COMPLEXITY OF VERIFICATION AND COMPUTATION FOR IBC PROBLEMS

H. Woźniakowski

Department of Computer Science, Columbia University
and
Institute of Informatics, University of Warsaw

CUCS-031-92

December, 1990

Abstract. We analyze the complexity of *verifying* whether a given element is within ε of the solution element. This may be contrasted with the complexity of *computing* an element that is within ε of the solution element.

For discrete problems with $\varepsilon = 0$ verification is no harder than computation in any setting. For IBC problems verification can be easier or harder than computation.

We will show that the worst case complexity of verification for IBC problems is often infinite. We therefore switch to the probabilistic case and study the probabilistic complexity of verification as a function of the error tolerance ε and the probability of failure δ. We assume that the solution element is specified by a linear continuous functional defined on a Banach space equipped with a Gaussian measure. For fixed δ and small ε, the complexity of verification is zero, whereas for fixed ε and small δ the complexity of verification is essentially a function of only δ and maybe exponentially harder than the complexity of computation.

1. Introduction

Information-based complexity, IBC for brevity, studies the minimal cost of computing an ε-approximation to the solution element in different settings, see e.g., TWW [88].[1] By

[1] By TWW [88] we mean Traub, Wasilkowski and Woźniakowski [88].

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

the setting we mean how *error* and *cost* are defined. The ε-*complexity* is then defined as the minimal cost of computing a solution element with error at most ε. The ε-complexity usually goes to infinity as ε goes to zero. We call such a problem the *computational problem* of IBC, and denote its ε-complexity by $\mathrm{comp}^{\mathrm{com}}(\varepsilon)$.

We deal with a different problem in this paper. Instead of computing an ε-approximation to the solution element, we wish to verify whether a given element is within ε of the solution element. We call such a problem the *verification problem* of IBC. By the ε-*complexity* of the verification problem, $\mathrm{comp}^{\mathrm{ver}}(\varepsilon)$, we mean the minimal cost needed to solve the verification problem in a given setting.

We now specify the computational and verification problems in more technical terms. The solution elements are defined by $S(f)$, where S is a given operator whose range is a normed linear space, and f belongs to a given set F. We assume that information about f is *partial* and provided by computing information operations from a class Λ.

For the computational problem we combine information operations from Λ to compute an element that is within ε of $S(f)$.

For the verification problem we wish to verify whether $\|g - S(f)\|$ is at most ε. Here g is from a given set G and f from F. That is, we wish to compute the *verification* operator $\mathrm{VER}(f,g)$ which is defined as $\mathrm{VER}(f,g) = \mathrm{YES}$ if $\|g - S(f)\| \le \varepsilon$, and $\mathrm{VER}(f,g) = \mathrm{NO}$ otherwise. We assume that the element g is known and the information about f is gathered as before by computing information operations from Λ.

Suppose first that $\varepsilon = 0$, i.e., we wish to verify whether $g = S(f)$. Obviously, if the solution elements can be computed with finite cost, then with one extra comparison we can solve the verification problem. In this case, the verification problem is no harder than the computation problem in any setting; see Garey and Johnson [79] where this point is explored for many combinatorial and discrete problems and serves as a basis for studying NP-problems.

But, as mentioned before, for a typical computational problem of IBC, the solution elements cannot be computed exactly with finite cost, and therefore to make complexity finite the error tolerance ε must be positive. In this paper we seek relations between the complexities of the verification and computational problems for positive ε.

We study this question for linear problems in the worst case and probabilistic settings. By a *linear* problem, see TWW [88, Chapter 4], we mean that S is a linear operator, $F = \{f : \|Tf\| < 1\}$ for a linear operator T, and Λ is a set of linear functionals. We also assume that G is an arbitrary nonempty subset of a normed linear space.

We first discuss the worst case setting in which we wish to compute the verification operator $\mathrm{VER}(f,g)$ for all considered f and g. Before we present a general result we present a simple example of a linear problem which should indicate why we may encounter troubles with the verification problem in the worst case setting.

Example 1.1

Suppose we are asked to verify whether a real number g is an ε-approximation of the integral of a function f, i.e., to verify whether

$$\left| g - \int_0^1 f(t)\,dt \right| \le \varepsilon. \tag{1.1}$$

Here g and a positive ε are known, and $f : [0,1] \to \mathbb{R}$ is a Lipschitz function with constant less than 1. For this example, G is an arbitrary nonempty subset of $\mathbb{R}$ and F is the class of Lipschitz functions with constant less than 1. Since Lipschitz functions are differentiable almost everywhere, the class F can be expressed as

$$F = \{ f : [0,1] \to \mathbb{R} : \ \|Tf\|_\infty < 1 \} \quad \text{with } Tf = f'$$

and $S(f) = \int_0^1 f(t)\,dt$.

We assume that the function f is not known but we can compute the values of f at any point from $[0,1]$, and we are charged a fixed positive cost c for each function evaluation. That is, Λ is now a set of linear functionals L of the form $L(f) = f(t)$, $\forall\, f \in F$ for some $t \in [0,1]$. We also assume that we can perform arithmetic operations and comparison of real numbers at unit cost.

We wish to verify inequality (1.1) for fixed ε and for any number g from G and any Lipschitz function with constant 1. Unfortunately, no matter whether ε is small or large, this cannot be done with finite cost. To see this, take any g from G and suppose that we compute function values at finitely many points and that the computed values are all equal to $g + \varepsilon$. Knowing these function values we have at least two different candidates for the function f. The first is $f(x) = f_1(x) \equiv g + \varepsilon$. Note that inequality (1.1) is true for f_1. The second candidate is to take $f(x) = f_2(x)$, where f_2 is a seesaw function between the computed points. The integral of f_2 is greater than $g + \varepsilon$, and therefore inequality (1.1) is false for f_2. Since our verification concerning inequality (1.1) is based on the function values computed, and because f_1 and f_2 have the same values, our verification must be wrong for either f_1 or f_2.

Thus, this verification problem cannot be solved in the worst case setting, and we have

$$\mathrm{comp}^{\mathrm{ver}}(\varepsilon) \;=\; +\infty \quad \forall\, \varepsilon \ge 0.$$

This should be compared to the ε-complexity of the computational problem which is

$$\mathrm{comp}^{\mathrm{com}}(\varepsilon) \;=\; \frac{c+a}{\min\{4\varepsilon, 1\}},$$

where $a \in [0,2]$, see e.g., Traub and Woźniakowski [80, Chapter 6].

In particular, if $\varepsilon \geq \frac{1}{4}$, then the computational problem is trivial; still the verification problem cannot be solved. $\square$

We now discuss general linear problems in the worst case setting. First of all observe that if the sets F and G are such that for any g from G we have $\mathrm{VER}(F,g) = \{\mathrm{YES}\}$ or $\mathrm{VER}(F,g) = \{\mathrm{NO}\}$ then obviously the verification problem is trivial and we can easily solve it.

Assume therefore that $\mathrm{VER}(F,g) = \{\mathrm{YES}, \mathrm{NO}\}$ at least for one element g from G. Then the negative result presented in Example 1.1 extends to *all* linear problems which cannot be computed exactly by a finite number of information operations provided that the range of the operator S has a strictly convex norm. That is, we have

$$\mathrm{comp}^{\mathrm{ver}}(\varepsilon) \;=\; +\infty \quad \forall\, \varepsilon \geq 0.$$

Thus, no matter if ε is large or small, the verification problem cannot be solved in the worst case setting. On the other hand, the ε-complexity of the computational problem can be essentially any nonincreasing function of ε, see Traub and Woźniakowski [80, Chapter 5]. In Section 3 we show that the assumptions that the linear problem cannot be computed exactly with a finite cost and that the norm of G is strictly convex are essential.

To make the worst case complexity finite we must relax the definition of the verification problem. This can be done by not requiring the computation of $\mathrm{VER}(f,g)$ whenever $\varepsilon < ||S(f) - g|| \leq (1+\alpha)\varepsilon$ for some positive α. Then the worst case complexity of the relaxed verification problem is at most equal to $\mathrm{comp}^{\mathrm{com}}(\varepsilon\,\alpha/2)$. This bound is sharp for some problems whereas for some other problems the dependence on α can be relaxed. For details, see Novak and Woźniakowski [91].

We now turn to the probabilistic setting. We assume that S is a linear continuous functional, and F is a separable Banach space equipped with a Gaussian measure μ whose mean is zero. Take the set Λ as a subset of the class of linear continuous functionals. Obviously, to make the verification problem interesting we must assume that $S \notin \Lambda$. We assume now that G is an arbitrary nonempty subset of $\mathbb{R}$.

In the probabilistic setting, we wish to compute the verification operator $\mathrm{VER}(\cdot, g)$ for each g from G modulo a set of elements f of measure at most δ. To stress the dependence on the parameter δ as well as the probabilistic setting, we denote the probabilistic complexity of the verification problem by $\mathrm{comp}^{\text{prob-ver}}(\varepsilon, \delta)$. Similarly, $\mathrm{comp}^{\text{prob-com}}(\varepsilon, \delta)$ will denote the probabilistic complexity of the computational problem in which the solution elements are not approximated to within ε on a set of measure at most δ.

To make both complexities finite for positive ε and δ we must assume that S can be approximated with arbitrarily small error (in norm generated by the Gaussian measure μ) by a linear combination of finitely many functionals from Λ. Equivalently, we assume that the minimal variance σ_n^* of the Gaussian measure of the solution elements which can be achieved after n information operations from Λ goes to zero as n approaches infinity.

Then the complexities of the computational and verification problems are finite, and the relations between the two complexities depend on the relations between ε and δ. We now state the results.

It is known, see TWW [88, Chapter 8], that the probabilistic complexity of the computational problem is given by

$$\text{comp}^{\text{prob-com}}(\varepsilon,\delta) \;=\; (c+a)\,\min\{n :\, \sqrt{\sigma_n^*} \le \varepsilon/\psi^{-1}(1-\delta)\}, \tag{1.2}$$

where $a \in [0,2]$ and ψ is the probability integral, $\psi(x) = \sqrt{2/\pi}\int_0^x e^{-t^2/2}dt$. For small δ we have $\psi^{-1}(1-\delta) \simeq \sqrt{2\ln 1/\delta}$.

For the verification problem, the probabilistic complexity depends on the relation between ε and δ. Assume first that

$$\psi\left(\frac{\varepsilon}{\sqrt{\sigma_0}}\right) \;\le\; \delta, \tag{1.3}$$

where $\sigma_0 \;=\; \int_F S^2(f)\mu(df)$ is the variance of the Gaussian measure μS^{-1}. Observe that (1.3) means that ε is small relative to δ. In this case the verification problem is trivial since

$$\text{comp}^{\text{prob-ver}}(\varepsilon,\delta) \;=\; 0. \tag{1.4}$$

To understand (1.4), note that

$$\mu\{f :\, |g - S(f)| \le \varepsilon\} \le \mu\{f :\, |S(f)| \le \varepsilon\} = \psi(\varepsilon/\sqrt{\sigma_0}) \le \delta,$$

and we can approximate the verification operator VER by NO with zero cost modulo a set of f of measure at most δ. In particular, we have

$$\text{comp}^{\text{prob-ver}}(0,\delta) \;=\; 0 \quad \forall\,\delta \ge 0.$$

Thus, as long as (1.3) holds the verification problem is trivial whereas the computational problem has probabilistic complexity which can increase quickly to infinity as ε goes to zero. This shows that the role of the error tolerance ε in the computational problem is quite different than in the verification problem.

Assume now that

$$\psi\left(\frac{\varepsilon}{\sqrt{\sigma_0}}\right) > \delta, \tag{1.5}$$

that is, δ is small relative to ε. Then, in particular, we prove that for $G = \mathbb{R}$ and δ going to zero, we have

$$\mathrm{comp}^{\text{prob-ver}}(\varepsilon,\delta) \;=\; (c+a)\,\min\left\{n:\ \sqrt{\frac{\sigma_n^*}{\sigma_0}} \;\le\; \frac{\pi}{a^*}\,\delta\,(1+o(1))\right\} \;+\; b, \tag{1.6}$$

where $a \in [0,2]$, $b \in [0,5]$ and $a^* \in [1,2]$.

Both complexities of the computational and verification problems are given by similar formulas, but the term

$$\varepsilon/\psi^{-1}(1-\delta) \simeq \varepsilon/\sqrt{2\ln 1/\delta}$$

in the complexity of the computational problem is replaced by $\sqrt{\sigma_0}\,\pi/a^*\,\delta$ in the complexity of the verification problem. We have

$$\mathrm{comp}^{\text{prob-ver}}(\varepsilon,\delta) \;=\; \mathrm{comp}^{\text{prob-com}}\left(\varepsilon, e^{-(\alpha\varepsilon/\delta)^2}\,(1+o(1))\right)$$

with $\alpha = a^*/(2\pi\sqrt{\sigma_0})$.

For $\delta = \Theta(\varepsilon)$ with ε tending to zero, and $\delta < \psi(\varepsilon/\sqrt{\sigma_0})$, we show that

$$\mathrm{comp}^{\text{prob-ver}}(\varepsilon,\delta) \;=\; (c+a)\,\min\left\{n:\ \sqrt{\frac{\sigma_n^*}{\sigma_0}} \;\le\; \Theta(\delta)\right\} \;+\; b,$$

with $a \in [0,2]$ and $b \in [0,5]$.

We thus see that for δ small relative to ε the verification problem corresponds to the computational problem with an exponentionally smaller δ. If, for example, $\sigma_n^* = \Theta(n^{-p})$ for a positive p then the verification problem is exponentionally harder than the computational problem.

We mention several open questions concerning the verification problem. It would be interesting to generalize the results of this paper for linear operators S in the probabilistic case, and for general operators S in the worst case setting.

Finally, we summarize the contents of this paper. In Section 2 we define the verification problem precisely. Section 3 deals with the worst case setting, where we prove that the verification problem cannot be solved for many linear problems. Section 4 deals with the probabilistic setting. We first analyze nonadaptive information and find sharp estimates on the probabilistic failure. We then show that adaptive information does not help and derive complexity bounds. The analysis of this section is illustrated by two integration examples. The first one is for scalar periodic smooth functions, and the second one for multivariate continuous functions.

2. Verification Problem

For a given set F and a given normed linear space G_1, consider an arbitrary operator

$$S : F \to G_1.$$

Let G be a nonempty subset of G_1.

By the *verification problem* we mean to compute the *verification operator* $\mathrm{VER} : F \times G \to \{\mathrm{YES}, \mathrm{NO}\}$ defined as

$$\mathrm{VER}(f,g) = \begin{cases} \mathrm{YES} & \text{if} \quad \|S(f) - g\| \le \varepsilon, \\ \mathrm{NO} & \text{if} \quad \|S(f) - g\| > \varepsilon. \end{cases}$$

That is, we wish to verify whether the elements $S(f)$ and g differ in norm by at most ε. Here ε is a given nonnegative number.

To compute $\mathrm{VER}(f,g)$ we assume that the element g is known, and the knowledge of the element f can be gathered by computing some information operations L about f. Let Λ be a given set of such information operations, see TWW [88, pp. 27-30].

The information which is computed about f may depend on the element g and is of the form

$$N(f;g) = [L_{1,g}(f), L_{2,g}(f), \dots, L_{n,g}(f)]. \tag{2.1}$$

As indicated the choice of $L_{i,g}$ may depend on the element g; additionally it may also depend on already computed values $L_{1,g}(f), L_{2,g}(f), \dots, L_{i-1,g}(f)$. That is,

$$L_{i,g} = L_{i,g}(\cdot; y_1, y_2, \dots, y_{i-1})$$

for $y_j = L_{j,g}(f)$. We assume that for each y_j, $L_{i,g}(\cdot; y_1, y_2, \dots, y_{i-1}) \in \Lambda$. The number $n = n(f,g)$ may depend on the element g and the computed values $L_{i,g}(f)$. The choice of $L_{i,g}$ and $n = n(f,g)$ may also depend on the fixed parameters of the verification problem such as the operator S and the error tolerance ε. We assume that $n(f,g) < +\infty$.

If the choice of $L_{1,g}, L_{2,g}, \dots, L_{n,g}$ or the number n varies with g or with the computed values, then N is called *adaptive*, otherwise (i.e., the choice of $L_{i,g}$ and the number n are fixed and do not vary with elements g nor with the computed values) N is called *nonadaptive*.

The operator N is called the *information* operator, and its *cardinality* is defined as the maximal number $n(f,g)$ of information operations,

$$\mathrm{card}(N) = \max_{f \in F,\ g \in G} n(f,g).$$

Knowing $N(f;g)$ and g, we combine this information by a mapping ϕ, called an *algorithm*, where $\phi : N(F \times G) \times G \to \{\text{YES}, \text{NO}\}$, to get

$$U(f,g) = \phi(N(f;g), g).$$

The *approximation* operator U, where $U : F \times G \to \{\text{YES}, \text{NO}\}$, may also depend on the fixed parameters of the verification problem such as the operator S and the error tolerance ε. It is sometimes convenient to identify the approximation U as the pair $U = (\phi, N)$.

In the *worst case* setting, we wish to find an approximation operator U such that

$$U(f,g) = \text{VER}(f,g) \quad \forall\, f \in F,\ g \in G.$$

In the *probabilistic* setting, assume that the set F is equipped with a probability measure μ. For a given $\delta \in [0,1]$, we wish to find an approximation operator U such that

$$\mu\{f \in F:\ U(f,g) = \text{VER}(f,g)\} \geq 1 - \delta, \quad \forall\, g \in G.$$

That is, for each g from G, we require that U differs from VER only on a set of f of measure at most δ.

3. Worst Case Setting

In this section we show that the verification problem cannot be solved for many linear problems in the worst case setting. That is, there exists no approximation operator U which is equal to VER.

By a *linear* problem, see TWW [88, Chapter 4], we mean a problem for which

(1) $F = \{f \in F_1 :\ \|Tf\| < 1\}$, where $T : F_1 \to X$ is a *linear* operator, and F_1 is a linear space and X is a normed linear space,
(2) S is a *linear operator*, $S : F_1 \to G_1$, and
(3) Λ is an arbitrary subclass of the class of *linear* functionals $L : F_1 \to \mathbb{R}$.

We will sometimes assume that the space G_1 is a strictly convex normed space over the real field, i.e.,

$$\|g_1\| = \|g_2\| \text{ and } \|g_1 + g_2\| = \|g_1\| + \|g_2\| \text{ implies that } g_1 = g_2.$$

We recall the concept of radius of information; see e.g., TWW [88, p. 44] for a detailed discussion of this concept. Let $U(f,g) = \phi(N(f;g), g)$. Define $N_g = N(\cdot; g)$. Usually the information operator N_g is many-to-one and knowing $y = N_g(f)$ it is impossible to recover f exactly. The set

$$A = SN_g^{-1}(y) = \{Sf : \ f \in F \text{ and } N_g(f) = y\}$$

is the set of indistinguishable solution elements with respect to the computed y. Let

$$\operatorname{rad}(A) = \inf_{x \in G_1} \sup_{a \in A} \|x - a\|$$

denote the radius of the set A. Then the *radius* of information N_g is given by

$$r(N_g) = \sup_{y \in N_g(F)} \operatorname{rad}\left(SN_g^{-1}(y)\right).$$

We also recall the *ε-cardinality number*,

$$m(\varepsilon) = \min\{n : \text{ there exists } M \text{ of the form (2.1) such that } \operatorname{card}(M) \le n \text{ and } r(M) \le \varepsilon\},$$

which is the minimal cardinality of information with radius no greater than ε, see e.g., TWW [88, p. 53] where the role of $m(\varepsilon)$ is explained in the worst case setting. Here we use the convention that $\min \emptyset = +\infty$.

Usually $m(\varepsilon)$ goes to infinity as ε goes to zero. How quickly $m(\varepsilon)$ goes to infinity tells us, in general, how hard it is to approximate the elements $S(f)$. For most S and Λ, we have $m(0) = +\infty$. On the other hand, if $m(0) < +\infty$ then the elements $S(f)$ can be recovered exactly by computing finitely many information operations from Λ.

The verification problem depends on a relation between the sets F and G. Clearly, if for any g from G we have $\mathrm{VER}(F, g) = \{\mathrm{YES}\}$ or $\mathrm{VER}(F, g) = \{\mathrm{NO}\}$ then the verification problem is trivial and can be easily computed by $U(\cdot, g) \equiv \mathrm{YES}$ or $U(\cdot, g) \equiv \mathrm{NO}$, respectively.

Therefore we assume that there exists an element g from G such that

$$\mathrm{VER}(F, g) = \{\mathrm{YES}, \mathrm{NO}\}. \tag{3.1}$$

Theorem 3.1. *Suppose that (3.1) holds. If $m(0) = +\infty$ and G_1 is strictly convex then the verification problem cannot be solved for linear problems in the worst case setting. That is,*

$$U(\cdot, g) \neq VER(\cdot, g)$$

for all approximation operators U and all g from G satisfying (3.1).

Proof. Observe that (3.1) means that there exist two elements f_1 and f_2 from F such that

$$\|g - S(f_2)\| \le \varepsilon \quad \text{and} \quad \|g - S(f_1)\| > \varepsilon.$$

Define $h(t) = \|g - S(tf_1 + (1-t)f_2)\|$ for $t \in [0,1]$. Since $h(0) \le \varepsilon$, $h(1) > \varepsilon$ and h is continuous, there exists $t \in [0,1)$ such that $h(t) = \varepsilon$. Let $f = tf_1 + (1-t)f_2$. Then

$$\|g - S(f)\| = \varepsilon \quad \text{and} \quad \|Tf\| < 1. \tag{3.2}$$

Take $U = (\phi, N)$, where N is an arbitrary information operator of finite cardinality and ϕ is an arbitrary algorithm. Apply the information N to the elements f and g of (3.2). Let $L_1^*, L_2^*, \ldots, L_{n^*}^*$ be the linear functionals used in the information N for f and g. Let $N^* = [L_1^*, L_2^*, \ldots, L_{n^*}^*]$. Note that, in general, $N^* \ne N$, but $N^*(f) = N(f;g)$ and $n^* = n(f,g) < +\infty$. The information N^* is nonadaptive and its radius is given by

$$r(N^*) = a \,\sup\{\|S(h)\| :\, h \in F,\, L_i^*(h) = 0,\, i = 1, 2, \ldots, n^*\},$$

with $a \in [1,2]$, see e.g., TWW [88, p. 58].

Since $m(0) = +\infty$, then $r(N^*) > 0$. Thus there exists an element h^* from F such that $N^*(h^*) = 0$ and $S(h^*) \ne 0$. Let $\beta = 1 - \|Tf\|$. Then $\beta > 0$. Note that for $\alpha \in \{-1, +1\}$, we have $f + \alpha\beta\, h^* \in F$, and $N^*(f + \alpha\beta\, h^*) = N^*(f)$.

Let $v = \beta\, S(h^*) \ne 0$. Consider the elements $g - S(f) - \alpha\, v$. Suppose that for both α, $\|g - S(f) - \alpha\, v\| \le \varepsilon$. Then

$$2\varepsilon = \|g - S(f) - v \,+\, g - S(f) + v\| \le \|g - S(f) - v\| \,+\, \|g - S(f) + v\| \le 2\varepsilon$$

implies that $\|g - S(f) - v\| = \|g - S(f) + v\| = \varepsilon$ and

$$\|(g - S(f) - v) \,+\, (g - S(f) + v)\| = \|g - S(f) - v\| \,+\, \|g - S(f) + v\|.$$

Due to strict convexity of the norm of G_1, $g - S(f) - v = g - S(f) + v$ and $v = 0$, which is a contradiction.

Hence, there exists $\alpha^* \in \{-1, +1\}$ such that

$$\|g - S(f) - S(\alpha^*\beta h^*)\| = \|g - S(f + \alpha^*\beta h^*)\| > \varepsilon.$$

Observe that

$$\mathrm{VER}(f, g) = \mathrm{YES}, \quad \mathrm{VER}(f + \alpha^*\beta h^*, g) = \mathrm{NO}.$$

On the other hand, $N(f+\alpha^*\beta h^*;g)=N(f;g)$ yields

$$U(f+\alpha^*\beta h^*,g)=\phi(N(f+\alpha^*\beta h^*;g),g)=\phi(N(f;g),g)=U(f,g).$$

Thus no matter how $U(f,g)$ is defined we have

$$U(f+\alpha^*\beta h^*,g)\neq \mathrm{VER}(f+\alpha^*\beta h^*,g) \quad \text{or} \quad U(f,g)\neq \mathrm{VER}(f,g).$$

This completes the proof. $\square$

We stress that Theorem 3.1 holds for arbitrary ε. Thus, even for large ε, the verification problem cannot be solved in the worst case setting. This holds despite the fact that the approximate computation of the solution elements $S(f)$ for large ε may be very easy.

This negative result immediately implies that the worst case complexity, $\mathrm{comp}^{\mathrm{ver}}(\varepsilon)$, which is the minimal cost of solving the verification problem in the worst case setting, is infinity for all ε. This holds under the assumption that each evaluation of $L(f)$ costs c, where c is positive. Indeed, we have to perform infinitely many evaluations $L(f)$ to get $U=\mathrm{VER}$.

In Theorem 3.1 we assume that $m(0)=+\infty$ and that G_1 is strictly convex. We now show that these two assumptions are essential. Indeed, if $m(0)<+\infty$ then there exist an information operator N of finite cardinality such that the elements $S(f)$ can be recovered exactly. Then $U(f,g)=\mathrm{VER}(\tilde{f},g)$, with $\tilde{f}$ from F chosen such that $N(\tilde{f};g)=N(f;g)$, is well defined and obviously $U=\mathrm{VER}$. In this case, the verification problem can be solved.

We now provide an example of a linear problem with a not strictly convex G_1 for which the verification problem can be solved for one particular G, and cannot be solved for some other G. Of course, (3.1) holds for both G's.

Example 3.1

Let F_1 be a linear space of infinite sequences $f=[f_1,f_2,\dots]$ such that $\sup_{1\le i<\infty}|a_if_i|<+\infty$ for some sequence of positive increasing numbers a_i which tend to infinity. Let $G_1=X$ be a normed linear space of infinite sequences $f=[f_1,f_2,\dots]$ such that $\|f\|=\sup_{1\le i<\infty}|f_i|<+\infty$. Obviously G_1 is *not* strictly convex. Define $Tf=[a_1f_1,a_2f_2,\dots]$. Thus

$$F=\left\{f\in F_1:\ \|Tf\|=\sup_{1\le i<\infty}|a_if_i|\ <\ 1\right\}.$$

Let $S(f)=f$ and let Λ consist of all linear functionals.

Consider first

$$G=\left\{g\in G_1:\ \|g\|_1=\sum_{i=1}^{+\infty}|g_i|\le 1\right\}.$$

Observe that (3.1) now holds, in particular, for $g = [\varepsilon, 0, 0, \dots]$ if $\varepsilon \le 1$.

We now show that there exists an approximation operator U that uses information of finite cardinality such that $U = \mathrm{VER}$. Let n be the smallest integer for which

$$\frac{1}{a_{n+1}} + \frac{1}{n+1} \le \varepsilon.$$

Define the information N of cardinality at most $2n$ as

$$N(f;g) = [f_1, f_2, \dots, f_n, f_{i_1}, f_{i_2}, \dots, f_{i_n}].$$

where i_j denote the indices of the n largest components of g,

$$|g_{i_1}| \ge |g_{i_2}| \ge \cdots \ge |g_{i_n}| \ge |g_i|$$

for all $i \ne i_j$, $j = 1, 2, \dots, n$. We stress that the information N is adaptive since the choice of the functionals depends on g.

Define the approximation U as

$$U(f,g) = \begin{cases} \text{YES} & \text{if } \ |f_j - g_j| \le \varepsilon, \ \forall\, j \in J, \\ \text{NO} & \text{otherwise,} \end{cases}$$

where $J = \{1, 2, \dots, n, i_1, i_2, \dots, i_n\}$.

We now check that $U \equiv \mathrm{VER}$. Indeed, if there exists $j \in J$ for which $|f_j - g_j| > \varepsilon$ then $U(f,g) = \mathrm{NO}$ and

$$\varepsilon < |f_j - g_j| \le \|f - g\| = \|S(f) - g\|$$

yields $\mathrm{VER}(f,g) = \mathrm{NO}$. Thus $U(f,g) = \mathrm{VER}(f,g)$ in this case.

Consider now the case when $|f_j - g_j| \le \varepsilon$, $\forall\, j \in J$. Then $U(f,g) = \mathrm{YES}$. Take $i \notin J$. We have

$$(n+1)|g_i| \le |g_i| + |g_{i_n}| + \cdots + |g_{i_1}| \le \sum_{j=1}^{+\infty} |g_j| \le 1.$$

Thus, $|g_i| \le 1/(n+1)$ and

$$|f_i - g_i| \le \frac{1}{a_{n+1}} + \frac{1}{n+1} \le \varepsilon, \quad \forall\, i \notin J.$$

Hence, $\|S(f) - g\| \le \varepsilon$ and $\mathrm{VER}(f,g) = \mathrm{YES}$. So also in this case $U(f,g) = \mathrm{VER}(f,g)$. Thus, $U \equiv \mathrm{VER}$, as claimed.

Consider now the same linear problem with a set G which contains an element $g = [\varepsilon, \varepsilon, \dots]$. Obviously (3.1) holds for g. It is easy to see that the verification problem cannot be now solved. Indeed, take $U = (\phi, N)$. For $f = 0$ denote $n = n(0, g)$. Since n is finite, there exists a nonzero h from F such that $N(h; g) = N(0, g) = 0$. Let h_i be its nonzero component. We may assume that $h_i < 0$. Then $\mathrm{VER}(0, g) = \mathrm{YES}$ and $\mathrm{VER}(h, g) = \mathrm{NO}$. On the other hand, $U(0, g) = \phi(0, g) = \phi(N(h; g), g) = U(h, g)$. Hence, there is no way to get $U(\cdot, g) = \mathrm{VER}(\cdot, g)$, as claimed. $\square$

Example 3.1 provides a linear problem with a not strictly convex space G_1 for which the verification problem may be solved in the worst case setting by using information of finite cardinality. This information is *adaptive* since the choice of functionals depends on g. This dependence is crucial since it is easy to show that if the choice of functionals does not depend on g then the verification problem cannot be solved. Furthermore, this holds for general problems, i.e., problems which are not necessarily linear. More precisely, we have the following.

For arbitrary S, F and Λ, let

$$N(f; g) \equiv [L_1(f), L_2(f), \dots, L_n(f)], \quad \forall\, g \in G,$$

with L_i from Λ which can depend only on the previously computed $L_j(f)$ for $j \leq i - 1$. Let $r(N)$ be the radius of information N.

Lemma 3.1. *Let $r(N) > 0$ and $S(F) + \{g \in G_1 : \|g\| = \varepsilon\} \subset G$. Then $U \neq VER$ for any approximation operator U that uses N.*

Proof. Let $N(\cdot) = N(\cdot; g)$. Since $r(N) > 0$, there exist two elements f_1 and f_2 from F such that $N(f_1) = N(f_2)$ and $h = S(f_1) - S(f_2) \neq 0$. Let

$$g = S(f_1) + \frac{\varepsilon}{\|h\|}\, h.$$

Then $g \in G$ and $\|S(f_1) - g\| = \varepsilon$ which implies that $\mathrm{VER}(f_1, g) = \mathrm{YES}$. On the other hand,

$$\|S(f_2) - g\| = \|h + \|h\|^{-1}\, \varepsilon\; h\| = \|h\| + \varepsilon > \varepsilon$$

which implies that $\mathrm{VER}(f_2, g) = \mathrm{NO}$.

Note that $N(f_1) = N(f_2)$ yields $U(f_1, g) = \phi(N(f_1), g) = \phi(N(f_2), g) = U(f_2, g)$. Thus, no matter how $U(f_1, g)$ is defined, we have

$$U(f_i, g) \neq \mathrm{VER}(f_i, g)$$

for $i = 1$ or for $i = 2$. This completes the proof. $\square$

Lemma 3.1 states that as long as G is large enough, the radius of information is positive and information does not depend on g, then there is no way to solve the verification problem.

4. Probabilistic Setting

In this section we show that the verification problem can be solved in the probabilistic setting under the following assumptions.

Let F be a separable Banach space equipped with a Gaussian measure μ whose mean element is zero and whose covariance operator is C_μ; for basic properties of Gaussian measures see e.g., Kuo [75] and Vakhania [81].

We assume that S is a linear continuous functional, $S : F \to G = \mathbb{R}$, and Λ is a subclass of continuous linear functionals, $\Lambda \subset F^*$. In this section we assume that G is an arbitrary nonempty subset of $\mathbb{R}$.

Let $\nu = \mu S^{-1}$. Then ν is a one dimensional Gaussian measure with mean zero and variance

$$\sigma_0 = S(C_\mu(S)) = \int_F |S(f)|^2 \mu(df).$$

By $\nu(g, r)$ we mean the Gaussian measure of the interval $[g - r, g + r]$, i.e.,

$$\nu(g, r) = \frac{1}{\sqrt{2\pi\sigma_0}} \int_{|t-g|\le r} e^{-t^2/(2\sigma_0)}\, dt.$$

In what follows, we will use the notion of probabilistic failure, or for brevity, failure, of the approximation U and information N. They are defined as follows. The *probabilistic failure* of the approximation U for g is defined as the measure of the set for which $U(\cdot, g)$ differs from $\mathrm{VER}(\cdot, g)$,

$$\mathrm{Fail}\,(U, g) = \mu\{f \in F :\ U(f, g) \ne \mathrm{VER}(f, g)\},$$

and the *probabilistic failure* of the approximation U is defined as the worst probabilistic failure of U with respect to g,

$$\mathrm{Fail}\,(U) = \sup_{g\in G} \mathrm{Fail}\,(U, g).$$

The *probabilistic failure* of the information N is defined as the minimal probabilistic failure of approximations $U = (\phi, N)$ that use the information N,

$$\mathrm{Fail}\,(N) = \inf_{\phi} \mathrm{Fail}\,(\phi, N).$$

Our goal is to find N with minimal cardinality for which $\mathrm{Fail}(N) \le \delta$.

4.1. Nonadaptive Information

In this section we analyze approximations U that use nonadaptive information N. Since $N(f;g)$ does not depend now on g, we write $N(f)$ instead of $N(f;g)$. Thus $U(f,g) = \phi(N(f),g)$ for nonadaptive N.

We find an algorithm ϕ^* which minimizes the probabilistic failure of U for each element g. Then we provide bounds on the probabilistic failure for $U^*(f,g) = \phi^*(N(f),g)$ for each g.

Consider nonadaptive information

$$N(f) = [L_1(f), L_2(f), \dots, L_n(f)], \tag{4.1}$$

with a priori given n and continuous linear functionals L_i from Λ. Without loss of generality we assume that L_i are normalized such that

$$\langle L_i, L_j \rangle_\mu = \delta_{i,j}, \ \ i,j = 1,2,\dots,n.$$

Here

$$\langle L, R \rangle_\mu \ = \ L(C_\mu(R)), \quad \text{for } L, R \in F^*.$$

Let ν_n be a one dimensional Gaussian measure with mean zero and variance

$$\sigma_n = \langle S, S \rangle_\mu - \sum_{i=1}^n \langle S, L_i \rangle_\mu^2 = \sigma_0 - \sum_{i=1}^n S(C_\mu(L_i))^2 \, = \inf_{\alpha_i \in \mathbb{R}} \left\| S - \sum_{i=1}^n \alpha_i L_i \right\|_\mu^2 .$$

As before we denote the ν_n-measure of the interval $[x-\varepsilon, x+\varepsilon]$ by $\nu_n(x,\varepsilon)$.

We are ready to define an algorithm ϕ^* such that the approximation U^* has minimal failure,

$$U^*(f,g) = \phi^*(N(f),g) = \begin{cases} \text{YES} & \text{if } \nu_n(x,\varepsilon) > \frac{1}{2}, \\ \text{NO} & \text{if } \nu_n(x,\varepsilon) \le \frac{1}{2}, \end{cases} \tag{4.2}$$

where $x = g - \sum_{i=1}^n L_i(f) S(C_\mu(L_i))$.

Lemma 4.1.1. *Among all approximations that use information N given by (4.1), the approximation U^* given by (4.2) has minimal failure for each g, and*

$$\mathit{Fail}(U^*, g) = \frac{1}{\sqrt{2\pi\alpha_n}} \int_{\mathbb{R}} \min\{\nu_n(g+t,\varepsilon), 1 - \nu_n(g+t,\varepsilon)\} e^{-t^2/(2\alpha_n)}\, dt,$$

where $\alpha_n = \sigma_0 - \sigma_n$. In particular,

$$\mathit{Fail}(N) = \mathit{Fail}(U^*).$$

Proof. We need to recall a few facts which can be found in, e.g., TWW [88, Chapter 6]. Let $\rho = \mu N^{-1}$. Then ρ is Gaussian on $\mathbb{R}^n$ with mean zero and correlation matrix identity. Given $y = [y_1, y_2, \ldots, y_n] = N(f)$, the conditional measure $\mu(\cdot|y)$ is defined on F and concentrated on $N^{-1}(y)$. The measure $\mu(\cdot|y)$ is also Gaussian with mean element m_y and covariance operator C_N given by

$$\begin{aligned} m_y =& \sum_{i=1}^n y_i C_\mu(L_i), \\ C_N =& C_\mu - \sum_{i=1}^n \langle \,\cdot\,, L_i \rangle_\mu \, C_\mu(L_i). \end{aligned}$$

Take an arbitrary U that uses the information N, $U(f,g) = \phi(N(f), g)$. We have

$$\mathrm{Fail}\,(U,g) = \int_{\mathbb{R}^n} \mu(\{f \in F:\ \phi(y,g) \neq \mathrm{VER}(f,g)\}|y)\,\rho(dy).$$

We wish to minimize $\mathrm{Fail}\,(U,g)$ by properly defining the element $\phi(y,g) \in \{\mathrm{YES}, \mathrm{NO}\}$. The formula above indicates that $\phi(y,g)$ should be chosen such that $\mu(\{f \in F:\ \phi(y,g) \neq \mathrm{VER}(f,g)\}|y)$ is minimized. Observe that

$$\mu(\{f \in F:\ \mathrm{VER}(f,g) = \mathrm{YES}\}|y) = \mu(\{f \in F:\ |S(f) - g| \le \varepsilon\}|y).$$

Changing variables by $t = S(f) - S(m_y)$ we get

$$\mu(\{f \in F:\ \mathrm{VER}(f,g) = \mathrm{YES}\}|y) = \nu_n(\{t \in \mathbb{R}:\ |t - (g - S(m_y))| \le \varepsilon\}) = \nu_n(g - S(m_y), \varepsilon),$$

and

$$\mu(\{f \in F:\ \mathrm{VER}(f,g) = \mathrm{NO}\}|y) = 1 - \nu_n(g - S(m_y), \varepsilon).$$

¿From this it follows that to minimize the failure we should set $\phi = \phi^*$ with

$$\phi^*(y,g) = \begin{cases} \mathrm{YES} & \text{if } \nu_n(g - S(m_y), \varepsilon) > \frac{1}{2}, \\ \mathrm{NO} & \text{if } \nu_n(g - S(m_y), \varepsilon) \le \frac{1}{2}. \end{cases}$$

Since $g - S(m_y) = g - \sum_{i=1}^n L_i(f)\, S(C_\mu(L_i)) = x$, this agrees with (4.2) and proves that U^* has minimal failure for each g.

To prove the formula for the failure of U^*, note that (4.2) yields

$$\text{Fail}\,(U^*,g) = \int_{\mathbb{R}^n} \min\{\nu_n(g-S(m_y),\varepsilon), 1-\nu_n(g-S(m_y),\varepsilon)\}\,\rho(dy).$$

Change variables by $t = Ay = -S(m_y)$. Then ρA^{-1} is a one dimensional Gaussian measure with mean zero and variance $\alpha_n = \sum_{i=1}^n S(C_\mu(L_i))^2 = \sigma_0 - \sigma_n$. Thus

$$\text{Fail}\,(U^*,g) = \frac{1}{\sqrt{2\pi\alpha_n}} \int_{\mathbb{R}} \min\{\nu_n(g+t,\varepsilon), 1-\nu_n(g+t,\varepsilon)\} e^{-t^2/(2\alpha_n)}\,dt,$$

which completes the proof of Lemma 4.1.1. $\square$

We now estimate the failure of U^*. Let

$$g^* = \inf\{|g| : \quad g \in G\}. \tag{4.3}$$

We first show that $\nu_n(0,\varepsilon) \le \frac{1}{2}$ implies $U^* \equiv \text{NO}$ and

$$\text{Fail}\,(U^*) = \nu(g^*,\varepsilon).$$

Indeed, we have $\nu_n(x,\varepsilon) \le \nu_n(0,\varepsilon) \le \frac{1}{2}$ which yields $U^* \equiv \text{NO}$. Then

$$\text{Fail}\,(U^*,g) = \mu\{f : \ |S(f)-g| \le \varepsilon\} = \nu([g-\varepsilon, g+\varepsilon]) \le \nu([g^*-\varepsilon, g^*+\varepsilon]) = \nu(g^*,\varepsilon)$$

and $\text{Fail}\,(U^*) = \nu(g^*,\varepsilon)$, as claimed.

Assume now that

$$\nu_n(0,\varepsilon) > \tfrac{1}{2}. \tag{4.4}$$

This means that variance σ_n of the measure ν_n must be sufficiently small relative to ε.

Let x_n be a positive number such that $\nu_n(x_n,\varepsilon) = \frac{1}{2}$. That is, x_n satisfies the equation

$$\frac{1}{\sqrt{2\pi}} \int_{(x_n-\varepsilon)/\sqrt{\sigma_n}}^{(x_n+\varepsilon)/\sqrt{\sigma_n}} e^{-t^2/2}\,dt \ = \ \tfrac{1}{2}. \tag{4.5}$$

Note that $x_n \in (0,\varepsilon)$. Denoting

$$x_n = \varepsilon - a_n\sqrt{\sigma_n} \tag{4.6}$$

we have $a_n \in (0, \varepsilon/\sqrt{\sigma_n})$ and $a_n \to 0$ as $\sqrt{\sigma_n} \to 0$. In what follows we shall use the probability integral,

$$\psi(x) = \sqrt{\frac{2}{\pi}} \int_0^x e^{-t^2/2}\,dt.$$

Recall that for large x, we have $\psi(x) = 1 - \sqrt{2/\pi}x^{-1}\exp(-x^2/2)(1+o(1))$.

Lemma 4.1.2. *Let $\nu_n(0,\varepsilon) > \frac{1}{2}$. The failure of U^* given by (4.2) satisfies the following estimates*

$$\begin{aligned} \mathit{Fail}(U^*,g) \le& \sqrt{\frac{\sigma_n}{\alpha_n}}\left(\frac{b(g,\varepsilon)+b(g,3\varepsilon)}{\pi} + a_n\, b(g,\varepsilon)\sqrt{\frac{2}{\pi}}\right) + \\ & \left(1-\psi\left(\frac{x_n+\varepsilon}{\sqrt{\sigma_n}}\right)\right)(b(g,\varepsilon)+1)\sqrt{\frac{2}{\pi\alpha_n}}\,\varepsilon \\ =& \frac{b(g,\varepsilon)+b(g,3\varepsilon)}{\pi}\sqrt{\frac{\sigma_n}{\sigma_0}}\,(1+o(1)) \quad \text{as } \sqrt{\sigma_n}\to 0, \\ \mathit{Fail}(U^*,g) =& \frac{a^*(g)}{\pi}\sqrt{\frac{\sigma_n}{\sigma_0}}\,(1+o(1)) \quad \text{as } \sqrt{\sigma_n}\to 0, \\ \mathit{Fail}(N) = \mathit{Fail}(U^*) =& \frac{a^*}{\pi}\sqrt{\frac{\sigma_n}{\sigma_0}}\,(1+o(1)) \quad \text{as } \sqrt{\sigma_n}\to 0, \end{aligned}$$

where $\alpha_n = \sigma_0 - \sigma_n$ and

$$\begin{aligned} a^*(g) &= e^{-(\varepsilon-g)^2/(2\sigma_0)} + e^{-(\varepsilon+g)^2/(2\sigma_0)} \\ a^* &= \max_{g\in G} a^*(g), \\ b(g,t) &= e^{-[(|g|-t)_+]^2/(2\alpha_n)}. \end{aligned}$$

Proof. We rewrite the failure of U^* using the fact that $\nu_n(x,\varepsilon) \ge \frac{1}{2}$ iff $|x| \le x_n$. We have $\mathrm{Fail}(U^*,g) = \mathrm{Fail}_1 + \mathrm{Fail}_2$, where

$$\begin{aligned} \mathrm{Fail}_1 &= \frac{1}{\sqrt{2\pi\alpha_n}}\int_{|t+g|\le x_n} (1-\nu_n(t+g,\varepsilon))\, e^{-t^2/(2\alpha_n)}\,dt, \\ \mathrm{Fail}_2 &= \frac{1}{\sqrt{2\pi\alpha_n}}\int_{|t+g|> x_n} \nu_n(t+g,\varepsilon)\, e^{-t^2/(2\alpha_n)}\,dt. \end{aligned}$$

We first estimate Fail_1. We have

$$\begin{aligned} \mathrm{Fail}_1 &= \frac{1}{\sqrt{2\pi\alpha_n}}\int_{|x|\le x_n} (1-\nu_n(x,\varepsilon))\, e^{-(x-g)^2/(2\alpha_n)}\,dx \\ &= \frac{1}{\sqrt{2\pi\alpha_n}}\int_{|x|\le x_n}\left(1-\frac{1}{\sqrt{2\pi}}\int_{(x-\varepsilon)/\sqrt{\sigma_n}}^{(x+\varepsilon)/\sqrt{\sigma_n}} e^{-t^2/2}\,dt\right) e^{-(x-g)^2/(2\alpha_n)}\,dx. \end{aligned}$$

Let $\chi(t;[a,b])$ denote the indicator function of the interval $[a,b]$, i.e., $\chi(t;[a,b]) = 1$ if $t\in[a,b]$ and 0 if $t\notin[a,b]$. Let $w=\sqrt{\sigma_n}$. Then

$$\begin{aligned} \mathrm{Fail}_1 =& \frac{1}{\sqrt{2\pi\alpha_n}}\int_{|x|\le x_n}\left[\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-t^2/2}\,dt - \right. \\ & \left. \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \chi(t;[(x-\varepsilon)/w,(x+\varepsilon)/w])\, e^{-t^2/2}\,dt\right] e^{-(x-g)^2/(2\alpha_n)}\,dx \\ =& \frac{1}{2\pi\sqrt{\alpha_n}}\int_{\mathbb{R}}\int_{|x|\le x_n} (1-\chi(t;[(x-\varepsilon)/w,(x+\varepsilon)/w]))\, e^{-(x-g)^2/(2\alpha_n)}\,dx\, e^{-t^2/2}\,dt. \end{aligned}$$

Denote $\mathrm{Fail}_1 = \mathrm{Fail}_{1,1} + \mathrm{Fail}_{1,2}$, where

$$\mathrm{Fail}_{1,1} = \frac{1}{2\pi\sqrt{\alpha_n}} \int_{|t|>(x_n+\varepsilon)/w} \int_{|x|\le x_n} e^{-(x-g)^2/(2\alpha_n)}\,dx\, e^{-t^2/2} dt,$$

$$\mathrm{Fail}_{1,2} = \frac{1}{2\pi\sqrt{\alpha_n}} \int_{|t|\le(x_n+\varepsilon)/w} \left[\int_{-x_n}^{-x_n+\max\{0,x_n-\varepsilon+tw\}} e^{-(x-g)^2/(2\alpha_n)}\,dx + \int_{x_n+\min\{0,\varepsilon-x_n+t\,w\}}^{x_n} e^{-(x-g)^2/(2\alpha_n)}\,dx \right] e^{-t^2/2} dt.$$

We now estimate $\mathrm{Fail}_{1,1}$. Note that

$$e^{-(x-g)^2/(2\alpha_n)} \le e^{[(|g|-\varepsilon)_+]^2/(2\alpha_n)} = b(g,\varepsilon) \qquad \text{for } |x| \le x_n \le \varepsilon.$$

Therefore

$$\begin{aligned} \mathrm{Fail}_{1,1} &= \frac{1}{\sqrt{2\pi\alpha_n}} \int_{-x_n}^{x_n} e^{-(x-g)^2/(2\alpha_n)}\,dx\, \frac{1}{\sqrt{2\pi}} \int_{|t|>(x_n+\varepsilon)/\sqrt{\sigma_n}} e^{-t^2/2} dt \\ &\le b(g,\varepsilon) \sqrt{\frac{2}{\pi\alpha_n}}\, \varepsilon \left(1 - \psi\left(\frac{x_n+\varepsilon}{\sqrt{\sigma_n}} \right) \right). \end{aligned}$$

Observe that

$$\mathrm{Fail}_{1,1} = o\left(\sqrt{\frac{\sigma_n}{\sigma_0}} \right) \quad \text{as } \sqrt{\sigma_n} \to 0.$$

To estimate $\mathrm{Fail}_{1,2}$, we change variables in the two inner integrals by $u = (x+x_n)/\sqrt{\sigma_n}$ and by $u = (x-x_n)/\sqrt{\sigma_n}$, respectively. Using $a_n = (\varepsilon - x_n)/\sqrt{\sigma_n}$ we get

$$\begin{aligned} \mathrm{Fail}_{1,2} &= \frac{1}{2\pi} \sqrt{\frac{\sigma_n}{\alpha_n}} \int_{|t|\le(x_n+\varepsilon)/\sqrt{\sigma_n}} \left[\int_0^{\max\{0,t-a_n\}} e^{-(\sqrt{\sigma_n}x - x_n - g)^2/(2\alpha_n)} dx + \int_{\min\{0,t+a_n\}}^0 e^{-(\sqrt{\sigma_n}x + x_n - g)^2/(2\alpha_n)} dx \right] e^{-t^2/2} dt \\ &\le \frac{b(g,\varepsilon)}{2\pi} \sqrt{\frac{\sigma_n}{\alpha_n}} \left[\int_{a_n}^{(x_n+\varepsilon)/\sqrt{\sigma_n}} \left(\int_0^{t-a_n} dx \right) e^{-t^2/2} dt + \int_{-(x_n+\varepsilon)/\sqrt{\sigma_n}}^{-a_n} \left(\int_{t+a_n}^0 dx \right) e^{-t^2/2} dt \right] \\ &\le \frac{b(g,\varepsilon)}{2\pi} \sqrt{\frac{\sigma_n}{\alpha_n}}\, 2 \int_0^{+\infty} t e^{-t^2/2} dt = \frac{b(g,\varepsilon)}{\pi} \sqrt{\frac{\sigma_n}{\alpha_n}}. \end{aligned}$$

Furthermore, if $\sqrt{\sigma_n} \to 0$ then $a_n \to 0$, $x_n \to \varepsilon$ and $\alpha_n \to \sigma_0$. Thus we have

$$\mathrm{Fail}_{1,2} = \frac{1}{2\pi} \sqrt{\frac{\sigma_n}{\sigma_0}} \left(e^{-(\varepsilon+g)^2/(2\sigma_0)} + e^{-(\varepsilon-g)^2/(2\sigma_0)} \right) (1 + o(1)), \quad \text{as } \sqrt{\sigma_n} \to 0.$$

Combining these estimates we obtain

$$\begin{aligned}
\mathrm{Fail}_1 &\le b(g,\varepsilon)\left(\sqrt{\frac{2}{\pi\alpha_n}}\,\varepsilon\,\left(1-\psi\left(\frac{x_n+\varepsilon}{\sqrt{\sigma_n}}\right)\right)\;+\;\frac{1}{\pi}\sqrt{\frac{\sigma_n}{\alpha_n}}\right),\\
\mathrm{Fail}_1 &= \frac{1}{2\pi}\sqrt{\frac{\sigma_n}{\sigma_0}}\,\left(e^{-(\varepsilon+g)^2/(2\sigma_0)}\;+\;e^{-(\varepsilon-g)^2/(2\sigma_0)}\right)(1+o(1)),\quad \text{as } \sqrt{\sigma_n}\to 0.
\end{aligned}$$

We now estimate the second term

$$\mathrm{Fail}_2=\frac{1}{\sqrt{2\pi\alpha_n}}\int_{x_n}^{+\infty}\nu_n(x,\varepsilon)\,h(x)\,dx,$$

where $h(x)=e^{-(x-g)^2/(2\alpha_n)}\;+\;e^{-(x+g)^2/(2\alpha_n)}$. Denote $\mathrm{Fail}_2=\mathrm{Fail}_{2,1}+\mathrm{Fail}_{2,2}$, where

$$\begin{aligned}
\mathrm{Fail}_{2,1} =&\frac{1}{\sqrt{2\pi\alpha_n}}\int_{x_n}^{+\varepsilon}\nu_n(x,\varepsilon)\,h(x)\,dx\\
\le& b(g,\varepsilon)\sqrt{\frac{2}{\pi\alpha_n}}\,(\varepsilon-x_n)\;=\;b(g,\varepsilon)\sqrt{\frac{2}{\pi}}\sqrt{\frac{\sigma_n}{\alpha_n}}\,a_n,\\
\mathrm{Fail}_{2,2} =&\frac{1}{\sqrt{2\pi\alpha_n}}\int_{\varepsilon}^{+\infty}\nu_n(x,\varepsilon)\,h(x)\,dx\\
=&\frac{1}{2\pi\sqrt{\alpha_n}}\int_{\varepsilon}^{+\infty}\left(\int_{(x-\varepsilon)/\sqrt{\sigma_n}}^{(x+\varepsilon)/\sqrt{\sigma_n}}e^{-t^2/2}dt\right)\,h(x)\,dx\\
=&\frac{1}{2\pi\sqrt{\alpha_n}}\int_{0}^{+\infty}\left(\int_{0}^{+\infty}\chi(t;[x/\sqrt{\sigma_n},(x+2\varepsilon)/\sqrt{\sigma_n}])e^{-t^2/2}dt\right)\,h(x+\varepsilon)\,dx\\
=&\frac{1}{2\pi\sqrt{\alpha_n}}\int_{0}^{+\infty}\left(\int_{0}^{+\infty}\chi(x;[-2\varepsilon+t\sqrt{\sigma_n},t\sqrt{\sigma_n}])h(x+\varepsilon)\,dx\right)\,e^{-t^2/2}dt\\
=&\frac{1}{2\pi\sqrt{\alpha_n}}\Bigg[\int_{0}^{2\varepsilon/\sqrt{\sigma_n}}\left(\int_{0}^{t\sqrt{\sigma_n}}h(x+\varepsilon)dx\right)\,e^{-t^2/2}dt\;+\\
&\qquad\int_{2\varepsilon/\sqrt{\sigma_n}}^{+\infty}\left(\int_{-2\varepsilon+t\sqrt{\sigma_n}}^{t\sqrt{\sigma_n}}h(x+\varepsilon)dx\right)\,e^{-t^2/2}dt\Bigg].
\end{aligned}$$

¿From this we obtain

$$\begin{aligned}
\mathrm{Fail}_{2,2}\le&\frac{1}{2\pi\sqrt{\alpha_n}}\left(2\,b(g,3\varepsilon)\,\sqrt{\sigma_n}\,\int_{0}^{+\infty}te^{-t^2/2}dt\;+\;4\varepsilon\int_{2\varepsilon/\sqrt{\sigma_n}}^{+\infty}e^{-t^2/2}dt\right)\\
=&\frac{b(g,3\varepsilon)}{\pi}\sqrt{\frac{\sigma_n}{\alpha_n}}\;+\;\sqrt{\frac{2}{\pi\alpha_n}}\,\varepsilon\,\left(1-\psi\left(\frac{2\varepsilon}{\sqrt{\sigma_n}}\right)\right).
\end{aligned}$$

Furthermore for $\sqrt{\sigma_n} \to 0$, we get

$$\mathrm{Fail}_{2,2} = \frac{1}{2\pi}\sqrt{\frac{\sigma_n}{\sigma_0}}\,\left(e^{-(\varepsilon-g)^2/(2\sigma_0)} \,+\, e^{-(\varepsilon+g)^2/(2\sigma_0)}\right)(1+o(1)).$$

Combining the estimates for Fail_1 and Fail_2 we finally obtain

$$\begin{aligned}\mathrm{Fail}\,(U^*,g) \le & \,(b(g,\varepsilon)\,+\,1)\,\sqrt{\frac{2}{\pi\alpha_n}}\,\varepsilon\,\left(1-\psi\left(\frac{x_n+\varepsilon}{\sqrt{\sigma_n}}\right)\right)\\ & + \left(\frac{b(g,\varepsilon)+b(g,3\varepsilon)}{\pi} + a_n\,b(g,\varepsilon)\sqrt{\frac{2}{\pi}}\right)\sqrt{\frac{\sigma_n}{\alpha_n}},\end{aligned}$$

which proves the first estimate of Lemma 4.1.2. For $\sqrt{\sigma_n} \to 0$ we have

$$\mathrm{Fail}\,(U^*,g) = \frac{1}{\pi}\sqrt{\frac{\sigma_n}{\sigma_0}}\,\left(e^{-(\varepsilon-g)^2/(2\sigma_0)} \,+\, e^{-(\varepsilon+g)^2/(2\sigma_0)}\right)(1+o(1))$$

which yields the rest of Lemma 4.1.2. $\square$

Remark 4.1.1 In Section 4.3 we will use the following estimate

$$\begin{aligned}\mathrm{Fail}\,(U^*) \ge & \mathrm{Fail}\,(U^*,g^*) \ge \mathrm{Fail}_{1,1} \ge \\ & x_n\sqrt{\frac{2}{\pi\alpha_n}}\,e^{-(x_n+g^*)^2/(2\alpha_n)}\left(1-\psi\left(\frac{x_n+\varepsilon}{\sqrt{\sigma_n}}\right)\right),\end{aligned}$$

where g^* is given by (4.3). This estimate follows from the first part of the proof of Lemma 4.1.2. $\square$

Lemma 4.1.2 presents bounds on the failure of nonadaptive information N. Note that the first bound on $\mathrm{Fail}\,(U^*,g)$ is constructive whereas the second one states that for small variance σ_n the failure is proportional to $\sqrt{\sigma_n}$, and, in particular, is independent of ε. The estimates of Lemma 4.1.2 explain the role of the elements g. For small variance σ_n, the element g affects the failure by $a^*(g)$. If $|g|$ is close to ε then $a^*(g)$ as well as $b(g,\varepsilon)$ are of order 1. If $|g|$ is large relative to 3ε then $b(g,\varepsilon)$, $b(g,3\varepsilon)$, and $a^*(g)$ are exponentially small. Observe also that $a^* \in [1,2]$ if $g=\varepsilon$ or $g=-\varepsilon$ belongs to G. In this case, if ε is small then a^* is close to 2, and if ε is large then a^* is close to 1.

In Section 4.2 we shall prove that the failure of information decreases with variance σ_n. Therefore to minimize the failure we should choose continuous linear functionals from the class Λ which lead to minimal variance. This is in full analogy with the problem of minimizing the probabilistic error for approximation of the solution elements $S(f)$, see TWW [88, Chapter 7].

Let σ_n^* denote the minimal variance which can be achieved after n information operations,

$$\sigma_n^* = \inf_{L_i \in \Lambda} \inf_{\alpha_i \in \mathbb{R}} \left\| S - \sum_{i=1}^{n} \alpha_i L_i \right\|_\mu^2 . \tag{4.7}$$

Without loss of generality assume that the first infimum is attained for functionals $L_i^* \in \Lambda$. Let

$$N_n^* = [L_1^*, L_2^*, \dots, L_n^*] \tag{4.8}$$

be nonadaptive information of cardinality n. It turns out that information N_n^* has minimal failure among *all* nonadaptive and adaptive information of cardinality n, see Section 4.2.

4.2. Adaptive Information

In this section we consider adaptive information N. As we shall see in Section 4.3 we may assume without loss of generality that the number $n(f,g)$ of information operations in N is always the same, $n(f,g) \equiv n$. We show that adaption does not help. That is, the failure of adaptive information is not smaller than the failure of nonadaptive information N_n^* defined by (4.8).

Theorem 4.2.1. *For any adaptive information N of cardinality n we have*

$$\mathit{Fail}(N) \geq \mathit{Fail}(N_n^*),$$

i.e., adaption does not help and the nonadaptive information N_n^* minimizes the failure in the class of all information of cardinality n.

Proof. Consider arbitrary adaptive information N. The information $y = [y_1, \dots, y_n] = N(f;g)$ is then given by

$$N(f;g) = [L_{1,g}(f), L_{2,g}(f;y_1), \dots, L_{n,g}(f;y_1,\dots,y_{n-1})].$$

Here we stress that the ith functional $L_{i,g}(\cdot;y_1,\dots,y_{i-1})$ from the class Λ may depend on the element g as well as on the already computed values $y_j = L_{j,g}(f;y_1,\dots,y_{j-1})$ for $j = 1,2,\dots,i-1$. For brevity we denote

$$L_{i,g,y} = L_{i,g}(\cdot;y_1,y_2,\dots,y_{i-1}).$$

As in Section 4.1 we may normalize the functionals such that

$$\langle L_{i,g,y}, L_{j,g,y}\rangle_\mu = \delta_{i,j} \text{ for all } g,\ y \text{ and } i,j = 1,2,\dots,n.$$

We know, see e.g., TWW [88, Chapter 6], that $\rho = \mu N^{-1}$ is Gaussian on $\mathbb{R}^n$ with mean zero and correlation matrix identity. The conditional measure $\mu(\cdot|y)$ is Gaussian with mean $m_{g,y}$ and covariance operator $C_{g,y}$ given by

$$\begin{aligned} m_{g,y} &= \sum_{i=1}^n y_i C_\mu(L_{i,g,y}),\\ C_{g,y} &= C_\mu - \sum_{i=1}^n \langle\,\cdot\,, C_\mu(L_{i,g,y})\rangle_\mu\, C_\mu(L_{i,g,y}). \end{aligned}$$

Let $\nu_{g,y}$ be a one dimensional Gaussian measure with mean zero and variance

$$\sigma_{g,y} = \langle S,S\rangle_\mu - \sum_{i=1}^n \langle S, L_{i,g,y}\rangle_\mu^2 = \inf_{\alpha_i\in\mathbb{R}} \|S - \sum_{i=1}^n \alpha_i L_{i,g,y}\|_\mu^2.$$

Proceeding as in Lemma 4.4.1 one can show that

$$U^*(f,g) = \phi^*(N(f;g),g) = \begin{cases} \text{YES} & \text{if } \nu_{g,y}(g - S(m_{g,y}),\varepsilon) > \frac{1}{2},\\ \text{NO} & \text{if } \nu_{g,y}(g - S(m_{g,y}),\varepsilon) \le \frac{1}{2} \end{cases}$$

minimizes the failure for each g. Furthermore this failure is equal to

$$\mathrm{Fail}(U^*,g) = \int_{\mathbb{R}^n} \min\{\nu_{g,y}(g - S(m_{g,y}),\varepsilon), 1 - \nu_{g,y}(g - S(m_{g,y}),\varepsilon)\}\,\rho(dy). \tag{4.9}$$

Observe that $\sigma_{g,y} \ge \sigma_n^*$, see (4.7), for all g and y. We claim that replacing the measure $\nu_{g,y}$ with variance $\sigma_{g,y}$ in (4.9) by the Gaussian measure ν_n^* with variance σ_n^* can only decrease the failure. Indeed, we can extend information N by adding n continuous linear functionals (not necessarily from the class Λ) such that the new variance of the measure $\nu_{g,y}$ is always equal to σ_n^*. In particular, if for some g and y we have $\sigma_{g,y} = \sigma_n^*$ then we add n functionals L such that $C_\mu(L)$ is in the kernel of S. Obviously, the failure of the extended information is no greater than the failure of the information N, as claimed.

Therefore it is enough to show the lower bound on the failure of the extended information of cardinality $2n$ for which $\sigma_{g,y} \equiv \sigma_n^*$. We may then rewrite (4.9) as

$$\mathrm{Fail}(U^*,g) \ge \int_{\mathbb{R}^{2n}} \min\{\nu_n^*(g - S(m_{g,y}),\varepsilon), 1 - \nu_n^*(g - S(m_{g,y}),\varepsilon)\}\,\rho(dy), \tag{4.10}$$

where $S(m_{g,y}) = \langle y, w(y)\rangle$ with $w(y) = [S(C_\mu(L_{1,g,y})), S(C_\mu(L_{2,g,y})), \dots, S(C_\mu(L_{2n,g,y}))]$. Then $\|w(y)\| = \sqrt{\sigma_n^* - \sigma_0}$ for all y. The integrand of (4.10) is a function of $\langle y, w(y)\rangle$ and we may apply Lemma 7.3 of Jackowski and Woźniakowski [87] which states that $w(y)$ in (4.10) can be replaced by any element z of the same norm as the common norm of $w(y)$. Choose $z = [S(C_\mu(L_1^*)), S(C_\mu(L_2^*)), \dots, S(C_\mu(L_n^*)), 0, \dots, 0]$, where L_n^* are given by (4.8). Then (4.10) takes the form

$$\mathrm{Fail}\,(U^*, g) \geq \int_{\mathbb{R}^{2n}} \min\{\nu_n^*(g - \langle y, z\rangle), \varepsilon), 1 - \nu_n^*(g - \langle y, z\rangle), \varepsilon)\}\, \rho(dy).$$

Changing variables by $t = -\langle y, z\rangle$ we conclude as in the proof of Lemma 4.1.1 that

$$\mathrm{Fail}\,(U^*, g) \geq \frac{1}{\sqrt{2\pi(\sigma_0 - \sigma_n^*)}} \int_{\mathbb{R}} \min\{\nu_n^*(g + t, \varepsilon), 1 - \nu_n^*(g + t), \varepsilon)\}\, e^{-t^2/(2(\sigma_0 - \sigma_n^*))}\, dt.$$

The right hand side is exactly the same as the formula for the failure for the nonadaptive information N_n^*. Taking the element g which maximizes the failure we conclude that $\mathrm{Fail}\,(N) \geq \mathrm{Fail}\,(N_n^*)$, which completes the proof. □

4.3. Complexity

We are now ready to address the main question of this paper, namely, the complexity of the verification problem in the probabilistic setting.

The model of computation is defined as in TWW [88]. That is, we assume that for each $L \in \Lambda$ and $f \in F$, the evaluation of $L(f)$ costs c, where c is positive and independent of L and f. We assume that we can perform arithmetic operations and comparisons of real numbers with unit cost. Usually $c \gg 1$. We also assume that we can use the integrals

$$\psi(a, b) \;=\; \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt.$$

Observe that $\psi(a, b)$ can be expressed by at most two values of the probability integral $\psi(\cdot)$ which is readily available. For simplicity, we assume that the use of $\psi(a, b)$ costs unity.

Let $U(f, g) = \phi(N(f; g), g)$. The cost of computing $U(f, g)$, $\mathrm{cost}(U, f, g)$, is defined as the cost of computing $y = N(f; g)$ plus the cost of computing $\phi(y, g)$. For nonadaptive information N of cardinality n, the cost of computing y is $c\,n$. For adaptive information N of cardinality n, the cost of computing y may be larger than $c\,n$ since adaptive selection of

$L_{i,g}$ and $n(f)$ may require additional operations. The cost of U is defined as the maximal cost of computing $\mathrm{cost}(U,f,g)$,

$$\mathrm{cost}(U) \;=\; \sup_{f\in F,\, g\in G} \;\mathrm{cost}(U,f,g).$$

Clearly, for $U=(\phi,N)$ we have $\mathrm{cost}(U)\ge c\,\mathrm{card}(N)$.

The complexity $\mathrm{comp}^{\text{prob-ver}}(\varepsilon,\delta)$ of the verification problem in the probabilistic setting is defined as the minimal cost of U with failure at most δ,

$$\mathrm{comp}^{\text{prob-ver}}(\varepsilon,\delta)=\inf\{\mathrm{cost}(U):\; U \text{ such that } \mathrm{Fail}\,(U)\le\delta\}.$$

We say that U^* is (ε,δ)-complexity optimal iff

$$\mathrm{cost}(U^*)=\mathrm{comp}^{\text{prob-ver}}(\varepsilon,\delta) \quad \text{and} \quad \mathrm{Fail}\,(U)\le\delta.$$

Using the results of Sections 4.1 and 4.2 we find tight bounds on the complexity and U^* which is (almost) (ε,δ)-complexity optimal.

Assume first that

$$\nu(g^*,\varepsilon) \;\le\; \delta$$

with g^* given by (4.3). Then the verification problem is trivial since, as already observed, $U\equiv\mathrm{NO}$ solves it and

$$\mathrm{comp}^{\text{prob-ver}}(\varepsilon,\delta) \;=\; 0.$$

Define

$$\delta_0(G,\varepsilon) \;=\; \sup_{g\in G}\min\{\nu(g,\varepsilon),1-\nu(g,\varepsilon)\}.$$

Note that $\delta_0(G,\varepsilon)\le\min\{\frac12,\nu(g^*,\varepsilon)\}$ and $\delta_0(G,\varepsilon)=\nu(g^*,\varepsilon)$ if $\nu(g^*,\varepsilon)\le\frac12$. Assume now that

$$\delta_0(G,\varepsilon) \;\le\; \delta.$$

Then the verification problem is still trivial. Indeed, it can be solved without computing any information N by

$$U(f,g)=\begin{cases}\mathrm{YES} & \text{if } \nu(g,\varepsilon)>\frac12,\\ \mathrm{NO} & \text{if } \nu(g,\varepsilon)\le\frac12.\end{cases} \tag{4.11}$$

Then $\mathrm{Fail}\,(U,g)=\min\{\nu(g,\varepsilon),1-\nu(g,\varepsilon)\}$ and $\mathrm{Fail}\,(U)=\delta_0(G,\varepsilon)$. Since $\nu(g,\varepsilon)=\psi((g-\varepsilon)/\sqrt{\sigma_0}),(g+\varepsilon)/\sqrt{\sigma_0})$ and $\sqrt{\sigma_0}$ can be precomputed, the value $U(f,g)$ can be computed by performing at most 6 operations. Then

$$\mathrm{comp}^{\text{prob-ver}}(\varepsilon,\delta) \;\le\; 6.$$

Consider thus the case

$$\delta \, < \, \delta_0(G, \varepsilon).$$

This inequality means that we have to compute information N with cardinality at least 1 to solve the verification problem.

Let $N_n^* = [L_1^*, L_2^*, \dots, L_n^*]$ be nonadaptive information of cardinality n given by (4.8). Without loss of generality assume that $\left\langle L_i^*, L_j^* \right\rangle_\mu = \delta_{i,j}$, for $i, j = 1, 2 \dots, n$. Let

$$m(\varepsilon, \delta) = \min\{n : \ \mathrm{Fail}\,(N_n^*) \le \delta\} \tag{4.12}$$

be the minimal cardinality of N_n^* with failure at most δ.

The first estimate of Lemma 4.1.2 can be used to compute an upper bound on $m(\varepsilon, \delta)$. Note that $m(\varepsilon, \delta)$ is finite for all positive (ε, δ) iff $\sigma_n^* \to 0$ as $n \to +\infty$. For $\delta \to 0$, the last estimate of Lemma 4.1.2 yields

$$m(\varepsilon, \delta) = \min\left\{ n : \ \sqrt{\sigma_n^*/\sigma_0} \, \le \frac{\pi}{a^*}\, \delta\, (1 + o(1)) \right\}. \tag{4.13}$$

On the other hand, if $\delta = \Theta(\varepsilon)$ as ε tends to zero and $\delta < \delta_0(G, \varepsilon)$ then the first estimate of Lemma 4.1.2 and Remark 4.1.1 yield

$$m(\varepsilon, \delta) = \min\left\{ n : \ \sqrt{\sigma_n^*/\sigma_0} \, \le \Theta(\delta) \right\}. \tag{4.14}$$

As in (4.2), define

$$U^*(f, g) = \phi^*(N_n^*(f; g), g) = \begin{cases} \text{YES} & \text{if } \psi(a, b) > \frac{1}{2}, \\ \text{NO} & \text{if } \psi(a, b) \le \frac{1}{2}, \end{cases} \tag{4.15}$$

where

$$a = \left(g - \varepsilon - \sum_{i=1}^{m(\varepsilon,\delta)} L_i^*(f) S(C_\mu(L_i^*)) \right) \sigma_{m(\varepsilon,\delta)}^{-1/2}, \quad b = a + 2\varepsilon / \sqrt{\sigma_{m(\varepsilon,\delta)}}.$$

We are ready to prove the main result of this paper.

Theorem 4.3.1.

(1) *Let $\nu(g^*, \varepsilon) \le \delta$. Then $U \equiv NO$ solves the verification problem and*

$$comp^{\text{prob-ver}}(\varepsilon, \delta) \; = \; 0.$$

(2) *Let $\delta_0(G, \varepsilon) \le \delta$. Then U given by (4.11) solves the verification problem and*

$$comp^{\text{prob-ver}}(\varepsilon, \delta) \; \le \; 6.$$

(3) *Let $\delta < \delta_0(G,\varepsilon)$. Then U^* given by (4.15) is (almost) (ε,δ)-complexity optimal,*

$$c\, m(\varepsilon,\delta) \le comp^{\text{prob-ver}}(\varepsilon,\delta) \le cost(U^*) \le (c+2)m(\varepsilon,\delta)+5.$$

For $\delta \to 0$ and $c \gg 1$, we have

$$comp^{\text{prob-ver}}(\varepsilon,\delta) \simeq cost(U^*) \simeq c \min\left\{n: \sqrt{\sigma_n^*/\sigma_0} \le \frac{\pi}{a^*}\,\delta\,(1+o(1))\right\}.$$

For $\delta = \Theta(\varepsilon)$ and $c \gg 1$, we have

$$comp^{\text{prob-ver}}(\varepsilon,\delta) \simeq cost(U^*) \simeq c \min\left\{n: \sqrt{\sigma_n^*/\sigma_0} \le \Theta(\delta)\right\}.$$

Proof. We only need to prove (iii). Take an arbitrary $U = (\phi, N)$ with adaptive information N of cardinality n and with failure at most δ. Then $\mathrm{cost}(U) \ge c\,n$ and $\mathrm{Fail}\,(N) \le \delta$.

If $n(f,g) < n$ then we can extend the information N by adding $n - n(f,g)$ linear functionals from Λ which obviously can only decrease the failure. Thus without loss of generality we may assume that $n(f,g) = n, \forall f \in F$. Then Theorem 4.2.1 yields $\mathrm{Fail}\,(N_n^*) \le \mathrm{Fail}\,(N) \le \delta$. Due to (4.12), $n \ge m(\varepsilon,\delta)$.

Thus, $\mathrm{cost}(U) \ge c\,m(\varepsilon,\delta)$ for any U with failure at most δ. This proves the needed lower bound on the complexity.

To get an upper bound, note that U^* given by (4.15) coincides with (4.2) for $N = N^*_{m(\varepsilon,\delta)}$ since $\psi(a,b) = \nu^*_{m(\varepsilon,\delta)}(x,\varepsilon)$. Then Lemma 4.1.1 states that $\mathrm{Fail}\,(U^*) = \mathrm{Fail}\,(N^*_{m(\varepsilon,\delta)})$ and (4.12) yields that $\mathrm{Fail}\,(U^*) \le \delta$. It is easy to check that the cost of U^* is at most $(c+2)m(\varepsilon,\delta)+5$ since the elements $S(C_\mu(L_i^*))$ as well as $\sqrt{\sigma_{m(\varepsilon,\delta)}}$ and $2\varepsilon/\sqrt{\sigma_{m(\varepsilon,\delta)}}$ can be precomputed. This yields the needed upper bound on the complexity. The formulas for $\delta \to 0$ and $\delta = \Theta(\varepsilon)$ follow directly from (4.13) and (4.14). The last formula was proven before. Hence, the proof is complete. $\square$

4.4. Examples

We illustrate analysis of the previous sections by two examples. The first one is integration of scalar periodic smooth functions equipped with a Gaussian measure, and the second one is integration of multivariate continuous functions equipped with the classical Wiener measure. In these two examples we assume that F is a class of functions $f : D \to \mathbb{R}$, and

Λ is a class of function evaluations, i.e., $L \in \Lambda$ means there exists a point $x \in D$ such that $L(f) = f(x)$ for $\forall\, f \in F$. Finally,

$$S(f) = \int_D f(t)\, dt, \quad G = \mathbb{R}\,.$$

Example 4.4.1

Let $D = [0, 2\pi]$ and let F be a Hilbert space of (2π)-periodic functions such that $f(0) = 0$ and whose $(r-1)$st derivative is absolutely continuous and rth derivative belongs to the space $\mathbb{L}_2(0, 2\pi)$, $r \geq 1$, equipped with the norm $\|f\|^2 = \int_0^{2\pi} (f^{(r)}(t))^2\, dt$.

We equipp the space F with a Gaussian measure μ with mean zero and correlation operator C_μ such that

$$C_\mu z_k \;=\; k^{-2s}\, z_k, \quad s \text{ a positive integer,}$$

where $z_{2k-1}(x) = (-1 + \cos\, kx)/(k^r \sqrt{\pi})$ and $z_{2k}(x) = (\sin\, kx)/(k^r \sqrt{\pi})$.

It is known, see TWW [88, p. 308], that

$$\sqrt{\frac{\sigma_n^*}{\sigma_0}} \;=\; \frac{1}{(n+1)^{r+s}},$$

and the optimal nonadaptive information N_n^* consists of function evaluations at equally spaced points,

$$N_n^*(f) = [f(2\pi/(n+1)), f(2\pi\, 2/(n+1)), \ldots, f(2\pi\, n/(n+1))].$$

For $\delta \to 0$ and large c, the complexity and the approximation U^* given by (4.15) are now of the form

$$\mathrm{comp}^{\text{prob-ver}}(\varepsilon, \delta) \;\simeq\; \mathrm{cost}(U^*) \simeq\; c\, m(\varepsilon, \delta) \;\simeq\; c\, \left(\frac{a^*}{\pi\, \delta}\right)^{1/(r+s)}$$

and

$$U^*(f, g) = \begin{cases} \text{YES} & \text{if } \psi(a, b) > \frac{1}{2}, \\ \text{NO} & \text{if } \psi(a, b) \leq \frac{1}{2}, \end{cases}$$

with

$$a = \beta^{-1} \left(g - \varepsilon - \frac{2\pi}{n+1} \sum_{i=1}^{m(\varepsilon,\delta)} f(2\pi\, i/(n+1)) \right) \quad b = a + 2\varepsilon/\beta,$$

where

$$\beta = \frac{\pi}{a^*} \sqrt{\frac{|B_{2(r+s)}|}{(2(r+s))!}}\, \delta,$$

and $B_{2(r+s)}$ is the $2(r+s)$ Bernoulli number. $\square$

Example 4.4.2

Let $D = [0,1]^d$ be the d dimensional unit cube. Let F be a Banach space of continuous functions such that $f(0) = 0$, and with the norm $\|f\| = \sup_{t \in D} |f(t)|$.

We equip the space F with the classical Wiener measure μ with mean zero and correlation operator C_μ such that

$$L_x(C_\mu(L_t)) = \prod_{i=1}^{d} \min\{x_i, t_i\},$$

where $L_x(f) = f(x)$ and $L_t(f) = f(t)$ for $x = [x_1, x_2, \ldots, x_d]$ and $t = [t_1, t_2, \ldots, t_d]$.

It is known, see Woźniakowski [91], that

$$\sqrt{\sigma_n^*} = \Theta\left(\frac{1}{n}\,(\ln\, n)^{(d-1)/2}\right) \quad \text{as } n \to +\infty.$$

This yields that

$$\begin{aligned} m(\varepsilon, \delta) &= \Theta\left(\frac{1}{\delta}\left(\ln\,\frac{1}{\delta}\right)^{(d-1)/2}\right) \quad \text{as } \delta \to 0, \\ \mathrm{comp}^{\text{prob-ver}}(\varepsilon, \delta) &= \Theta\left(c\,\frac{1}{\delta}\left(\ln\,\frac{1}{\delta}\right)^{(d-1)/2}\right) \quad \text{as } \delta \to 0. \end{aligned}$$

Let $n = m(\varepsilon, \delta)$. The (almost) optimal nonadaptive information N_n^* consists of function evaluations at t_i^* which are related to *Hammersley* points,

$$N_n^*(f) = [f(t_1^*)), f(t_2^*)), \ldots, f(t_n^*))].$$

The approximation U^* given by (4.15) is now of the form

$$U^*(f,g) = \begin{cases} \text{YES} & \text{if } \psi(a,b) > \frac{1}{2}, \\ \text{NO} & \text{if } \psi(a,b) \le \frac{1}{2}, \end{cases}$$

with

$$a = \beta^{-1}\left(g - \varepsilon - \frac{1}{n}\sum_{i=1}^{n} f(t_i^*))\right), \quad b = a + 2\varepsilon/\beta,$$

where $\beta = \Theta(\delta)$. Note that the complexity of the verification problem depends weakly on the dimension d. $\square$

Acknowledgments

I thank B. Z. Kacewicz, L. Plaskota, K. Ritter, G. W. Wasilkowski and A. G. Werschulz for valuable comments. I would like also to thank R. M. Karp for helpful discussions regarding this topic.

References

Garey, M. R. and Johnson D. S.,, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, New York, 1979.

Jackowski, T. and Woźniakowski, H., *Complexity of Approximation with Relative Error Criterion in Worst, Average, and Probabilistic Settings*, J. of Complexity **3** (1987), 114–134.

Novak, E. and Woźniakowski, H., *A Relaxed Verification Problem for Continuous Problems*, in progress, 1991..

Kuo, H.-H., *Gaussian Measures in Banach Spaces*, Lectures Notes in Mathematics, 463, Springer Verlag, Berlin, 1975.

Traub, J.F., Wasilkowski, G.W. and Woźniakowski, H., *Information-based Complexity*, Academic Press, New York, 1988.

Traub, J. F. and H. Woźniakowski, *A General Theory of Optimal Algorithms*, Academic Press, New York, 1980.

Vakhania, N. N., *Probability Distributions on Linear Spaces*, North-Holland, New York, 1981.

Woźniakowski, H., *Average Case Complexity of Multivariate Integration*, Bull. AMS (new series) **24** (1991), 185–194.