On the Composition of Randomized Query Complexity and Approximate Degree

Sourav Chakraborty
Indian Statistical Institute, Kolkata, India

Chandrima Kayal
Indian Statistical Institute, Kolkata, India

Rajat Mittal
Indian Institute of Technology Kanpur, India

Manaswi Paraashar
Aarhus University, Denmark

Swagato Sanyal
Indian Institute of Technology Kharagpur, India

Nitin Saurabh
Indian Institute of Technology Hyderabad, India

Abstract

For any Boolean functions f and g, the question whether $\mathrm{R}(f \circ g) = \widetilde{\Theta}(\mathrm{R}(f) \cdot \mathrm{R}(g))$, is known as the composition question for the randomized query complexity. Similarly, the composition question for the approximate degree asks whether $\widetilde{\deg}(f \circ g) = \widetilde{\Theta}(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g))$. These questions are two of the most important and well-studied problems in the field of analysis of Boolean functions, and yet we are far from answering them satisfactorily.

It is known that the measures compose if one assumes various properties of the outer function f (or inner function g). This paper extends the class of outer functions for which R and $\widetilde{\deg}$ compose.

A recent landmark result (Ben-David and Blais, 2020) showed that $\mathrm{R}(f \circ g) = \Omega(\mathrm{noisyR}(f) \cdot \mathrm{R}(g))$. This implies that composition holds whenever $\mathrm{noisyR}(f) = \widetilde{\Theta}(\mathrm{R}(f))$. We show two results:

1. When $\mathrm{R}(f) = \Theta(n)$, then $\mathrm{noisyR}(f) = \Theta(\mathrm{R}(f))$. In other words, composition holds whenever the randomized query complexity of the outer function is full.
2. If R composes with respect to an outer function, then noisyR also composes with respect to the same outer function.

On the other hand, no result of the type $\widetilde{\deg}(f \circ g) = \Omega(M(f) \cdot \widetilde{\deg}(g))$ (for some non-trivial complexity measure $M(\cdot)$) was known to the best of our knowledge. We prove that

$$\widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\sqrt{\mathrm{bs}(f)} \cdot \widetilde{\deg}(g)),$$

where $\mathrm{bs}(f)$ is the block sensitivity of f. This implies that $\widetilde{\deg}$ composes when $\widetilde{\deg}(f)$ is asymptotically equal to $\sqrt{\mathrm{bs}(f)}$.

It is already known that both R and $\widetilde{\deg}$ compose when the outer function is symmetric. We also extend these results to weaker notions of symmetry with respect to the outer function.

2012 ACM Subject Classification Theory of computation

Keywords and phrases Approximate degree, Boolean functions, Composition Theorem, Partial functions, Randomized Query Complexity

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2023.63

Category RANDOM

Related Version *Full Version*: `https://eccc.weizmann.ac.il/report/2023/099/`

Funding *Swagato Sanyal*: supported by an ISIRD grant by Sponsored Research and Industrial Consultancy, IIT Kharagpur.

1 Introduction

For studying the complexity of Boolean functions, a number of simple complexity measures (like decision tree complexity, randomized query complexity, degree, certificate complexity and so on) have been studied over the years. (Refer to the survey [15] for an introduction to complexity measures of Boolean functions.) Understanding how these measures are related to each other [1, 2, 4, 27], and how they behave for various classes of Boolean functions has been a central area of research in complexity theory [34, 23, 42].

A crucial step towards understanding a complexity measure is: how does the complexity measure behave when two Boolean functions are combined to obtain a new function (i.e., what is the relationship between the measure of the resulting function and the measures of the two individual functions) [16, 12, 25, 44]? One particularly natural combination of functions is called *composition*, and it takes a central role in complexity theory.

For any two Boolean functions $f : \{0,1\}^n \to \{0,1\}$ and $g : \{0,1\}^m \to \{0,1\}$, the composed function $f \circ g : \{0,1\}^{nm} \to \{0,1\}$ is defined as the function

$$f \circ g(x_1, \ldots, x_n) = f(g(x_1), \ldots, g(x_n)),$$

where $x_i \in \{0,1\}^m$ for $i \in [n]$. For the function $f \circ g$, the function f is called the outer function and g is called the inner function. See Definition 13 for a natural extension to partial functions.

Let $M(\cdot)$ be a complexity measure of Boolean functions. A central question in complexity theory is the following.

▶ **Question 1** (Composition question for M). *Is the following true for all Boolean functions f and g:*

$$M(f \circ g) = \widetilde{\Theta}(M(f) \cdot M(g))?$$

The notation $\widetilde{\Theta}(\cdot)$ hides poly-logarithmic factors of the arity of the outer function f.

Composition of Boolean functions with respect to different complexity measures is a very important and useful tool in areas like communication complexity, circuit complexity and many more. To take an example, a popular application of composition is to create new functions demonstrating better separations (refer to [33, 44, 3, 25] for some other results of similar flavour).

It is known that the answer to the composition question is in the affirmative for complexity measures like deterministic decision tree complexity [37, 44, 31], degree [44] and quantum query complexity [35, 29, 28]. While it is well understood how several complexity measures behave under composition, there are two important measures (even though well studied) for which this problems remains wide open: randomized query complexity (denoted by R) and approximate degree (denoted by $\widetilde{\deg}$) [38, 33, 3, 39, 17, 40]. (See Definition 30 and Definition 31 for their respective formal definitions.)

For both R and $\widetilde{\deg}$ the upper bound inequality is known, i.e., $\mathrm{R}(f \circ g) = \widetilde{O}(\mathrm{R}(f) \cdot \mathrm{R}(g))$ (folklore) and $\widetilde{\deg}(f \circ g) = O(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g))$ [41]. Thus it is enough to prove the lower bound on the complexity of composed function in terms of the individual functions. Most of the attempts to prove this direction of the question have focused on special cases when the outer function has certain special properties[1].

[1] We note that some works have also studied the composition of R and $\widetilde{\deg}$ when the inner functions have special properties, for example, [1, 13, 6, 24, 30, 10].

The initial steps taken towards answering the composition question for R were to show a lower bound by using a weaker measure of outer function than the randomized query complexity. In particular, it was shown that $\mathrm{R}(f \circ g) = \Omega(\mathrm{s}(f) \cdot \mathrm{R}(g))$ [26, 5], where $\mathrm{s}(f)$ is the sensitivity of f (Definition 32). Since $\mathrm{s}(f) = \Theta(\mathrm{R}(f))$ for any symmetric function[2] f, these results show that R composes when the outer function is a symmetric function (like OR, AND, Majority, Parity, etc.). The lower bound was later improved to obtain $\mathrm{R}(f \circ g) = \Omega(\mathrm{fbs}(f) \cdot \mathrm{R}(g))$ [7, 8], where $\mathrm{fbs}(f)$ is the fractional block sensitivity of f (Definition 33).

Unfortunately, at this stage, there could even be a cubic gap between R and fbs; the best known bound is $\mathrm{R}(f) = O(\mathrm{fbs}(f)^3)$ [2]. Given that there are already known Boolean functions with quadratic gap between $\mathrm{fbs}(f)$ and $\mathrm{R}(f)$ (e.g., BKK function [1]), it is natural to consider composition question for randomized query complexity when R is *big* but fbs is *small*. We take a step towards this problem by showing that composition for R holds when the outer function has full randomized query complexity, i.e., $\mathrm{R}(f) = \Theta(n)$, where n is the arity of the outer function f.

For composition of $\widetilde{\deg}$, Sherstov [38] already showed that $\widetilde{\deg}(f \circ g)$ composes when the approximate degree of the outer function f is $\Theta(n)$, where n is the arity of the outer function. Thus approximate degree composes for several symmetric functions (having approximate degree $\Theta(n)$, like Majority and Parity). Though, until recently it was not even clear if $\widetilde{\deg}(\mathsf{OR} \circ \mathsf{AND}) = \Omega(\widetilde{\deg}(\mathsf{OR})\,\widetilde{\deg}(\mathsf{AND}))$ (arguably the simplest of composed functions). OR has approximate degree $O(\sqrt{n})$, and thus the result of [38] does not prove $\widetilde{\deg}$ composition when the outer function is OR (similarly for AND). In a long series of work by [33, 3, 39, 17, 40], the question was finally resolved; it was later generalized to the case when the outer function is symmetric [11].

In contrast to R composition, no lower bound on the approximate degree of composed function is known with a weaker measure on the outer function. It is well known that for any Boolean function f, $\sqrt{\mathrm{s}(f)} \leq \sqrt{\mathrm{bs}(f)} = O(\widetilde{\deg}(f))$ [33]. So a natural step towards proving $\widetilde{\deg}$ composition is: prove a lower bound on $\widetilde{\deg}(f \circ g)$ by $\sqrt{\mathrm{bs}(f)} \cdot \widetilde{\deg}(g)$. We show this result by generalizing the approach of [11].

While the techniques used for the composition of R and $\widetilde{\deg}$ are quite different, one can still observe similarities between the classes of outer functions for which the composition of R and $\widetilde{\deg}$ is known to hold respectively. We further expand these similarities, by extending the classes of outer functions for which the composition theorem hold.

2 Our Results and Techniques

It is well-known, by amplification, that $\mathrm{R}(f \circ g) = O(\mathrm{R}(f) \cdot \mathrm{R}(g) \cdot \log \mathrm{R}(f))$. In the case of approximate degree, Shrestov [41] showed that $\widetilde{\deg}(f \circ g) = O(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g))$. So, to answer the composition question for R (or $\widetilde{\deg}$), we are only concerned about proving a lower bound on $\mathrm{R}(f \circ g)$ (or $\widetilde{\deg}(f \circ g)$) in terms of $\mathrm{R}(f)$ and $\mathrm{R}(g)$ (or $\widetilde{\deg}(f)$ and $\widetilde{\deg}(g)$) respectively.

We split our results into three parts. In the first part we prove the tight lower bound on $\mathrm{R}(f \circ g)$ when the outer function has full randomized complexity. In the second part we provide a tight lower bound on $\widetilde{\deg}(f \circ g)$ in terms of $\mathrm{bs}(f)$ and $\widetilde{\deg}(g)$. Our results on the lower bound of $\mathrm{R}(f \circ g)$ and $\widetilde{\deg}(f \circ g)$ are summarized in Table 1. Finally, we also prove composition theorems for R and $\widetilde{\deg}$ when the outer functions have a slightly relaxed notion of symmetry.

[2] Functions that depend only on the Hamming weight of their input.

Table 1 Composition of R and $\widetilde{\deg}$ depending on the complexity of the outer function in terms of block-sensitivity and arity.

	In terms of $\mathrm{bs}(f)$	In terms of arity of f
R	$\mathrm{R}(f \circ g) = \widetilde{\Omega}(\mathrm{bs}(f) \cdot \mathrm{R}(g))$ [26]	$\mathrm{R}(f \circ g) = \widetilde{\Omega}(\mathrm{R}(f) \cdot \mathrm{R}(g))$ when $\mathrm{R}(f) = \Theta(n)$ Theorem 2
$\widetilde{\deg}$	$\widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\sqrt{\mathrm{bs}(f)} \cdot \widetilde{\deg}(g))$ Theorem 7	$\widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g))$ when $\widetilde{\deg}(f) = \Theta(n)$ [38]

2.1 Lower bounds on $\mathrm{R}(f \circ g)$ when the outer function has full randomized query complexity

Sherstov [38] proved that $\widetilde{\deg}(f \circ g) = \Omega(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g))$ when the approximate degree of the outer function f is $\Theta(n)$, where n is the arity of f. Though, a corresponding result for the case of randomized query complexity was not known. Our main result is to prove the corresponding theorem for randomized query complexity.

▶ **Theorem 2.** *Let f be a partial Boolean function on n-bits such that $\mathrm{R}(f) = \Theta(n)$. Then for all partial functions g, we have*

$$\mathrm{R}(f \circ g) = \Omega(\mathrm{R}(f) \cdot \mathrm{R}(g)).$$

The proof of this theorem is given in Section 4. Notice, since $\mathrm{R}(f \circ g) = O(\mathrm{R}(f) \cdot \mathrm{R}(g) \log \mathrm{R}(f))$ (by error reduction), Theorem 2 says that composition of R holds when the randomized query complexity of the outer function, f, is $\Theta(n)$. Next, we give main ideas behind the proof of the above theorem.

Ideas behind proof of Theorem 2

A crucial complexity measure that we use for the proof of Theorem 2 is called the *noisy randomized query complexity*, first introduced by Ben-David and Blais [9] in order to study randomized query complexity. Noisy randomized query complexity can be seen as a generalization of randomized query complexity where the algorithm can query a bit with any bias and only pays proportionally to the square of the bias in terms of cost (see Definition 16). They give the following characterization of $\mathrm{noisyR}(f)$ (the noisy randomized query complexity of f).

▶ **Theorem 3** (Ben-David and Blais [9]). *For all partial functions f on n-bits, we have*

$$\mathrm{noisyR}(f) = \Theta\left(\frac{\mathrm{R}(f \circ \mathrm{GapMaj}_n)}{n}\right), \tag{1}$$

where GapMaj_n is the Gap-Majority function on n bits whose input is promised to have Hamming weight either $(n/2 + 2\sqrt{n})$ (in which case the output is -1) or $(n/2 - 2\sqrt{n})$ (in which case the output is 1).

We want to point out that the arity of f and Gap-Majority is the *same* in Theorem 3. Towards a proof of Theorem 2, we first make the following crucial observation.

▶ **Observation 4.** *Let f be a partial Boolean function on n bits. If $t(n) \geq 1$ is a non-decreasing function of n and*

$$\mathrm{noisyR}(f) = \Omega\left(\frac{\mathrm{R}(f \circ \mathrm{GapMaj}_{t(n)})}{t(n)}\right),$$

then $\mathrm{R}(f \circ g) = \Omega((\mathrm{R}(f) \cdot \mathrm{R}(g))/t(n))$ for all partial functions g.

In particular choosing $t(n)$ to be $(\log n)$, if the outer function f satisfies

$$\mathrm{noisyR}(f) = \Omega\left(\frac{\mathrm{R}(f \circ \mathrm{GapMaj}_{\log n})}{\log n}\right). \tag{2}$$

then the above observation gives $\mathrm{R}(f \circ g) = \Omega((\mathrm{R}(f) \cdot \mathrm{R}(g))/(\log n))$ for all partial functions g.

The Observation 4 allows us to approach the composition question for randomized query complexity in a conceptually fresh manner. The goal of proving that randomized query complexity composes for a function or a class of functions, say upto $(\log n)$-factor, reduces to showing that Equation 2 holds for that function or class of functions for $t(n) = \log n$.

We are able to show that Equation 2 holds for all non-decreasing functions $t(n)$ with a slight overhead.

▶ **Theorem 5.** *Let f be a partial function on n bits and let $t \geq 1$, then* $\mathrm{R}(f \circ \mathrm{GapMaj}_t) = O(t \cdot \mathrm{noisyR}(f) + n)$.

Notice that this is a generalization of Ben-David and Blais' characterization of noisyR given by Theorem 3 in one direction. To give an idea of the proof, their characterization (Theorem 3) shows that any noisy oracle algorithm for f can be simulated using only two biases, 1 and $1/\sqrt{n}$ (where n is the arity of f), with only constant overhead. We generalize this by showing that the same simulation works with a slight overhead even when the bias $1/\sqrt{n}$ is replaced by a bias $1/\sqrt{t}$, for some $t \geq 1$. A detailed proof the above theorem has been included in the full version of this paper [20].

This seemingly inconsequential generalization allows us to complete the proof of Theorem 2, i.e. if for an n-bit partial function f, $\mathrm{R}(f) = \Theta(n)$, then $\mathrm{R}(f \circ g) = \widetilde{\Theta}(\mathrm{R}(f) \cdot \mathrm{R}(g))$ for all partial functions g (see Section 4 for details).

Furthermore, Theorem 5 even sheds light on *the composition question for* noisyR. A corollary of this theorem is that if R composes with respect to an outer function, then noisyR also composes with respect to the same outer function (see Section 4 for a proof).

▶ **Corollary 6.** *Let f be a partial Boolean function. If* $\mathrm{R}(f \circ g) = \widetilde{\Theta}(\mathrm{R}(f) \cdot \mathrm{R}(g))$ *for all partial functions g then* $\mathrm{noisyR}(f \circ g) = \widetilde{\Theta}(\mathrm{noisyR}(f) \cdot \mathrm{noisyR}(g))$.

2.2 Lower bound on $\widetilde{\deg}(f \circ g)$ in terms of block sensitivity of f and $\widetilde{\deg}(g)$

As discussed in the introduction, the composition question for $\widetilde{\deg}$ is only known to hold when the outer function f is symmetric [11] or has high approximate degree [38]. There are also no known lower bounds on $\widetilde{\deg}(f \circ g)$ in terms of weaker measures of f and $\widetilde{\deg}(g)$. Compare this with the situation with respect to composition of R. It was shown in [26] that $\mathrm{R}(f \circ g) = \Omega(\mathrm{s}(f)\,\mathrm{R}(g))$, where $\mathrm{s}(f)$ denotes the sensitivity of f. This was later strengthened to $\Omega(\mathrm{fbs}(f)\,\mathrm{R}(g))$ [7, 8], where $\mathrm{fbs}(f)$ is the fractional block sensitivity of f.

In this second part we show analogous lower bounds on approximate degree of composed function $f \circ g$. Our main result here is the following.

▶ **Theorem 7.** *For all non-constant (possibly partial)*[3] *Boolean functions* $f : \{0,1\}^n \to \{0,1\}$ *and* $g : \{0,1\}^m \to \{0,1\}$*, we have*

$$\widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\sqrt{\mathrm{bs}(f)} \cdot \widetilde{\deg}(g)).$$

[3] For definitions of block sensitivity and approximate degree in the context of partial functions, please see Definitions 32 and 17.

We first note that the above theorem is tight in terms of block sensitivity, i.e., we cannot have $\widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\mathrm{bs}(f)^c \cdot \widetilde{\deg}(g))$ for any $c > 1/2$. This is because the OR function over n bits witnesses the tight quadratic separation between $\widetilde{\deg}$ and bs, i.e., $\widetilde{\deg}(\mathsf{OR}_n) = \Theta(\sqrt{n}) = \Theta(\sqrt{\mathrm{bs}(\mathsf{OR}_n)})$ [33].

We also get the following composition theorem as a corollary. It says that the composition for $\widetilde{\deg}$ holds when the outer function has minimal approximate degree with respect to its block sensitivity. Recall, $\widetilde{\deg}(f) = \Omega(\sqrt{\mathrm{bs}(f)})$ [33].

▶ **Corollary 8.** *For all Boolean function $f : \{0,1\}^n \to \{0,1\}$ with $\widetilde{\deg}(f) = \Theta(\sqrt{\mathrm{bs}(f)})$ and for all $g : \{0,1\}^m \to \{0,1\}$, we have $\widetilde{\deg}(f \circ g) = \widetilde{\Theta}(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g))$.*

This complements a result of Sherstov [38, Theorem 6.6], which shows that composition of $\widetilde{\deg}$ holds when the outer function has maximal $\widetilde{\deg}$ with respect to its arity.

We further note that Corollary 8 covers new set of composed functions $f \circ g$ for which the composition theorem for $\widetilde{\deg}$ doesn't follow from the known results [11, 38]. For example, consider the Rubinstein function RUB with arity n(see [36] for Definition) as the outer function f. It is clearly not a symmetric function. It also doesn't have high approximate degree, i.e., $\widetilde{\deg}(\mathsf{RUB}) = O(\sqrt{n}\log n)$ (see Lemma A.7 from [20]). Therefore, the composition of $\widetilde{\deg}(\mathsf{RUB} \circ g)$ doesn't follow from the existing results. However, it follows from Corollary 8, since $\mathrm{bs}(\mathsf{RUB}) = \Omega(n)$ and so $\widetilde{\deg}(\mathsf{RUB}) = \widetilde{\Theta}(\sqrt{\mathrm{bs}(\mathsf{RUB})})$.

Another example is the sink function SINK over $\binom{n}{2}$ variables ([22]), which is also not a symmetric function. Furthermore, its approximate degree is $O(\sqrt{n}\log n)$ (Lemma A.7 from [20]). Therefore, the composition of $\widetilde{\deg}(\mathsf{SINK} \circ g)$ also doesn't follow from the existing results. Again, it follows from Corollary 8, since $\mathrm{bs}(\mathsf{SINK}) = \Theta(n)$ (Observation A.4, [20]) and $\widetilde{\deg}(\mathsf{SINK}) = \widetilde{\Theta}(\sqrt{n})$.

Ideas behind proof of Theorem 7

We will first sketch the proof ideas in the case when f and g are total Boolean functions, and then explain how to extend it to partial functions too.

Our starting point is the well known Nisan-Szegedy's embedding of PrOR (see Definition 18) over $\mathrm{bs}(f)$ many bits in a Boolean function f [33]. Carrying out this transformation in $f \circ g$ embeds $\mathsf{PrOR}_{\mathrm{bs}(f)} \circ (g_1, \ldots, g_{\mathrm{bs}(f)})$ into $f \circ g$, where $g_1, \ldots, g_{\mathrm{bs}(f)}$ are different partial functions such that $\widetilde{\mathrm{bdeg}}(g_i) \geq \widetilde{\deg}(g)$ for all $i \in [\mathrm{bs}(f)]$[4]. Since the transformation is just substitutions of variables by constants, we further have

$$\widetilde{\deg}(f \circ g) \geq \widetilde{\mathrm{bdeg}}(\mathsf{PrOR}_{\mathrm{bs}(f)} \circ (g_1, \ldots, g_{\mathrm{bs}(f)})). \tag{3}$$

It now looks like that we can appeal to the composition theorem for PrOR (Theorem 21) [11] to obtain our theorem. However, there is a technical difficulty – Theorem 21 doesn't hold for *different* inner *partial* functions. It only deals with a single total inner function. We therefore generalize the proof of Theorem 21 to obtain the following general version of the composition theorem for PrOR.

▶ **Theorem 9.** *For any partial Boolean functions $g_1, g_2, \ldots, g_n$, we have*

$$\widetilde{\mathrm{bdeg}}\left(\mathsf{PrOR}_n \circ (g_1, g_2, \ldots, g_n)\right) = \Omega\left(\frac{\sqrt{n} \cdot \min_{i=1}^{n} \widetilde{\mathrm{bdeg}}(g_i)}{\log n}\right).$$

[4] $\widetilde{\mathrm{bdeg}}$ is the notion of approximate degree in the context of partial functions. For a formal definition, see Definition 17.

We can now obtain Theorem 7 from Equation (3) and Theorem 9. The proof of Theorem 9 is a generalization of a result due to [11] (see Theorem 21). For lack of space, we present the proof of this theorem in the complete version of this paper [20].

2.3 Composition results when the outer functions has some symmetry

The class of symmetric functions capture many important function like OR, AND, Parity and Majority. Recall that a function is symmetric when the function value only depends on the Hamming weight of the input; in other words, a function is symmetric iff its value on an input remains unchanged even after permuting the bits of the input. As noted earlier, both for R and $\widetilde{\deg}$, composition was known to hold when the outer function was symmetric.

A natural question is, whether one can prove composition theorems when the outer function is *weakly* symmetric (it is symmetric with respect to a weaker notion of symmetry). In this paper we consider one such notion of symmetry – junta-symmetric functions.

▶ **Definition 10** (k-junta symmetric function). *A function $f : \{0,1\}^n \to \{0,1\}$ is called a k-junta symmetric function if there exists a set $\mathcal{J}$ of size k of variables such that the function value depends on assignments to the variables in $\mathcal{J}$ as well as on the Hamming weight of the whole input.*

k-junta symmetric functions can be seen as a mixture of symmetric functions and k-juntas. This class of functions has been considered previously in literature, particularly in [19, 14] where these functions plays a crucial role. [19] even presents multiple characterisations of k-junta symmetric functions for constant k. Note that by definition an arbitrary k-junta (i.e., a function that depend on k variables) is also a k-junta symmetric function, since we can consider the dependence on Hamming weight to be trivial. Thus, this notion loses out on the symmetry of the function considered. We, therefore, consider the class of *strongly* k-junta symmetric functions.

▶ **Definition 11** (Strongly k-junta symmetric function). *A k-junta symmetric function is called strongly k-junta symmetric if every variable is influential. In other words, there exists a setting to the junta variables such that the function value depends on the Hamming weight of the whole input in a non-trivial way.*

We prove that if the outer function is strongly $\sqrt{n}$-junta symmetric ("strongly" indicating that the dependence on the Hamming weight is non-trivial) then $\widetilde{\deg}$ composes. On the other hand, Theorem 2 implies that R composes for any strongly k-junta symmetric functions (as long as $n - k = \Theta(n)$).

▶ **Theorem 12.** *For any strongly k-junta symmetric function $f : \{0,1\}^n \to \{0,1\}$ and any Boolean function $g : \{0,1\}^m \to \{0,1\}$, we have*

- $\widetilde{\deg}(f \circ g) = \widetilde{\Theta}(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g))$ *where* $k = O(\sqrt{n})$.
- $\mathrm{R}(f \circ g) = \widetilde{\Theta}(\mathrm{R}(f) \cdot \mathrm{R}(g))$ *where* $n - k = \Theta(n)$.

For the lack of space, the proof of the above theorem is given in Appendix C. Note that if one is able to prove the above theorem for k-junta-symmetric functions (without the requirement of "strongly") for any non-constant k then we would have the full composition theorem.

Organization of the paper

We have formally defined complexity measures and Boolean functions needed for our results in Section 3 and Appendix A. Section 4 contains proofs of our results related to the composition of randomized query complexity (Theorem 2). In Section 5 we give the proof of our result for the composition of approximate degree (Theorem 7). Finally, for the sake of space, the results about the composition of functions with the weak notion of symmetry are in Appendix C.

3 Preliminaries

Notations: We will use $[n]$ to represent the set $\{1, \dots, n\}$. For any (possibly partial) Boolean function $f : \{0,1\}^n \to \{0,1,*\}$ we will denote by $\mathrm{Dom}(f)$ the set $f^{-1}(\{0,1\})$. The arity of f is the number of variables - in this case n. A Boolean function $f : \{0,1\}^n \to \{0,1,*\}$ is said to be total if $\mathrm{Dom}(f) = \{0,1\}^n$. Any function (not otherwise stated) will be a total Boolean function.

For any $x \in \{0,1\}^n$, we will use $|x|$ to denote the number of 1s in x, that is, the Hamming weight of the string x. The string x^i denotes the modified string x with the i-th bit flipped. Similarly, x^B is defined to be the string such that all the bits whose index is contained in the set $B \subseteq [n]$ are flipped in x.

Following is a formal definition of (partial) function composition.

▶ **Definition 13** (Generalized composition of functions). *For any (possibly partial) Boolean function $f : \{0,1\}^n \to \{0,1,*\}$ and n (possibly partial) Boolean functions $g_1, g_2, \dots, g_n$, define the (possibly partial) composed function*

$$f \circ (g_1, g_2, \dots, g_n)(x_1, x_2, \dots, x_n) = f(g_1(x_1), g_2(x_2), \dots, g_n(x_n)),$$

where g_i's can have different arities and, moreover, if $x_i \notin \mathrm{Dom}(g_i)$ for any $i \in [n]$ or the string $(g_1(x_1), g_2(x_2), \dots, g_n(x_n)) \notin \mathrm{Dom}(f)$, then the function $f \circ g$ outputs $*$.

In this paper we use the standard definitions of various complexity measures like randomized query complexity, sensitivity, block-sensitivity, fractional block sensitivity and approximate degree. We present the formal definitions in Appendix A.

3.1 Standard definitions and functions for the composition of R

The function Gap-Majority has played an important role in the study of composition of R.

▶ **Definition 14** (Gap-Majority). *The function $\mathrm{GapMaj}_t : \{0,1\}^t \to \{0,1,*\}$ is a partial function with arity t such that*

$$\mathrm{GapMaj}_t(x) = \begin{cases} 1 & \textit{if } |x| = t/2 + 2\sqrt{t}, \\ 0 & \textit{if } |x| = t/2 - 2\sqrt{t}, \\ * & \textit{otherwise.} \end{cases}$$

It can be shown that $\mathrm{R}(\mathrm{GapMaj}_t) = \Theta(t)$ [9].

In regards to the composition question of R, one of the most significant complexity measures (defined by Ben-David and Blais [9]) is that of noisyR. We first define the noisy oracle model.

▶ **Definition 15** (Noisy Oracle Model and Noisy Oracle Access to a String ([9])). *For $b \in \{0,1\}$, a noisy oracle to b takes a parameter $-1 \le \gamma \le 1$ as input and returns a bit b' such that* $\Pr[b' = b] = (1+\gamma)/2$. *The cost of one such query is γ^2. Each query to noisy oracle returns independent bits.*

For $x = (x_1, \ldots, x_n) \in \{0,1\}^n$, noisy oracle access to x is access to n independent noisy oracles, one for each bit x_i, $i \in [n]$.

Next, we define the noisy oracle model of computation.

▶ **Definition 16** (Noisy Oracle Model of Computation ([9])). *Let $f : \{0,1\}^n \to \{0,1,*\}$ be a partial Boolean function. A* noisyR *query algorithm A computes f if for all $x \in \mathrm{Dom}(f)$,* $\Pr[A(x) \ne f(x)] \le 1/3$, *where A is a randomized algorithm given noisy oracle access to x, and the probability is over both noisy oracle calls and the internal randomness of the algorithm A. The cost of the algorithm A for an input x is the sum of the cost of all noisy oracle calls made by A on x, and the cost of A, cost(A), is the maximum cost over all $x \in \mathrm{Dom}(f)$. The* noisyR *randomized query complexity of f, denoted by* $\mathrm{noisyR}(f)$, *is defined as*

$$\mathrm{noisyR}(f) = \min_{A \text{ computes } f} cost(A).$$

Again, $1/3$ in the above definition can be replaced by any constant $< 1/2$. If only queries with $\gamma = 1$ are allowed in the noisy query model, then we obtain the usual randomized algorithm for f, thus $\mathrm{noisyR}(f) = O(\mathrm{R}(f))$.

3.2 Standard definitions and functions for the composition of $\widetilde{\deg}$

The definition of $\widetilde{\deg}$ can naturally be extended to partial functions f by restricting the definition to hold only for inputs in $\mathrm{Dom}(f)$, but the approximating polynomial can take arbitrarily large values on points outside the domain. However, for the purpose of understanding the composition of approximate degree of Boolean functions (or even total Boolean functions) one need a measure of approximate degree of partial Boolean functions which is bounded on all the points of the Boolean cube.

▶ **Definition 17** (Bounded approximate degree ($\widetilde{\mathrm{bdeg}}$)). *For a partial Boolean function $f : \{0,1\}^n \to \{0,1,*\}$, the bounded approximate degree of f ($\widetilde{\mathrm{bdeg}}(f)$) is the minimum possible degree of a polynomial p such that*

- $|p(x) - f(x)| \le 1/3, \quad \forall x \in \mathrm{Dom}(f)$, *and*
- $0 \le p(x) \le 1 \quad \forall x \in \{0,1\}^n$.

In other words, we take the minimum possible degree of a polynomial which is bounded for all possible inputs ($p(x) \in [0,1]$ for all $x \in \{0,1\}^n$), and it approximates f in the usual sense over $\mathrm{Dom}(f)$.

Over the years people have tried to study the composition of $\widetilde{\deg}$ with different outer functions. In this context the following restriction of OR is an important partial function:

▶ **Definition 18** (Promise-OR). *Promise-*OR *(denoted by* PrOR_n*) is the function* $\mathsf{PrOR}_n : \{0,1\}^n \to \{0,1,*\}$ *such that* $\mathsf{PrOR}_n(x) = 0$ *if* $|x| = 0$, *equals to 1 if* $|x| = 1$, *and* $*$ *otherwise.*

Some useful previous results. We will also be crucially using a few results from prior works in our proofs. The following are a couple of useful results on noisyR.

▶ **Lemma 19** ([9]). *Let f be a non-constant partial Boolean function then* $\mathrm{noisyR}(f) = \Omega(1)$.

▶ **Theorem 20** ([9]). *For all partial functions f and g,* $\mathrm{R}(f \circ g) = \Omega(\mathrm{noisyR}(f) \cdot \mathrm{R}(g))$.

We will also be using the following theorem of [11] regarding the composition question of $\widetilde{\mathrm{bdeg}}$ when the outer function is PrOR_n. Informally, we will call it the Promise-OR composition theorem.

▶ **Theorem 21** ([11]). *For any Boolean function $g : \{0,1\}^m \to \{0,1\}$ we have,*

$$\widetilde{\mathrm{bdeg}}\left(\mathsf{PrOR}_n \circ g\right) = \Omega\left(\sqrt{n} \cdot \widetilde{\deg}(g)/\log n\right).$$

4 Results about composition of R

This section is devoted to the results related to the composition of randomized query complexity. Our main result states that composition of R holds if the outer function has full randomized query complexity (Theorem 2). As mentioned in the proof idea, the proof critically depends on the notion of noisy randomized query complexity and its properties (introduced by Ben-David and Blais [9]).

Recall the definition of noisy randomized query complexity of a function f from Definition 16. As mentioned in the introduction (Theorem 3), Ben-David and Blais [9] proved that

$$\mathrm{noisyR}(f) = \Theta\left(\frac{\mathrm{R}(f \circ \mathrm{GapMaj}_n)}{n}\right), \tag{4}$$

where GapMaj_n is the Gap-Majority function on n bits. Note that Ben-David and Blais proved Equation 4 when the arity of functions f and Gap-Majority is the same. We show that if Equation 4 can be generalized for Gap-Majority functions of arbitrary arity for some outer function f, then randomized query complexity composes for the function f. We restate the following observation from the introduction.

▶ **Observation 4.** *Let f be a partial Boolean function on n bits. If $t(n) \geq 1$ is a non-decreasing function of n and*

$$\mathrm{noisyR}(f) = \Omega\left(\frac{\mathrm{R}(f \circ \mathrm{GapMaj}_{t(n)})}{t(n)}\right),$$

then $\mathrm{R}(f \circ g) = \Omega((\mathrm{R}(f) \cdot \mathrm{R}(g))/t(n))$ *for all partial functions g.*

Proof. Suppose $\mathrm{noisyR}(f) = \Omega\left(\frac{\mathrm{R}(f \circ \mathrm{GapMaj}_{t(n)})}{t(n)}\right)$, since $\mathrm{R}(f \circ \mathrm{GapMaj}_t) \geq \mathrm{R}(f)$, we have $\mathrm{noisyR}(f) = \Omega(\mathrm{R}(f)/(t(n)))$. Theorem 20 implies that a lower bound on noisyR translates to a lower bound on $\mathrm{R}(f \circ g)$. We have,

$$\begin{aligned} \mathrm{R}(f \circ g) &= \Omega(\mathrm{noisyR}(f) \cdot \mathrm{R}(g)) && \text{(Theorem 20)} \\ &= \Omega\left(\frac{\mathrm{R}(f) \cdot \mathrm{R}(g)}{t(n)}\right). \end{aligned}$$ ◀

Observation 4 follows from the above observation by choosing $t(n)$ to be a small function of n.

We restate from Section 1 our generalized characterization of noisyR (i.e., generalization of Equation 4). For a complete proof of Theorem 5 we refer to the full version of this paper [20].

▶ **Theorem 5.** *Let f be a partial function on n bits and let $t \geq 1$, then* $\mathrm{R}(f \circ \mathrm{GapMaj}_t) = O\left(t \cdot \mathrm{noisyR}(f) + n\right)$.

This allows us to show that if for an n-bit partial function f, $\mathrm{R}(f) = \Theta(n)$, then $\mathrm{R}(f \circ g) = \widetilde{\Theta}(\mathrm{R}(f) \cdot \mathrm{R}(g))$ for all partial functions g (Theorem 2).

The proof of Theorem 2 is discussed in Section 4.1. A corollary of this theorem is that if R composes with respect to an outer function, then noisyR also composes with respect to the same outer function (Corollary 6).

We give proof of Theorem 2 in the next section and prove Corollary 6 in Section 4.3. We need the following theorem for these proofs, which lower bounds $R(f \circ g)$ in terms of $R(f)$ and $R(g)$.

▶ **Theorem 22** ([24])**.** *Let f and g be partial functions then* $\mathrm{R}(f \circ g) = \Omega(\mathrm{R}(f) \cdot \sqrt{\mathrm{R}(g)})$.

4.1 Composition for functions with $\mathrm{R}(f) = \Theta(n)$

We restate the theorem below.

▶ **Theorem 2.** *Let f be a partial Boolean function on n-bits such that* $\mathrm{R}(f) = \Theta(n)$*. Then for all partial functions g, we have*

$$\mathrm{R}(f \circ g) = \Omega(\mathrm{R}(f) \cdot \mathrm{R}(g)).$$

Proof. From Theorem 22 we have a lower bound on the randomized query complexity of $(f \circ \mathrm{GapMaj}_t)$:

$$\mathrm{R}(f \circ \mathrm{GapMaj}_t) = \Omega(\mathrm{R}(f) \cdot \sqrt{t}). \tag{5}$$

On the other hand, Theorem 5 gives an upper bound of $O\left(t \cdot \mathrm{noisyR}(f) + n\right)$ on $\mathrm{R}(f \circ \mathrm{GapMaj}_t)$. Thus, choosing $t = \left(\frac{C \cdot n}{\mathrm{noisyR}(f)}\right)$ for a large enough constant C, we have

$$\mathrm{R}(f) \cdot \sqrt{\frac{n}{\mathrm{noisyR}(f)}} = O\left(\frac{n}{\mathrm{noisyR}(f)} \cdot \mathrm{noisyR}(f) + n\right).$$

This implies that

$$\mathrm{R}(f) = O\left(\sqrt{n \cdot \mathrm{noisyR}(f)}\right). \tag{6}$$

Thus, if $\mathrm{R}(f) = \Theta(n)$, then $\mathrm{noisyR}(f) = \Omega(\mathrm{R}(f))$, which implies composition from Theorem 20. ◀

Notice that Equation 6 is equivalent to the following observation.

▶ **Observation 23.** *Let f be a partial Boolean function on n-bits. Then,* $\mathrm{noisyR}(f) = \Omega\left(\frac{\mathrm{R}(f)^2}{n}\right)$.

When $\mathrm{R}(f) = \Theta(n)$, we have already seen that Observation 23 implies composition of randomized query complexity when the outer function is f.

Though, Observation 23 implies a more general result. When $\mathrm{R}(f)$ is close to n (arity of f), Observation 23 places a limit on the gap between $\mathrm{R}(f)$ and $\mathrm{noisyR}(f)$ (consequently on the violation of composition with outer function being f). These implications are formally discussed in Appendix 4.2.

Another implication of Theorem 2 is that composition of R for an outer function f implies the composition of noisyR for outer function being f (Corollary 6).

4.2 Additional implications of Observation 23

Without loss of generality we can assume $\mathrm{R}(f \circ g) = \Omega(\mathrm{R}(g))$ (note that this is true when f is non-constant).

Ben-David and Blais [9] gave a counterexample for composition, but the arity of the used function was very high compared to the randomized query complexity. They observed that the composition can still be true in the weaker sense:

$$\mathrm{R}(f \circ g) = \Omega\left(\frac{\mathrm{R}(f) \cdot \mathrm{R}(g)}{\log n}\right).$$

Observation 23 shows that a much weaker composition result is true.

▶ **Corollary 24.** *Let f and g be partial functions on n and m bits respectively. If $\mathrm{R}(f \circ g) = \Omega(\mathrm{R}(g))$, then*

$$\mathrm{R}(f \circ g) = \Omega\left(\frac{\mathrm{R}(f) \cdot \mathrm{R}(g)}{\sqrt{n}}\right).$$

Proof.

$$\begin{aligned}\mathrm{R}(f \circ g) &= \Omega(\mathrm{noisyR}(f) \cdot \mathrm{R}(g)) && \text{(Theorem 20)}\\ &= \Omega\left(\frac{\mathrm{R}(f)^2 \cdot \mathrm{R}(g)}{n}\right). && (7)\end{aligned}$$

Where the last equality follows from Observation 23 [5] Now there are two cases:

- **Case 1.** $\mathrm{R}(f) = O(\sqrt{n})$. In this case $\mathrm{R}(f)/\sqrt{n} = O(1)$ and since we assumed $\mathrm{R}(f \circ g) = \Omega(\mathrm{R}(g))$, the claim follows from Equation 7.
- **Case 2.** $\mathrm{R}(f) = \Theta(n^{1/2} \cdot t(n))$ where $t(n)$ is a strictly increasing function of n. Thus,

$$\frac{\mathrm{R}(f)^2 \cdot \mathrm{R}(g)}{n} = \Omega\left(t(n)^2 \cdot \mathrm{R}(g)\right) = \Omega\left(\frac{\mathrm{R}(f) \cdot \mathrm{R}(g)}{\sqrt{n}}\right).$$

Again, the claim follows from Equation 7. ◀

The weaker composition, Corollary 24, implies that if $\mathrm{R}(f)$ and $\mathrm{R}(g)$ are comparable to the arity of these functions, the randomized query complexity of $f \circ g$ is "not far" from the conjectured randomized query complexity $\mathrm{R}(f) \cdot \mathrm{R}(g)$. In other words, if there is a large polynomial separation between $\mathrm{R}(f \circ g)$ and $(\mathrm{R}(f) \cdot \mathrm{R}(g))$, then $\mathrm{R}(f)$ and $\mathrm{R}(g)$ can not be too large.

▶ **Corollary 25.** *Let f and g be partial functions such that f is a function on n-bits and g is a function on $t(n)$-bits where $t(n)$ is a strictly increasing function of n. If $\mathrm{R}(f) = \Theta(n^\beta)$, $\mathrm{R}(g) = \Theta(n^\gamma)$ and $\mathrm{R}(f \circ g) = O((\mathrm{R}(f) \cdot \mathrm{R}(g))^\alpha)$, where $\alpha < 1$ is a constant, then $(1-\alpha)(\alpha+\beta) < 1/2$.*

Proof. For some constants A and B we have

$$A \cdot \frac{\mathrm{R}(f) \cdot \mathrm{R}(g)}{\sqrt{n}} \leq \mathrm{R}(f \circ g) \leq B \cdot (\mathrm{R}(f) \cdot \mathrm{R}(g))^\alpha,$$

[5] Sherstov [38] proved that for Boolean functions f and g, $\widetilde{\deg}(f \circ g) = \Omega((\widetilde{\deg}(f)^2 \widetilde{\deg}(g))/n)$. Thus in Equation 7 we prove the same result but in the randomized world.

where the first inequality follows from Corollary 24 and second from assumption. Assigning the values of $\mathrm{R}(f)$ and $\mathrm{R}(g)$ in terms on n we have,

$$A \cdot n^{\beta+\gamma-1/2} \leq B \cdot n^{\alpha(\beta+\gamma)}$$
$$n^{(1-\alpha)(\beta+\gamma)-1/2} \leq \frac{B}{A}.$$

which implies, for large enough n, $(1-\alpha)(\beta+\gamma) \leq 1/2$. ◀

A special case of the above corollary is when arity and randomized query complexity of g are superpolynomial in n. In this case a polynomial gap between $\mathrm{R}(f \circ g)$ and $(\mathrm{R}(f) \cdot \mathrm{R}(g)))$ is not possible.

4.3 Proof of Corollary 6

First, we need the following lemma which follows from Theorem 5, Theorem 20 and Lemma 19.

▶ **Lemma 26.** *Let f be a partial function on n bits and let $t = \Omega(n)$. Then*

$$\mathrm{noisyR}(f) = \Theta\left(\frac{\mathrm{R}(f \circ \mathrm{GapMaj}_t)}{t}\right).$$

Proof. From Theorem 5 we have for all $t \geq 1$, $\mathrm{R}(f \circ \mathrm{GapMaj}_t) = O(t \cdot \mathrm{noisyR}(f) + n)$. Since we have assumed $t = \Omega(n)$ and $\mathrm{noisyR}(f) = \Omega(1)$ (Lemma 19), we get $\mathrm{R}(f \circ \mathrm{GapMaj}_t) = O(t \cdot \mathrm{noisyR}(f))$. Thus, $\mathrm{noisyR}(f) = \Omega\left(\frac{\mathrm{R}(f \circ \mathrm{GapMaj}_t)}{t}\right)$.

The upper bound $\mathrm{noisyR}(f) = O\left(\frac{\mathrm{R}(f \circ \mathrm{GapMaj}_t)}{t}\right)$ follows from Theorem 20 and the fact that $\mathrm{R}(\mathrm{GapMaj}_t) = \Theta(t)$. ◀

Now we prove that if R composes for f then noisyR composes for that f. For convenience, we recall the statement of the corollary from the introduction.

▶ **Corollary 6.** *Let f be a partial Boolean function. If $\mathrm{R}(f \circ g) = \widetilde{\Theta}(\mathrm{R}(f) \cdot \mathrm{R}(g))$ for all partial functions g then $\mathrm{noisyR}(f \circ g) = \widetilde{\Theta}(\mathrm{noisyR}(f) \cdot \mathrm{noisyR}(g))$.*

Proof. From Theorem 3, we have

$$\mathrm{noisyR}(f \circ g) = \Theta\left(\frac{\mathrm{R}\left((f \circ g) \circ \mathrm{GapMaj}_{mn}\right)}{mn}\right).$$

Since $(f \circ g) \circ h = f \circ (g \circ h)$, the right hand side of the above expression is equal to

$$\Theta\left(\frac{\mathrm{R}\left(f \circ (g \circ \mathrm{GapMaj}_{mn})\right)}{mn}\right).$$

The proof follows from the assumption that R composes and Lemma 26.

$$\begin{aligned}
\mathrm{noisyR}(f \circ g) &= \Theta\left(\frac{\mathrm{R}\left(f\right) \cdot \mathrm{R}\left(g \circ \mathrm{GapMaj}_{mn}\right)}{mn}\right) && \text{(assuming R composes)}\\
&= \Theta\left(\mathrm{R}(f) \cdot \mathrm{noisyR}(g)\right) && \text{(from Lemma 26)}\\
&= \Theta\left(\mathrm{noisyR}(f) \cdot \mathrm{noisyR}(g)\right). && \text{(assuming R composes)}
\end{aligned}$$

◀

5 Composition of approximate degree in terms of block sensitivity

In this section we study the composition question for approximate degree. Recall that the composition question asks: whether for all Boolean functions f and g

$$\widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\widetilde{\deg}(f)\,\widetilde{\deg}(g))?$$

Following our discussion from the introduction, we know that the above composition is known to hold for only two sub-classes of outer functions, namely symmetric functions [11] and functions with high approximate degree [38]. It is thus natural to seek weaker lower bounds to make progress towards the composition question. One way to weaken the expression on the right-hand side would be to replace the measure $\widetilde{\deg}(f)$ by a weaker measure (like $\sqrt{\mathrm{s}(f)}$, $\sqrt{\mathrm{bs}(f)}$ or $\sqrt{\mathrm{fbs}(f)}$). Here we will establish one such lower bound of $\sqrt{\mathrm{bs}(f)}\,\widetilde{\deg}(g)$.

We restate our theorem now.

▶ **Theorem 7.** *For all non-constant (possibly partial)*[6] *Boolean functions $f : \{0,1\}^n \to \{0,1\}$ and $g : \{0,1\}^m \to \{0,1\}$, we have*

$$\widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\sqrt{\mathrm{bs}(f)} \cdot \widetilde{\deg}(g)).$$

We note that many analogous results are known in the setting of composition of R; see, for example, [26, 7, 8, 13, 6, 24, 10]. To the best of our knowledge, this is the first such result in the setting of $\widetilde{\deg}$. We present only a proof sketch here; most of the technical parts of the proof appear in Appendix B.

Further, we present the sketch of the proof in two parts. For simplicity, in the first part we sketch a proof of the lower bound $\sqrt{\mathrm{s}(f)}\,\widetilde{\deg}(g)$ for total function f, and then in the second part we modify the arguments to obtain Theorem 7.

We begin with a proof sketch for a lower bound of $\sqrt{\mathrm{s}(f)}\,\widetilde{\deg}(g)$. Let $x \in \{0,1\}^n$ be an input having the maximum sensitivity with respect to f, and $S \subseteq [n]$ be the set of sensitive bits at x ($|S| = \mathrm{s}(f)$). Consider the subfunction f' obtained from f by fixing the set of variables *not* in S according to x. By construction, f' is defined over $\mathrm{s}(f)$ many variables and is fully sensitive at the input $x|_S$ given by x restricted to the indices in S. Since f' is a subfunction of f and g is non-constant, we have $\widetilde{\deg}(f \circ g) \geq \widetilde{\deg}(f' \circ g)$.

Notice that f' at the neighbourhood of x, in the Boolean cube, is the partial function PrOR (Definition 18) or its negation. Therefore, we have $\widetilde{\deg}(f \circ g) \geq \widetilde{\deg}(f' \circ g) \geq \widetilde{\mathrm{bdeg}}(\mathsf{PrOR}_{|S|} \circ g)$ (see Definition 17 for a definition of the bounded approximate degree). We can now invoke the composition theorem for PrOR (Theorem 21) [11] to obtain our lower bound:

$$\widetilde{\deg}(f \circ g) \geq \widetilde{\deg}(f' \circ g) \geq \widetilde{\mathrm{bdeg}}(\mathsf{PrOR}_{|S|} \circ g) = \widetilde{\Omega}(\sqrt{\mathrm{s}(f)}\,\widetilde{\deg}(g)).$$

However, there is a technical issue with our argument above. When we claimed that f' looks like a PrOR function we were not quite correct. Technically, it is a Shifted-PrOR function $\mathsf{PrOR}_{|S|}^{x|_S}$, where $\mathsf{PrOR}_n^a(y_1, y_2, \ldots, y_n) := \mathsf{PrOR}_n(y_1 \oplus a_1, y_2 \oplus a_2, \ldots, y_n \oplus a_n)$ for $a \in \{0,1\}^n$. Formally, we have

$$\widetilde{\deg}(f \circ g) \geq \widetilde{\deg}(f' \circ g) \geq \widetilde{\mathrm{bdeg}}(\mathsf{PrOR}_{|S|}^{x|_S} \circ g) = \widetilde{\mathrm{bdeg}}(\mathsf{PrOR}_{|S|} \circ (g_1, \ldots, g_{|S|})), \tag{8}$$

where $g_i = g$ or $\neg g$ depending on the corresponding i-th bit in $x|_S$.

[6] For definitions of block sensitivity and approximate degree in the context of partial functions, please see Definitions 32 and 17.

We, therefore, need a composition theorem for PrOR with *different* inner functions, while Theorem 21 requires that all the inner functions be same. In fact, we would need a more general composition theorem with *different* inner *partial* functions, which we restate below. This generalization is crucially used when dealing with block sensitivity.

▶ **Theorem 9.** *For any partial Boolean functions $g_1, g_2, \ldots, g_n$, we have*

$$\widetilde{\mathrm{bdeg}}\left(\mathsf{PrOR}_n \circ (g_1, g_2, \ldots, g_n)\right) = \Omega\left(\frac{\sqrt{n} \cdot \min_{i=1}^{n} \widetilde{\mathrm{bdeg}}(g_i)}{\log n}\right).$$

The proof of Theorem 9 is a generalization of proof of Theorem 21. For the sake of completeness we have added a proof in the full version of the paper [20].

Now returning to Equation (8) and using Theorem 9, we obtain the desired lower bound:

$$\widetilde{\deg}(f \circ g) \geq \widetilde{\deg}(f' \circ g) \geq \widetilde{\mathrm{bdeg}}(\mathsf{PrOR}_{|S|}^{x|_S} \circ g) = \widetilde{\Omega}(\sqrt{\mathrm{s}(f)}\, \widetilde{\deg}(g)).$$

We are now ready to present the modifications required to improve the lower bound to $\widetilde{\Omega}(\sqrt{\mathrm{bs}(f)}\, \widetilde{\deg}(g))$.

Proof of Theorem 7. Let $b = \mathrm{bs}(f)$ and $a = (a_1, a_2, \ldots, a_n)$ be an input where f achieves the maximum block sensitivity. Further, let $B_1, B_2, \ldots, B_b$ be disjoint minimal sets of variables that achieves the block sensitivity at a, i.e., $f(a) \neq f(a^{B_i})$ for all $i \in [b]$. Recall, a^{B_i} denotes the Boolean string obtained from a by flipping the bits at all the indices given by B_i. Define a partial function $f' : \{0,1\}^n \to \{0,1,*\}$ such that,

$$f'(x) = \begin{cases} 0 & \text{if } x = a, \\ 1 & \text{if } x = a^{B_i}, \text{ for some } i \in [b], \\ * & \text{otherwise.} \end{cases}$$

Note that f contains f' or its negation as a sub function. Thus, $\widetilde{\deg}(f \circ g) \geq \widetilde{\mathrm{bdeg}}(f' \circ g)$.

Since g is non-constant, we can fix the indices *not* in $\bigcup_{i=1}^{b} B_i$ according to a to obtain $f'' \circ g$. We would now like to embed PrOR_b over the remaining variables in f''. For this purpose we define the following partial functions: for every $i \in [b]$, let $I_i : \{0,1\}^{B_i} \to \{0,1,*\}$ be such that

$$I_i(x) = \begin{cases} 0 & \text{if } x = a|_{B_i}, \\ 1 & \text{if } x = a^{B_i}|_{B_i}, \\ * & \text{otherwise.} \end{cases}$$

Now observe that $f'' \circ g$ can be rewritten as $\mathsf{PrOR}_b \circ (I_1 \circ g, \ldots, I_b \circ g)$. We therefore have

$$\begin{aligned} \widetilde{\deg}(f \circ g) \geq \widetilde{\mathrm{bdeg}}(f' \circ g) \geq \widetilde{\mathrm{bdeg}}(f'' \circ g) &= \widetilde{\mathrm{bdeg}}(\mathsf{PrOR}_b \circ (I_1 \circ g, \ldots, I_b \circ g)) \\ &= \Omega\left(\frac{\sqrt{b} \cdot \min_i \widetilde{\mathrm{bdeg}}(I_i \circ g)}{\log b}\right) = \widetilde{\Omega}\left(\sqrt{b} \cdot \widetilde{\deg}(g)\right), \end{aligned}$$

where the second last equality follows from Theorem 9 and the last equality uses the fact $\widetilde{\mathrm{bdeg}}(I_i \circ g) \geq \widetilde{\deg}(g)$ for all i, which in turn follows from each I_i being non-constant. ◀

We end this section with few final remarks. As a corollary to Theorem 7 we have the following composition for $\widetilde{\deg}$ when the outer function has minimal approximate degree with respect to its block sensitivity.

▶ **Corollary 8.** *For all Boolean function* $f : \{0,1\}^n \to \{0,1\}$ *with* $\widetilde{\deg}(f) = \Theta(\sqrt{\mathrm{bs}(f)})$ *and for all* $g : \{0,1\}^m \to \{0,1\}$*, we have* $\widetilde{\deg}(f \circ g) = \widetilde{\Theta}(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g))$*.*

We also note that the set of Boolean functions with $\widetilde{\deg}(f) = \Theta(\sqrt{\mathrm{bs}(f)})$ includes examples of *non-symmetric* functions f with *low* approximate degree. In other words, when such functions are outer function in a composed function then the composition of $\widetilde{\deg}$ doesn't follow from the known results [11, 38]. Such examples are described in Subsection 2.2.

As stated in the introduction, we recall that Theorem 7 is tight in terms of block-sensitivity, i.e., the lower bound can not be improved to $\widetilde{\Omega}(\mathrm{bs}(f)^c \cdot \widetilde{\deg}(g))$ for some $c > 1/2$.

6 Conclusion

While our work makes progress on the composition problem for R and $\widetilde{\deg}$, the main problems of whether $\widetilde{\deg}$ and R composes for any pair of Boolean functions still remains open. In this light, we would like to highlight some questions that can be useful stepping stones towards the main questions.

We showed that the composition question for R is equivalent to the following open question (which is a generalization of Ben-David and Blais [9] result):

▶ **Open question 27.** *Let* $f : \{0,1\}^n \to \{0,1,*\}$ *be a Boolean function. Then, is it true that for arbitrary* t*,* $\mathrm{noisyR}(f) = \Theta\left(\mathrm{R}(f \circ \mathrm{GapMaj}_t)/t\right)$*?*

In case of approximate degree composition, a natural question is whether $\sqrt{\mathrm{bs}(f)}$ can be replaced by some other complexity measures. In this regards we state the following open problems:

▶ **Open question 28.** *For all Boolean functions* f *and* g*, can we prove either of the following:*

- $\widetilde{\deg}(f \circ g) = \Omega(\sqrt{\widetilde{\deg}(f)} \cdot \widetilde{\deg}(g))$*?*
- $\widetilde{\deg}(f \circ g) = \Omega(\sqrt{\mathrm{fbs}(f)} \cdot \widetilde{\deg}(g))$*?*

Recently, in [43, 42, 23, 18, 21], the classes of transitive functions got a lot of attention as natural generalization of the classes of symmetric functions. Can the result for symmetric functions be extended to transitive functions?

▶ **Open question 29.** *Can we prove that* $\widetilde{\deg}$ *and* R *compose when the outer function is transitive?*

References

1 Scott Aaronson, Shalev Ben-David, and Robin Kothari. Separations in query complexity using cheat sheets. In *STOC*, pages 863–876, 2016. doi:10.1145/2897518.2897644.

2 Scott Aaronson, Shalev Ben-David, Robin Kothari, Shravas Rao, and Avishay Tal. Degree vs. approximate degree and quantum implications of Huang's sensitivity theorem. In *STOC*, pages 1330–1342, 2021. doi:10.1145/3406325.3451047.

3 Andris Ambainis. Polynomial degree and lower bounds in quantum complexity: Collision and element distinctness with small range. *Theory Comput.*, 1(1):37–46, 2005. doi:10.4086/toc.2005.v001a003.

4 Andris Ambainis, Kaspars Balodis, Aleksandrs Belovs, Troy Lee, Miklos Santha, and Juris Smotrovs. Separations in query complexity based on pointer functions. *Journal of the ACM*, 64(5):32:1–32:24, 2017. doi:10.1145/3106234.

5 Andris Ambainis, Martins Kokainis, and Robin Kothari. Nearly optimal separations between communication (or query) complexity and partitions. In *CCC*, volume 50, pages 4:1–4:14, 2016. doi:10.4230/LIPIcs.CCC.2016.4.

6 Anurag Anshu, Dmitry Gavinsky, Rahul Jain, Srijita Kundu, Troy Lee, Priyanka Mukhopadhyay, Miklos Santha, and Swagato Sanyal. A composition theorem for randomized query complexity. In *FSTTCS*, volume 93, pages 10:1–10:13, 2017. doi:10.4230/LIPIcs.FSTTCS.2017.10.

7 Andrew Bassilakis, Andrew Drucker, Mika Göös, Lunjia Hu, Weiyun Ma, and Li-Yang Tan. The power of many samples in query complexity. In *ICALP*, volume 168, pages 9:1–9:18, 2020. doi:10.4230/LIPIcs.ICALP.2020.9.

8 Shalev Ben-David and Eric Blais. A tight composition theorem for the randomized query complexity of partial functions. *arXiv:2002.10809*, 2020.

9 Shalev Ben-David and Eric Blais. A tight composition theorem for the randomized query complexity of partial functions: Extended abstract. In *FOCS*, pages 240–246, 2020. doi:10.1109/FOCS46700.2020.00031.

10 Shalev Ben-David, Eric Blais, Mika Göös, and Gilbert Maystre. Randomised composition and small-bias minimax. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 624–635. IEEE, 2022.

11 Shalev Ben-David, Adam Bouland, Ankit Garg, and Robin Kothari. Classical lower bounds from quantum upper bounds. In *FOCS*, pages 339–349, 2018. doi:10.1109/FOCS.2018.00040.

12 Shalev Ben-David, Mika Göös, Robin Kothari, and Thomas Watson. When Is Amplification Necessary for Composition in Randomized Query Complexity? In *APPROX/RANDOM 2020*, volume 176, pages 28:1–28:16, 2020. doi:10.4230/LIPIcs.APPROX/RANDOM.2020.28.

13 Shalev Ben-David and Robin Kothari. Randomized query complexity of sabotaged and composed functions. *Theory Comput.*, 14(1):1–27, 2018. doi:10.4086/toc.2018.v014a005.

14 Eric Blais, Amit Weinstein, and Yuichi Yoshida. Partially symmetric functions are efficiently isomorphism testable. *SIAM J. Comput.*, 44(2):411–432, 2015. doi:10.1137/140971877.

15 Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: a survey. *Theoretical Computer Science*, 288(1):21–43, 2002. doi:10.1016/S0304-3975(01)00144-X.

16 Mark Bun, Robin Kothari, and Justin Thaler. Quantum algorithms and approximating polynomials for composed functions with shared inputs. In Timothy M. Chan, editor, *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019, San Diego, California, USA, January 6-9, 2019*, pages 662–678. SIAM, 2019. doi:10.1137/1.9781611975482.42.

17 Mark Bun and Justin Thaler. Dual lower bounds for approximate degree and Markov-Bernstein inequalities. In *ICALP*, volume 7965, pages 303–314, 2013. doi:10.1007/978-3-642-39206-1_26.

18 Sourav Chakraborty. On the sensitivity of cyclically-invariant boolean functions. *Discret. Math. Theor. Comput. Sci.*, 13(4):51–60, 2011.

19 Sourav Chakraborty, Eldar Fischer, David García-Soriano, and Arie Matsliah. Junto-symmetric functions, hypergraph isomorphism and crunching. In *CCC*, pages 148–158, 2012. doi:10.1109/CCC.2012.28.

20 Sourav Chakraborty, Chandrima kayal, Rajat Mittal, Manaswi Paraashar, Swagato Sanyal, and Nitin Saurabh. On the composition of randomized query complexity and approximate degree. *Electron. Colloquium Comput. Complex.*, 1(1):37–46, 2023. doi:TR23-099.

21 Sourav Chakraborty, Chandrima Kayal, and Manaswi Paraashar. Separations between combinatorial measures for transitive functions. In *ICALP*, volume 229, pages 36:1–36:20, 2022. doi:10.4230/LIPIcs.ICALP.2022.36.

22 Arkadev Chattopadhyay, Nikhil S. Mande, and Suhail Sherif. The log-approximate-rank conjecture is false. *Journal of the ACM*, 67, 2020. URL: https://eccc.weizmann.ac.il/report/2018/176/.

23 Andrew Drucker. Block sensitivity of minterm-transitive functions. *Theor. Comput. Sci.*, 412(41):5796–5801, 2011. doi:10.1016/j.tcs.2011.06.025.

24 Dmitry Gavinsky, Troy Lee, Miklos Santha, and Swagato Sanyal. A Composition Theorem for Randomized Query Complexity via Max-Conflict Complexity. In *ICALP*, volume 132, pages 64:1–64:13, 2019. doi:10.4230/LIPIcs.ICALP.2019.64.

25 Justin Gilmer, Michael E. Saks, and Srikanth Srinivasan. Composition limits and separating examples for some Boolean function complexity measures. *Combinatorica*, 36(3):265–311, 2016. doi:10.1007/s00493-014-3189-x.

26 Mika Göös, T. S. Jayram, Toniann Pitassi, and Thomas Watson. Randomized communication versus partition number. *ACM Trans. Comput. Theory*, 10(1), 2018. doi:10.1145/3170711.

27 Hao Huang. Induced subgraphs of hypercubes and a proof of the sensitivity conjecture. *Annals of Mathematics*, 190(3):949–955, 2019. doi:10.4007/annals.2019.190.3.6.

28 Shelby Kimmel. Quantum adversary (upper) bound. *Chicago Journal of Theoretical Computer Science*, 2013(4), April 2013. doi:10.4086/cjtcs.2013.004.

29 Troy Lee, Rajat Mittal, Ben W. Reichardt, Robert Špalek, and Mario Szegedy. Quantum query complexity of state conversion. In *FOCS*, pages 344–353, 2011. doi:10.1109/FOCS.2011.75.

30 Yaqiao Li. Conflict complexity is lower bounded by block sensitivity. *Theor. Comput. Sci.*, 856:169–172, 2021. doi:10.1016/j.tcs.2020.12.038.

31 Ashley Montanaro. A composition theorem for decision tree complexity. *Chicago Journal of Theoretical Computer Science*, 2014(6), July 2014.

32 Noam Nisan. Crew prams and decision trees. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 327–335, 1989.

33 Noam Nisan and Mario Szegedy. On the degree of boolean functions as real polynomials. *Comput. Complex.*, 4:301–313, 1994. doi:10.1007/BF01263419.

34 Ramamohan Paturi. On the degree of polynomials that approximate symmetric boolean functions. In *STOC*, pages 468–474, 1992. doi:10.1145/129712.129758.

35 Ben W. Reichardt. *Reflections for quantum query algorithms*, pages 560–569. SODA. SIAM, 2011. doi:10.1137/1.9781611973082.44.

36 David Rubinstein. Sensitivity vs. block sensitivity of Boolean functions. *Combinatorica*, 15(2):297–299, 1995. doi:10.1007/BF01200762.

37 Petr Savický. On determinism versus unambiquous nondeterminism for decision trees. *Electron. Colloquium Comput. Complex.*, 9(009), 2002. URL: https://eccc.weizmann.ac.il//report/2002/009/.

38 A. A. Sherstov. Strong direct product theorems for quantum communication and query complexity. *SIAM J. Comput.*, 41(5):1122–1165, 2012. doi:10.1137/110842661.

39 Alexander A. Sherstov. Approximating the AND-OR tree. *Theory Comput.*, 9:653–663, 2013. doi:10.4086/toc.2013.v009a020.

40 Alexander A. Sherstov. The intersection of two halfspaces has high threshold degree. *SIAM J. Comput.*, 42(6):2329–2374, 2013. doi:10.1137/100785260.

41 Alexander A. Sherstov. Making polynomials robust to noise. *Theory Comput.*, 9:593–615, 2013. doi:10.4086/toc.2013.v009a018.

42 Xiaoming Sun. Block sensitivity of weakly symmetric functions. *Theor. Comput. Sci.*, 384(1):87–91, 2007. doi:10.1016/j.tcs.2007.05.020.

43 Xiaoming Sun, Andrew Chi-Chih Yao, and Shengyu Zhang. Graph properties and circular functions: How low can quantum query complexity go? In *CCC*, pages 286–293, 2004. doi:10.1109/CCC.2004.1313851.

44 Avishay Tal. Properties and applications of boolean function composition. In *ITCS*, pages 441–454, 2013. doi:10.1145/2422436.2422485.

A Standard definition of complexity measures

We look at many different complexity measures in the paper, let us start with the formal definition of R and $\widetilde{\deg}$.

▶ **Definition 30** (Randomized query complexity (R)). *Let $f : \{0,1\}^n \to \{0,1,*\}$ be a (possibly partial) Boolean function. A randomized query algorithm A computes f if $\forall x \in \mathrm{Dom}(f), \Pr[A(x) \neq f(x)] \leq 1/3$, where the probability is over the internal randomness of the algorithm. The cost of the algorithm A, cost(A), is the number of queries made in the worst case over any input as well as internal randomness. The randomized query complexity of f, denoted by $\mathrm{R}(f)$, is defined as*

$$\mathrm{R}(f) = \min_{A \text{ computes } f} cost(A).$$

▶ **Definition 31** (Approximate degree ($\widetilde{\deg}$)). *A polynomial $p : \mathbb{R}^n \to \mathbb{R}$ is said to approximate a Boolean function $f : \{0,1\}^n \to \{0,1\}$ if $|p(x) - f(x)| \leq 1/3, \quad \forall x \in \{0,1\}^n$. The approximate degree of f, $\widetilde{\deg}(f)$, is the minimum possible degree of a polynomial which approximates f.*

Note that the constant $1/3$ in the above definitions can be replaced by any constant strictly smaller than $1/2$ which changes $\widetilde{\deg}(f)$ by only a constant factor.

Other than R and $\widetilde{\deg}$, two important related measures are sensitivity ($\mathrm{s}(f)$) and block sensitivity ($\mathrm{bs}(f)$). While the sensitivity and block sensitivity of a total function is well defined, we note that for the case of partial functions there are at least two valid ways of extending the definition from total functions to partial functions. All our results in this paper will hold for partial functions with the following definitions of sensitivity and block sensitivity.

▶ **Definition 32.** *The sensitivity $\mathrm{s}(f,x)$ of a function $f : \{0,1\} \to \{0,1,*\}$ on x is the maximum number s such that there are indices $i_1, i_2, \ldots, i_s \in [n]$ with $f(x^{i_j}) = 1 - f(x)$, for all $1 \leq j \leq s$. Here x^i is obtained from x by flipping the i^{th} bit. The sensitivity of f is defined to be $\mathrm{s}(f) = \max_{x \in \mathrm{Dom}(f)} \mathrm{s}(f,x)$.*

Similarly, the block sensitivity $\mathrm{bs}(f,x)$ of a function $f : \{0,1\} \to \{0,1,*\}$ on x is the maximum number b such that there are disjoint sets $B_1, B_2, \ldots, B_b \subseteq [n]$ with $f(x^{B_j}) = 1 - f(x)$ for all $1 \leq j \leq b$. Recall x^{B_j} is obtained from x by flipping all bits inside B_j. The block sensitivity of f is defined to be $\mathrm{bs}(f) = \max_{x \in \mathrm{Dom}(f)} \mathrm{bs}(f,x)$.

In the definition of block sensitivity, the constraint that the blocks has to be disjoint can be relaxed by extending the definition to "fractional blocks". This gives the measure of fractional block sensitivity.

▶ **Definition 33.** *The fractional block sensitivity $\mathrm{fbs}(f,x)$ of a function $f : \{0,1\} \to \{0,1,*\}$ on x is the maximum value of $\sum_{j=i}^{b} p_j$ such that there are sets $B_1, B_2, \ldots, B_b \subseteq [n]$ and $p_1, \ldots, p_b \in (0,1]$ satisfying the following two conditions.*

- *For each $1 \leq j \leq b$, $f(x^{B_j}) = 1 - f(x)$, and*
- *For each $1 \leq i \leq n$, $\sum_{j:\, i \in B_j} p_j \leq 1$.*

The fractional block sensitivity of f is defined to be $\mathrm{fbs}(f) = \max_{x \in \mathrm{Dom}(f)} \mathrm{fbs}(f,x)$.

B Approximate degree of Promise-OR composed with different inner functions

In this section we show that the approximate degree composes when the outer function is PrOR and the inner functions are (possibly) different partial functions. The proof is essentially a straightforward generalization of the proof of Theorem 21 [11, Theorem 16 (arXiv version)]. However, for the sake of completeness and reader's convenience, we give an overview of the proof here. We will need some definitions and theorems from [11] which we state now. We start with the definition of a problem called "singleton combinatorial group testing". It generalizes the combinatorial group testing problem.

▶ **Definition 34** (Singleton CGT). *Let D be the set of all $w \in \{0,1\}^{2^n}$ for which there exists an $x \in \{0,1\}^n$ such that for all $S \subseteq [n]$ satisfying $\sum_{i \in S} x_i \in \{0,1\}$, we have $\sum_{i \in S} x_i = w_S$. Note that for all $w \in D$, the string x is uniquely defined by $x_i = w_{\{i\}}$. Let us denote this string by $x(w)$. we then define the partial function $\mathsf{SCGT}_{2^n} : D \to \{0,1\}^n$ by $\mathsf{SCGT}_{2^n}(w) = x(w)$.*

▶ **Theorem 35** ([11, Theorem 19 (arXiv version)]). *The bounded-error quantum query complexity of SCGT_{2^n} is $\Theta(\sqrt{n})$.*

For a formal Definition of bounded error quantum query complexity we refer the survey by [15]. Before we state the next result that we need from [11] we are defining robustness of a polynomial to input noise.

▶ **Definition 36** (Robustness to input noise). *For any function $f : \{0,1\}^n \to \{0,1,*\}$ we say a polynomial $p : \{0,1\}^n \to \mathrm{R}$ approximately computes f with δ-robustness where $\delta \in [0, \frac{1}{2})$ if for any $x \in \mathrm{Dom}(f)$ and $\Delta \in [-\delta, \delta]^n$, we have $|f(x) - p(\Delta + x)| \leq \frac{1}{3}$.*

Now we are ready to state the next result.

▶ **Theorem 37** ([11, Theorem 17 (arXiv version)]). *For a partial Boolean function f, there exists a bounded multilinear polynomial p of degree $O(\mathsf{Q}(f))$ that approximates f with robustness $\Omega(1/\mathsf{Q}(f)^2)$ where $\mathsf{Q}(f)$ is the bounded error quantum query complexity of the function f.*

We refer [11] for more details about robustness of a polynomial induces by quantum algorithm. We also need the existence of a multilinear robust polynomial for $\mathsf{XOR}_n \circ \mathsf{SCGT}_{2^n}$, which follows from Theorems 35 and 37 above, where $\mathsf{XOR}_n \circ \mathsf{SCGT}_{2^n}$ is the parity of n output bits of SCGT_{2^n}.

▶ **Theorem 38** ([11, Theorem 20 (arXiv version)]). *There is a real polynomial p of degree $O(\sqrt{n})$ over 2^n variables $\{w_S\}_{S \subseteq [n]}$ and a constant $c \geq 10^{-5}$ such that for any input $w \in \{0,1\}^{2^n}$ with $\mathsf{XOR}_n \circ \mathsf{SCGT}_{2^n}(w) \neq *$ and any $\Delta \in [-c/n, c/n]^{2^n}$,*

$$|p(w + \Delta) - \mathsf{XOR}_n \circ \mathsf{SCGT}_{2^n}(w)| \leq 1/3.$$

Furthermore, p is multilinear and for all $w \in \{0,1\}^{2^n}$, $p(w) \in [0,1]$.

We also need the following result of Sherstov that shows composition holds for the approximate degree of the parity of n different functions.

▶ **Theorem 39** ([38, Theorem 5.9]). *For any partial Boolean functions $f_1, \ldots, f_n$, we have*

$$\widetilde{\mathrm{bdeg}}(\mathsf{XOR} \circ (f_1, \ldots, f_n)) = \Omega\left(\sum_{i=1}^{n} \widetilde{\mathrm{bdeg}}(f_i)\right).$$

Theorem 9 can be proved in the similar line of [11, Theorem 20 (arXiv version)], which we are restating below. For the sake of completeness a detailed proof has been added in the full version of the paper [20].

▶ **Theorem 40.** *For any partial Boolean functions $f_1, f_2, \ldots, f_n$, we have*

$$\widetilde{\mathrm{bdeg}}\left(\mathsf{PrOR}_n \circ (f_1, f_2, \ldots, f_n)\right) = \Omega\left(\frac{\sqrt{n} \cdot \min_{i=1}^{n} \widetilde{\mathrm{bdeg}}(f_i)}{\log n}\right).$$

Furthermore the following upper bound also holds,

$$\widetilde{\mathrm{bdeg}}\left(\mathsf{PrOR}_n \circ (f_1, f_2, \ldots, f_n)\right) = O\left(\sqrt{n} \cdot \max_{i=1}^{n} \widetilde{\mathrm{bdeg}}(f_i) \cdot \log n\right).$$

We will now use this weak bound to establish nearly optimal bound for the approximate degree of PrOR composed with n different *partial* functions. This will again be a simple generalization of OR composed with different functions [11, Theorem 37]. For the sake of completeness, we work out some of the details.

▶ **Theorem 41.** *For any partial Boolean functions $f_1, f_2, \ldots, f_n$, we have*

$$\widetilde{\mathrm{bdeg}}\left(\mathsf{PrOR}_n \circ (f_1, f_2, \ldots, f_n)\right) = \widetilde{\Theta}\left(\sqrt{\sum_{i=1}^{n} \widetilde{\mathrm{bdeg}}(f_i)^2}\right),$$

when the lcm of $\widetilde{\mathrm{bdeg}}(f_i)^2$ for $i \in [n]$ is $\Theta(\max_i \ \widetilde{\mathrm{bdeg}}(f_i)^2)$.

Proof. As mentioned before, the proof is merely working out the details of [11, Theorem 37] while keeping in mind that we are working with *partial* functions.

Let $F = \mathsf{PrOR}_n \circ (f_1, f_2, \ldots, f_n)$, $d_i = \widetilde{\mathrm{bdeg}}(f_i)^2$ for $i \in [n]$, and ℓ be the lcm of d_i's. Now consider the function $G = \mathsf{PrOR}_\ell \circ F$. From Theorem 40, we have the following bounds on $\widetilde{\mathrm{bdeg}}(G)$ up to constants

$$\frac{\sqrt{\ell} \cdot \widetilde{\mathrm{bdeg}}(F)}{\log \ell} \leq \widetilde{\mathrm{bdeg}}(G) \leq \sqrt{\ell} \cdot \widetilde{\mathrm{bdeg}}(F) \cdot \log \ell. \tag{9}$$

Now using the associativity of PrOR we can rewrite G as

$$G = \mathsf{PrOR}_{n\ell} \circ (\underbrace{f_1, \ldots, f_1}_{\ell \text{ times}}, \ldots, \underbrace{f_n, \ldots, f_n}_{\ell \text{ times}}). \tag{10}$$

Further regrouping f_i's, we can rewrite G as follows

$$G = \mathsf{PrOR}_d \circ (\underbrace{\mathsf{PrOR}_{\ell/d_1} \circ f_1, \ldots, \mathsf{PrOR}_{\ell/d_1} \circ f_1}_{d_1 \text{ times}}, \ldots, \underbrace{\mathsf{PrOR}_{\ell/d_n} \circ f_n, \ldots, \mathsf{PrOR}_{\ell/d_n} \circ f_n}_{d_n \text{ times}}), \tag{11}$$

where $d = \sum_{i=1}^{n} d_i$. Now using Theorem 40 and $\sqrt{d_i} = \widetilde{\mathrm{bdeg}}(f_i)$, we obtain following bounds for $\mathsf{PrOR}_{\ell/d_i} \circ f_i$ (up to constants)

$$\frac{\sqrt{\ell}}{\log(\ell/d_i)} \leq \widetilde{\mathrm{bdeg}}(\mathsf{PrOR}_{\ell/d_i} \circ f_i) \leq \sqrt{\ell} \cdot \log(\ell/d_i). \tag{12}$$

Now consider (11) and using Theorem 40 along with (12), we obtain

$$\frac{\sqrt{d\ell}}{\log d \cdot \log \ell} \leq \widetilde{\mathrm{bdeg}}(G) \leq \sqrt{d\ell} \cdot \log \ell \cdot \log d \tag{13}$$

Now from (13) and (9) it follows

$$\frac{\sqrt{d}}{\log d \cdot \log^2 \ell} \leq \widetilde{\mathrm{bdeg}}(F) \leq \sqrt{d} \cdot \log^2 \ell \cdot \log d.$$ ◀

C Composition theorems for strongly-k-junta symmetric outer functions

In this section we will prove the composition result of $\widetilde{\deg}$ and R when the outer function has some amount of symmetry. Of course, there are various notion of symmetry. Traditionally a function is said to have the maximum amount of symmetry when the function value is invariant under any permutation of the variables. Such functions are called symmetric. Symmetric functions are very well studied in the literature of Boolean function analysis. In the terms of composition theorems of $\widetilde{\deg}$ and R it was proved in [11] and [26] that $\widetilde{\deg}$ and R respectively composes when the outer function is symmetric.

In terms of weaker notions of symmetry there are various possible definitions. In this paper we consider the case of strongly-k-junta symmetric functions. The composition theorem for $\widetilde{\deg}$ when the outer function is strongly-k-junta symmetric (Theorem 12(Part(i)) is presented in Appendix C.1. The proof of the composition theorem for R when the outer function is strongly-k-junta symmetric (Theorem 12(Part(ii)) follows easily from Theorem 2.

▶ **Observation 42.** *For any strongly $k-$junta symmetric function $f : \{0,1\}^n \to \{0,1\}$ and any Boolean function $g : \{0,1\}^m \to \{0,1\}$, we have $\mathrm{R}(f \circ g) = \widetilde{\Omega}(\mathrm{R}(f) \cdot \mathrm{R}(g))$ where $n - k = \Theta(n)$.*

Proof. There exists an assignment of the k-bits such that the resulting function is a non-constant symmetric function on $(n - k)$ bits. Since the sensitivity of the restricted function is $\Omega(n)$, the randomized query complexity is also $\Omega(n)$ (see [32]). Hence, from Theorem 2 the result follows. ◀

C.1 Composition of approximate degree for $\sqrt{n}$-junta symmetric functions

In this section, first, we define Multiplexer Function or Addressing Function that will be useful is the analysis.

▶ **Definition 43** (Multiplexer Function or Addressing Function).
The function MUX $: \{0,1\}^{k+2^k} \to \{0,1\}$ *with input $(x_0, \ldots, x_{k-1}, y_0, \ldots, y_{2^k-1})$ outputs the bit y_t, where $t = \sum_{i=0}^{k-1} x_i 2^i$.*

A crucial result that we use in the prove of composition theorem of $\widetilde{\deg}$ is the following result from [34].

▶ **Theorem 44** ([34]). *For any non-constant symmetric function $f : \{0,1\}^n \to \{0,1\}$, let k be the closest integer to $n/2$ such that f takes different values on inputs of Hamming weight k and $k+1$. Define,*

$$\gamma(f) = \begin{cases} k & \text{if } k \leq n/2, \\ n-k & \text{otherwise.} \end{cases}$$

Then

$$\widetilde{\deg}(f) = \Theta\left(\sqrt{n(\gamma(f)+1)}\right).$$

Using the result of [34] we prove the following proposition about the approximate degree of a k-junta symmetric function. Recall the multiplexer function from Definition 43.

▶ **Proposition 45.** *For any k-junta symmetric function $f : \{0,1\}^n \to \{0,1\}$, we have* $\widetilde{\deg}(f) = \Omega\left(\sqrt{(n-k)\gamma_{\max}}\right)$ *and* $\widetilde{\deg}(f) = O\left(\max\{k, \sqrt{(n-k)\gamma_{\max}}\}\right)$*, where* $\gamma_{\max} = \max_{i\in\{0,1\}^k}\{\gamma(f_i)\}$ *such that f_i is the symmetric function obtained by restricting the junta variables according to i.*

Proof. Fixing the junta variables in f we obtain a symmetric function on $n-k$ variables with approximate degree $\Omega(\sqrt{(n-k)\gamma_{\max}})$ (Theorem 44), which in turn implies the same lower bound on $\widetilde{\deg}(f)$.

For the upper bound, we obtain an approximating polynomial for f by composing the (exact) polynomial for the multiplexer function $\mathsf{MUX} : \{0,1\}^{k+2^k} \to \{0,1\}$ with the approximating polynomials for different symmetric functions obtained by restricting the k junta variables. Therefore, $\widetilde{\deg}(f) = k + O(\sqrt{(n-k)\gamma_{\max}}) = O\left(\max\{k, \sqrt{(n-k)\gamma_{\max}}\}\right)$. ◀

As mentioned earlier, the composition of $\widetilde{\deg}$ when the outer function is symmetric was proved in [11]. The following is their result that we crucially use in the proof of Theorem 12.

▶ **Theorem 46** ([11]). *For any symmetric Boolean function $f : \{0,1\}^n \to \{0,1\}$ and any Boolean function $g : \{0,1\}^m \to \{0,1\}$ we have,*

$$\widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g))$$

We now present the proof of Part (i) of Theorem 12, that the proof of composition of $\widetilde{\deg}$ when the outer function is strongly-k-junta symmetric.

Proof of Theorem 12(Part (i)). Since f is a strongly-k-junta symmetric function so there exists a setting of the k junta variables such that the resulting function is a non-constant symmetric function. Let f' be the symmetric function obtained by restricting the junta variables of f so that f' is non-constant. Then by Theorem 44 the approximate degree of f' is $\Omega(\sqrt{(n-k)\gamma_{\max}})$. Then clearly we have

$$\widetilde{\deg}(f \circ g) \geq \widetilde{\deg}(f' \circ g) = \widetilde{\Omega}(\widetilde{\deg}(f') \cdot \widetilde{\deg}(g)) = \widetilde{\Omega}(\sqrt{(n-k)\gamma_{\max}} \cdot \widetilde{\deg}(g)), \tag{14}$$

where the first equality follows from Theorem 46. Now from Proposition 45 we know that $\widetilde{\deg}(f) = O(\sqrt{(n-k)\gamma_{\max}})$ if $k = O(\sqrt{(n-k)\gamma_{\max}})$, which is satisfied when $k = O(\sqrt{n})$. Thus from (14) we obtain

$$\widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g)).$$ ◀