) = \operatorname{comp}_s(\varepsilon) = \Theta \left(\left(\frac{d}{sr} \right)^{s} \left(\frac{C(rs + d)}{d\varepsilon} \right)^{d/r + s} \right). \tag{7.8}$$

Simplifying a bit, we see that the optimal δ is proportional to ε, and that

$$\operatorname{comp}(\varepsilon) = \Theta \left(\left(\frac{1}{\varepsilon} \right)^{d/r + s} \right).$$

Recall that the complexity of this problem using exact information is

$$\operatorname{comp}^{\text{exact}}(\varepsilon) = \Theta\left(\left(\frac{1}{\varepsilon}\right)^{d/r}\right),$$

see [10, Section 5.5]. Let us compare the results for noisy and exact information.

First, note that $\lim_{s\to 0} c_s(\delta) = 1$, i.e., the cost of obtaining δ-accurate samples becomes a constant, independent of δ, when s tends to zero. Using (7.8), we see that $\lim_{s\to 0} \operatorname{comp}_s(\varepsilon) = \Theta\big(\operatorname{comp}^{\text{exact}}(\varepsilon)\big)$. Thus as the (varying) cost of noisy information approaches the (fixed) cost of exact information, the problem complexity for noisy information approaches that for exact information.

Moreover, we can determine the penalty that must be paid when noisy information is used for the elliptic problem, instead of exact information. As mentioned in the Introduction to this paper, one way of measuring this penalty is to write

$$\operatorname{comp}(\varepsilon) = \Theta\left(\left(\frac{1}{\varepsilon}\right)^{d/r'}\right),$$

where

$$r' = \frac{d}{d+rs} r.$$

Hence, the complexity of our problem using noisy information of smoothness r is the same as the complexity using exact information of lesser smoothness r'.

Bibliography

1. Babuška, I. and Aziz, A. K., *Survey lectures on the mathematical foundations of the finite element method*, The Mathematical Foundations of the Finite Element Method With Applications to Partial Differential Equations, (A. K. Aziz, ed.), Academic Press, New York, 1972, pp. 3–359.
2. Bramble, J., *Multigrid Methods*, Pitman Research Notes in Mathematics, Vol. 294, Wiley, New York, 1993.
3. Brenner, S. C. and Scott, L. R., *The Mathematical Theory of Finite Element Methods*, Springer, New York, 1994.
4. Ciarlet, P. G., *The Finite Element Method For Elliptic Problems*, North-Holland, Amsterdam, 1978.
5. Friedman, A., *Partial Differential Equations*, Krieger, Malabar, FL, 1976.
6. Oden, J. T. and Carey, G. F., *Finite elements: Mathematical aspects*, Prentice-Hall, Englewood Cliffs, 1983.
7. Plaskota, L., *Noisy Information and Computational Complexity*, Cambridge University Press, Cambridge, 1996.
8. Plaskota, L., *Worst case complexity of problems with random information noise*, J. Complexity (1996) (to appear).
9. Traub, J. F., Wasilkowski, G. W., and Woźniakowski, H., *Information-Based Complexity*, Academic Press, New York, 1988.
10. Werschulz, A. G., *The Computational Complexity of Differential and Integral Equations: An Information-Based Approach*, Oxford University Press, Oxford, 1991.
11. ———, *The complexity of multivariate elliptic problems with analytic data*, J. Complexity **11** (1995), 154–173.
12. ———, *The complexity of the Poisson problem for spaces of bounded mixed derivatives*, Lectures in Applied Mathematics: Proceedings of the Joint AMS-IMS-SIAM Summer Seminar on "Mathematics of Numerical Analysis: Real Number Algorithms," Park City, UT, July 17–August 11, 1995 (J. Renegar, M. Shub, S. Smale, eds.), 1996 (to appear).
13. ———, *The complexity of indefinite elliptic problems with noisy data*, J. Complexity (to appear).