Complexity of multivariate problems based on binary information

Simon Foucart
Department of Mathematics, Texas A&M University, College Station, Texas 77843–3368
Email: foucart@tamu.edu

***Abstract*—This note suggests some investigations about the complexity of multivariate problems based on quantized information rather than standard information. The extreme case of binary information is studied on two classical examples: the integration of multivariate Lipschitz functions, for which it is shown that adaptivity of the quantization process is beneficial, and the integration of multivariate trigonometric polynomials, for which it is hinted that special binary information is superior to standard binary information.**

I. Introduction

An important question in compressive sensing pertains to the minimal number of linear samples allowing for exact reconstruction of sparse vectors. A similar concern is central to the field of information-based complexity (IBC), except that the objects to be reconstructed are not sparse vectors but multivariate functions and that approximate reconstruction instead of exact reconstruction is sought. Let us recall a few key concepts from IBC relevant to this note. We refer to the encyclopedic volumes [5] for full details. We use the conventional IBC notation: d stands for the number of variables, ϵ for the approximation accuracy, and n for the number of samples. The generic task consists in approximating a quantity of interest $Q(f)$ with accuracy ϵ for all d-variate functions f in a class $\mathcal{F}$ using the n pieces of information of type Λ. To do so, one uses a rule $R = R_n$ (ideally a practical algorithm) taking $\lambda_1(f), \ldots, \lambda_n(f)$ as inputs and returning an approximant $R_n(\lambda(f))$ as output. The process may be nonadaptive, i.e., the information functionals $\lambda_j : \mathcal{F} \to \mathbb{R}$ are chosen once and for all, or it may be adaptive, i.e., the information functionals depend on the information previously acquired, so that $\lambda_j = \lambda_{j;\{(\lambda_i,\lambda_i(f)),i<j\}}$. The difficulty of the task is assessed by the relationship between the required accuracy and the amount of information to gather. This relationship is quantified through

$$e^{\Lambda}_{\rm (non)ada}(n,d)_{\mathcal{F}} := \inf_{R_n} \left\{ \sup_{f \in \mathcal{F}} |Q(f) - R_n(\lambda(f))| \right\},$$

where the infimum is taken over all possible (non)adaptative rules R_n, or 'equivalently' through

$$n^{\Lambda}_{\rm (non)ada}(\epsilon,d)_{\mathcal{F}} = \inf\{n\text{: there is a (non)adaptive rule } R_n\text{: } |Q(f) - R_n(\lambda(f))| \le \epsilon \text{ for all } f \in \mathcal{F}\}.$$

The 'equivalence' alluded to above reflects the fact that, loosely speaking, $e^{\Lambda}_{\rm (non)ada}(\cdot,d)_{\mathcal{F}}$ and $n^{\Lambda}_{\rm (non)ada}(\cdot,d)_{\mathcal{F}}$ are inverse functions of one another. The major inquiry in IBC asks whether the task at hand suffers from the curse of dimensionality, meaning that $n^{\Lambda}_{\rm (non)ada}(\epsilon,d)_{\mathcal{F}}$ grows exponentially with ϵ or d. This question is very well studied for many different scenarios.

The effect of error in the samples is rather less studied in IBC ([7] is the reference on the subject), whereas it appears as a predominant theme in compressive sensing. In fact, quantization of the samples has also been considered in compressive sensing. Notably, the extreme quantization scenario where each sample contains only one bit of information was introduced in [1] and gained popularity under the name of one-bit compressive sensing. The novel contribution of this note consists in attaching this extreme quantization scenario to IBC problems. More precisely, instead of standard information of the form $\lambda_i(f) = f(x_i)$ for some $x_1, \ldots, x_n \in \mathbb{R}^d$, we assume that we can only gather binary information resulting from comparing the $f(x_i)$, or some of their linear combinations, to some threshold y_i. Thus, we only have access to what we shall call standard binary information, i.e.,

$$\lambda_i(f) = \mathrm{sgn}(f(x_i) - y_i) = \begin{Bmatrix} +1 & \text{if } f(x_i) \ge y_i \\ -1 & \text{if } f(x_i) < y_i \end{Bmatrix}, \qquad (1)$$

or to what we shall call special binary information, i.e.,

$$\lambda_i(f) = \mathrm{sgn}\Big(\sum_j a_{i,j} f(x_j) - y_i \Big). \qquad (2)$$

It is also assumed that we have complete freedom in choosing the parameters (x_i, y_i) or $(x_j, a_{i,j}, y_i)$.

The initial investigations presented in this note deal with arguably the two most straightforward IBC problems, i.e., the integration of Lipschitz functions and of trigonometric polynomials — these are the first two examples put forward in [5, Volume 1]. In Section II, we consider the worst-case integration error over a set of Lipschitz functions observed via standard binary information, which we estimate rather precisely in both the adaptive and nonadaptive case. In contrast to several IBC results, adaptivity does help here (adaptivity in the quatization process, that is). This is reminiscent of the message of [2], which showed that, in one-bit compressive sensing, adaptively selecting thresholds allows the reconstruction error to decay exponentially with the number of samples. In Section III, our results are not as compelling, because upper bounds on the worst-case error over a set of multivariate trigonometric polynomials are not accompanied by matching lower bounds. Moreover, only the nonadaptive case is considered. But we give upper bounds (speculated to be sharp) for standard binary information and for special binary information and we stress that the latter is substantially smaller than the former. In Section IV, we conclude by isolating some of the many questions left unanswered by this initial work on quantization in IBC.

II. Integration of Lipschitz functions

We consider here the set of Lipschitz functions

$$\mathcal{L} := \mathcal{L}_d := \left\{ f : [0,1]^d \to \mathbb{R} : \sup_{x \in [0,1]^d} |f(x)| \le \frac{1}{2}, \right. \\ \left. \sup_{x \neq x' \in [0,1]^d} \frac{|f(x) - f(x')|}{\|x - x'\|_\infty} \le 1 \right\}.$$

For standard information of the type $\lambda_i(f) = f(x_i)$, the fundamental result of [8] states that, when $n = m^d$ for some integer m,

$$e^{\rm sta}_{\rm ada}(n,d)_{\mathcal{L}} = e^{\rm sta}_{\rm nonada}(n,d)_{\mathcal{L}} = \frac{d}{2(d+1)} \left(\frac{1}{n}\right)^{1/d},$$

or equivalently, when $\epsilon = 1/(2(d+1)m)$ for some integer m,

$$n^{\rm sta}_{\rm ada}(\epsilon,d)_{\mathcal{L}} = n^{\rm sta}_{\rm nonada}(\epsilon,d)_{\mathcal{L}} = \frac{1}{2^d(1+1/d)^d} \left(\frac{1}{\epsilon}\right)^{d}.$$

This exponential dependence in d makes the problem that uses standard information intractable, hence the harder problem that uses standard binary information is also intractable. We are nonetheless interested in finding the exact order of the worst-case integration errors $e^{\rm st.bin}_{\rm nonada}(n,d)_{\mathcal{L}}$ and $e^{\rm st.bin}_{\rm ada}(n,d)_{\mathcal{L}}$ to show that adaptivity helps (ever so slightly).

Theorem 1. *When $n = m^{d+1}$ for some integer m,*

$$\frac{d}{2(d+1)} \left(\frac{1}{n}\right)^{\frac{1}{d+1}} \le e^{\rm st.bin}_{\rm nonada}(n,d)_{\mathcal{L}} \le \frac{2d+1}{2(d+1)} \left(\frac{1}{n}\right)^{\frac{1}{d+1}}, \quad (3)$$

and when $n = 3m^d$, say, for some integer m,

$$e^{\rm st.bin}_{\rm ada}(n,d)_{\mathcal{L}} \le \frac{2d+1}{2(d+1)} \left(\frac{3}{n}\right)^{\frac{1}{d}}. \quad (4)$$

Proof. We start by proving the most significant result, i.e., the lower bound in (3). With $n = m^{d+1}$, we want to prove that, for any $z_1 := (x_1,y_1), \ldots, z_n := (x_n,y_n) \in [0,1]^d \times [-1/2,1/2]$ and for any map $R_n : \mathbb{R}^n \to \mathbb{R}$,

$$\sigma := \sup_{f \in \mathcal{L}} \left| \int_{[0,1]^d} f - R_n(\lambda(f)) \right| \ge \frac{d}{2(d+1)} \frac{1}{m},$$

where $\lambda(f) = [\lambda_1(f), \ldots, \lambda_n(f)]$ denotes some standard binary information of type (1). So given such points $z_1, \ldots, z_n$ in $[0,1]^d \times [-1/2,1/2]$, there is a 'horizontal' strip $[0,1]^d \times [y - 1/(2m), y + 1/(2m)]$ containing $k \le n/m = m^d$ of these points, say, $z_{i_1}, \ldots, z_{i_k}$. Let us now define a distance on $[0,1]^d$ by $\mathrm{dist}(x,x') = \min\{\|x - x'\|_\infty, 1/(2m)\}$ and in turn let us define two functions f_- and f_+ by

$$f_\pm(x) := y \pm \mathrm{dist}(x, \{x_{i_1}, \ldots, x_{i_k}\}), \quad x \in [0,1]^d.$$

Note that f_- and f_+ are Lipschitz functions with constant one satisfying $f_\pm(x) \in [y - 1/(2m), y + 1/(2m)] \subseteq [-1/2, 1/2]$ for all $x \in [0,1]^d$. Thus, both f_- and f_+ belong to $\mathcal{L}$. Furthermore, note that $\mathrm{sgn}(f_-(x_i) - y_i) = \mathrm{sgn}(f_+(x_i) - y_i)$ for all $i \in [\![1:n]\!]$ (separate the cases $i \in \{i_1, \ldots, i_k\}$ and $i \notin \{i_1, \ldots, i_k\}$), that is to say $\lambda(f_-) = \lambda(f_+)$. Therefore,

$$\sigma \ge \frac{1}{2} \left| \int_{[0,1]^d} f_+ - R_n(\lambda(f_+)) \right| + \frac{1}{2} \left| \int_{[0,1]^d} f_- - R_n(\lambda(f_-)) \right| \\ \ge \frac{1}{2} \int_{[0,1]^d} (f_+ - f_-) = \int_{[0,1]^d} \mathrm{dist}(\cdot, \{x_{i_1}, \ldots, x_{i_k}\}) =: \sigma'.$$

We estimate the latter integral as follows:

$$\begin{aligned} \sigma' &= \int_0^\infty \left| \{x \in [0,1]^d : \mathrm{dist}(x, \{x_{i_1}, \ldots, x_{i_k}\}) > t\} \right| dt \\ &= \int_0^{\frac{1}{2m}} \left(1 - \left| \cup_{j=1}^k \{x \in [0,1]^d : \|x - x_{i_j}\|_\infty \le t\} \right| \right) dt \\ &\ge \frac{1}{2m} - k \int_0^{\frac{1}{2m}} (2t)^d dt = \frac{1}{2m} - \frac{k}{2(d+1)m^{d+1}} \\ &\ge \frac{d}{2(d+1)} \frac{1}{m}. \end{aligned}$$

This immediately implies the desired lower bound on σ.

Let us now turn to the upper bound in (3). Consider the m^d grid points indexed by $i = (i_1, \ldots, i_d) \in [\![1:m]\!]^d$ and defined by

$$x_i = \left(\frac{i_1 - 1/2}{m}, \ldots, \frac{i_d - 1/2}{m} \right) \in [0,1]^d. \quad (5)$$

Given $f \in \mathcal{L}$, by comparing $f(x_i)$ for each $i \in [\![1:m]\!]^d$ to

$$y_1 = -\frac{1}{2} + \frac{1}{2m}, \quad y_2 = -\frac{1}{2} + \frac{3}{2m}, \quad \ldots, \\ y_m = -\frac{1}{2} + \frac{2m-1}{2m} = \frac{1}{2} - \frac{1}{2m}, \quad (6)$$

i.e., by taking the m binary samples $\lambda_{i,j}(f) = \mathrm{sgn}(f(x_i) - y_j)$, $j \in [\![1:m]\!]$, we can approximate $f(x_i)$ by some $f_i \in \mathbb{R}$ in such a way that $|f(x_i) - f_i| \le 1/(2m)$. In total, we are taking $m^d \times m = n$ standard binary samples. The rule for integrating $f \in \mathcal{L}$ from these binary samples consists in outputting

$$R_n(\lambda(f)) := \frac{1}{m^d} \sum_{i \in [\![1:m]\!]^d} f_i. \quad (7)$$

The integration error then satisfies

$$\begin{aligned} \left| \int_{[0,1]^d} f - R_n(\lambda(f)) \right| &\le \left| \int_{[0,1]^d} f - \frac{1}{m^d} \sum_{i \in [\![1:m]\!]^d} f(x_i) \right| \\ &\quad + \frac{1}{m^d} \sum_{i \in [\![1:m]\!]^d} |f(x_i) - f_i| \\ &\le \frac{d}{2(d+1)} \frac{1}{m} + \frac{1}{2m} \\ &= \frac{2d+1}{2(d+1)} \frac{1}{m}, \end{aligned} \quad (8)$$

where the last inequality made use of the result of [8] for the integration of $f \in \mathcal{L}$ based on standard information. This yields the desired upper bound.

Our last task is to establish the upper bound (4). To do so, we consider the same grid points x_i, $i \in [\![1 : m]\!]^d$, as in (5) and the same thresholds y_j, $j \in [\![1 : m]\!]$, as in (6). Given $f \in \mathcal{L}$, for each mutliindex $i = (i_1, \ldots, i_{d-1}, 1)$ whose last entry is fixed, we can again approximate $f(x_i)$ by some $f_i \in \mathbb{R}$ with accuracy $1/(2m)$ from the m standard binary samples $\lambda_{i,j}(f) = \operatorname{sgn}(f(x_i) - y_j)$, $j \in [\![1 : m]\!]$. Then we can propagate the approximation in the $(0, \ldots, 0, 1)$-direction using the Lipschitz condition. Indeed, suppose that $f(x_{(i_1,\ldots,i_{d-1},h)})$ is approximated with accuracy $1/(2m)$, i.e., we know that it belongs to some $[y_j, y_{j+1}]$. In view of

$$|f(x_{(i_1,\ldots,i_{d-1},h+1)}) - f(x_{(i_1,\ldots,i_{d-1},h)})| \le 1/m,$$

we know that $f(x_{(i_1,\ldots,i_{d-1},h+1)})$ belongs to $[y_{j-1}, y_{j+2}]$. With the two standard binary samples

$$\operatorname{sgn}(f(x_{(i_1,\ldots,i_{d-1},h+1)}) - y_j),\ \operatorname{sgn}(f(x_{(i_1,\ldots,i_{d-1},h+1)}) - y_{j+1}),$$

we can locate $f(x_{(i_1,\ldots,i_{d-1},h+1)})$ in $[y_{j-1}, y_j]$, $[y_j, y_{j+1}]$, or $[y_{j+1}, y_{j+2}]$, hence approximate if with accuracy $1/(2m)$. Thus, the $f(x_{(i_1,\ldots,i_{d-1},h)})$, $h \in [\![1 : m]\!]$, can be approximated from $m + 2(m-1) \le 3m$ standard binary samples. In total, all the $f(x_i)$ can be approximated by some f_i with accuracy $1/(2m)$ using at most $m^{d-1} \times 3m = n$ binary measurements. With the same integration rule as in (7), we derive in exactly the same way as (8) that

$$\left| \int_{[0,1]^d} f - R_n(\lambda(f)) \right| \le \frac{2d+1}{2(d+1)} \frac{1}{m},$$

which is precisely the required upper bound. Note that the adaptive quantization procedure is improvable in the sense that we can certainly approximate all the $f(x_i)$ with accuracy $1/(2m)$ with less than $3m^d$ standard binary samples (e.g. $2m^d$) but that the extra effort was unnecessary here. □

Remark. *The integration rules yielding the upper bounds in* (3) *and* (4) *rely on the production of grid points and thresholds (see* (5) *and* (6)*) that can only be carried out from the knowledge that $f \in \mathcal{L}$. Ideally, we would prefer procedures that do not exploit an a priori knowledge of the function class, but that supply better error bounds when the functions happen to belong to specific classes.*

III. Integration of trigonometric polynomials

We consider here the tensor-product space of trigonometric polynomials of degree at most one, i.e.,

$$\begin{aligned}\mathcal{T} := \mathcal{T}_d &:= \bigotimes_{\ell=1}^{d} \operatorname{span}\{1, \cos(2\pi x_\ell), \sin(2\pi x_\ell)\} \\ &=: \bigotimes_{\ell=1}^{d} \operatorname{span}\{e_1(x_\ell), e_2(x_\ell), e_3(x_\ell)\},\end{aligned}$$

equipped with the Euclidean norm that makes the system $\{x \in [0,1]^d \mapsto \prod_{\ell=1}^d e_{i_\ell}(x_\ell), i_1, \ldots, i_d \in \{1,2,3\}\}$ into an orthonormal basis. The space $\mathcal{T}$ is then a reproducing kernel Hilbert space with kernel given by

$$K(x, x') := \prod_{\ell=1}^{d} (1 + \cos(2\pi(x_\ell - x'_\ell))), \quad x, x' \in [0,1]^d.$$

We are interested in the worst-case integration error over the unit ball of $\mathcal{T}$ based on binary information. We shall only be looking at the nonadaptive case and the estimates we shall prove lack lower bounds, but we point out that, even for standard information, the optimal worst-case integration error is not known with certainty. However, it is believed that

$$e^{\rm sta}_{\rm ada}(n,d)_{\mathcal{T}} = e^{\rm sta}_{\rm nonada}(n,d)_{\mathcal{T}} = \max\left\{0, 1 - \frac{n}{2^d}\right\},$$

see Open Problem 3 in [5, Volume 1]. These inequalities do hold when the number n of standard samples equals 2^d, because the equal-weight quadrature formula relative to the points $\{\xi_1, \ldots, \xi_{2^d}\} := \{0, 1/2\}^d$ is exact on $\mathcal{T}$. Thus, the integration rule

$$R_n(f_1, \ldots, f_{2^d}) = \frac{1}{2^d} \sum_{i=1}^{2^d} f_i$$

relying on approximations f_i of $f(\xi_i)$ obtained from n binary samples yields the worst-case integration error

$$\begin{aligned} e(R_n) &= \sup_{\substack{f \in \mathcal{T} \\ \|f\| \le 1}} \left| \int_{[0,1]^d} f - R_n(f_1, \ldots, f_{2^d}) \right| \\ &= \sup_{\substack{f \in \mathcal{T} \\ \|f\| \le 1}} \frac{1}{2^d} \left| \sum_{i=1}^{2^d} (f(\xi_i) - f_i) \right| \le \sup_{\substack{f \in \mathcal{T} \\ \|f\| \le 1}} \frac{1}{2^d} \sum_{i=1}^{2^d} |f(\xi_i) - f_i|. \end{aligned} \tag{9}$$

This leads to the task of approximating $2^{-d}[f(\xi_1), \ldots, f(\xi_{2^d})]$ in ℓ_1-norm from binary information.[1] The proposition below hints that the strategy of quantizing each $f(\xi_i)$ separately may not be optimal.

Proposition 1. *The optimal worst-case ℓ_1-approximation errors of $2^{-d}[f(\xi_1), \ldots, f(\xi_{2^d})]$ over the unit ball of $\mathcal{T}$ satisfy*

$$e^{\rm st.bin}_{\rm nonada}(n,d)_{B_{\mathcal{T}}} \le \frac{2^{3d/2}}{n}, \tag{10}$$

$$e^{\rm sp.bin}_{\rm nonada}(n,d)_{B_{\mathcal{T}}} \lesssim \left(\frac{d}{n}\right)^{1/2}, \tag{11}$$

in the cases of standard and of special binary information, respectively. It follows that

$$n^{\rm st.bin}_{\rm nonada}(\epsilon,d)_{B_{\mathcal{T}}} \le \epsilon^{-1} 2^{3d/2}, \qquad n^{\rm sp.bin}_{\rm nonada}(\epsilon,d)_{B_{\mathcal{T}}} \lesssim \epsilon^{-2} d.$$

The latter improves upon the former in the high-dimensional regime $d \gtrsim \ln(1/\epsilon)$.

Proof. Given $f \in \mathcal{T}$ with $\|f\| \le 1$, we start by noticing that

$$|f(x)| = |\langle f, K(\cdot, x)\rangle| \le \|f\| \|K(\cdot, x)\| \le 1 \times 2^{d/2}$$

for any $x \in [0,1]^d$, where the last inequality resulted from

$$\|K(\cdot, x)\|^2 = \langle K(\cdot, x), K(\cdot, x)\rangle = K(x, x) = 2^d.$$

[1] For the original integration problem, special binary information of type (2) is too powerful, since it includes information of the type $\operatorname{sgn}\left(\int_{[0,1]^d} f - y_i\right)$.

Therefore, as in the proof of Theorem 1, we can approximate each $f(\xi_i)$ for a fixed $i \in [\![1 : 2^d]\!]$ by some $f_i \in \mathbb{R}$ with accuracy $\epsilon = 2^{d/2}/m$ by taking m binary samples of the type $\mathrm{sgn}(f(\xi_i) - y_j)$, $j \in [\![1 : m]\!]$. The total number of standard binary samples to approximate $f(\xi_1), \ldots, f(\xi_{2^d})$ simultaneously with accuracy ϵ is then $n = 2^d m$. We derive that the quantity in (9) satisfies

$$\sup_{\substack{f \in \mathcal{T} \\ \|f\| \le 1}} \frac{1}{2^d} \sum_{i=1}^{2^d} |f(\xi_i) - f_i| \le \epsilon = \frac{2^{d/2}}{m} = \frac{2^{3d/2}}{n}.$$

This proves the estimate (10).

To prove the estimate (11), we introduce the set

$$S := \left\{ \frac{1}{2^{d/2}} [f(\xi_1), \ldots, f(\xi_{2^d})] : f \in \mathcal{T},\ \|f\| \le 1 \right\} \subseteq \mathbb{R}^{2^d}.$$

We claim that elements of S have ℓ_2-norm bounded by one. Indeed, given $f \in \mathcal{T}$ with $\|f\| \le 1$, we remark that

$$\begin{aligned} \sum_{i=1}^{2^d} f(\xi_i)^2 &= \sum_{i=1}^{2^d} f(\xi_i) \langle f, K(\cdot, \xi_i) \rangle \\ &= \langle f, \sum_{i=1}^{2^d} f(\xi_i) K(\cdot, \xi_i) \rangle \\ &\le \left\| \sum_{i=1}^{2^d} f(\xi_i) K(\cdot, \xi_i) \right\|. \end{aligned} \tag{12}$$

Next, in view of

$$\begin{aligned} \left\| \sum_{i=1}^{2^d} f(\xi_i) K(\cdot, \xi_i) \right\|^2 &= \sum_{i,j=1}^{2^d} f(\xi_i) f(\xi_j) \langle K(\cdot, \xi_i), K(\cdot, \xi_j) \rangle \\ &= \sum_{i,j=1}^{2^d} f(\xi_i) f(\xi_j) K(\xi_i, \xi_j) \end{aligned}$$

and the facts that $K(\xi_i, \xi_i) = 2^d$ and $K(\xi_i, \xi_j) = 0$ for $i \ne j$, we obtain

$$\left\| \sum_{i=1}^{2^d} f(\xi_i) K(\cdot, \xi_i) \right\| = 2^{d/2} \left[\sum_{i=1}^{2^d} f(\xi_i)^2 \right]^{1/2}. \tag{13}$$

Putting (12) and (13) together gives the ℓ_2-norm estimate $\|[f(\xi_1), \ldots, f(\xi_{2^d})]\|_2 \le 2^{d/2}$, which confirms our claim. We also point out that the set S is embedded in a linear space of dimension $\le 3d$, since $\mathcal{T}$ itself has dimension $= 3d$. Therefore, the mean width of S satisfies (see e.g. [6])

$$w(S) \le \sqrt{3d}.$$

Therefore, we know (from probabilistic arguments, see e.g. [4, Theorem 2.2][2]) that there exist $a_{i,j}$, $i \in [\![1 : n]\!]$, $j \in [\![1 : 2^d]\!]$, such that nonadaptive special binary information of type (2) allows any element of S to be approximated in ℓ_2-norm with accuracy δ provided that $n \gtrsim \delta^{-2} w(S)^2$, hence as soon as $n \asymp \delta^{-2}\, d$. This means that, for each $f \in \mathcal{T}$ with $\|f\| \le 1$, the vector $v = [f(\xi_1), \ldots, f(\xi_{2^d})] \in \mathbb{R}^{2^d}$ can be approximated by a vector $v^\sharp = [f_1, \ldots, f_{2^d}] \in \mathbb{R}^{2^d}$ with $\|v - v^\sharp\|_2 \le 2^{d/2}\delta$. In turn, we derive that

$$\begin{aligned} \frac{1}{2^d} \sum_{i=1}^{2^d} |f(\xi_i) - f_i| &\le \frac{1}{2^d} \|v - v^\sharp\|_1 \le \frac{1}{2^{d/2}} \|v - v^\sharp\|_2 \\ &\le \delta \asymp \left(\frac{d}{n} \right)^{1/2}, \end{aligned}$$

which yields the estimate (11) on the quantity in (9). □

[2] Strictly speaking, the theorem applies only to a subset of the unit sphere, not to a subset of the unit ball, but the lifting trick presented in [3, Section 4] takes care of a full explanation.

IV. Perspectives

We have initiated a study of the complexity of multivariate problems based on binary information by considering two rather simple problems (for which we did not even give a complete picture). There are countless other IBC problems that had been investigated in the context of standard information and that could be revisited in this novel context. In particular, it would be intriguing to uncover a situation where a problem is intractable based on nonadaptive binary information but for which adaptivity would enable tractability. This search should be preceded by a serious discussion of what tractability should mean now. Indeed, since acquiring a binary sample is so much cheaper than acquiring a standard sample, the numbers of samples of both kinds cannot be placed on an equal footing, so one should wonder if it is still pertinent to declare a d-variate problem (polynomially) tractable under the condition that the number of binary samples to achieve accuracy ϵ depends polynomially on ϵ^{-1} and d. Finally, binary information sets the scene for neat mathematical problems, but the one-bit quantization attached to it may be regarded as a toy process. In practice, more sophisticated quantization processes are used and ideally the complexity of multivariate problems should be examined in the context of the quantized information resulting from these processes.

Acknowledgment

S.F. is partially supported by the NSF grant DMS-1622134.

References

[1] P. Boufounos and R. Baraniuk. *1-bit compressive sensing.* 42nd Annual Conference on Information Sciences and Systems. IEEE (2008).

[2] R. Baraniuk, S. Foucart, D. Needell, Y. Plan, and M. Wootters. *Exponential decay of reconstruction error from binary measurements of sparse signals.* IEEE Transactions on Information Theory. To appear.

[3] R. Baraniuk, S. Foucart, D. Needell, Y. Plan, and M. Wootters. *One-bit compressive sensing of dictionary-sparse signals.* Preprint.

[4] S. Oymak and B. Recht. *Near-optimal bounds for binary embeddings of arbitrary sets.* arXiv preprint arXiv:1512.04433 (2015).

[5] E. Novak and H. Woźniakowski. *Tractability of Multivariate Problems (Volumes I, II, III).* European Mathematical Society Publishing House (2008, 2010, 2012).

[6] Y. Plan and R. Vershynin. *Robust 1-bit compressed sensing and sparse logistic regression: a convex programming approach.* IEEE Transactions on Information Theory 59.1 (2013): 482–494.

[7] L. Plaskota. *Noisy Information and Computational Complexity.* Cambridge University Press (1996).

[8] A. G. Sukharev. *Optimal numerical integration formulas for some classes of functions.* Sov. Math. Dokl. 20 (1979): 472–475.