

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 $R(f \circ g) = \tilde{\Theta}(R(f) \cdot R(g))$, is known as the composition question for the randomized query complexity. Similarly, the composition question for the approximate degree asks whether $\widetilde{\deg}(f \circ g) = \tilde{\Theta}(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g))$. 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 $R(f \circ g) = \Omega(\text{noisyR}(f) \cdot R(g))$. This implies that composition holds whenever $\text{noisyR}(f) = \tilde{\Theta}(R(f))$. We show two results:

1. When $R(f) = \Theta(n)$, then $\text{noisyR}(f) = \Theta(R(f))$. 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

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

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

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 On the Composition of Randomized Query Complexity and Approximate Degree

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

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

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 $f : \{0, 1\}^n \rightarrow \{0, 1\}$ and $g : \{0, 1\}^m \rightarrow \{0, 1\}$, the composed function $f \circ g : \{0, 1\}^{nm} \rightarrow \{0, 1\}$ is defined as the function

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

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 :*

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

The notation $\tilde{\Theta}(\cdot)$ 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 problem remains wide open: randomized query complexity (denoted by R) and approximate degree (denoted by \deg) [38, 33, 3, 39, 17, 40]. (See Definition 30 and Definition 31 for their respective formal definitions.)

For both R and \deg the upper bound inequality is known, i.e., $R(f \circ g) = \tilde{O}(R(f) \cdot R(g))$ (folklore) and $\deg(f \circ g) = O(\deg(f) \cdot \deg(g))$ [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 properties¹.

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

84 The initial steps taken towards answering the composition question for R were to show
 85 a lower bound by using a weaker measure of outer function than the randomized query
 86 complexity. In particular, it was shown that $R(f \circ g) = \Omega(s(f) \cdot R(g))$ [26, 5], where $s(f)$
 87 is the sensitivity of f (Definition 32). Since $s(f) = \Theta(R(f))$ for any symmetric function²
 88 f , these results show that R composes when the outer function is a symmetric function
 89 (like OR, AND, Majority, Parity, etc.). The lower bound was later improved to obtain
 90 $R(f \circ g) = \Omega(\text{fbs}(f) \cdot R(g))$ [7, 8], where $\text{fbs}(f)$ is the fractional block sensitivity of f
 91 (Definition 33).

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

99 For composition of $\widetilde{\text{deg}}$, Sherstov [38] already showed that $\widetilde{\text{deg}}(f \circ g)$ composes when the
 100 approximate degree of the outer function f is $\Theta(n)$, where n is the arity of the outer function.
 101 Thus approximate degree composes for several symmetric functions (having approximate
 102 degree $\Theta(n)$, like Majority and Parity). Though, until recently it was not even clear if
 103 $\widetilde{\text{deg}}(\text{OR} \circ \text{AND}) = \Omega(\widetilde{\text{deg}}(\text{OR}) \widetilde{\text{deg}}(\text{AND}))$ (arguably the simplest of composed functions). OR
 104 has approximate degree $O(\sqrt{n})$, and thus the result of [38] does not prove $\widetilde{\text{deg}}$ composition
 105 when the outer function is OR (similarly for AND). In a long series of work by [33, 3, 39, 17, 40],
 106 the question was finally resolved; it was later generalized to the case when the outer function
 107 is symmetric [11].

108 In contrast to R composition, no lower bound on the approximate degree of composed
 109 function is known with a weaker measure on the outer function. It is well known that for any
 110 Boolean function f , $\sqrt{s(f)} \leq \sqrt{\text{bs}(f)} = O(\widetilde{\text{deg}}(f))$ [33]. So a natural step towards proving
 111 $\widetilde{\text{deg}}$ composition is: prove a lower bound on $\widetilde{\text{deg}}(f \circ g)$ by $\sqrt{\text{bs}(f)} \cdot \widetilde{\text{deg}}(g)$. We show this
 112 result by generalizing the approach of [11].

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

117 2 Our Results and Techniques

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

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

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

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

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

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

Sherstov [38] proved that $\widetilde{\text{deg}}(f \circ g) = \Omega(\widetilde{\text{deg}}(f) \cdot \widetilde{\text{deg}}(g))$ 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 $R(f) = \Theta(n)$. Then for all partial functions g , we have*

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

The proof of this theorem is given in Section 4. Notice, since $R(f \circ g) = O(R(f) \cdot R(g) \log R(f))$ (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 $\text{noisyR}(f)$ (the noisy randomized query complexity of f).

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

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

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*

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

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

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

$$162 \quad \text{noisyR}(f) = \Omega\left(\frac{R(f \circ \text{GapMaj}_{\log n})}{\log n}\right). \quad (2)$$

163
164 then the above observation gives $R(f \circ g) = \Omega((R(f) \cdot R(g))/(\log n))$ for all partial functions
165 g .

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

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

172 ► **Theorem 5.** *Let f be a partial function on n bits and let $t \geq 1$, then $R(f \circ \text{GapMaj}_t) =$
173 $O(t \cdot \text{noisyR}(f) + n)$.*

174 Notice that this is a generalization of Ben-David and Blais' characterization of noisyR
175 given by Theorem 3 in one direction. To give an idea of the proof, their characterization
176 (Theorem 3) shows that any noisy oracle algorithm for f can be simulated using only two
177 biases, 1 and $1/\sqrt{n}$ (where n is the arity of f), with only constant overhead. We generalize
178 this by showing that the same simulation works with a slight overhead even when the bias
179 $1/\sqrt{n}$ is replaced by a bias $1/\sqrt{t}$, for some $t \geq 1$. A detailed proof the above theorem has
180 been included in the full version of this paper [20].

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

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

187 ► **Corollary 6.** *Let f be a partial Boolean function. If $R(f \circ g) = \tilde{\Theta}(R(f) \cdot R(g))$ for all
188 partial functions g then $\text{noisyR}(f \circ g) = \tilde{\Theta}(\text{noisyR}(f) \cdot \text{noisyR}(g))$.*

189 2.2 Lower bound on $\widetilde{\text{deg}}(f \circ g)$ in terms of block sensitivity of f and 190 $\widetilde{\text{deg}}(g)$

191 As discussed in the introduction, the composition question for $\widetilde{\text{deg}}$ is only known to hold
192 when the outer function f is symmetric [11] or has high approximate degree [38]. There
193 are also no known lower bounds on $\widetilde{\text{deg}}(f \circ g)$ in terms of weaker measures of f and $\widetilde{\text{deg}}(g)$.
194 Compare this with the situation with respect to composition of R . It was shown in [26] that
195 $R(f \circ g) = \Omega(s(f) R(g))$, where $s(f)$ denotes the sensitivity of f . This was later strengthened
196 to $\Omega(\text{fbs}(f) R(g))$ [7, 8], where $\text{fbs}(f)$ is the fractional block sensitivity of f .

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

199 ► **Theorem 7.** *For all non-constant (possibly partial)³ Boolean functions $f : \{0, 1\}^n \rightarrow \{0, 1\}$*

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

200 and $g : \{0, 1\}^m \rightarrow \{0, 1\}$, we have

$$201 \quad \widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\sqrt{\text{bs}(f)} \cdot \widetilde{\deg}(g)).$$

202 We first note that the above theorem is tight in terms of block sensitivity, i.e., we
203 cannot have $\widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\text{bs}(f)^c \cdot \widetilde{\deg}(g))$ for any $c > 1/2$. This is because the OR
204 function over n bits witnesses the tight quadratic separation between $\widetilde{\deg}$ and bs , i.e.,
205 $\widetilde{\deg}(\text{OR}_n) = \Theta(\sqrt{n}) = \Theta(\sqrt{\text{bs}(\text{OR}_n)})$ [33].

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

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

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

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

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

225 Ideas behind proof of Theorem 7

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

228 Our starting point is the well known Nisan-Szegedy's embedding of PrOR (see Definition 18)
229 over $\text{bs}(f)$ many bits in a Boolean function f [33]. Carrying out this transformation in
230 $f \circ g$ embeds $\text{PrOR}_{\text{bs}(f)} \circ (g_1, \dots, g_{\text{bs}(f)})$ into $f \circ g$, where $g_1, \dots, g_{\text{bs}(f)}$ are different partial
231 functions such that $\widetilde{\text{bdeg}}(g_i) \geq \widetilde{\deg}(g)$ for all $i \in [\text{bs}(f)]^4$. Since the transformation is just
232 substitutions of variables by constants, we further have

$$233 \quad \widetilde{\deg}(f \circ g) \geq \widetilde{\text{bdeg}}(\text{PrOR}_{\text{bs}(f)} \circ (g_1, \dots, g_{\text{bs}(f)})). \quad (3)$$

235 It now looks like that we can appeal to the composition theorem for PrOR (Theorem 21)
236 [11] to obtain our theorem. However, there is a technical difficulty – Theorem 21 doesn't
237 hold for *different* inner *partial* functions. It only deals with a single total inner function. We
238 therefore generalize the proof of Theorem 21 to obtain the following general version of the
239 composition theorem for PrOR.

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

240 ► **Theorem 9.** For any partial Boolean functions g_1, g_2, \dots, g_n , we have

$$241 \quad \widetilde{\text{bdeg}}(\text{PrOR}_n \circ (g_1, g_2, \dots, g_n)) = \Omega\left(\frac{\sqrt{n} \cdot \min_{i=1}^n \widetilde{\text{bdeg}}(g_i)}{\log n}\right).$$

242
243 We can now obtain Theorem 7 from Equation (3) and Theorem 9. The proof of Theorem 9
244 is a generalization of a result due to [11] (see Theorem 21). For lack of space, we present the
245 proof of this theorem in the complete version of this paper [20].

246 2.3 Composition results when the outer functions has some symmetry

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

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

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

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

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

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

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

- 277 ■ $\widetilde{\text{deg}}(f \circ g) = \widetilde{\Theta}(\widetilde{\text{deg}}(f) \cdot \widetilde{\text{deg}}(g))$ where $k = O(\sqrt{n})$.
- 278 ■ $\text{R}(f \circ g) = \widetilde{\Theta}(\text{R}(f) \cdot \text{R}(g))$ where $n - k = \Theta(n)$.

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

283 Organization of the paper

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

289 3 Preliminaries

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

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

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

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

$$303 \quad f \circ (g_1, g_2, \dots, g_n)(x_1, x_2, \dots, x_n) = f(g_1(x_1), g_2(x_2), \dots, g_n(x_n)),$$

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

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

309 3.1 Standard definitions and functions for the composition of R

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

311 **► Definition 14** (Gap-Majority). *The function $\text{GapMaj}_t : \{0, 1\}^t \rightarrow \{0, 1, *\}$ is a partial*
 312 *function with arity t such that*

$$313 \quad \text{GapMaj}_t(x) = \begin{cases} 1 & \text{if } |x| = t/2 + 2\sqrt{t}, \\ 0 & \text{if } |x| = t/2 - 2\sqrt{t}, \\ * & \text{otherwise.} \end{cases}$$

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

315 In regards to the composition question of R, one of the most significant complexity
 316 measures (defined by Ben-David and Blais [9]) is that of noisyR. We first define the noisy
 317 oracle model.
 318

319 ► **Definition 15** (Noisy Oracle Model and Noisy Oracle Access to a String ([9])). For $b \in \{0, 1\}$,
 320 a noisy oracle to b takes a parameter $-1 \leq \gamma \leq 1$ as input and returns a bit b' such that
 321 $\Pr[b' = b] = (1 + \gamma)/2$. The cost of one such query is γ^2 . Each query to noisy oracle returns
 322 independent bits.

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

325 Next, we define the noisy oracle model of computation.

326 ► **Definition 16** (Noisy Oracle Model of Computation ([9])). Let $f : \{0, 1\}^n \rightarrow \{0, 1, *\}$ be a
 327 partial Boolean function. A noisyR query algorithm A computes f if for all $x \in \text{Dom}(f)$,
 328 $\Pr[A(x) \neq f(x)] \leq 1/3$, where A is a randomized algorithm given noisy oracle access to x ,
 329 and the probability is over both noisy oracle calls and the internal randomness of the algorithm
 330 A . The cost of the algorithm A for an input x is the sum of the cost of all noisy oracle calls
 331 made by A on x , and the cost of A , $\text{cost}(A)$, is the maximum cost over all $x \in \text{Dom}(f)$. The
 332 noisyR randomized query complexity of f , denoted by $\text{noisyR}(f)$, is defined as

$$333 \quad \text{noisyR}(f) = \min_{A \text{ computes } f} \text{cost}(A).$$

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

338 3.2 Standard definitions and functions for the composition of $\widetilde{\text{deg}}$

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

345 ► **Definition 17** (Bounded approximate degree ($\widetilde{\text{bdeg}}$)). For a partial Boolean function
 346 $f : \{0, 1\}^n \rightarrow \{0, 1, *\}$, the bounded approximate degree of f ($\widetilde{\text{bdeg}}(f)$) is the minimum
 347 possible degree of a polynomial p such that

$$348 \quad \begin{aligned} & \text{--- } |p(x) - f(x)| \leq 1/3, \quad \forall x \in \text{Dom}(f), \text{ and} \\ 349 & \text{--- } 0 \leq p(x) \leq 1 \quad \forall x \in \{0, 1\}^n. \end{aligned}$$

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

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

355 ► **Definition 18** (Promise-OR). Promise-OR (denoted by PrOR_n) is the function $\text{PrOR}_n : \{0, 1\}^n \rightarrow \{0, 1, *\}$ such that $\text{PrOR}_n(x) = 0$ if $|x| = 0$, equals to 1 if $|x| = 1$, and $*$
 356 otherwise.
 357

358

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

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

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

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

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

$$367 \quad \widetilde{\text{bdeg}}(\text{PrOR}_n \circ g) = \Omega\left(\sqrt{n} \cdot \widetilde{\text{deg}}(g) / \log n\right).$$

369 4 Results about composition of R

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

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

$$378 \quad \text{noisyR}(f) = \Theta\left(\frac{\text{R}(f \circ \text{GapMaj}_n)}{n}\right), \quad (4)$$

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

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

$$387 \quad \text{noisyR}(f) = \Omega\left(\frac{\text{R}(f \circ \text{GapMaj}_{t(n)})}{t(n)}\right),$$

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

390 **Proof.** Suppose $\text{noisyR}(f) = \Omega\left(\frac{\text{R}(f \circ \text{GapMaj}_{t(n)})}{t(n)}\right)$, since $\text{R}(f \circ \text{GapMaj}_t) \geq \text{R}(f)$, we have
391 $\text{noisyR}(f) = \Omega(\text{R}(f)/t(n))$. Theorem 20 implies that a lower bound on noisyR translates to
392 a lower bound on $\text{R}(f \circ g)$. We have,

$$393 \quad \begin{aligned} \text{R}(f \circ g) &= \Omega(\text{noisyR}(f) \cdot \text{R}(g)) && (\text{Theorem 20}) \\ &= \Omega\left(\frac{\text{R}(f) \cdot \text{R}(g)}{t(n)}\right). \end{aligned} \quad \blacktriangleleft$$

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

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

401 ► **Theorem 5.** *Let f be a partial function on n bits and let $t \geq 1$, then $R(f \circ \text{GapMaj}_t) =$
402 $O(t \cdot \text{noisyR}(f) + n)$.*

403 This allows us to show that if for an n -bit partial function f , $R(f) = \Theta(n)$, then
404 $R(f \circ g) = \tilde{\Theta}(R(f) \cdot R(g))$ for all partial functions g (Theorem 2).

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

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

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

412 **4.1 Composition for functions with $R(f) = \Theta(n)$**

413 We restate the theorem below.

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

$$416 \quad R(f \circ g) = \Omega(R(f) \cdot R(g)).$$

417 **Proof.** From Theorem 22 we have a lower bound on the randomized query complexity of
418 $(f \circ \text{GapMaj}_t)$:

$$419 \quad R(f \circ \text{GapMaj}_t) = \Omega(R(f) \cdot \sqrt{t}). \quad (5)$$

421 On the other hand, Theorem 5 gives an upper bound of $O(t \cdot \text{noisyR}(f) + n)$ on $R(f \circ$
422 $\text{GapMaj}_t)$. Thus, choosing $t = \left(\frac{C \cdot n}{\text{noisyR}(f)}\right)$ for a large enough constant C , we have

$$423 \quad R(f) \cdot \sqrt{\frac{n}{\text{noisyR}(f)}} = O\left(\frac{n}{\text{noisyR}(f)} \cdot \text{noisyR}(f) + n\right).$$

425 This implies that

$$426 \quad R(f) = O\left(\sqrt{n \cdot \text{noisyR}(f)}\right). \quad (6)$$

428 Thus, if $R(f) = \Theta(n)$, then $\text{noisyR}(f) = \Omega(R(f))$, which implies composition from The-
429 orem 20. ◀

430 Notice that Equation 6 is equivalent to the following observation.

431 ► **Observation 23.** *Let f be a partial Boolean function on n -bits. Then, $\text{noisyR}(f) =$
432 $\Omega\left(\frac{R(f)^2}{n}\right)$.*

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

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

439 Another implication of Theorem 2 is that composition of R for an outer function f implies
440 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 $R(f \circ g) = \Omega(R(g))$ (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:

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

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 $R(f \circ g) = \Omega(R(g))$, then*

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

Proof.

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

Where the last equality follows from Observation 23⁵ Now there are two cases:

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

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

Again, the claim follows from Equation 7.

The weaker composition, Corollary 24, implies that if $R(f)$ and $R(g)$ 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 $R(f) \cdot R(g)$. In other words, if there is a large polynomial separation between $R(f \circ g)$ and $(R(f) \cdot R(g))$, then $R(f)$ and $R(g)$ 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 $R(f) = \Theta(n^\beta)$, $R(g) = \Theta(n^\gamma)$ and $R(f \circ g) = O((R(f) \cdot R(g))^\alpha)$, where $\alpha < 1$ is a constant, then $(1 - \alpha)(\alpha + \beta) < 1/2$.*

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.

474 **Proof.** For some constants A and B we have

$$475 \quad A \cdot \frac{R(f) \cdot R(g)}{\sqrt{n}} \leq R(f \circ g) \leq B \cdot (R(f) \cdot R(g))^\alpha,$$

476

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

$$479 \quad A \cdot n^{\beta+\gamma-1/2} \leq B \cdot n^{\alpha(\beta+\gamma)}$$

$$480 \quad n^{(1-\alpha)(\beta+\gamma)-1/2} \leq \frac{B}{A}.$$

481

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

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

486 4.3 Proof of Corollary 6

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

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

$$489 \quad \text{noisyR}(f) = \Theta \left(\frac{R(f \circ \text{GapMaj}_t)}{t} \right).$$

490 **Proof.** From Theorem 5 we have for all $t \geq 1$, $R(f \circ \text{GapMaj}_t) = O(t \cdot \text{noisyR}(f) + n)$. Since
491 we have assumed $t = \Omega(n)$ and $\text{noisyR}(f) = \Omega(1)$ (Lemma 19), we get $R(f \circ \text{GapMaj}_t) =$
492 $O(t \cdot \text{noisyR}(f))$. Thus, $\text{noisyR}(f) = \Omega \left(\frac{R(f \circ \text{GapMaj}_t)}{t} \right)$.

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

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

497 **► Corollary 6.** *Let f be a partial Boolean function. If $R(f \circ g) = \tilde{\Theta}(R(f) \cdot R(g))$ for all
498 partial functions g then $\text{noisyR}(f \circ g) = \tilde{\Theta}(\text{noisyR}(f) \cdot \text{noisyR}(g))$.*

499 **Proof.** From Theorem 3, we have

$$500 \quad \text{noisyR}(f \circ g) = \Theta \left(\frac{R((f \circ g) \circ \text{GapMaj}_{mn})}{mn} \right).$$

501

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

$$503 \quad \Theta \left(\frac{R(f \circ (g \circ \text{GapMaj}_{mn}))}{mn} \right).$$

504

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

$$506 \quad \text{noisyR}(f \circ g) = \Theta \left(\frac{R(f) \cdot R(g \circ \text{GapMaj}_{mn})}{mn} \right) \quad (\text{assuming } R \text{ composes})$$

$$507 \quad = \Theta(R(f) \cdot \text{noisyR}(g)) \quad (\text{from Lemma 26})$$

$$508 \quad = \Theta(\text{noisyR}(f) \cdot \text{noisyR}(g)). \quad (\text{assuming } R \text{ composes})$$

509

510 \blacktriangleleft

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 $\sqrt{s(f)}$, $\sqrt{\text{bs}(f)}$ or $\sqrt{\text{fbs}(f)}$). Here we will establish one such lower bound of $\sqrt{\text{bs}(f)} \widetilde{\deg}(g)$.

We restate our theorem now.

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

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

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 $\sqrt{s(f)} \widetilde{\deg}(g)$ 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 $\sqrt{s(f)} \widetilde{\deg}(g)$. 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 ($|S| = s(f)$). 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 $s(f)$ 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 $\widetilde{\deg}(f \circ g) \geq \widetilde{\deg}(f' \circ g) \geq \widetilde{\text{bdeg}}(\text{PrOR}_{|S|} \circ g)$ (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:

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

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 $\text{PrOR}_{|S|}^{x|_S}$, where $\text{PrOR}_n^a(y_1, y_2, \dots, y_n) := \text{PrOR}_n(y_1 \oplus a_1, y_2 \oplus a_2, \dots, y_n \oplus a_n)$ for $a \in \{0, 1\}^n$. Formally, we have

$$\widetilde{\deg}(f \circ g) \geq \widetilde{\deg}(f' \circ g) \geq \widetilde{\text{bdeg}}(\text{PrOR}_{|S|}^{x|_S} \circ g) = \widetilde{\text{bdeg}}(\text{PrOR}_{|S|} \circ (g_1, \dots, g_{|S|})), \quad (8)$$

where $g_i = g$ or $-g$ 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, see Definitions 32 and 17.

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

554 ► **Theorem 9.** *For any partial Boolean functions g_1, g_2, \dots, g_n , we have*

$$555 \quad \widetilde{\text{bdeg}}(\text{PrOR}_n \circ (g_1, g_2, \dots, g_n)) = \Omega \left(\frac{\sqrt{n} \cdot \min_{i=1}^n \widetilde{\text{bdeg}}(g_i)}{\log n} \right).$$

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

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

$$560 \quad \widetilde{\text{deg}}(f \circ g) \geq \widetilde{\text{deg}}(f' \circ g) \geq \widetilde{\text{bdeg}}(\text{PrOR}_{|S|}^{x|_S} \circ g) = \widetilde{\Omega}(\sqrt{s(f)} \widetilde{\text{deg}}(g)).$$

561 We are now ready to present the modifications required to improve the lower bound to
 562 $\widetilde{\Omega}(\sqrt{\text{bs}(f)} \widetilde{\text{deg}}(g))$.

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

$$568 \quad f'(x) = \begin{cases} 0 & \text{if } x = a, \\ 1 & \text{if } x = a^{B_i}, \text{ for some } i \in [b], \\ * & \text{otherwise.} \end{cases}$$

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

571 Since g is non-constant, we can fix the indices *not* in $\bigcup_{i=1}^b B_i$ according to a to obtain
 572 $f'' \circ g$. We would now like to embed PrOR_b over the remaining variables in f'' . For this
 573 purpose we define the following partial functions: for every $i \in [b]$, let $I_i : \{0, 1\}^{B_i} \rightarrow \{0, 1, *\}$
 574 be such that

$$575 \quad I_i(x) = \begin{cases} 0 & \text{if } x = a|_{B_i}, \\ 1 & \text{if } x = a^{B_i}|_{B_i}, \\ * & \text{otherwise.} \end{cases}$$

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

$$578 \quad \widetilde{\text{deg}}(f \circ g) \geq \widetilde{\text{bdeg}}(f' \circ g) \geq \widetilde{\text{bdeg}}(f'' \circ g) = \widetilde{\text{bdeg}}(\text{PrOR}_b \circ (I_1 \circ g, \dots, I_b \circ g)) \\ 579 \quad = \Omega \left(\frac{\sqrt{b} \cdot \min_i \widetilde{\text{bdeg}}(I_i \circ g)}{\log b} \right) = \widetilde{\Omega} \left(\sqrt{b} \cdot \widetilde{\text{deg}}(g) \right),$$

581 where the second last equality follows from Theorem 9 and the last equality uses the fact
 582 $\widetilde{\text{bdeg}}(I_i \circ g) \geq \widetilde{\text{deg}}(g)$ for all i , which in turn follows from each I_i being non-constant. ◀

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

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

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

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

594 6 Conclusion

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

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

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

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

606 ▶ **Open question 28.** For all Boolean functions f and g , can we prove either of the following:
 607 ■ $\widetilde{\text{deg}}(f \circ g) = \Omega(\sqrt{\text{deg}(f)} \cdot \widetilde{\text{deg}}(g))$? ■ $\widetilde{\text{deg}}(f \circ g) = \Omega(\sqrt{\text{fbs}(f)} \cdot \widetilde{\text{deg}}(g))$?

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

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

613 ——— References ———

- 614 1 Scott Aaronson, Shalev Ben-David, and Robin Kothari. Separations in query complexity using
 615 cheat sheets. In *STOC*, pages 863–876, 2016. doi:10.1145/2897518.2897644.
- 616 2 Scott Aaronson, Shalev Ben-David, Robin Kothari, Shrawas Rao, and Avishay Tal. Degree
 617 vs. approximate degree and quantum implications of Huang's sensitivity theorem. In *STOC*,
 618 pages 1330–1342, 2021. doi:10.1145/3406325.3451047.
- 619 3 Andris Ambainis. Polynomial degree and lower bounds in quantum complexity: Collision and
 620 element distinctness with small range. *Theory Comput.*, 1(1):37–46, 2005. doi:10.4086/toc.
 621 2005.v001a003.
- 622 4 Andris Ambainis, Kaspars Balodis, Aleksandrs Belovs, Troy Lee, Miklos Santha, and Juris
 623 Smotrovs. Separations in query complexity based on pointer functions. *Journal of the ACM*,
 624 64(5):32:1–32:24, 2017. doi:10.1145/3106234.
- 625 5 Andris Ambainis, Martins Kokainis, and Robin Kothari. Nearly optimal separations between
 626 communication (or query) complexity and partitions. In *CCC*, volume 50, pages 4:1–4:14,
 627 2016. doi:10.4230/LIPIcs.CCC.2016.4.

- 628 **6** Anurag Anshu, Dmitry Gavinsky, Rahul Jain, Srijita Kundu, Troy Lee, Priyanka Mukho-
629 padhyay, Miklos Santha, and Swagato Sanyal. A composition theorem for randomized query
630 complexity. In *FSTTCS*, volume 93, pages 10:1–10:13, 2017. doi:10.4230/LIPIcs.FSTTCS.
631 2017.10.
- 632 **7** Andrew Basilakis, Andrew Drucker, Mika Göös, Lunjia Hu, Weiyun Ma, and Li-Yang Tan.
633 The power of many samples in query complexity. In *ICALP*, volume 168, pages 9:1–9:18, 2020.
634 doi:10.4230/LIPIcs.ICALP.2020.9.
- 635 **8** Shalev Ben-David and Eric Blais. A tight composition theorem for the randomized query
636 complexity of partial functions. *arXiv:2002.10809*, 2020.
- 637 **9** Shalev Ben-David and Eric Blais. A tight composition theorem for the randomized query
638 complexity of partial functions: Extended abstract. In *FOCS*, pages 240–246, 2020. doi:
639 10.1109/FOCS46700.2020.00031.
- 640 **10** Shalev Ben-David, Eric Blais, Mika Göös, and Gilbert Maystre. Randomised composition
641 and small-bias minimax. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer
642 Science (FOCS)*, pages 624–635. IEEE, 2022.
- 643 **11** Shalev Ben-David, Adam Bouland, Ankit Garg, and Robin Kothari. Classical lower bounds
644 from quantum upper bounds. In *FOCS*, pages 339–349, 2018. doi:10.1109/FOCS.2018.00040.
- 645 **12** Shalev Ben-David, Mika Göös, Robin Kothari, and Thomas Watson. When Is Amplification
646 Necessary for Composition in Randomized Query Complexity? In *APPROX/RANDOM 2020*,
647 volume 176, pages 28:1–28:16, 2020. doi:10.4230/LIPIcs.APPROX/RANDOM.2020.28.
- 648 **13** Shalev Ben-David and Robin Kothari. Randomized query complexity of sabotaged and
649 composed functions. *Theory Comput.*, 14(1):1–27, 2018. doi:10.4086/toc.2018.v014a005.
- 650 **14** Eric Blais, Amit Weinstein, and Yuichi Yoshida. Partially symmetric functions are efficiently
651 isomorphism testable. *SIAM J. Comput.*, 44(2):411–432, 2015. doi:10.1137/140971877.
- 652 **15** Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: a
653 survey. *Theoretical Computer Science*, 288(1):21–43, 2002. doi:https://doi.org/10.1016/
654 S0304-3975(01)00144-X.
- 655 **16** Mark Bun, Robin Kothari, and Justin Thaler. Quantum algorithms and approxim-
656 ating polynomials for composed functions with shared inputs. In Timothy M. Chan, ed-
657 itor, *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms,*
658 *SODA 2019, San Diego, California, USA, January 6-9, 2019*, pages 662–678. SIAM, 2019.
659 doi:10.1137/1.9781611975482.42.
- 660 **17** Mark Bun and Justin Thaler. Dual lower bounds for approximate degree and Markov-Bernstein
661 inequalities. In *ICALP*, volume 7965, pages 303–314, 2013. doi:10.1007/978-3-642-39206-1\
662 _26.
- 663 **18** Sourav Chakraborty. On the sensitivity of cyclically-invariant boolean functions. *Discret.*
664 *Math. Theor. Comput. Sci.*, 13(4):51–60, 2011.
- 665 **19** Sourav Chakraborty, Eldar Fischer, David García-Soriano, and Arie Matsliah. Junto-symmetric
666 functions, hypergraph isomorphism and crunching. In *CCC*, pages 148–158, 2012. doi:
667 10.1109/CCC.2012.28.
- 668 **20** Sourav Chakraborty, Chandrima kayal, Rajat Mittal, Manaswi Paraashar, Swagato Sanyal,
669 and Nitin Saurabh. On the composition of randomized query complexity and approximate
670 degree. *Electron. Colloquium Comput. Complex.*, 1(1):37–46, 2023. URL: https://eccc.
671 weizmann.ac.il/report/2023/099/, doi:TR23-099.
- 672 **21** Sourav Chakraborty, Chandrima Kayal, and Manaswi Paraashar. Separations between com-
673 binatorial measures for transitive functions. In *ICALP*, volume 229, pages 36:1–36:20, 2022.
674 doi:10.4230/LIPIcs.ICALP.2022.36.
- 675 **22** Arkadev Chattopadhyay, Nikhil S. Mande, and Suhail Sherif. The log-approximate-rank
676 conjecture is false. *Journal of the ACM*, 67, 2020. URL: https://eccc.weizmann.ac.il/
677 report/2018/176/.
- 678 **23** Andrew Drucker. Block sensitivity of minterm-transitive functions. *Theor. Comput. Sci.*,
679 412(41):5796–5801, 2011. doi:10.1016/j.tcs.2011.06.025.

- 680 24 Dmitry Gavinsky, Troy Lee, Miklos Santha, and Swagato Sanyal. A Composition Theorem for
681 Randomized Query Complexity via Max-Conflict Complexity. In *ICALP*, volume 132, pages
682 64:1–64:13, 2019. doi:10.4230/LIPIcs.ICALP.2019.64.
- 683 25 Justin Gilmer, Michael E. Saks, and Srikanth Srinivasan. Composition limits and separating
684 examples for some Boolean function complexity measures. *Combinatorica*, 36(3):265–311, 2016.
685 doi:10.1007/s00493-014-3189-x.
- 686 26 Mika Göös, T. S. Jayram, Toniann Pitassi, and Thomas Watson. Randomized communication
687 versus partition number. *ACM Trans. Comput. Theory*, 10(1), 2018. doi:10.1145/3170711.
- 688 27 Hao Huang. Induced subgraphs of hypercubes and a proof of the sensitivity conjecture. *Annals
689 of Mathematics*, 190(3):949–955, 2019. doi:10.4007/annals.2019.190.3.6.
- 690 28 Shelby Kimmel. Quantum adversary (upper) bound. *Chicago Journal of Theoretical Computer
691 Science*, 2013(4), April 2013. doi:10.4086/cjtcs.2013.004.
- 692 29 Troy Lee, Rajat Mittal, Ben W. Reichardt, Robert Špalek, and Mario Szegedy. Quantum query
693 complexity of state conversion. In *FOCS*, pages 344–353, 2011. doi:10.1109/FOCS.2011.75.
- 694 30 Yaqiao Li. Conflict complexity is lower bounded by block sensitivity. *Theor. Comput. Sci.*,
695 856:169–172, 2021. doi:10.1016/j.tcs.2020.12.038.
- 696 31 Ashley Montanaro. A composition theorem for decision tree complexity. *Chicago Journal of
697 Theoretical Computer Science*, 2014(6), July 2014.
- 698 32 Noam Nisan. Crew prams and decision trees. In *Proceedings of the twenty-first annual ACM
699 symposium on Theory of computing*, pages 327–335, 1989.
- 700 33 Noam Nisan and Mario Szegedy. On the degree of boolean functions as real polynomials.
701 *Comput. Complex.*, 4:301–313, 1994. doi:10.1007/BF01263419.
- 702 34 Ramamohan Paturi. On the degree of polynomials that approximate symmetric boolean
703 functions. In *STOC*, page 468–474, 1992. doi:10.1145/129712.129758.
- 704 35 Ben W. Reichardt. *Reflections for quantum query algorithms*, pages 560–569. SODA. SIAM,
705 2011. doi:10.1137/1.9781611973082.44.
- 706 36 David Rubinfeld. Sensitivity vs. block sensitivity of Boolean functions. *Combinatorica*,
707 15(2):297–299, 1995. doi:10.1007/BF01200762.
- 708 37 Petr Savický. On determinism versus unambiguous nondeterminism for decision trees. *Electron.
709 Colloquium Comput. Complex.*, 9(009), 2002. URL: [https://eccc.weizmann.ac.il/report/
710 2002/009/](https://eccc.weizmann.ac.il/report/2002/009/).
- 711 38 A. A. Sherstov. Strong direct product theorems for quantum communication and query
712 complexity. *SIAM J. Comput.*, 41(5):1122–1165, 2012. doi:10.1137/110842661.
- 713 39 Alexander A. Sherstov. Approximating the AND-OR tree. *Theory Comput.*, 9:653–663, 2013.
714 doi:10.4086/toc.2013.v009a020.
- 715 40 Alexander A. Sherstov. The intersection of two halfspaces has high threshold degree. *SIAM J.
716 Comput.*, 42(6):2329–2374, 2013. doi:10.1137/100785260.
- 717 41 Alexander A. Sherstov. Making polynomials robust to noise. *Theory Comput.*, 9:593–615,
718 2013. doi:10.4086/toc.2013.v009a018.
- 719 42 Xiaoming Sun. Block sensitivity of weakly symmetric functions. *Theor. Comput. Sci.*,
720 384(1):87–91, 2007. doi:10.1016/j.tcs.2007.05.020.
- 721 43 Xiaoming Sun, Andrew Chi-Chih Yao, and Shengyu Zhang. Graph properties and circular
722 functions: How low can quantum query complexity go? In *CCC*, pages 286–293, 2004.
723 doi:10.1109/CCC.2004.1313851.
- 724 44 Avishay Tal. Properties and applications of boolean function composition. In *ITCS*, page
725 441–454, 2013. doi:10.1145/2422436.2422485.

726 **A** Standard definition of complexity measures

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

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

$$735 \quad R(f) = \min_{A \text{ computes } f} \text{cost}(A).$$

736

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

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

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

749 **► Definition 32.** *The sensitivity $s(f, x)$ of a function $f : \{0, 1\} \rightarrow \{0, 1, *\}$ on x is the*
750 *maximum number s such that there are indices $i_1, i_2, \dots, i_s \in [n]$ with $f(x^{i_j}) = 1 - f(x)$,*
751 *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*
752 *defined to be $s(f) = \max_{x \in \text{Dom}(f)} s(f, x)$.*

753 *Similarly, the block sensitivity $\text{bs}(f, x)$ of a function $f : \{0, 1\} \rightarrow \{0, 1, *\}$ on x is the*
754 *maximum number b such that there are disjoint sets $B_1, B_2, \dots, B_b \subseteq [n]$ with $f(x^{B_j}) =$*
755 *$1 - f(x)$ for all $1 \leq j \leq b$. Recall x^{B_j} is obtained from x by flipping all bits inside B_j . The*
756 *block sensitivity of f is defined to be $\text{bs}(f) = \max_{x \in \text{Dom}(f)} \text{bs}(f, x)$.*

757 In the definition of block sensitivity, the constraint that the blocks has to be disjoint can be
758 relaxed by extending the definition to “fractional blocks”. This gives the measure of fractional
759 block sensitivity.

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

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

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

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 $\text{SCGT}_{2^n} : D \rightarrow \{0, 1\}^n$ by $\text{SCGT}_{2^n}(w) = x(w)$.*

► **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 $f : \{0, 1\}^n \rightarrow \{0, 1, *\}$ we say a polynomial $p : \{0, 1\}^n \rightarrow \mathbb{R}$ approximately computes f with δ -robustness where $\delta \in [0, \frac{1}{2}]$ if for any $x \in \text{Dom}(f)$ 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 $O(Q(f))$ that approximates f with robustness $\Omega(1/Q(f)^2)$ where $Q(f)$ is the bounded error quantum query complexity of the function f .*

We refer [11] for more details about robustness of a polynomial induced by quantum algorithm. We also need the existence of a multilinear robust polynomial for $\text{XOR}_n \circ \text{SCGT}_{2^n}$, which follows from Theorems 35 and 37 above, where $\text{XOR}_n \circ \text{SCGT}_{2^n}$ 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 $\text{XOR}_n \circ \text{SCGT}_{2^n}(w) \neq *$ and any $\Delta \in [-c/n, c/n]^{2^n}$,*

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

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, \dots, f_n , we have*

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

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

807 ► **Theorem 40.** *For any partial Boolean functions f_1, f_2, \dots, f_n , we have*

$$808 \quad \widetilde{\text{bdeg}}(\text{PrOR}_n \circ (f_1, f_2, \dots, f_n)) = \Omega\left(\frac{\sqrt{n} \cdot \min_{i=1}^n \widetilde{\text{bdeg}}(f_i)}{\log n}\right).$$

809 *Furthermore the following upper bound also holds,*

$$810 \quad \widetilde{\text{bdeg}}(\text{PrOR}_n \circ (f_1, f_2, \dots, f_n)) = O\left(\sqrt{n} \cdot \max_{i=1}^n \widetilde{\text{bdeg}}(f_i) \cdot \log n\right).$$

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

815 ► **Theorem 41.** *For any partial Boolean functions f_1, f_2, \dots, f_n , we have*

$$816 \quad \widetilde{\text{bdeg}}(\text{PrOR}_n \circ (f_1, f_2, \dots, f_n)) = \tilde{\Theta}\left(\sqrt{\sum_{i=1}^n \widetilde{\text{bdeg}}(f_i)^2}\right),$$

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

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

821 Let $F = \text{PrOR}_n \circ (f_1, f_2, \dots, f_n)$, $d_i = \widetilde{\text{bdeg}}(f_i)^2$ for $i \in [n]$, and ℓ be the lcm of d_i 's. Now
 822 consider the function $G = \text{PrOR}_\ell \circ F$. From Theorem 40, we have the following bounds on
 823 $\widetilde{\text{bdeg}}(G)$ up to constants

$$824 \quad \frac{\sqrt{\ell} \cdot \widetilde{\text{bdeg}}(F)}{\log \ell} \leq \widetilde{\text{bdeg}}(G) \leq \sqrt{\ell} \cdot \widetilde{\text{bdeg}}(F) \cdot \log \ell. \quad (9)$$

825
 826 Now using the associativity of PrOR we can rewrite G as

$$827 \quad G = \text{PrOR}_{n\ell} \circ \underbrace{(f_1, \dots, f_1)}_{\ell \text{ times}}, \dots, \underbrace{(f_n, \dots, f_n)}_{\ell \text{ times}}. \quad (10)$$

828
 829 Further regrouping f_i 's, we can rewrite G as follows

$$830 \quad G = \text{PrOR}_d \circ \underbrace{(\text{PrOR}_{\ell/d_1} \circ f_1, \dots, \text{PrOR}_{\ell/d_1} \circ f_1)}_{d_1 \text{ times}}, \dots, \underbrace{(\text{PrOR}_{\ell/d_n} \circ f_n, \dots, \text{PrOR}_{\ell/d_n} \circ f_n)}_{d_n \text{ times}}, \quad (11)$$

831
 832 where $d = \sum_{i=1}^n d_i$. Now using Theorem 40 and $\sqrt{d_i} = \widetilde{\text{bdeg}}(f_i)$, we obtain following bounds
 833 for $\text{PrOR}_{\ell/d_i} \circ f_i$ (up to constants)

$$834 \quad \frac{\sqrt{\ell}}{\log(\ell/d_i)} \leq \widetilde{\text{bdeg}}(\text{PrOR}_{\ell/d_i} \circ f_i) \leq \sqrt{\ell} \cdot \log(\ell/d_i). \quad (12)$$

835

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

$$837 \frac{\sqrt{d\ell}}{\log d \cdot \log \ell} \leq \widetilde{\text{bdeg}}(G) \leq \sqrt{d\ell} \cdot \log \ell \cdot \log d \quad (13)$$

839 Now from (13) and (9) it follows

$$840 \frac{\sqrt{d}}{\log d \cdot \log^2 \ell} \leq \widetilde{\text{bdeg}}(F) \leq \sqrt{d} \cdot \log^2 \ell \cdot \log d.$$

841

842 **C** Composition theorems for strongly- k -junta symmetric outer 843 functions

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

851 In terms of weaker notions of symmetry there are various possible definitions. In this paper
852 we consider the case of strongly- k -junta symmetric functions. The composition theorem for
853 $\widetilde{\text{deg}}$ when the outer function is strongly- k -junta symmetric (Theorem 12(Part(i))) is presented
854 in Appendix C.1. The proof of the composition theorem for R when the outer function is
855 strongly- k -junta symmetric (Theorem 12(Part(ii))) follows easily from Theorem 2.

856 **► Observation 42.** For any strongly k -junta symmetric function $f : \{0, 1\}^n \rightarrow \{0, 1\}$
857 and any Boolean function $g : \{0, 1\}^m \rightarrow \{0, 1\}$, we have $\text{R}(f \circ g) = \widetilde{\Omega}(\text{R}(f) \cdot \text{R}(g))$ where
858 $n - k = \Theta(n)$.

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

863 C.1 Composition of approximate degree for \sqrt{n} -junta symmetric 864 functions

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

867 **► Definition 43** (Multiplexer Function or Addressing Function). The function $\text{MUX} : \{0, 1\}^{k+2^k} \rightarrow$
868 $\{0, 1\}$ with input $(x_0, \dots, x_{k-1}, y_0, \dots, y_{2^k-1})$ outputs the bit y_t , where $t = \sum_{i=0}^{k-1} x_i 2^i$.

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

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

$$874 \gamma(f) = \begin{cases} k & \text{if } k \leq n/2, \\ n - k & \text{otherwise.} \end{cases}$$

875 Then

$$876 \quad \widetilde{\deg}(f) = \Theta\left(\sqrt{n(\gamma(f) + 1)}\right).$$

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

879 ► **Proposition 45.** *For any k -junta symmetric function $f : \{0, 1\}^n \rightarrow \{0, 1\}$, we have*
880 $\widetilde{\deg}(f) = \Omega\left(\sqrt{(n-k)\gamma_{\max}}\right)$ *and* $\widetilde{\deg}(f) = O\left(\max\{k, \sqrt{(n-k)\gamma_{\max}}\}\right)$, *where* $\gamma_{\max} =$
881 $\max_{i \in \{0, 1\}^k} \{\gamma(f_i)\}$ *such that* f_i *is the symmetric function obtained by restricting the junta*
882 *variables according to* i .

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

886 For the upper bound, we obtain an approximating polynomial for f by composing
887 the (exact) polynomial for the multiplexer function $\text{MUX} : \{0, 1\}^{k+2^k} \rightarrow \{0, 1\}$ with the
888 approximating polynomials for different symmetric functions obtained by restricting the k
889 junta variables. Therefore, $\widetilde{\deg}(f) = k + O(\sqrt{(n-k)\gamma_{\max}}) = O\left(\max\{k, \sqrt{(n-k)\gamma_{\max}}\}\right)$.
890 ◀

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

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

$$895 \quad \widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g))$$

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

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

$$904 \quad \widetilde{\deg}(f \circ g) \geq \widetilde{\deg}(f' \circ g) = \widetilde{\Omega}(\widetilde{\deg}(f') \cdot \widetilde{\deg}(g)) = \widetilde{\Omega}(\sqrt{(n-k)\gamma_{\max}} \cdot \widetilde{\deg}(g)), \quad (14)$$

906 where the first equality follows from Theorem 46. Now from Proposition 45 we know that
907 $\widetilde{\deg}(f) = O(\sqrt{(n-k)\gamma_{\max}})$ if $k = O(\sqrt{(n-k)\gamma_{\max}})$, which is satisfied when $k = O(\sqrt{n})$.
908 Thus from (14) we obtain

$$909 \quad \widetilde{\deg}(f \circ g) = \widetilde{\Omega}(\widetilde{\deg}(f) \cdot \widetilde{\deg}(g)).$$

910 ◀