




# Noncommutative Bohnenblust–Hille inequalities

Alexander Volberg<sup>1,2</sup> · Haonan Zhang<sup>3,4</sup> 

Received: 4 December 2022 / Revised: 7 July 2023 / Accepted: 10 July 2023 /  
Published online: 24 July 2023  
© The Author(s) 2023

## Abstract

Bohnenblust–Hille inequalities for Boolean cubes have been proven with dimension-free constants that grow subexponentially in the degree (Defant et al. in *Math Ann* 374(1):653–680, 2019). Such inequalities have found great applications in learning low-degree Boolean functions (Eskenazis and Ivanisvili in *Proceedings of the 54th annual ACM SIGACT symposium on theory of computing*, pp 203–207, 2022). Motivated by learning quantum observables, a qubit analogue of Bohnenblust–Hille inequality for Boolean cubes was recently conjectured in Rouzé et al. (*Quantum Talagrand, KKL and Friedgut’s theorems and the learnability of quantum Boolean functions*, 2022. arXiv preprint [arXiv:2209.07279](https://arxiv.org/abs/2209.07279)). The conjecture was resolved in Huang et al. (*Learning to predict arbitrary quantum processes*, 2022. arXiv preprint [arXiv:2210.14894](https://arxiv.org/abs/2210.14894)). In this paper, we give a new proof of these Bohnenblust–Hille inequalities for qubit system with constants that are dimension-free and of exponential growth in the degree. As a consequence, we obtain a junta theorem for low-degree polynomials. Using similar ideas, we also study learning problems of low degree quantum observables and Bohr’s radius phenomenon on quantum Boolean cubes.

**Mathematics Subject Classification** 47A30 · 81P45 · 06E30

---

The research of A.V. is supported by NSF DMS-1900286, DMS-2154402 and by Hausdorff Center for Mathematics. H.Z. is supported by the Lise Meitner fellowship, Austrian Science Fund (FWF) M3337. This work is partially supported by NSF DMS-1929284 while both authors were in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, RI, during the Harmonic Analysis and Convexity program.

---

✉ Haonan Zhang  
haonanzhangmath@gmail.com

Alexander Volberg  
volberg@math.msu.edu

<sup>1</sup> Department of Mathematics, MSU, East Lansing, MI 48823, USA

<sup>2</sup> Hausdorff Center of Mathematics, Bonn, Germany

<sup>3</sup> Department of Mathematics, University of California, Irvine, CA 92617, USA

<sup>4</sup> Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400 Klosterneuburg, Austria

## 1 Introduction

In 1930, Littlewood [19] proved that for any  $n \geq 1$  and any bilinear form  $B : \mathbb{C}^n \otimes \mathbb{C}^n \rightarrow \mathbb{C}$ , we have

$$\left( \sum_{i,j} |B(e_i, e_j)|^{4/3} \right)^{3/4} \leq \sqrt{2} \|B\|, \quad (1.1)$$

where  $\{e_j, 1 \leq j \leq n\}$  is the canonical basis of  $\mathbb{C}^n$  and  $\|B\|$  denotes the norm of the bilinear form, i.e.

$$\|B\| := \sup\{|B(x, y)| : x, y \in \mathbb{C}^n, \|x\|_\infty \leq 1, \|y\|_\infty \leq 1\}.$$

Here  $4/3$  is optimal and (1.1) is known as *Littlewood's 4/3 inequality*. Right after Littlewood's proof of (1.1), Bohnenblust and Hille [1] extended this result to multilinear forms: For any  $d \geq 1$ , there exists a constant  $C_d > 0$  depending only on  $d$  such that for any  $n \geq 1$  and any  $d$ -linear form  $B : \mathbb{C}^n \times \dots \times \mathbb{C}^n \rightarrow \mathbb{C}$  we have

$$\left( \sum_{i_1, \dots, i_d=1}^n |B(e_{i_1}, \dots, e_{i_d})|^{2d/d+1} \right)^{\frac{d+1}{2d}} \leq C_d \|B\|, \quad (1.2)$$

where  $\|B\|$  is defined in a similar way as above and the exponent  $2d/(d+1)$  is optimal. The inequalities (1.2) have played a key role in Bohnenblust and Hille's solution [1] to *Bohr's strip problem* [4] concerning the convergence of Dirichlet series. Such multilinear form inequalities (1.2) and their polynomial variants (which we shall recall for Boolean cubes) are known as *Bohnenblust–Hille inequalities*.

Since then, Bohnenblust–Hille inequalities have been extended to different contexts. Recent years have seen great progress in improving the constants in Bohnenblust–Hille inequalities (e.g.  $C_d$  in (1.2)) and this has led to the resolution of a number of open problems in harmonic analysis. See for example [7, 9, 10, 13] and references therein.

In [3] Blei extended Bohnenblust–Hille inequalities to polynomials on Boolean cubes with dimension-free constants. Recently, this result was revisited by Defant et al. [12]. Moreover, they proved that the dimension-free Bohnenblust–Hille constants for Boolean cubes actually grow at most subexponentially in the degree. To state their results, recall that any function  $f : \{-1, 1\}^n \rightarrow \mathbb{C}$  has the *Fourier–Walsh expansion*:

$$f(x) = \sum_{S \subset [n]} \widehat{f}(S) \chi_S(x),$$

where for each  $S \subset [n] := \{1, \dots, n\}$ ,  $\widehat{f}(S) \in \mathbb{C}$  and

$$\chi_S(x) := \prod_{j \in S} x_j, \quad x = (x_1, \dots, x_n).$$

The function  $f$  is said to be of *degree at most  $d$*  if  $\widehat{f}(S) = 0$  whenever  $|S| > d$ ; and is said to be  *$d$ -homogeneous* if  $\widehat{f}(S) = 0$  whenever  $|S| \neq d$ . Defant, Mastysłó and Pérez proved the following theorem (they considered real-valued functions but the proof works for complex-valued case as well).

**Theorem 1.1** [12, Theorem 1] *For any  $d \geq 1$ , there exists  $C_d > 0$  such that for any  $n \geq 1$  and any  $f : \{-1, 1\}^n \rightarrow \mathbb{C}$  of degree at most  $d$ , we have*

$$\left( \sum_{|S| \leq d} |\widehat{f}(S)|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leq C_d \|f\|_\infty. \tag{1.3}$$

Denoting  $BH_{\{\pm 1\}}^{\leq d}$  the best constant  $C_d$  such that (1.3) holds, then there exists  $C > 0$  such that  $BH_{\{\pm 1\}}^{\leq d} \leq C^{\sqrt{d \log d}}$ . So the Bohnenblust–Hille constant  $BH_{\{\pm 1\}}^{\leq d}$  is of subexponential growth.

As we mentioned, Blei [3] first proved  $BH_{\{\pm 1\}}^{\leq d} < \infty$ . One of main contributions of Defant et al. [12] lies in this subexponential bound  $BH_{\{\pm 1\}}^{\leq d} \leq C^{\sqrt{d \log d}}$ . Recently, this Bohnenblust–Hille inequality for Boolean cubes (1.3) has found great applications in learning bounded low-degree functions on Boolean cubes [14], which will be explained in more detail in Sect. 4. In view of this, an analogue of (1.3) for qubit systems was conjectured in [20], motivated by learning quantum observables following the work of Eskenazis and Ivanišvili [14]. But actually this quantum analogue of Bohnenblust–Hille inequality was already contained in a result of Huang, Chen and Preskill in the preprint [16] that was not online available when the conjecture was made. Their motivation is to predict any quantum process. In this paper, we provide another proof that is simpler and gives better constants. Moreover, our method is more general which allows us to reduce many problems on the qubit systems to classical Boolean cubes. We refer to Sect. 2 for more discussions and to Sect. 6 for the comparison on our results and the work in [16].

In our quantum setup, the Boolean cubes  $\{-1, 1\}^n$  are replaced by  $M_2(\mathbb{C})^{\otimes n}$ , the  $n$ -fold tensor product of 2-by-2 complex matrix algebras. Recall that Pauli matrices and the identity matrix

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

form a basis of  $M_2(\mathbb{C})$ . For  $\mathbf{s} = (s_1, \dots, s_n) \in \{0, 1, 2, 3\}^n$ , we put

$$\sigma_{\mathbf{s}} := \sigma_{s_1} \otimes \dots \otimes \sigma_{s_n}.$$

All the  $\sigma_{\mathbf{s}}$ ,  $\mathbf{s} \in \{0, 1, 2, 3\}^n$  form a basis of  $M_2(\mathbb{C})^{\otimes n}$  and play the role of characters  $\chi_S$ ,  $S \in [n]$  in the classical case. So any  $A \in M_2(\mathbb{C})^{\otimes n}$  has the unique Fourier expansion

$$A = \sum_{\mathbf{s} \in \{0, 1, 2, 3\}^n} \widehat{A}_{\mathbf{s}} \sigma_{\mathbf{s}}$$

with  $\widehat{A}_{\mathbf{s}} \in \mathbb{C}$  being the Fourier coefficient. For any  $\mathbf{s} = (s_1, \dots, s_n) \in \{0, 1, 2, 3\}^n$ , we denote by  $|\mathbf{s}|$  the number of non-zero  $s_j$ 's. Similar to the classical setting,  $A \in M_2(\mathbb{C})^{\otimes n}$  is of degree at most  $d$  if  $\widehat{A}_{\mathbf{s}} = 0$  whenever  $|\mathbf{s}| > d$ , and it is  $d$ -homogeneous if  $\widehat{A}_{\mathbf{s}} = 0$  whenever  $|\mathbf{s}| \neq d$ .

In the sequel, we always use  $\|A\|$  to denote the operator norm of  $A$ . Our main result is the following:

**Theorem 1.2** *For any  $d \geq 1$ , there exists  $C_d > 0$  such that for all  $n \geq 1$  and all  $A = \sum_{|\mathbf{s}| \leq d} \widehat{A}_{\mathbf{s}} \sigma_{\mathbf{s}} \in M_2(\mathbb{C})^{\otimes n}$  of degree at most  $d$ , we have*

$$\left( \sum_{|\mathbf{s}| \leq d} |\widehat{A}_{\mathbf{s}}|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leq C_d \|A\|. \tag{1.4}$$

Moreover, denote  $\text{BH}_{M_2(\mathbb{C})}^{\leq d}$  the best constant  $C_d > 0$ . Then we have  $\text{BH}_{M_2(\mathbb{C})}^{\leq d} \leq 3^d \text{BH}_{\{\pm 1\}}^{\leq d}$ , so that it is at most of exponential growth.

A special choice of polynomials yields noncommutative analogues of Bohnenblust–Hille inequalities for multilinear forms. For this we use the following notation. Fix  $n \geq 1$ . For  $\kappa \in \{1, 2, 3\}$  and  $i \in [n]$ , we write  $\sigma_i^{(\kappa)}$  for  $\sigma_{\mathbf{s}}$  where  $\mathbf{s} = (s_1, \dots, s_n) \in \{0, 1, 2, 3\}^n$  with  $s_i = \kappa$  and  $s_j = 0$  whenever  $j \neq i$ .

**Corollary 1.3** *Fix  $d \geq 1$ . Then there exists  $C_d > 0$  such that for any  $n \geq 1$  and any (each  $\sigma_{i_j}^{(\kappa_j)} \in M_2(\mathbb{C})^{\otimes n}$ )*

$$A := \sum_{\kappa_1, \dots, \kappa_d \in \{1, 2, 3\}} \sum_{i_1, \dots, i_d = 1}^n a_{i_1, \dots, i_d}^{\kappa_1, \dots, \kappa_d} \sigma_{i_1}^{(\kappa_1)} \otimes \dots \otimes \sigma_{i_d}^{(\kappa_d)},$$

we have

$$\left( \sum_{\kappa_1, \dots, \kappa_d \in \{1, 2, 3\}} \sum_{i_1, \dots, i_d = 1}^n |a_{i_1, \dots, i_d}^{\kappa_1, \dots, \kappa_d}|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leq C_d \|A\|.$$

Moreover,  $C_d \leq 3^d \text{BH}_{\{\pm 1\}}^{\leq d}$ , and it becomes a noncommutative analogue of Littlewood’s 4/3 inequality when  $d = 2$ .

**Remark 1.4** Note that the algebra of function on  $\{-1, 1\}^n$  can be viewed as a commutative subalgebra of  $M_2(\mathbb{C})^{\otimes n}$  spanned by  $\sigma_{\mathbf{s}}, \mathbf{s} \in \{0, 3\}^n$ . So (1.3) is a special case of (1.4)

and we always have  $\text{BH}_{\{\pm 1\}}^{\leq d} \leq \text{BH}_{M_2(\mathbb{C})}^{\leq d}$ . Our main result Theorem 1.2 states that the converse holds up to a factor  $3^d$ .

**Remark 1.5** The main theorem of Huang et al. [16, Theorem 5] is actually more general which admits (1.4) as a corollary [16, Corollary 3]. Their proof is different from ours and the constant they obtained is  $C_d \sim d^{\mathcal{O}(d)}$  which is worse.

Recall that a function  $f : \{-1, 1\}^n \rightarrow \mathbb{C}$  is called a  $k$ -junta if it depends on at most  $k$  coordinates. Similarly, a matrix  $A \in M_2(\mathbb{C})^{\otimes n}$  is a  $k$ -junta if it acts non-trivially on at most  $k$  qubits, that is,

$$|\{1 \leq j \leq n : \exists \mathbf{s} \in \{0, 1, 2, 3\}^n \text{ s.t. } \widehat{A}_{\mathbf{s}} \neq 0 \ \& \ s_j \neq 0\}| \leq k.$$

It is known that [6, 8] if a bounded function  $f$  over  $\{-1, 1\}^n$  is of low degree, then it is close to a junta. In the next corollary we derive such a result in a quantum setting. We refer to [20, Theorem 3.9] to another quantum junta type theorem related to the influences instead of the degree.

**Corollary 1.6** *Suppose that  $A \in M_2(\mathbb{C})^{\otimes n}$  is of degree at most  $d$  and  $\|A\| \leq 1$ . Then for any  $\epsilon > 0$ , there exists a  $k$ -junta  $B \in M_2(\mathbb{C})^{\otimes n}$  such that*

$$\|A - B\|_2 \leq \epsilon \quad \text{with} \quad k \leq \frac{d \left( \text{BH}_{M_2(\mathbb{C})}^{\leq d} \right)^{2d}}{\epsilon^{2d}}.$$

Here  $\|\cdot\|_2$  denoted the Hilbert–Schmidt norm with respect to the normalized trace, that is,  $\|A\|_2^2 = 2^{-n} \text{tr}[A^*A]$ . In particular, we may choose  $k \leq dC^{2d^2} \epsilon^{-2d}$  for some universal  $C > 0$ .

**Remark 1.7** The results in [6, 8] are more general. However, in the case when polynomials are of low degree, the proof presented here that uses Bohnenblust–Hille inequalities is simpler. We are grateful to Alexandros Eskenazis for pointing out to us this proof.

To prove Theorem 1.2, we reduce the problem to the (commutative) Boolean cube case, at a price of an extra factor  $3^d$ . In fact, our main contribution is a general method that reduces many problems in the quantum setting to the classical setting. This method will be explained in Sect. 2, while the proof of Theorem 1.2 and Corollary 1.6 will be presented in Sect. 3.

We shall illustrate the strength of this reduction method with two more applications. The first one concerns learning bounded low-degree quantum observables which will be Sect. 4 and the main result is Theorem 4.1. The second one is Theorem 5.2 on Bohr’s radius phenomenon in the context of quantum Boolean cubes which will be discussed in Sect. 5.

In Sect. 6, we briefly compare our results with the work of Huang et al. [16].

**Notation.** We shall use  $\text{tr}$  to denote the usual (unnormalized) trace on matrix algebras, and  $\langle \cdot, \cdot \rangle$  the inner product on  $\mathbb{C}^n$  that is linear in the second argument. By  $\|A\|_p$  of a  $k$ -by- $k$  matrix  $A$  we always mean the normalized Schatten- $p$  norm, i.e.  $\|A\|_p^p = 2^{-k} \text{tr}|A|^p$ . For any unit vector  $\eta \in \mathbb{C}^n$ , we use  $|\eta\rangle\langle\eta|$  to denote the associated rank one projection operator. Sometimes people use the convention  $\eta \otimes \eta$  instead. By a density matrix we mean a positive semi-definite matrix with unit trace.

## 2 Reduction to the commutative case

In this section, we present a general reduction method. For this let us collect a few facts about Pauli matrices. For each  $j = 1, 2, 3$ ,  $\sigma_j$  is self-adjoint unitary, and has 1 and  $-1$  as eigenvalues. We denote by  $e_1^j$  and  $e_{-1}^j$  the corresponding unit eigenvectors, respectively. Pauli matrices  $\sigma_j, j = 1, 2, 3$  satisfy the following anticommutation relation:

$$\sigma_j \sigma_k + \sigma_k \sigma_j = 0, \quad j \neq k \in \{1, 2, 3\}. \tag{2.1}$$

We record the following simple fact as a lemma.

**Lemma 2.1** *For any  $j, k \in \{1, 2, 3\}$  and  $\epsilon \in \{-1, 1\}$ , we have*

$$\langle \sigma_j e_\epsilon^k, e_\epsilon^k \rangle = \delta_{jk} \epsilon. \tag{2.2}$$

**Proof** When  $j = k$ , (2.2) is trivial by definition of  $e_\epsilon^k$ . When  $j \neq k$ , by (2.1) we have

$$\epsilon \langle \sigma_j e_\epsilon^k, e_\epsilon^k \rangle = \langle \sigma_j \sigma_k e_\epsilon^k, e_\epsilon^k \rangle = -\langle \sigma_k \sigma_j e_\epsilon^k, e_\epsilon^k \rangle = -\epsilon \langle \sigma_j e_\epsilon^k, e_\epsilon^k \rangle.$$

This gives (2.2) for  $j \neq k$  since  $\epsilon \neq 0$ . □

Recall that  $A \in M_2(\mathbb{C})^{\otimes n}$  has the form

$$A = \sum_{\mathbf{s} \in \{0,1,2,3\}^n} \widehat{A}_{\mathbf{s}} \sigma_{\mathbf{s}}. \tag{2.3}$$

It will be more convenient for us to rewrite it as

$$A = \sum_{l \geq 0} \sum_{\kappa_1, \dots, \kappa_l \in \{1,2,3\}} \sum_{1 \leq i_1 < \dots < i_l \leq n} a_{i_1, \dots, i_l}^{\kappa_1, \dots, \kappa_l} \sigma_{i_1, \dots, i_l}^{\kappa_1, \dots, \kappa_l} \tag{2.4}$$

where to each  $(l; \kappa_1, \dots, \kappa_l; i_1 < \dots < i_l)$ , we associate it with  $\mathbf{s} = (s_1, \dots, s_n) \in \{0, 1, 2, 3\}^n$  of length  $|\mathbf{s}| = l$  with

$$s_k = \begin{cases} \kappa_j, & k = i_j, 1 \leq j \leq l \\ 0, & \text{otherwise} \end{cases}$$

so that

$$\widehat{A}_{\mathbf{s}} = a_{i_1, \dots, i_l}^{\kappa_1, \dots, \kappa_l} \quad \text{and} \quad \sigma_{\mathbf{s}} = \sigma_{i_1, \dots, i_l}^{\kappa_1, \dots, \kappa_l}.$$

In other words,  $\sigma_{i_1, \dots, i_l}^{\kappa_1, \dots, \kappa_l}$  is defined as

$$\sigma_{i_1, \dots, i_l}^{\kappa_1, \dots, \kappa_l} := \dots \otimes \sigma_{\kappa_l} \otimes \dots \otimes \sigma_{\kappa_1} \otimes \dots,$$

where  $\sigma_{\kappa_j}$  appears in the  $i_j$ -th place for each  $1 \leq j \leq l$ , and all other  $(n-l)$  components are simply identity matrices  $\sigma_0$ .

The main ingredient of our method is the following Proposition 2.2. To make the statement brief, we shall use the map

$$q = q_n : \{\mathbf{s} \in \{0, 1, 2, 3\}^n\} \rightarrow \{S \subset [3n]\},$$

that realizes the above identification of (2.3) and (2.4). If  $\mathbf{s} = \mathbf{0}$  is the 0 vector, then we define  $q(\mathbf{0}) := \emptyset$ . Each  $\mathbf{s} = (s_1, \dots, s_n) \in \{0, 1, 2, 3\}^n$  with  $1 \leq |\mathbf{s}| = l \leq n$  is assigned to a tuple  $(\kappa_1, \dots, \kappa_l; i_1 < \dots < i_l) \in [3]^l \times [n]^l$  with  $s_{i_j} = \kappa_j$ . Then we define

$$q(\mathbf{s}) := \{n(\kappa_j - 1) + i_j : 1 \leq j \leq l\} =: S.$$

For example, if  $n = 5$  and  $\mathbf{s} = (0, 1, 2, 3, 1)$ , then

$$q(\mathbf{s}) = S = \{2, 5 + 3, 10 + 4, 5\} \subset [15].$$

Note that  $|S| = |\mathbf{s}| = l$ . So

$$q(\{\mathbf{s} \in \{0, 1, 2, 3\}^n : |\mathbf{s}| \leq d\}) \subset \{S \subset [3n] : |S| \leq d\}.$$

Moreover, this map  $q : \{0, 1, 2, 3\}^n \rightarrow \{S \subset [3n]\}$  is injective but not surjective. We denote by  $p = p_n$  its inverse over  $\text{Im}(q) \subset 2^{[3n]}$  (here and in what follows, we use  $2^{[3n]}$  to denote the family of subsets of  $[3n]$ )

$$\mathbf{s} = p(S), \quad S \in \text{Im}(q) \subset 2^{[3n]}.$$

Therefore, the above formulae (2.3) and (2.4) can also be rewritten as

$$A = \sum_{S \in \text{Im}(q) \subset 2^{[3n]}} \widehat{A}_{p(S)} \sigma_{p(S)}. \tag{2.5}$$

**Proposition 2.2** Fix  $n \geq 1$ . There exists a family

$$\mathcal{S} = \mathcal{S}_n = \{\rho(\vec{\epsilon}) : \vec{\epsilon} \in \{-1, 1\}^{3n}\}$$

of density matrices in  $M_2(\mathbb{C})^{\otimes n}$  such that the following holds: For any  $A \in M_2(\mathbb{C})^{\otimes n}$ , the function  $f_A : \{-1, 1\}^{3n} \rightarrow \mathbb{C}$  defined by

$$f_A(\vec{\epsilon}) = \text{tr}[A\rho(\vec{\epsilon})], \quad \vec{\epsilon} \in \{-1, 1\}^{3n}$$

satisfies

$$\widehat{f}_A(S) = \begin{cases} 3^{-|S|}\widehat{A}_\rho(S), & S \in \text{Im}(q) \subset 2^{[3n]} \\ 0, & S \in 2^{[3n]} \setminus \text{Im}(q) \end{cases} \tag{2.6}$$

In particular,  $\|f_A\|_\infty \leq \|A\|$ ,  $\text{deg}(f_A) = \text{deg}(A)$  and for all  $d \geq 0$

$$\left( \sum_{s \in \{0,1,2,3\}^n: |s| \leq d} |\widehat{A}_s|^r \right)^{1/r} \leq 3^d \left( \sum_{S \subset [3n]: |S| \leq d} |\widehat{f}_A(S)|^r \right)^{1/r}, \quad 0 < r < \infty. \tag{2.7}$$

**Proof** For any

$$\vec{\epsilon} := \left( \epsilon_1^{(1)}, \dots, \epsilon_n^{(1)}, \epsilon_1^{(2)}, \dots, \epsilon_n^{(2)}, \epsilon_1^{(3)}, \dots, \epsilon_n^{(3)} \right) \in \{-1, 1\}^{3n},$$

the matrix  $\rho(\vec{\epsilon})$  is defined as follows

$$\rho := \rho(\vec{\epsilon}) = \rho_1 \otimes \dots \otimes \rho_n \in M_2(\mathbb{C})^{\otimes n},$$

where for each  $1 \leq j \leq n$

$$\rho_j := \rho_j(\vec{\epsilon}) = \frac{1}{3} |e_{\epsilon_j^{(1)}}^1\rangle\langle e_{\epsilon_j^{(1)}}^1| + \frac{1}{3} |e_{\epsilon_j^{(2)}}^2\rangle\langle e_{\epsilon_j^{(2)}}^2| + \frac{1}{3} |e_{\epsilon_j^{(3)}}^3\rangle\langle e_{\epsilon_j^{(3)}}^3|.$$

Recall for any  $\kappa \in \{1, 2, 3\}$  and  $\epsilon \in \{-1, 1\}$ ,  $e_\epsilon^\kappa$  is a unit vector. Hence each  $\rho_j(\vec{\epsilon})$  is positive semi-definite with trace 1. So  $\rho(\vec{\epsilon})$  is also a density matrix.

To prove (2.6), let us employ the notation (2.4). Then for  $\rho = \rho(\vec{\epsilon})$  defined as above

$$\text{tr}[A\rho] = \sum_{l \geq 0} \sum_{\kappa_1, \dots, \kappa_l \in \{1,2,3\}} \sum_{1 \leq i_1 < \dots < i_l \leq n} d_{i_1, \dots, i_l}^{\kappa_1, \dots, \kappa_l} \text{tr} \left[ \sigma_{i_1, \dots, i_l}^{\kappa_1, \dots, \kappa_l} \rho \right].$$

By definition,

$$\text{tr} \left[ \sigma_{i_1, \dots, i_l}^{\kappa_1, \dots, \kappa_l} \rho \right] = \text{tr}[\sigma_{\kappa_1} \rho_{i_1}] \cdots \text{tr}[\sigma_{\kappa_l} \rho_{i_l}] \prod_{j \notin \{i_1, \dots, i_l\}} \text{tr}[\sigma_0 \rho_j]$$



$$= \text{tr}[\sigma_{\kappa_1} \rho_{i_1}] \cdots \text{tr}[\sigma_{\kappa_l} \rho_{i_l}].$$

By Lemma 2.1, for each  $1 \leq \kappa \leq 3$  and  $\epsilon \in \{-1, 1\}$

$$\text{tr}[\sigma_{\kappa_j} |e_\epsilon^\kappa\rangle\langle e_\epsilon^\kappa|] = \langle \sigma_{\kappa_j} e_\epsilon^\kappa, e_\epsilon^\kappa \rangle = \epsilon \delta_{\kappa_j \kappa}.$$

Thus (recall that  $\kappa_j \in \{1, 2, 3\}$ )

$$\text{tr}[\sigma_{\kappa_j} \rho_{i_j}] = \frac{1}{3} \sum_{\kappa=1}^3 \text{tr} \left[ \sigma_{\kappa_j} |e_{\epsilon_{i_j}^{(\kappa)}}^\kappa\rangle\langle e_{\epsilon_{i_j}^{(\kappa)}}^\kappa| \right] = \frac{1}{3} \sum_{\kappa=1}^3 \epsilon_{i_j}^{(\kappa)} \delta_{\kappa_j \kappa} = \frac{1}{3} \epsilon_{i_j}^{(\kappa_j)}.$$

So we have shown that

$$\text{tr} \left[ \sigma_{i_1, \dots, i_l}^{\kappa_1, \dots, \kappa_l} \rho \right] = \text{tr}[\sigma_{\kappa_1} \rho_{i_1}] \cdots \text{tr}[\sigma_{\kappa_l} \rho_{i_l}] = \frac{1}{3^l} \epsilon_{i_1}^{(\kappa_1)} \cdots \epsilon_{i_l}^{(\kappa_l)},$$

and thus

$$f_A(\vec{\epsilon}) = \sum_{l \geq 0} \sum_{\kappa_1, \dots, \kappa_l \in \{1, 2, 3\}} \sum_{1 \leq i_1 < \dots < i_l \leq n} \frac{1}{3^l} a_{i_1, \dots, i_l}^{\kappa_1, \dots, \kappa_l} \epsilon_{i_1}^{(\kappa_1)} \cdots \epsilon_{i_l}^{(\kappa_l)}.$$

This is nothing but (2.6). To see the rest of statements, note first that by definition and Hölder’s inequality:

$$|f_A(\vec{\epsilon})| \leq \text{tr}|\rho(\vec{\epsilon})| \cdot \|A\| = \|A\|.$$

The desired  $\text{deg}(f_A) = \text{deg}(A)$  and the inequality (2.7) follow immediately from (2.6) and the fact that  $|p(S)| = |S|$ . □

### 3 Proof of Theorem 1.2 and Corollary 1.6

In this section we prove Theorem 1.2 and Corollary 1.6.

**Proof of Theorem 1.2** Using the notation (2.5) we need to prove

$$\left( \sum_{S \in \text{Im}(q) \subset 2^{[3n]}: |S| \leq d} |\widehat{A}_{p(S)}|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leq 3^d \text{BH}_{\pm 1}^{\leq d} \|A\|$$

for all

$$A = \sum_{S \in \text{Im}(q) \subset 2^{[3n]}: |S| \leq d} \widehat{A}_{p(S)} \sigma_{p(S)} \in M_2(\mathbb{C})^{\otimes n}.$$

Let  $f_A : \{-1, 1\}^{3n} \rightarrow \mathbb{C}$  be the function associated to  $A \in M_2(\mathbb{C})^{\otimes n}$  constructed in Proposition 2.2. So it is also of degree at most  $d$ . Then the desired result follows from the following chain of inequalities:

$$\begin{aligned} \left( \sum_{S \in \text{Im}(q) \subset 2^{[3n]}; |S| \leq d} |\widehat{A}_{p(S)}|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} &\leq 3^d \left( \sum_{S \subset [3n]; |S| \leq d} |\widehat{f}_A(S)|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \\ &\leq 3^d \text{BH}_{\pm 1}^{\leq d} \|f_A\|_{\infty} \\ &\leq 3^d \text{BH}_{\pm 1}^{\leq d} \|A\|, \end{aligned}$$

where in the first and last inequalities we used Proposition 2.2, and in the second inequality we used the Bohnenblust–Hille inequality on  $\{-1, 1\}^n$ .  $\square$

Note that Corollary 1.3 follows immediately from Theorem 1.2. Now we prove Corollary 1.6.

**Proof of Corollary 1.6** Fix  $\eta > 0$ . Consider the following set

$$\mathcal{A}_{\eta} := \{s \in \{0, 1, 2, 3\}^n : |\widehat{A}_s| > \eta\}.$$

Recall that  $\|A\| \leq 1$ . Then by Markov’s inequality and noncommutative Bohnenblust–Hille inequality (denoting  $|\mathcal{A}_{\eta}|$  the cardinality of  $\mathcal{A}_{\eta}$ )

$$|\mathcal{A}_{\eta}| \leq \eta^{-\frac{2d}{d+1}} \sum_{s \in \mathcal{A}_{\eta}} |\widehat{A}_s|^{\frac{2d}{d+1}} \leq \eta^{-\frac{2d}{d+1}} \left( \text{BH}_{M_2(\mathbb{C})}^{\leq d} \right)^{\frac{2d}{d+1}}.$$

Then  $B_{\eta} := \sum_{s \in \mathcal{A}_{\eta}} \widehat{A}_s \sigma_s$  depends on at most  $k$  coordinates with

$$k \leq d|\mathcal{A}_{\eta}| \leq d\eta^{-\frac{2d}{d+1}} \left( \text{BH}_{M_2(\mathbb{C})}^{\leq d} \right)^{\frac{2d}{d+1}}.$$

Again, by noncommutative Bohnenblust–Hille inequality and the fact that  $\|A\| \leq 1$

$$\|A - B_{\eta}\|_2^2 = \sum_{s \notin \mathcal{A}_{\eta}} |\widehat{A}_s|^2 \leq \eta^{\frac{2}{d+1}} \sum_{s \notin \mathcal{A}_{\eta}} |\widehat{A}_s|^{\frac{2d}{d+1}} \leq \eta^{\frac{2}{d+1}} \left( \text{BH}_{M_2(\mathbb{C})}^{\leq d} \right)^{\frac{2d}{d+1}}.$$

Now choose  $B = B_{\eta}$  with

$$\eta := \epsilon^{d+1} \left( \text{BH}_{M_2(\mathbb{C})}^{\leq d} \right)^{-d}.$$

Then  $\|A - B\|_2 \leq \epsilon$ , where  $B$  is a  $k$ -junta with

$$k \leq d\eta^{-\frac{2d}{d+1}} \left( \text{BH}_{M_2(\mathbb{C})}^{\leq d} \right)^{\frac{2d}{d+1}} = \epsilon^{-2d} d \left( \text{BH}_{M_2(\mathbb{C})}^{\leq d} \right)^{2d}.$$

This finishes the proof, since  $\text{BH}_{M_2(\mathbb{C})}^{\leq d} \leq C^d$  for some universal  $C > 0$  by Theorem 1.2. □

### 4 Learning quantum observables of low degree

Let us review the classical learning model first. Suppose that we want to learn a class  $\mathcal{F}$  of functions on  $\{-1, 1\}^n$  using the random query model which we shall explain. For  $N \geq 1$ , let  $X_1, \dots, X_N$  be  $N$  i.i.d. random variables uniformly distributed on  $\{-1, 1\}^n$ . Then how many random queries do we need to recover functions  $f \in \mathcal{F}$  nicely? More precisely, fix the error parameters  $\epsilon, \delta \in (0, 1)$ , what is the least number  $N = N(\mathcal{F}, \epsilon, \delta) > 0$  such that for any  $f \in \mathcal{F}$  together with  $N$  random queries

$$(X_1, f(X_1)), \dots, (X_N, f(X_N))$$

one can construct a random function  $h : \{-1, 1\}^n \rightarrow \mathbb{R}$  such that

$$\|f - h\|_2^2 \leq \epsilon$$

with probability at least  $1 - \delta$ ? Here  $\|f\|_2$  denotes the  $L^2$ -norm of  $f$  with respect to the uniform probability measure on  $\{-1, 1\}^n$ . When  $\mathcal{F} = \mathcal{F}_n^{\leq d}$  consists of functions  $f : \{-1, 1\}^n \rightarrow [-1, 1]$  of degree at most  $d$ , Eskenazis and Ivanisvili [14] proved that

$$N(\mathcal{F}_n^{\leq d}, \epsilon, \delta) \leq \frac{1}{\epsilon^{d+1}} \left( \log \frac{n}{\delta} \right) C(d),$$

where  $C(d)$  depends on the Bohnenblust–Hille constant  $\text{BH}_{\{\pm 1\}}^{\leq d}$  for functions on  $\{-1, 1\}^n$  of degree at most  $d$ . So logarithmic number  $\mathcal{O}_{\epsilon, \delta, d}(\log(n))$  of random queries is sufficient to learn bounded low-degree polynomials in the above sense. This improves significantly the previous work [17, 18] on the dimension-dependence of  $N(\mathcal{F}_n^{\leq d}, \epsilon, \delta)$ . Later on, Eskenazis et al. [15] proved that  $\mathcal{O}_{\epsilon, \delta, d}(\log(n))$  is also necessary.

Now suppose that we want to learn quantum observable  $A$  over  $n$ -qubits, i.e.  $A \in M_2(\mathbb{C})^{\otimes n}$ , of degree at most  $d$  with  $\|A\| \leq 1$ .

Our learning model is similar to the classical setting. The random queries are now replaced with

$$\text{tr}[A\rho], \quad \rho \sim \mathcal{S},$$

where  $\rho$  samples uniformly in some set  $\mathcal{S}$  of density matrices in  $M_2(\mathbb{C})^{\otimes n}$ . Our hope is to build another (random) observable  $\tilde{A}$  out of  $N$  random queries such that

$$\|\tilde{A} - A\|_2^2 \leq \epsilon \tag{4.1}$$

with probability at least  $1 - \delta$ . Again, the question is, how many random queries  $N = N(\epsilon, \delta, d, n)$  do we need to accomplish this, and how does  $N$  depend on  $n$ ?

In the remaining part of this section we provide one answer to this question with  $\mathcal{S} = \mathcal{S}_n$  constructed in Proposition 2.2.

**Theorem 4.1** *Suppose that  $A \in M_2(\mathbb{C})^{\otimes n}$  is of degree at most  $d$  and  $\|A\| \leq 1$ . Fix  $\delta, \epsilon \in (0, 1)$  and*

$$N \geq \frac{C^d \sqrt{d \log d}}{\epsilon^{d+1}} \log \left( \frac{n}{\delta} \right),$$

with  $C > 0$  large enough. Then given any  $N$  random density matrices  $\rho^{(m)}, 1 \leq m \leq N$  independently and uniformly sampled in  $\mathcal{S}_n$ , as well as random queries

$$\left( \rho^{(m)}, \text{tr}[A\rho^{(m)}] \right), \quad \rho^{(m)} \sim \mathcal{S}_n$$

we can construct a random  $\tilde{A} \in M_2(\mathbb{C})^{\otimes n}$  such that  $\|A - \tilde{A}\|_2^2 \leq \epsilon$  with probability at least  $1 - \delta$ .

**Proof** Again, we will reduce the problem to commutative case, which we shall explain how. To any

$$A = \sum_{S \in \text{Im}(q) \subset 2^{[3n]}; |S| \leq d} \hat{A}_{p(S)} \sigma_{p(S)} \in M_2(\mathbb{C})^{\otimes n},$$

we associate it with the function  $f_A : \{-1, 1\}^{3n} \rightarrow \mathbb{C}$  constructed in Proposition 2.2. Recall that  $f_A$  is of degree at most  $d$ ,  $\|f_A\|_\infty \leq 1$ , and

$$f_A = \text{tr}[A\rho(\cdot)] = \sum_{S \in \text{Im}(q) \subset 2^{[3n]}; |S| \leq d} 3^{-|S|} \hat{A}_{p(S)} \chi_S.$$

Suppose that we have  $\rho(\vec{x}(m)) \in \mathcal{S}_n, 1 \leq m \leq N$  as our independent random density matrices and

$$\left( \rho(\vec{x}(m)), \text{tr}[A\rho(\vec{x}(m))] \right), \quad 1 \leq m \leq N$$

as random queries, where  $\vec{x}(m), 1 \leq m \leq N$  are i.i.d. random variables uniformly distributed on  $\{-1, 1\}^{3n}$ . Similar to the commutative setting [15], we approximate

$$\hat{f}_A(S) = 3^{-|S|} \hat{A}_{p(S)}, \quad S \in \text{Im}(q), \quad |S| \leq d$$

with the empirical Walsh coefficients

$$\alpha_S := \frac{1}{N} \sum_{m=1}^N f_A(\vec{x}(m)) \chi_S(\vec{x}(m)) = \frac{1}{N} \sum_{m=1}^N \text{tr}[A\rho(m)] \chi_S(\vec{x}(m)).$$

Of course  $\mathbf{E}\alpha_S = \widehat{f}_A(S)$  where  $\mathbf{E}$  is with respect to the uniform distribution. Since  $\alpha_S$  is the sum of i.i.d. random variables, we get by  $\|f_A\|_\infty \leq 1$  and the Chernoff bound that for  $b > 0$

$$\mathbf{P}\{|\alpha_S - \widehat{f}_A(S)| > b\} \leq 2 \exp\left(-Nb^2/2\right), \forall S \in \text{Im}(q), |S| \leq d. \tag{4.2}$$

Note that

$$\begin{aligned} |\{S \in \text{Im}(q) : |S| \leq d\}| &= |\{\mathbf{s} \in \{0, 1, 2, 3\}^n : |\mathbf{s}| \leq d\}| \\ &= \sum_{l=0}^d 3^l \binom{n}{l} \leq 3^d \sum_{l=0}^d \binom{n}{l}. \end{aligned}$$

Then by (4.2) and the union bound

$$\begin{aligned} \mathbf{P}\{\exists S \in \text{Im}(q), |S| \leq d : |\alpha_S - \widehat{f}_A(S)| > b\} \\ \leq 2e^{-Nb^2/2} |\{S \in \text{Im}(q) : |S| \leq d\}| \\ \leq 2e^{-Nb^2/2} 3^d \sum_{l=0}^d \binom{n}{l}. \end{aligned}$$

Choosing

$$N \geq \frac{2}{b^2} \log\left(\frac{2 \cdot 3^d}{\delta} \sum_{l=0}^d \binom{n}{l}\right), \tag{4.3}$$

one achieves

$$\mathbf{P}\{|\alpha_S - \widehat{f}_A(S)| \leq b, \forall S \in \text{Im}(q), |S| \leq d\} \geq 1 - \delta. \tag{4.4}$$

We continue to copycat [14] and introduce the random sets

$$\mathcal{S}_b := \{S \in \text{Im}(q), |S| \leq d : |\alpha_S| \geq 2b\}.$$

In view of (4.4), with probability  $\geq 1 - \delta$ :

$$\begin{cases} |\widehat{f}_A(S)| \leq |\alpha_S| + |\alpha_S - \widehat{f}_A(S)| < 3b, & \text{if } S \notin \mathcal{S}_b \\ |\widehat{f}_A(S)| \geq |\alpha_S| - |\alpha_S - \widehat{f}_A(S)| \geq b, & \text{if } S \in \mathcal{S}_b \end{cases}. \tag{4.5}$$

The second line of (4.5), together with  $\|f_A\|_\infty \leq 1$  and the commutative Bohnenblust–Hille inequality (1.3), yields

$$|\mathcal{S}_b| \leq b^{-\frac{2d}{d+1}} \sum_{S \in \mathcal{S}_b} |\widehat{f}_A(S)|^{\frac{2d}{d+1}} \leq b^{-\frac{2d}{d+1}} \left(\text{BH}_{\{\pm 1\}}^{\leq d}\right)^{\frac{2d}{d+1}}.$$

Fix  $b$  as above. Consider the random polynomial in  $M_2(\mathbb{C})^{\otimes n}$

$$A_b := \sum_{S \in \mathcal{S}_b} 3^{|S|} \alpha_S \sigma_{p(S)} = \sum_{q(s) \in \mathcal{S}_b} 3^{|s|} \alpha_{q(s)} \sigma_s.$$

All combined, we have with probability at least  $1 - \delta$  that

$$\begin{aligned} \|A - A_b\|_2^2 &\leq 3^{2d} \sum_{S \in \mathcal{S}_b} |\alpha_S - \widehat{f}_A(S)|^2 + 3^{2d} \sum_{S \notin \mathcal{S}_b} |\widehat{f}_A(S)|^2 \\ &\leq 3^{2d} b^2 |\mathcal{S}_b| + 3^{2d} (3b)^{\frac{2}{d+1}} \sum_{S \notin \mathcal{S}_b} |\widehat{f}_A(S)|^{\frac{2d}{d+1}} \\ &\leq \left(3^{d+1} \text{BH}_{\{\pm 1\}}^{\leq d}\right)^{\frac{2d}{d+1}} \left(b^2 \cdot b^{-\frac{2d}{d+1}} + (3b)^{\frac{2}{d+1}}\right) \\ &\leq 10 \left(3^{d+1} \text{BH}_{\{\pm 1\}}^{\leq d}\right)^{\frac{2d}{d+1}} b^{\frac{2}{d+1}}. \end{aligned}$$

To get an error bound  $\|A - A_b\|_2^2 \leq \epsilon$ , we choose

$$b = 10^{-\frac{d+1}{2}} \left(3^{d+1} \text{BH}_{\{\pm 1\}}^{\leq d}\right)^{-d} \epsilon^{\frac{d+1}{2}}.$$

Inserting this into (4.3), we choose  $N$  such that

$$N \geq \frac{2 \cdot 10^{d+1} \left(3^{d+1} \text{BH}_{\{\pm 1\}}^{\leq d}\right)^{2d}}{\epsilon^{d+1}} \log \left( \frac{2 \cdot 3^d}{\delta} \sum_{l=0}^d \binom{n}{l} \right).$$

Noting moreover that (see for example [14])

$$\sum_{l=0}^d \binom{n}{l} \leq \left(\frac{en}{d}\right)^d,$$

we may choose

$$N \geq \frac{C^{d\sqrt{d\log d}}}{\epsilon^{d+1}} \log \left(\frac{n}{\delta}\right),$$

for some  $C > 0$  large enough. Here, the Bohenblust–Hille constant  $\text{BH}_{\{\pm 1\}}^{\leq d}$  is contained in  $C^{d\sqrt{d\log d}}$ . With this choice of  $N$ , the random polynomial  $\widetilde{A} := A_b$  above satisfies

$$\|A - \widetilde{A}\|_2^2 \leq \epsilon,$$

with probability  $\geq 1 - \delta$ . □

### 5 Bohr’s radius phenomenon on quantum Boolean cubes

One important application of classical Bohnenblust–Hille inequalities is to study *Bohr’s radius problem* [5]. The original problem [2] concerns the  $n$ -dimensional torus  $\mathbb{T}^n$  with  $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$  and the exact asymptotic behaviour of Bohr’s radius was obtained by Bayart et al. [7] using the polynomial version of Bohnenblust–Hille inequalities (1.2) with the best constants (denoted by  $BH_{\mathbb{T}}^{\leq d}$ ) of subexponential growth in the degree  $d$ . See also [9]. A Boolean analogue of the problem was studied by Defant et al. in [11], where Bohr’s radius is replaced by *Boolean’s radius*.

**Definition 5.1** *Boolean’s radius* of a function  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$  is the positive real number  $Br_n(f)$  such that

$$\sum_{S \subset [n]} |\widehat{f}(S)| Br_n(f)^{|S|} = \|f\|_{\infty}.$$

Given a class  $\mathcal{F}_n$  of functions on  $\{-1, 1\}^n$ , the Boolean radius of  $\mathcal{F}_n$  is defined as

$$Br_n(\mathcal{F}_n) := \inf\{Br_n(f) : f \in \mathcal{F}_n\}.$$

Of particular interests to us are the following four classes of functions

(1)  $\mathcal{F}_n(\text{all})$ : all real functions on  $\{-1, 1\}^n$  with

$$Br_n(\text{all}) := Br_n(\mathcal{F}_n(\text{all}));$$

(2)  $\mathcal{F}_n(\text{hom})$ : all real homogeneous functions on  $\{-1, 1\}^n$  with

$$Br_n(\text{hom}) := Br_n(\mathcal{F}_n(\text{hom}));$$

(3)  $\mathcal{F}_n(= d)$ : all  $d$ -homogeneous real functions on  $\{-1, 1\}^n$  with

$$Br_n(= d) := Br_n(\mathcal{F}_n(= d));$$

(4)  $\mathcal{F}_n(\leq d)$ : all real functions on  $\{-1, 1\}^n$  of degree at most  $d$  with

$$Br_n(\leq d) := Br_n(\mathcal{F}_n(\leq d)).$$

The problem of Boolean’s radius is to determine the right order of decay of  $Br_n(\mathcal{F}_n)$  as  $n \rightarrow \infty$ . Among others, Defant, Mastyło and Pérez proved the following.

**Theorem 5.2** [11, Theorems 2.1 & 3.1 & 4.1 and Corollary 3.2] *For any  $1 \leq d \leq n$ , we have*

(1)  $Br_n(\text{all}) = 2^{1/n} - 1$  and thus  $\lim_{n \rightarrow \infty} n Br_n(\text{all}) = \log 2$ ;

(2) there exists  $C > 1$  such that for all

$$c_d n^{\frac{1}{2n}} \binom{n}{d}^{-\frac{1}{2d}} \leq \text{Br}_n(= d) \leq C_d n^{\frac{1}{2n}} \binom{n}{d}^{-\frac{1}{2d}},$$

where

$$c_d = \frac{1}{d^{\frac{1}{2d}} C \sqrt{\frac{\log d}{d}}} \quad \text{and} \quad C_d = C^{\frac{1}{d}};$$

(3)  $\lim_{n \rightarrow \infty} \sqrt{\frac{n}{\log n}} \text{Br}_n(\text{hom}) = 1$ ;

(4) there exist  $c'_d, C'_d > 0$  such that

$$\frac{c'_d}{n^{\frac{1}{2}}} \leq \text{Br}_n(\leq d) \leq \frac{C'_d}{n^{\frac{d-1}{2d}}}.$$

We refer to [11] for more discussions on the similarity and differences between the Boolean cube case and torus case.

The concept of Boolean radius carries over to the quantum setting without any difficulties.

**Definition 5.3** The *quantum Boolean radius* of self-adjoint  $A \in M_2(\mathbb{C})^{\otimes n}$  is the positive real number  $\text{qBr}_n(A)$  such that

$$\sum_{s \in \{0,1,2,3\}^n} |\widehat{A}_s| \text{qBr}_n(A)^{|s|} = \|A\|.$$

Similarly, one may define for a class  $\mathcal{F}'_n$  of self-adjoint matrices in  $M_2(\mathbb{C})^{\otimes n}$

$$\text{qBr}_n(\mathcal{F}'_n) := \inf\{\text{qBr}_n(A) : A \in \mathcal{F}'_n\}.$$

**Remark 5.4** In [11] only the real-valued functions were considered and some arguments rely on the real structure. We are not going to discuss the possible extension to the complex-valued functions here. It is for this reason we require  $A$  to be self-adjoint. In fact, when  $A$  is self-adjoint, the function  $f_A = \text{tr}[A\rho(\cdot)]$  on  $\{-1, 1\}^{3n}$  constructed in Proposition 2.2 is real-valued and has real Fourier coefficients, so that we can use the results in [11] directly. We leave the problem for general  $A$  to future study.

If  $\mathcal{F}_n$  denotes one of the four classes of functions (1–4) on  $\{-1, 1\}^n$  listed as above, then we use  $\mathcal{F}_n^q$  to denote the quantum counterpart of class of matrices in  $M_2(\mathbb{C})^{\otimes n}$ . For example, if  $\mathcal{F}_n = \mathcal{F}_n(\leq d)$  is the class of polynomials on  $\{-1, 1\}^n$  of degree at most  $d$ , then  $\mathcal{F}_n^q = \mathcal{F}_n^q(\leq d)$  denotes the class of self-adjoint  $A \in M_2(\mathbb{C})^{\otimes n}$  of degree at most  $d$ . Then our main result on quantum Boolean radius is the following.



**Theorem 5.5** *For any  $1 \leq d \leq n$  and any  $\mathcal{F}_n$  of the four classes of functions (1–4) listed above, we have*

$$\text{qBr}_n(\mathcal{F}_n^q) \leq \text{Br}_n(\mathcal{F}_n) \quad \text{and} \quad \text{Br}_{3n}(\mathcal{F}_{3n}) \leq 3\text{qBr}_n(\mathcal{F}_n^q).$$

**Proof** The first inequality is trivial, as  $\mathcal{F}_n$  can be viewed as a subset of  $\mathcal{F}_n^q$  with all the relevant structures (norm, degree etc.) preserved. In fact, for any  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$  with Fourier expansion

$$f(x_1, \dots, x_n) = \sum_{0 \leq l \leq n} \sum_{1 \leq i_1 < \dots < i_l \leq n} a_{i_1, \dots, i_l} x_{i_1} \dots x_{i_l},$$

consider the self-adjoint matrix

$$A_f = \sum_{0 \leq l \leq n} \sum_{1 \leq i_1 < \dots < i_l \leq n} a_{i_1, \dots, i_l} \sigma_{i_1, \dots, i_l}^{3, \dots, 3}.$$

Clearly,  $\text{deg}(f) = \text{deg}(A_f)$  and  $A_f \in \mathcal{F}_n^q$  whenever  $f \in \mathcal{F}_n$ . Note that the canonical basis of  $M_2(\mathbb{C})$  are eigenvectors of  $\sigma_3$  corresponding to eigenvalues  $1, -1$ , respectively. Let us denote this basis  $\{e_1, e_{-1}\}$ . Then under the basis  $\{e_{x_1} \otimes \dots \otimes e_{x_n} : x_j \in \{-1, 1\}, 1 \leq j \leq n\}$ ,  $A_f$  is a diagonal matrix with diagonal entries being

$$\langle A_f(e_{x_1} \otimes \dots \otimes e_{x_n}), e_{x_1} \otimes \dots \otimes e_{x_n} \rangle = \sum_{0 \leq l \leq n} \sum_{1 \leq i_1 < \dots < i_l \leq n} a_{i_1, \dots, i_l} x_{i_1} \dots x_{i_l}.$$

So  $\|f\|_\infty = \|A_f\|$ . By definition, we have  $\text{Br}_n(f) = \text{qBr}_n(A_f)$  and this proves the first inequality.

To prove the second inequality, we appeal to our reduction method again. In fact, take any  $A = A^* \in \mathcal{F}_n^q \subset M_2(\mathbb{C})^{\otimes n}$ . Consider the function  $f_A : \{-1, 1\}^{3n} \rightarrow \mathbb{R}$  constructed in Proposition 2.2 which belongs to  $\mathcal{F}_{3n}$  whenever  $\mathcal{F}_n$  is one of four aforementioned classes (1–4). Take the class (3)  $\mathcal{F}_n (= d)$  for example. By construction,  $f_A$  is  $d$ -homogeneous if  $A$  is. Recall that  $\|f_A\|_\infty \leq \|A\|$ , and by definition of  $\text{qBr}_n(A)$ :

$$\sum_s |\widehat{A}_s| \text{qBr}_n(A)^{|s|} = \sum_{S \in \text{Im}(q) \subset 2^{[3n]}} |\widehat{A}_{p(S)}| \text{qBr}_n(A)^{|S|} = \|A\|.$$

Then for  $f_A : \{-1, 1\}^{3n} \rightarrow \mathbb{R}$ , we have

$$\begin{aligned} \sum_{S \in \text{Im}(q) \subset 2^{[3n]}} |\widehat{f}_A(S)| (3\text{qBr}_n(A))^{|S|} &= \sum_{S \in \text{Im}(q) \subset 2^{[3n]}} |\widehat{A}_{p(S)}| \text{qBr}_n(A)^{|S|} \\ &= \|A\| \geq \|f_A\|_\infty. \end{aligned}$$

Therefore, by definition of  $\text{Br}_{3n}(f_A)$ :

$$\text{Br}_{3n}(f_A) \leq 3\text{qBr}_n(A).$$

So for any  $A = A^* \in \mathcal{F}_n^q$  we find  $f_A \in \mathcal{F}_{3n}$  such that the above inequality holds. Therefore we get

$$\text{Br}_{3n}(\mathcal{F}_{3n}) \leq 3\text{qBr}_n(\mathcal{F}_n^q)$$

by definitions of  $\text{Br}_{3n}(\mathcal{F}_{3n})$  and  $\text{qBr}_n(\mathcal{F}_n^q)$ . This concludes the proof of the second inequality. □

**Remark 5.6** It is still possible to improve our estimates using the reduction method. For example, we have shown that for the class  $\mathcal{F}_n^q(\text{all})$  of all self-adjoint matrices in  $M_2(\mathbb{C})^{\otimes n}$

$$\frac{2^{1/(3n)} - 1}{3} \leq \text{qBr}_n(\mathcal{F}_n^q(\text{all})) \leq 2^{1/n} - 1.$$

But we can actually prove

$$\frac{2^{1/n} - 1}{9} \leq \text{qBr}_n(\mathcal{F}_n^q(\text{all})) \leq 2^{1/n} - 1.$$

In fact, for any  $A = A^* \in M_2(\mathbb{C})^{\otimes n}$  with  $\|A\| = 1$ , suppose  $f_A : \{-1, 1\}^{3n} \rightarrow [-1, 1]$  is constructed as before. Then for any  $\mathbf{0} \neq \mathbf{s} = p(S) \in \{0, 1, 2, 3\}^n$ , we have by [11, Lemma 2.2] that

$$|\widehat{A}_{\mathbf{0}}| + 3^{-|\mathbf{S}|}|\widehat{A}_{p(\mathbf{S})}| = |\widehat{f}_A(\emptyset)| + |\widehat{f}_A(S)| \leq 1.$$

So we get

$$\begin{aligned} \sum_{\mathbf{s} \in \{0,1,2,3\}^n} |\widehat{A}_{\mathbf{s}}| r^{|\mathbf{s}|} &\leq |\widehat{A}_{\mathbf{0}}| + (1 - |\widehat{A}_{\mathbf{0}}|) \sum_{\mathbf{s} \neq \mathbf{0}} (3r)^{|\mathbf{s}|} \\ &\leq |\widehat{A}_{\mathbf{0}}| + (1 - |\widehat{A}_{\mathbf{0}}|) \sum_{k=1}^n \binom{n}{k} (9r)^k \\ &= |\widehat{A}_{\mathbf{0}}| + (1 - |\widehat{A}_{\mathbf{0}}|) ((1 + 9r)^n - 1). \end{aligned}$$

This establishes the bound  $\text{qBr}_n(\mathcal{F}_n^q(\text{all})) \geq \frac{2^{1/n}-1}{9}$  by definition.

## 6 Discussions

We briefly compare our results with the work of Huang et al. [16].

- (1) In the work [16], the noncommutative Bohnenblust–Hille inequality (1.4) [16, Corollary 3] follows from a more general result [16, Theorem 5] which we do not have. The noncommutative Bohnenblust–Hille constant ( $C_d$  in (1.4)) we obtained is of exponential growth  $\sim C^d$ , which is better than theirs  $\sim d^{\mathcal{O}(d)}$ . Remark that

the best known bound for the (commutative) Boolean cubes is subexponential  $\sim C\sqrt{d\log d}$  [12].

- (2) Our proof of (1.4) is different from theirs. We use an argument that reduces the problem to the commutative case and the proof looks simpler, while their proof that is more self-contained, combines several technical estimates that can be useful to other problems.
- (3) Our better Bohnenblust–Hille constant yields an immediate improvement for learning quantum observable up to a small prediction error in their work. In particular, this allows us to remove the  $\log \log(1/\epsilon)$  factor in the exponent of [16, Eq. (A17)]. It also yields an immediate improvement in the sample complexity for learning arbitrary quantum process. This is because the learning is achieved by considering the unknown observable to be the observable after Heisenberg evolution under the unknown quantum process.

**Acknowledgements** We are grateful to Hsin-Yuan Huang, Sitan Chen and John Preskill for sharing their preprint and for their helpful comments on an earlier version of this paper, especially Sect. 6. We also would like to thank Alexandros Eskenazis for Corollary 1.6 and the anonymous referee for valuable comments and remarks.

**Data availability** All data generated or analysed during this study are included in this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Bohnenblust, H.F., Hille, E.: On the absolute convergence of Dirichlet series. *Ann. Math.* **32**, 600–622 (1931)
2. Boas, H.P., Khavinson, D.: Bohr's power series theorem in several variables. *Proc. Am. Math. Soc.* **125**(10), 2975–2979 (1997)
3. Blei, R.: *Analysis in Integer and Fractional Dimensions*, vol. 71. Cambridge University Press, Cambridge (2001)
4. Bohr, H.: Über die Bedeutung der Potenzreihen unendlich vieler Variablen in der Theorie der Dirichletschen reihen  $\sum \frac{a_n}{n^s}$ . *Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl.* 1913:441–488 (1913)
5. Bohr, H.: A theorem concerning power series. *Proc. Lond. Math. Soc.* **2**(13), 1–5 (1914)
6. Bourgain, J.: On the distribution of the Fourier spectrum of Boolean functions. *Isr. J. Math.* **131**(1), 269–276 (2002)
7. Bayart, F., Pellegrino, D., Seoane-Sepúlveda, J.B.: The Bohr radius of the  $n$ -dimensional polydisk is equivalent to  $\sqrt{(\log n)/n}$ . *Adv. Math.* **264**, 726–746 (2014)
8. Dinur, I., Friedgut, E., Kindler, G., O'Donnell, R.: On the Fourier tails of bounded functions over the discrete cube. *Isr. J. Math.* **160**(1), 389–412 (2007)
9. Defant, A., Frerick, L., Ortega-Cerdà, J., Ounaies, M., Seip, K.: The Bohnenblust–Hille inequality for homogeneous polynomials is hypercontractive. *Ann. Math.* **174**, 485–497 (2011)
10. Defant, A., García, D., Maestre, M., Sevilla-Peris, P.: *Dirichlet Series and Holomorphic Functions in High Dimensions*. New Mathematical Monographs, vol. 37. Cambridge University Press, Cambridge (2019)

11. Defant, A., Mastysłó, M., Pérez, A.: Bohr's phenomenon for functions on the Boolean cube. *J. Funct. Anal.* **275**(11), 3115–3147 (2018)
12. Defant, A., Mastysłó, M., Pérez, A.: On the Fourier spectrum of functions on Boolean cubes. *Math. Ann.* **374**(1), 653–680 (2019)
13. Defant, A., Sevilla-Peris, P.: The Bohnenblust–Hille cycle of ideas from a modern point of view. *Functiones et Approximatio Commentarii Mathematici* **50**(1), 55–127 (2014)
14. Eskenazis, A., Iwanisvili, P.: Learning low-degree functions from a logarithmic number of random queries. In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pp. 203–207 (2022)
15. Eskenazis, A., Iwanisvili, P., Streck, L.: Low-degree learning and the metric entropy of polynomials. *arXiv preprint [arXiv:2203.09659](https://arxiv.org/abs/2203.09659)* (2022)
16. Huang, H.-Y., Chen, S., Preskill, J.: Learning to predict arbitrary quantum processes. *arXiv preprint [arXiv:2210.14894](https://arxiv.org/abs/2210.14894)* (2022)
17. Iyer, S., Rao, A., Reis, V., Rothvoss, T., Yehudayoff, A.: Tight bounds on the Fourier growth of bounded functions on the hypercube. In: *ECCC 2021*. Preprint [arXiv: 2107.06309](https://arxiv.org/abs/2107.06309) (2021)
18. Linial, N., Mansour, Y., Nisan, N.: Constant depth circuits Fourier transform, and learnability. *J. Assoc. Comput. Mach.* **40**, 607–620 (1993)
19. Littlewood, J.E.: On bounded bilinear forms in an infinite number of variables. *Q. J. Math. Oxf. Ser.* **1**, 164–174 (1930)
20. Rouzé, C., Wirth, M., Zhang, H.: Quantum Talagrand, KKL and Friedgut's theorems and the learnability of quantum Boolean functions. *arXiv preprint [arXiv:2209.07279](https://arxiv.org/abs/2209.07279)* (2022)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.