Quantum Search With Generalized Wildcards

Arjan Cornelissen* Nikhil S. Mande† Subhasree Patro‡ Nithish Raja§
Swagato Sanyal¶

Abstract

In the "search with wildcards" problem [Ambainis, Montanaro, Quantum Inf. Comput.'14], one's goal is to learn an unknown bit-string $x \in \{-1,1\}^n$. An algorithm may, at unit cost, test equality of any subset of the hidden string with a string of its choice. Ambainis and Montanaro showed a quantum algorithm of cost $O(\sqrt{n}\log n)$ and a near-matching lower bound of $\Omega(\sqrt{n})$. Belovs [Comput. Comp.'15] subsequently showed a tight $O(\sqrt{n})$ upper bound.

We consider a natural generalization of this problem, parametrized by a subset $\mathcal{Q} \subseteq 2^{[n]}$, where an algorithm may test whether $x_S = b$ for an arbitrary $S \in \mathcal{Q}$ and $b \in \{-1,1\}^S$ of its choice, at unit cost. We show the following:

- For all $k \in [n]$, when $\mathcal{Q}$ is the collection of all sets of size at most k, the quantum query complexity is $\Theta(n/\sqrt{k})$. In particular when $k = n$, this corresponds to the standard search with wildcards setting. This recovers and generalizes the tight characterization of Belovs, and Ambainis and Montanaro, using completely different techniques.
- When $\mathcal{Q}$ is the collection of contiguous blocks, the quantum query complexity is $\widetilde{\Theta}(n)$.
- When $\mathcal{Q}$ is the collection of prefixes, the quantum query complexity is $\Theta(n)$.

All of these results are derived using a framework that we develop. We apply a symmetry reduction to the *primal* version of the negative-weight adversary bound, and show that the quantum query complexity of learning x is characterized, up to a constant factor, by a particular optimization program, which can be succinctly described as follows: 'maximize over all odd functions $f : \{-1,1\}^n \to \mathbb{R}$ the ratio of the maximum value of f to the maximum (over $T \in \mathcal{Q}$) standard deviation of f on a subcube whose free variables are exactly T.'

To the best of our knowledge, ours is the first work to use the primal version of the negative-weight adversary bound (which is a *maximization* program typically used to show lower bounds) to show new quantum query *upper* bounds without explicitly resorting to SDP duality.

*Simons Institute for the Theory of Computing, University of California, Berkeley, United States of America ajcornelissen@outlook.com

†University of Liverpool, UK mande@liverpool.ac.uk

‡Eindhoven University of Technology, Netherlands patrofied@gmail.com

§Eindhoven University of Technology, Netherlands n.r.raja@tue.nl

¶University of Sheffield, UK swagato.sanyal@sheffield.ac.uk

Contents

1 Introduction

Reconstructing an unknown string from limited access to its substrings is a natural and widely studied task across many areas of science and engineering. In computational biology, for example, genome assembly and DNA sequencing rely on reconstructing a long unknown strand from local substring measurements (see, for example, [SS95, RSC+10, CWE15, LD24] and the references therein). Beyond applied motivations, the problem provides a clean and mathematically tractable model for understanding the complexity of learning a hidden string via substring queries.

Most relevant to our work is the *search with wildcards* problem, introduced by Ambainis and Montanaro [AM14]. In this problem, the goal is to learn a hidden bitstring $x \in \{-1,1\}^n$ (i.e., an algorithm must correctly output x with high probability) using the minimum number of queries in the following *wildcard query* model. At unit cost, an algorithm may "check its guess for x" on any subset of the inputs. Formally, at unit cost, an algorithm may make a query (S, b) where $S \subseteq [n], b \in \{-1,1\}^S$. The query returns 1 if $x_S = b$, and 0 otherwise. Wildcard queries are a generalization of the usual query model, which only allows for $|S| = 1$. Classically even in the wildcard model every query provides at most 1 bit of information, yielding an information-theoretic lower bound of $\Omega(n)$ for learning x. In the usual quantum query model with $|S| = 1$, even though roughly a factor-2 speedup is possible [vD98], this is essentially the best speedup one can get [FGGS99].

Ambainis and Montanaro [AM14] showed that in the wildcard model, quantum algorithms can learn x using only $O(\sqrt{n}\log n)$ queries, exhibiting roughly a quadratic quantum speedup over the classical lower bound. However, their analysis and algorithm crucially rely on the ability to query *all subsets* of coordinates. In many practical settings, this assumption can be unrealistic. For instance, in *de novo* genome sequencing and related problems in computational biology [RSC+10], one has no promise on the input string, and the goal is to reconstruct it. In practical sequencing settings, the accessible measurements are limited to contiguous segments of the underlying strand; one cannot, for instance, simultaneously probe all odd positions at unit cost. Likewise, in more general settings one might be restricted to a fixed family of allowed subsets $\mathcal{Q} \subseteq 2^{[n]}$, and the structure of $\mathcal{Q}$ can significantly affect the power of the algorithm.

In this work, we study the quantum query complexities of learning an unknown string when the allowed query subsets are drawn from an arbitrary fixed collection $\mathcal{Q}$. This unifies previously studied models, including the standard query model $\mathcal{Q} = \{\{i\} : i \in [n]\}$, and the search-with-wildcards model $\mathcal{Q} = 2^{[n]}$ under a single framework.

1.1 Our Results

We refer the reader to Section 2 for formal definitions. We summarize here the notions that we need to state our main results. For an arbitrary collection of sets $\mathcal{Q} \subseteq 2^{[n]}$ and a function $F : \{-1,1\}^n \to E$, let $\mathsf{Q}^{\mathcal{Q}}(F)$ denote the quantum query query complexity of computing F, where an algorithm can make unit-cost queries of the form $\mathbb{I}[x_S = b_S]$ for an arbitrary $S \in \mathcal{Q}$ and $b \in \{-1,1\}^S$ of its choice. We call $\mathsf{Q}^{\mathcal{Q}}(F)$ the *quantum $\mathcal{Q}$-substring query complexity of F*. Let $\oplus(\cdot)$ denote the parity function, i.e., $\oplus(x) = \prod_{i=1}^n x_i$ for $x \in \{-1,1\}^n$. A function $f : \{-1,1\}^n \to \mathbb{R}$ is said to be *odd* if $f(x) = -f(-x)$ for all $x \in \{-1,1\}^n$. We write $\mathbf{Var}(f) := \mathbf{Var}_{\mathbf{x}\sim\{-1,1\}^n}[f(\mathbf{x})]$ for the variance of the random variable $f(\mathbf{x})$, where $\mathbf{x}$ is sampled uniformly at random from $\{-1,1\}^n$.

Our main technical contribution is to show that the quantum $\mathcal{Q}$-substring query complexity of parity is characterized by the optimal value of an abstract analytic optimization program. Below,

$f_{S|b}$ refers to the function f restricted to the bits in $\bar{S}$ set to b.

Theorem 1.1. *Let n be a positive integer and $\mathcal{Q} \subseteq 2^{[n]}$. Then $\mathsf{Q}^{\mathcal{Q}}(\oplus) = \Theta(\mathsf{val}_{\mathcal{Q}})$, where*

$$\mathsf{val}_{\mathcal{Q}} := \left(\max_{\substack{f:\{-1,1\}^n \to \mathbb{R} \\ f \text{ is an odd function}}} \frac{\|f\|_\infty}{\max\limits_{\substack{S \in \mathcal{Q}, \\ b \in \{-1,1\}^{\bar{S}}}} \sqrt{\mathbf{Var}(f_{S|b})}} \right). \tag{1}$$

In other words, this characterization can concisely be stated as:

> $\mathsf{Q}^{\mathcal{Q}}(\oplus)$ is asymptotically equal to the maximum, over all odd functions $f : \{-1,1\}^n \to \mathbb{R}$, of the ratio of $\|f\|_\infty$ to the largest standard deviation of f in a subcube whose free variables are a valid query set.

In order to show this, we use the characterization of quantum query complexity by the negative-weight adversary bound [HLŠ07, Rei11]. We show that for all $\mathcal{Q}$, the corresponding adversary bound *equals* $\mathsf{val}_{\mathcal{Q}}$ as defined above. We next analyze $\mathsf{val}_{\mathcal{Q}}$, as defined in Equation (1) for various interesting and natural settings of $\mathcal{Q}$ such as bounded-size sets, contiguous blocks, prefixes, and only the full set.

Theorem 1.2. *Let n be a positive integer. Then, the following bounds hold for $\mathsf{val}_{\mathcal{Q}}$ for various classes of $\mathcal{Q} \subseteq 2^{[n]}$.*

Query Set	$\mathcal{Q}$	$\mathsf{val}_{\mathcal{Q}}$	*Reference*
Bounded-size sets	$\{S \subseteq [n] : \lvert S\rvert \leq k\}$	$\Theta\left(\frac{n}{\sqrt{k}}\right)$	*Theorem 4.2*
Contiguous blocks	$\{S = [i,j] : i \leq j \in [n]\}$	$\widetilde{\Theta}(n)$	*Theorem 4.4*
Prefixes	$\{[1,i] : i \in [n]\}$	$\Theta(n)$	*Theorem 4.5*
Only full set	$\{[n]\}$	$2^{(n-1)/2}$	*Theorem 4.7*

These results have immediate implications for the query complexity of the string-recovery problem in these query models. Let $\mathsf{rec} : \{-1,1\}^n \to \{-1,1\}^n$ be the (string recovery/learning) function defined by $\mathsf{rec}(x) = x$. Using Theorem 1.1, Theorem 1.2, and the Bernstein-Vazirani algorithm [BV97] to recover a bitstring in one query given query access to arbitrary parities of it, we obtain the following corollary.[1]

Corollary 1.3. *Let n be a positive integer. Then, the following bounds hold on the quantum*

[1]The upper bounds for contiguous and prefix queries follow from the naive upper bound of querying one bit at a time (or an increasingly long prefix in the prefix query model) to learn an input string in $O(n)$ queries. The upper bound for full set queries follows from [Gro96].

$\mathcal{Q}$-*substring query complexity of learning an input string.*

Query Set	$\mathcal{Q}$	$\mathrm{Q}^{\mathcal{Q}}(\mathrm{rec})$
Bounded-size sets	$\{S \subseteq [n] : \lvert S\rvert \leq k\}$	$\Theta\left(\frac{n}{\sqrt{k}}\right)$
Contiguous blocks	$\{S = [i,j] : i \leq j \in [n]\}$	$\widetilde{\Theta}(n)$
Prefixes	$\{[1,i] : i \in [n]\}$	$\Theta(n)$
Only full set	$\{[n]\}$	$\Theta(2^{n/2})$

In particular, setting $k = n$ in the first bound above tightens the $O(\sqrt{n}\log n)$ algorithm of Ambainis and Montanaro [AM14] for quantum search with wildcards, and improves it to the optimal $O(\sqrt{n})$. We note that the optimal $O(\sqrt{n})$ bound in this case has already been shown by Belovs [Bel15] using completely different techniques from ours. Additionally, our results in the prefix-query model prove that recovering any bitstring cannot be done with a sublinear number of prefix queries, which resolves an open question raised in [XZL23]. The result in the only full set case recovers the well-known tight $\Omega(\sqrt{N})$ lower bound for searching for a marked element in a database with N items [BBHT98] and the matching upper bound is given by [Gro96].

1.2 Our Techniques

We now discuss the techniques that we use to show Theorems 1.1 and 1.2. We then make a comparison with related work.

1.2.1 Sketch of Proof of Theorem 1.1

Our analysis begins from the characterization of quantum query complexity by the negative-weight adversary bound [Rei11]. This characterization considers the quantum query complexity of a function $F : D \to E$, with respect to a query set $\mathcal{R}$, where every query $q \in \mathcal{R}$ is a function mapping D to a set of query outcomes. Reichardt [Rei11] expresses the query complexity of F w.r.t. $\mathcal{R}$, up to a constant factor, in terms of the optimum value of a certain semidefinite optimization program known as the general adversary bound, which we refer to as the *primal adversary bound*, defined as

$$\mathrm{ADV}^{\mathcal{R},\pm}(F) := \max_{\Gamma \neq 0} \frac{\|\Gamma\|}{\max_{q \in \mathcal{R}} \|\Gamma_q\|}. \tag{2}$$

Here the maximization is over all matrices $\Gamma \in \mathbb{R}^{D \times D}$ with $F(x) = F(y) \implies \Gamma[x,y] = 0$, and the maximization in the denominator is over all available queries $\mathcal{R}$. For a query $q \in \mathcal{R}$, the matrix Γ_q is defined to be equal to Γ, with all those $[x,y]$ entries zeroed out where the outcome of the query q is the same on inputs x and y. This quantity was defined in [HLŠ07], where it was shown to be a lower bound on quantum query complexity. Reichardt [Rei11] subsequently showed that the query complexity of F with respect to query set $\mathcal{R}$ is equal to $\mathrm{ADV}^{\mathcal{R},\pm}(F)$ up to a constant factor.

It is well-known (see e.g., [Rei09, Theorem 1.2]) that the negative-weight adversary bound can be written as a semidefinite program (SDP). We show in Theorem 3.3 that the feasible solutions of the maximization version of this SDP are directly related to the choices of Γ in Equation (2). It then follows that symmetry reductions on the SDP-level can be translated into additional symmetry constraints on the matrix Γ in Equation (2), without changing the optimal objective value. Note that attempting to naively apply such a symmetrization to the formulation in Equation (2) does not

immediately work since, for example, the numerator could potentially reduce after symmetrizing. This is why all our symmetrization is done at the SDP level.

For the specific case where $F = \oplus$ and $\mathcal{R} = \{(S, b) : S \in \mathcal{Q}, b \in \{-1, 1\}^S\}$, with $\mathcal{Q} \subseteq 2^{[n]}$, we subsequently characterize in Theorems 3.7 and 3.8 the symmetries in the primal adversary bound SDP, from which we deduce in Theorem 3.9 that a maximizing Γ can be assumed to be the communication matrix of an XOR function. That is, we can assume that $\Gamma = M_{f \circ \mathsf{XOR}}$ for some $f : \{-1, 1\}^n \to \mathbb{R}$, where $M_{f \circ \mathsf{XOR}}(x, y) := f(x \oplus y)$, i.e., the (x, y)'th entry only depends on the bitwise XOR of x and y. It is well known that such matrices are simultaneously diagonalizable by the Hadamard matrix, and hence diagonal in the Fourier basis, which motivates analyzing the numerator and denominators of Equation (2) using Fourier analysis.

It is known that the numerator of Equation (2) in this case, which is the spectral norm of $M_{f \circ \mathsf{XOR}}$, equals the (appropriately scaled) maximum Fourier coefficient of f [BC02]. For the denominator, we analyze the sparsity structure of the matrix Γ_q, for any query $(S, b) = q \in \mathcal{R}$. Through a series of calculations involving Fourier analysis of restricted Boolean functions, we show that the matrix Γ_q's spectral norm equals the (appropriately scaled) maximum, over all fixings of bits in $\bar{S}$, standard deviation of the restriction of f to the subcube that fixes these bits in $\bar{S}$. Putting everything together, we obtain $\mathsf{val}_{\mathcal{Q}} = \mathrm{ADV}^{\mathcal{R},\pm}(\oplus)$, which together with Reichardt's characterization [Rei11] implies Theorem 1.1.

We provide a schematic illustration of the proof overview for Theorem 1.1 in Figure 1.

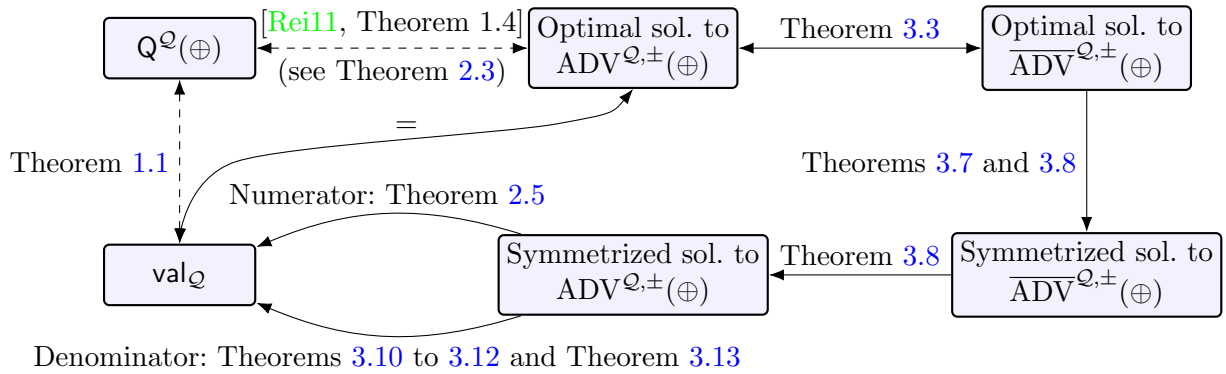

Figure 1: A schematic illustration to obtain our framework, which enables us to transform the problem of computing $\mathsf{Q}^{\mathcal{Q}}(\oplus)$, for an arbitrary collection of sets $\mathcal{Q} \subseteq 2^{[n]}$, into the task of analyzing an abstract analytic optimization program $\mathsf{val}_{\mathcal{Q}}$. The solid lines represent exact equalities, whereas the dashed lines represents characterizations up to constants.

1.2.2 Sketch of Proof of Theorem 1.2

Theorem 1.2 proves matching lower and upper bounds on $\mathsf{val}_{\mathcal{Q}}$ for four different classes of sets $\mathcal{Q}$. Note that $\mathsf{val}_{\mathcal{Q}}$ involves a maximization over all odd real-valued functions over the Boolean hypercube. The lower bounds are proven by exhibiting explicit functions for which every valid (for the class of $\mathcal{Q}$ considered) subcube restriction has low standard deviation with respect to the maximum absolute value of the function. The upper bounds are proven by showing that every odd real-valued function over the Boolean hypercube has a high standard deviation (with respect

to its maximum absolute value) when restricted to some subcube that is valid for the class of $\mathcal{Q}$ considered. We now briefly discuss the proofs we give for various classes of $\mathcal{Q}$.

Bounded size Theorem 4.2 deals with query sets of bounded size. The valid subcubes are the ones that fix at least $n-k$ variables. The lower bound is witnessed by the function that maps each bit-string to its Hamming weight minus $n/2$ (note that this is an odd function). For the upper bound, we first prove by an induction on the number of variables, that for every odd function there is some subcube (not necessarily valid) which simultaneously has a high variance and the absolute values of all its degree-1 Fourier coefficients are large. If this subcube is not valid (i.e., if it fixes less than $n-k$ variables), we argue via the probabilistic method and Fourier analysis that the function can be further restricted by appropriately setting additional variables to a valid subcube, while simultaneously ensuring that the variance does not drop too much.

Contiguous blocks Theorem 4.4 deals with query sets comprising consecutive variables (the same theorem statement holds true even if one is allowed wraparound). The valid subcubes are the ones that restrict consecutive variables. The upper bound of n follows from the observation that the query sets include all singletons, and Proposition 4.1 which verifies via the framework of Theorem 1.1 the obvious fact that if all the single bit queries are allowed, then the quantum query complexity is $O(n)$.

The lower bound is witnessed by a function f defined as follows. Given an input bit string of size n, view it as consisting of n/k disjoint blocks of size k, where $k = \log n - \log\log n + O(1)$. Let the blocks be denoted by $z^{(1)}, \ldots, z^{(n/k)}$. First, each $z^{(i)}$ is subjected to a function g that maps the string 1^k to 1, $(-1)^k$ to -1 and maps all other strings to 0. Finally f is defined to be $\sum_{i=1}^{n/k} g(z^{(i)})$. Note that each subcube setting contiguous variables (even if wraparound is allowed) touches, i.e., the endpoints of the block of queried indices are contained in, at most two of the blocks, so the restricted function is essentially a reduced version of f on a smaller number of blocks plus a bounded function on at most $2k$ bits. Thus, up to an additive shift by a bounded function, the (random variable corresponding to the) restriction of f is essentially the sum of independent random variables taking values in $\{0, 1, -1\}$ (the independence is due to the disjointness of the blocks), each of which is 0 except with probability $1/2^{k-1}$. Furthermore, the additional bounded summand is also independent of this sum and has $O(1)$ variance. We can thus apply linearity of variance for independent random variables to finish our argument.

Prefixes Theorem 4.5 deals with query sets which are prefixes, i.e., contain the first i variables for some positive integer i. The theorem shows that for this class $\mathcal{Q}$ of queries $\mathsf{val}_{\mathcal{Q}}$ is $\Theta(n)$. Valid subcubes in this setting fix a set of variables that is a suffix, i.e., contains the last i variables for some non-negative integer i.

The lower bound is witnessed by a function that is defined first by constructing a *decision list*: if $x_{n-1} = -1$ then output 1, else if $x_{n-2} = -1$ then output 2, ..., else if $x_1 = 1$ then output $n-1$, else output n. We then turn this decision list into an odd function by reflecting it about an additional appended variable and inverting the reflection (both the input variables and the function output). The maximum value of this function is n. At a high level, inside subcubes that only fix suffixes, the values that the function takes 'grow linearly' but their probabilities 'drop exponentially'. This allows us to show that every valid subcube has $O(1)$ variance.

For the upper bound we consider an arbitrary odd real function f, and consider the sequence of subcubes that are obtained by fixing the last i variables consistent with an input that maximizes f, for $i = 0, 1, \ldots, n$. Note that the average of f over the first subcube (which is $\{-1, 1\}$) is 0 (as the function is odd and hence balanced), and over the last subcube is $\|f\|_\infty$. Thus, we have that there is a point where the function average jumps by at least $\|f\|_\infty/n$. This implies that switching the value of the variable that is being restricted at that step leads to a large change in the function average. We could quantify this phenomenon in the form of a lower bound of $\|f\|_\infty/n$ on the absolute value of the degree-1 Fourier coefficient of the restricted function (note that potentially multiple variables have been fixed before we reach this point) corresponding to this variable. A lower bound on the variance, and hence an upper bound of $\mathsf{val}_\mathcal{Q}$, then follows easily by using the Fourier formula for variance.

Remark 1.4. *In Section B we provide a proof directly using the negative-weight adversary bound without resorting to our framework, of the stronger result that computing the last bit of an input string (which can only be an easier task than learning the whole string) requires $\Omega(n)$ prefix queries. We choose to still retain Theorem 4.5 and its proof to demonstrate the versatility of our framework.*

Only full set Theorem 4.7 deals with the case when $\mathcal{Q}$ consists of the full set of variables. In this case the denominator in the definition of $\mathsf{val}_\mathcal{Q}$ is just the standard deviation of the function over all inputs. For the lower bound, we construct a function f that takes only two non-zero values: value $+1$ on the all-1 input, value -1 on the all--1 input, and 0 everywhere else. From this, its standard deviation turns out to be extremely small, exactly $2^{-(n-1)/2}$, giving the lower bound.

For the upper bound, consider any odd function f. Because f is odd, it attains values $\|f\|_\infty$ and $-\|f\|_\infty$ on some pair of opposite inputs. These two points alone already contribute enough to the second moment to guarantee that the standard deviation in the whole cube is at least $\|f\|_\infty \cdot 2^{-(n-1)/2}$. Dividing $\|f\|_\infty$ by this quantity, as required in the definition of $\mathsf{val}_\mathcal{Q}$, shows that no odd function can achieve a value above about $2^{(n-1)/2}$. This gives the upper bound.

Together, the lower and upper bounds match, so the value for this setting of $\mathcal{Q}$ is exactly $2^{(n-1)/2}$.

1.2.3 Comparison With Related Work

The quantum $\mathcal{Q}$-substring query complexity of rec with $\mathcal{Q} = 2^{[n]}$ was introduced and studied by Ambainis and Montanaro [AM14], who dubbed this problem the "search with wildcards" problem and gave an $O(\sqrt{n}\log n)$ upper bound and a near matching $\Omega(\sqrt{n})$ lower bound. The algorithm by Ambainis and Montanaro uses the Pretty Good Measurement to iteratively learn $O(\sqrt{n})$ bits and fixes any errors made via binary search. In the search with wildcards setting, binary search can be performed with just $O(\log n)$ queries. Therefore, the algorithm learns the entire string in $O(\sqrt{n}\log n)$ queries. Setting $k = n$ in our Corollary 1.3 gives a tight $O(\sqrt{n})$ bound for this task. Ambainis and Montanaro also gave quantum query algorithms for the combinatorial group testing problem, which they observed to be a generalization of the $\mathcal{Q}$-substring query complexity of rec with $\mathcal{Q} = 2^{[n]}$. Subsequently, Belovs [Bel15] gave, among other results, tight bounds on the quantum query complexity of combinatorial group testing, which implies a tight $O(\sqrt{n})$ upper bound on the quantum $2^{[n]}$-substring query complexity of rec. In contrast with the algorithms of Ambainis and Montanaro, Belovs showed a $O(\sqrt{n})$ upper bound on the dual of the negative-weight adversary

bound. Since the negative-weight adversary bound is known to be asymptotically equal to quantum query complexity [Rei11], this immediately gives an upper bound on $\mathsf{Q}^{\mathcal{Q}}(\mathsf{rec})$.

In contrast to the work of Ambainis and Montanaro [AM14] who construct an explicit algorithm, and that of Belovs [Bel15] who constructs an explicit solution to the dual adversary bound, we deal, perhaps counter-intuitively, completely with the *primal* version of the negative-weight adversary bound [HLŠ07]. Our work does have the drawback that it does not give explicit algorithms for any of our upper bounds. However, as far as we are aware, ours is the first work to prove novel quantum query *upper* bounds using the *primal* version of the negative-weight adversary bound (which is primarily used as a lower bound technique) without explicitly resorting to SDP duality.

Another work of relevance to ours is [BBGK18], who showed classical lower bounds for composed functions using quantum upper bounds for combinatorial group testing [Bel15]. It is not immediately clear to us if our results have any new consequences for classical lower bounds using their framework, but our work can be viewed as partial progress towards answering their question [BBGK18, Section V] "Are there nontrivial quantum algorithms that use g-queries to learn x for some function $g \notin \{\mathsf{OR}, \mathsf{XOR}\}$? Are there nontrivial quantum algorithms that use g-queries to compute some other function $h(x)$ of the input?" Theorem 1.1 gives a complete framework that allows one to easily analyze the complexity of computing Parity (which is often equivalent to the complexity of learning the full input string) when g is restricted to subcube queries. Our framework gives provably tight bounds in this setting for all classes of allowed subcubes to query (with the constraint that if the algorithm is allowed to query if x is in a subcube whose fixed variables are S, all fixing of variables in S are allowed queries). The key advantage of our framework is that it replaces quantum and SDP-based reasoning with a simple, combinatorial analysis of real-valued functions on the Boolean hypercube. We justify this simpler viewpoint via Theorem 1.2, which establishes tight bounds within this framework for several natural classes of allowed subcube query sets.

1.2.4 Outlook

While our results do not apply in the combinatorial group testing setting, where only OR queries are allowed and the input string is promised to have bounded Hamming weight, we consider a generalization of the search with wildcards problem in a different direction, namely when a query algorithm is restricted on the support of the substring queries it can make. Our results are specifically of interest in the context of biological applications such as genome reconstruction, DNA sequencing, and RNA sequencing: In these models, a unit-cost task is often assumed to be observing a contiguous substring of the input string [SS95, RSC+10, CWE15, LD24]. Our second bound in Corollary 1.3 shows that quantum algorithms cannot provide an advantage over classical ones in this setting.

We view our work as an invitation to explore upper bounds through the *primal* adversary formulation itself, by studying the structural properties that optimal adversary matrices must satisfy. While this formulation can tend to be regarded as less convenient than its dual, our results illustrate that, with sufficient symmetrization and structural insight, the primal characterization can in fact yield tight bounds and reveal new patterns in how optimal Γ matrices arise. This suggests that it may be a more practical tool for upper bounds than previously thought.

2 Preliminaries

All logarithms in this paper are taken base 2. For a positive integer n and integers $i \leq j \in [n]$, we use the notation $[i,j] := \{k \in [n] : i \leq k \leq j\}$. For strings, $x, y \in \{-1,1\}^n$, we use the notation $x \oplus y$ to denote the bitwise XOR of x and y, or equivalently, the string obtained by taking component-wise multiplication of x and y. For a string $x \in \{-1,1\}^n$, we use the notation $|x|$ to denote the Hamming weight of x, i.e., the number of -1's in x. For a string $x \in \{-1,1\}^n$ and $S \subseteq [n]$, we use x_S to denote the string x restricted to the bits in S. For an expression X, we define $\mathbb{I}[X]$ to equal 1 if X is true, and 0 if X is false. For a string $x = (x_1, x_2, \ldots, x_n) \in \{-1,1\}^n$, we use the notation $-x$ to denote the string $(-x_1, -x_2, \ldots, -x_n)$. We denote vectors with the ket-notation, $|v\rangle$, and refer to their conjugate transpose with the bra-notation, i.e., $\langle v|$.

A function $f : \{-1,1\}^n \to \mathbb{R}$ is said to be an *odd function* if $f(x) = -f(-x)$ for all $x \in \{-1,1\}^n$. We abuse notation and use f to interchangeably mean a function as well as a random variable corresponding to picking a value of f uniformly at random. That is, for $f : \{-1,1\}^n \to \mathbb{R}$, we use $\mathbb{E}[f] := \mathbb{E}_{x \in \{-1,1\}^n}[f(x)]$, $\mathbb{E}[f^2] := \mathbb{E}_{x \in \{-1,1\}^n}[f(x)^2]$, and $\mathbf{Var}(f) := \mathbb{E}[f^2] - \mathbb{E}[f]^2$. We use $\|f\|_\infty$ to denote $\max_{x \in \{-1,1\}^n} |f(x)|$. For a positive integer n and a 2^n-dimensional real matrix M whose rows and columns are indexed by n-bit strings, we say that a M is the *communication matrix of an* XOR *function* if there exists a function $f : \{-1,1\}^n \to \mathbb{R}$ such that $M(x,y) = f(x \oplus y)$. We use the notation $M_{f \circ \mathsf{XOR}}$ to denote the communication matrix of $f \circ \mathsf{XOR}$.

A collection of sets $\mathcal{Q} \subseteq 2^{[n]}$ is said to be *downward closed* if for all $Q \in \mathcal{Q}$, all subsets of Q are also in $\mathcal{Q}$.

For an n-dimensional real square matrix M, its spectral norm, denoted $\|M\|$, is defined by

$$\|M\| := \sup_{x \in \mathbb{R}^n : \|x\|_2 = 1} \|Mx\|_2 ,$$

where $\|\cdot\|_2$ denotes the usual Euclidean 2-norm. Equivalently it is well known that the following characterization also holds true.

Fact 2.1. *For an n-dimensional real square matrix M,*

$$\|M\| = \sqrt{\lambda_{\max}(M^T M)},$$

where $\lambda_{\max}(\cdot)$ equals the maximum eigenvalue.

2.1 Quantum Query Complexity

Let $\mathcal{H}$ be a finite-dimensional Hilbert space. Any quantum algorithm $\mathcal{A}$ can be represented as a sequence of unitary operators $\{U_i\}_{i=0}^T \in \mathcal{L}(\mathcal{H})$ on an initial state $|\psi_0\rangle \in \mathcal{H}$. The algorithm is terminated by a measurement on $|\phi\rangle = U_T U_{T-1} \cdots U_1 U_0 |\psi_0\rangle$ to observe an outcome.

Let D and E be finite sets and $F : D \to E$ be a function. For every $x \in D$, let O_x be a unitary operator that depends on x, referred to as the *oracle*. A quantum query algorithm $\mathcal{A}$ alternates input-independent unitary operators with calls to the oracle, i.e., it applies the operations U_0, O_x, U_1, O_x, etc. The oracle calls are referred to as *queries* to the input x. The algorithm $\mathcal{A}$ is a bounded-error, quantum query algorithm for computing F if $\forall x \in D$ the output of $\mathcal{A}$ is $F(x)$ with probability at least $2/3$. The quantum query complexity of computing F is the minimum number of queries required to compute F with probability $2/3$. This is denoted by $\mathsf{Q}(F)$.

In this work, we consider algorithms with access to different *query sets* and therefore use $\mathsf{Q}^{\mathcal{R}}(F)$ to refer to the quantum query complexity of computing F using query set $\mathcal{R}$. Formally a query set $\mathcal{R}$ is a set consisting of functions $r : \{-1,1\}^n \to \mathcal{O}$ for some discrete range $\mathcal{O}$. The usual query model is captured by the query set $\mathcal{R} = \{r_i : \{-1,1\}^n \to \{-1,1\} \mid i \in [n]\}$ where $r_i(x) := x_i$. The final state of a quantum query algorithm with query set $\mathcal{R}$ looks like

$$|\phi\rangle = U_T O_x \dots U_2 O_x U_1 O_x U_0 |\psi_0\rangle$$

where the action of O_x on each basis state is

$$O_x : |r\rangle |b\rangle \to |r\rangle |b + r(x) \mod |\mathcal{O}|\rangle \qquad \forall r \in \mathcal{R}, b \in \{0, 1, \dots, |\mathcal{O}|\}.$$

Let $x \in \{-1,1\}^n$ be an input string and $\mathcal{Q} \subseteq 2^{[n]}$. We abuse notation and use $\mathcal{Q}$ to be the query set consisting of all queries $q = (S, b_S)$ of the form $q(x) = \mathbb{I}[x_S = b_S]$ where $S \in \mathcal{Q}$, $b_S \in \{-1,1\}^S$ and x_S is the substring of x restricted to indices in S. Throughout this work, we only work with queries of this form. Thus, for $\mathcal{Q} \subseteq 2^{[n]}$, we abuse notation and use $\mathcal{Q}$ to also denote the query set $\left\{q = (S,b) : S \in \mathcal{Q}, b \in \{-1,1\}^S\right\}$.

Definition 2.2 (Positive semi-definiteness)**.** *A matrix $\Gamma \in \mathbb{R}^{N \times N}$ is said to be* positive semi-definite*, denoted by $\Gamma \succeq 0$, if*

$$\langle v| \Gamma |v\rangle \geq 0 \qquad \forall |v\rangle \in \mathbb{R}^N.$$

Standard formulation of primal adversary bound The primal adversary bound is a semi-definite program that characterizes the quantum query complexity of computing function $F : D \to E$, relative to a query set $\mathcal{R}$. To that end, $\forall q \in \mathcal{R}, \Delta_q \in \{0,1\}^{D \times D}$ is the matrix defined as:

$$\Delta_q[x,y] = \begin{cases} 1, & q(x) \neq q(y), \\ 0, & \text{otherwise.} \end{cases}$$

That is, $\Delta_q[x,y] = 1$ if and only if the query q can distinguish between inputs x and y. Throughout this work, we use the notation $\mathrm{ADV}^{\mathcal{R},\pm}$ to denote the adversary bound relative to query set $\mathcal{R}$. Now we state the primal version of the general adversary bound, which we will refer to as the *primal adversary bound*.

Primal adversary bound

$$\begin{aligned} \mathrm{ADV}^{\mathcal{R},\pm}(F) = \max_{\Gamma} \quad & \frac{\|\Gamma\|}{\max\limits_{q \in \mathcal{R}} \|\Gamma \circ \Delta_q\|} && (3) \\ \text{subject to} \quad & \Gamma \in \mathbb{R}^{D \times D}, \\ & \Gamma \text{ is symmetric}, \\ & \Gamma[x,y] = 0. && (\forall\, x,y \in D \text{ with } F(x) = F(y)) \end{aligned}$$

Since the objective function is invariant under multiplying Γ with a constant factor, we can without loss of generality assume that the denominator is equal to 1. Thus, we can equivalently write the SDP computing $\mathrm{ADV}^{\mathcal{R},\pm}$ with objective function $\max_\Gamma \|\Gamma\|$, and the additional constraint that $\forall q \in \mathcal{R}, \|\Gamma \circ \Delta_q\| \leq 1$.

Reichardt [Rei11] showed that the primal adversary bound is a tight characterization of quantum query complexity.

Theorem 2.3 ([Rei11, Theorem 1.4]). *Let D and E be finite sets and $F : D \to E$ be a function. The bounded-error, quantum query complexity of computing F using query set $\mathcal{R}$ is characterized by the adversary bound.*

$$\mathsf{Q}^{\mathcal{R}}(F) = \Theta(\mathrm{ADV}^{\mathcal{R},\pm}(F)).$$

Although [Rei11, Theorem 1.4] states the adversary bound for index queries, it can be observed that the adversary bound characterizes the quantum query complexity for any query set and therefore we present it as such. Let $\mathcal{R}$ be a query set (assume for simplicity that all queries output values in $\{-1,1\}$. The argument below can easily be generalized if this is not the case). For every $x \in D$, define the string $\mathcal{R}(x) = (r(x))_{r\in\mathcal{R}}$. Define the partial function $G : \{-1,1\}^{\mathcal{R}} \to E$ on the input domain $\left\{z \in \{-1,1\}^{\mathcal{R}} : z = \mathcal{R}(x) \text{ for some } x \in D\right\}$ as $G(\mathcal{R}(x)) := F(x)$. Observe that querying a input variable, say $R \in \mathcal{R}$, to G on input $\mathcal{R}(x)$ is exactly the same as making the query R on input x. Thus, $\mathsf{Q}^{\mathcal{R}}(F) = \mathsf{Q}(G)$ and $\mathrm{ADV}^{\mathcal{R},\pm}(F) = \mathrm{ADV}^{\pm}(G)$, and therefore the result follows by applying [Rei11, Theorem 1.4] to G.

When $\mathcal{Q} \subseteq 2^{[n]}$, and the query set consists of all queries of the form $\mathbb{I}[x_S = b]$ where $S \in \mathcal{Q}$ and $b \in \{-1,1\}^S$ are arbitrary, recall that we abuse notation and use $\mathcal{Q}$ to also denote the query set $\left\{q = (S,b) : S \in \mathcal{Q}, b \in \{-1,1\}^S\right\}$. Thus $\mathrm{ADV}^{\mathcal{Q},\pm}(\cdot)$ and $\mathsf{Q}^{\mathcal{Q}}(\cdot)$ denote the negative-weight adversary bound and the quantum query complexity, respectively, with respect to this query set.

2.2 Fourier Analysis of Boolean Functions and Restrictions

Below are some relevant formal definitions and Fourier-analytic properties of restricted functions, adapted from [O'D14].

Consider the vector space of functions from $\{-1,1\}^n$ to $\mathbb{R}$, equipped with the following inner product.

$$\langle f, g\rangle := \mathbb{E}_{x\in\{-1,1\}^n}[f(x)g(x)] = \frac{1}{2^n}\sum_{x\in\{-1,1\}^n} f(x)g(x).$$

For a set $S \subseteq [n]$, define the parity function $\chi_S : \{-1,1\}^n \to \mathbb{R}$ by $\chi_S(x) = \prod_{i\in S} x_i$. It is not hard to show that the set of all parities $\{\chi_S : S \subseteq [n]\}$ forms an orthonormal basis of the space above. Thus, every function $f : \{-1,1\}^n \to \mathbb{R}$ has a unique expansion with respect to this basis, $f = \sum_{S\subseteq[n]} \widehat{f}(S)\chi_S$. This expansion is called the Fourier expansion of f and the coefficients $\left\{\widehat{f}(S) : S \subseteq [n]\right\}$ are called the Fourier coefficients of f.

Fact 2.4 (Parseval's Theorem). *Let $f : \{-1,1\}^n \to \mathbb{R}$ be a function. Then*

$$\mathbb{E}_{x\in\{-1,1\}^n} f(x)^2 = \langle f, f\rangle = \sum_{S\subseteq[n]} \widehat{f}(S)^2.$$

Lemma 2.5 ([BC02], also see [Man18, Lemma 2.2.3]). *Let $f : \{-1,1\}^n \to \mathbb{R}$ be a function and let M be the communication matrix of $f \circ \mathsf{XOR}$. Then,*

$$\|M\| = 2^n \max_{S\subseteq[n]} |\widehat{f}(S)|.$$

Specifically, we require the following fact. Below, H represents the 2^n-dimensional Hadamard matrix defined by $H(S,T) = \frac{1}{2^{n/2}}(-1)^{|S\cap T|}$.

Fact 2.6. *Let $f : \{-1,1\}^n \to \mathbb{R}$ be a function. Then,*

$$M_{f\circ \mathsf{XOR}} = H D_{2^n \widehat{f}} H^{-1},$$

where $D_{2^n \widehat{f}}$ is a diagonal matrix with $D_{2^n \widehat{f}}(S) = 2^n \widehat{f}(S)$ for all $S \subseteq [n]$.

Definition 2.7. *Let $f : \{-1,1\}^n \to \mathbb{R}$ be a function, and let $(J, \bar{J})$ be a partition of $[n]$. Let $z \in \{-1,1\}^{\bar{J}}$. Then we write $f_{J|z} : \{-1,1\}^J \to \mathbb{R}$ (pronounced "the restriction of f to J using z") for the subfunction of f obtained by fixing the coordinates in $\bar{J}$ to the bit values z, i.e.,*

$$\forall x \in \{-1,1\}^J, f_{J|z}(x) := f(x_J z_{\bar{J}})$$

where $x_J z_{\bar{J}}$ represents the string that equals x on J and z on $\bar{J}$.

Proposition 2.8 ([O'D14, Proposition 3.21])**.** *Let $f : \{-1,1\}^n \to \mathbb{R}$ and let $(J, \bar{J})$ be a partition of $[n]$. For $z \in \{-1,1\}^{\bar{J}}$ and $S \subseteq J$ we have*

$$\widehat{f_{J|z}}(S) = \sum_{T \subseteq \bar{J}} \widehat{f}(S \cup T)\chi_T(z).$$

Setting $z = 1^{\bar{J}}$ in the proposition above, we obtain the following corollary.

Corollary 2.9. *Let $f : \{-1,1\}^n \to \mathbb{R}$ and let $(J, \bar{J})$ be a partition of $[n]$. For $S \subseteq J$ we have*

$$\widehat{f_{J|1^{\bar{J}}}}(S) = \sum_{T \subseteq \bar{J}} \widehat{f}(S \cup T).$$

We also obtain the following general corollary.

Corollary 2.10 ([O'D14, Corollary 3.22])**.** *Let $(J, \bar{J})$ be a partition of $[n]$ and fix $S \subseteq J$. Suppose $z \sim \{-1,1\}^{\bar{J}}$ is chosen uniformly at random. Then,*

$$\mathbb{E}_z[\widehat{f_{J|z}}(S)^2] = \sum_{T \subseteq \bar{J}} \widehat{f}(S \cup T)^2.$$

Lemma 2.11. *For $f : \{-1,1\}^n \to \mathbb{R}$, the following conditions are equivalent:*

- *For all x with $|x|$ even, $f(x) = 0$.*
- *For all $S \subseteq [n]$, we have $\widehat{f}(S) = -\widehat{f}(\bar{S})$.*

Proof. For all $x \in \{-1,1\}^n$,

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S)\chi_S(x) = \frac{1}{2} \sum_{S \subseteq [n]} \left[\widehat{f}(S)\chi_S(x) + \widehat{f}(\bar{S})\chi_{\bar{S}}(x) \right].$$

Assume that $\forall S \subseteq [n]$, $\widehat{f}(S) = -\widehat{f}(\bar{S})$. Let $x \in \{-1,1\}^n$ such that $|x|$ is even. Now we note that if $|x_S|$ is even, then $|x_{\bar{S}}|$ is also even. Similarly, when $|x_S|$ is odd, $|x_{\bar{S}}|$ is also odd. Thus, for such x we have

$$f(x) = \frac{1}{2} \sum_{S \subseteq [n]} \left[\widehat{f}(S) + \widehat{f}(\bar{S}) \right] = \frac{1}{2} \sum_{S \subseteq [n]} \left[\widehat{f}(S) - \widehat{f}(S) \right] = 0.$$

To prove the other direction, we begin with the assumption that $f(x) = 0$ for all $x \in \{-1,1\}^n$ such that $|x|$ is even. By orthogonality of the characters, we have

$$\widehat{f}(S) = \mathbb{E}[f(x)\chi_S(x)] = -\mathbb{E}[f(x)\chi_S(x) \cdot -1] = -\mathbb{E}[f(x)\chi_S(x) \cdot \chi_{[n]}(x)] = -\mathbb{E}[f(x)\chi_{\bar{S}}(x)] = -\widehat{f}(\bar{S}),$$

where the third equality holds since for all x in the support of the expectation, we have $|x|$ odd and therefore $\chi_{[n]}(x) = -1$, and the fourth equality uses $\chi_A(x) \cdot \chi_B(x) = \chi_{A \triangle B}(x)$ where $A \triangle B$ denotes the symmetric difference of A and B. □

We require the following well-known expression for the variance of a function. See, for example, [O'D14, Proposition 1.13] for a simple proof.

Lemma 2.12. *Let $f : \{-1,1\}^n \to \mathbb{R}$ be a function. Then,*

$$\mathbf{Var}(f) = \sum_{S \neq \emptyset} \widehat{f}(S)^2.$$

Finally, we recall this fact.

Fact 2.13. *Let $v \in \mathbb{R}^d$ be a vector, and*

$$A = \begin{bmatrix} 0 & v^T \\ v & 0 \end{bmatrix} \in \mathbb{R}^{(d+1)\times(d+1)}.$$

Then, $\|A\| = \|v\|_2$, the Euclidean norm of v.

3 The Framework

In this section we build our framework that allows us to analyze the $\mathcal{Q}$-substring query complexity of learning input strings, for arbitrary $\mathcal{Q}$.

3.1 Learning vs. Parity

We first make the simple observation that for every collection of sets $\mathcal{Q} \subseteq 2^{[n]}$, the quantum query complexity of learning x is bounded from below by the quantum query complexity of computing the parity of all of the bits of x.

Observation 3.1. *Let $\mathcal{Q} \subseteq 2^{[n]}$. Then $\mathsf{Q}^{\mathcal{Q}}(\mathsf{rec}) \geq \mathsf{Q}^{\mathcal{Q}}(\oplus)$.*

This immediately follows from the fact that once a string has been learned, the parity of its input bits can be computed with no extra queries. We remark that the same argument in fact yields $\mathsf{Q}^{\mathcal{Q}}(\mathsf{rec}) \geq \max_{S \subseteq [n]} \mathsf{Q}^{\mathcal{Q}}(\oplus_S)$, but we do not use this.

Next we observe that in the case when $\mathcal{Q}$ is downward closed, these complexities are actually equal.

Lemma 3.2. *For all $\mathcal{Q} \subseteq 2^{[n]}$ that is downward closed, we have $\mathsf{Q}^{\mathcal{Q}}(\mathsf{rec}) = \mathsf{Q}^{\mathcal{Q}}(\oplus)$.*

Proof. Theorem 3.1 provides the lower bound on $\mathsf{Q}^{\mathcal{Q}}(\mathsf{rec})$, so it remains to prove the upper bound. Note that we can recover a bit string if we are able to compute the parity on each of the subsets in superposition, by the Bernstein-Vazirani algorithm. Thus, it remains to prove that $\mathsf{Q}^{\mathcal{Q}}(\oplus_S) \leq \mathsf{Q}^{\mathcal{Q}}(\oplus)$, for all $S \subseteq [n]$.

To that end, let $S \subseteq [n]$. For any $x \in \{-1,1\}^n$, let $y(x) \in \{-1,1\}^n$ be the input that is equal to x on S, and identically 1 on $\bar{S}$. Then, $\oplus_S(x) = \oplus(y(x))$, so it remains to run the algorithm on $y(x)$. To that end, note that we can simulate a substring query (T, b) to $y(x)$ as follows. First, we check if b is identically 1 on $T \setminus S$, else we simply output 1 (False). Otherwise, we perform the query $(S \cap T, b_S)$ on x, and we observe that $y(x)_T = b$ if and only if $y(x)_{S\cap T} = b_S$. Since $\mathcal{Q}$ is downward-closed and $S \cap T \subseteq T$, we indeed have access to this query, and so we can simulate the parity-algorithm on $y(x)$ without additional overhead. □

We conclude that in order to characterize the $\mathcal{Q}$-complexity of bit string recovery, it suffices to characterize the $\mathcal{Q}$-complexity of computing the parity function, as long as $\mathcal{Q}$ is downward closed.

3.2 Simplification by Symmetry Reduction

In this subsection we perform a symmetry reduction on the primal adversary bound for the quantum $\mathcal{Q}$-substring quantum query complexity of computing parity, for an arbitrary $\mathcal{Q} \subseteq 2^{[n]}$. We show that the maximizing matrix Γ can be assumed to be the communication matrix of a XOR function, i.e., we may assume that $\Gamma = M_{f\circ\mathsf{XOR}}$ for some $f : \{-1,1\}^n \to \mathbb{R}$. Recall that $M_{f\circ\mathsf{XOR}}[x,y] = f(x \oplus y)$ for all $x, y \in \{-1,1\}^n$.

Alternate formulation of adversary bound for functions with Boolean output Here we use an alternative formulation of the adversary bound that was shown by Cornelissen [Cor23, Theorem 6.2.4]. This formulation allows us to average over optimal solutions without any loss in the optimal value. An important detail to note is that this formulation holds only for functions with Boolean output.[2]

Alternate formulation of primal adversary bound for functions with Boolean output

$$
\begin{aligned}
\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F) = \quad & \max_{\Gamma,\beta} && \sum_{x,y\in D} \Gamma[x,y] && (4)\\
& \text{subject to} && \Gamma \in \mathbb{R}^{D\times D}, \beta \in \mathbb{R}^D_{\geq 0}, \\
& && \Gamma \text{ is symmetric}, \\
& && \Gamma[x,y] = 0, && (\forall\, x,y \in D \text{ with } F(x) = F(y)) \\
& && \mathsf{diag}(\beta) - \Gamma \circ \Delta_q \succeq 0, && (\forall q \in \mathcal{R}) \\
& && \sum_{x\in f^{-1}(1)} \beta[x] = \sum_{x\in f^{-1}(-1)} \beta[x] = \frac{1}{2}.
\end{aligned}
$$

[2] A similar optimization program can be derived for functions with non-Boolean output. Since its presentation is slightly more cumbersome and we don't need it in this work, we merely present the Boolean version here.

Lemma 3.3. *Let $D \subseteq \{-1,1\}^n$ and $F : D \to \{-1,1\}$ be a Boolean function. Then* $\mathrm{ADV}^{\mathcal{R},\pm}(F) = \overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F)$. *Moreover, if (β, Γ) is an optimal solution for* $\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F)$, *then Γ', defined as*[3]

$$\Gamma'[x,y] = \frac{\Gamma[x,y]}{\sqrt{\beta[x]\beta[y]}},$$

is an optimal solution for $\mathrm{ADV}^{\mathcal{R},\pm}(F)$.

The proof of this lemma essentially appears in the proof of [Cor23, Theorem 6.2.4], but not with the exact formulation above. We provide a proof of Theorem 3.3 in Section A for completeness. The benefit of this alternate formulation of the primal adversary bound is that it is a convex optimization problem. Therefore, it is always possible to take a linear combination of two feasible solutions, say with objective values c_1 and c_2, and obtain a new feasible solution with an objective value that is at least as big as $\min\{c_1, c_2\}$. We will use this to perform a symmetry reduction on the level of the SDP. In particular, we show how symmetries in the function and in the query set translate to additional constraints that we can impose on the optimal solutions to the alternate formulation of the primal adversary bound.

Let $\sigma : D \to D$ be a permutation of the domain D. We abuse notation to use σ to refer to the permutation function as well as the corresponding permutation matrix. It will be clear based on context whether σ is treated as a function or a matrix. When treated as a matrix, right-multiplying by σ^T corresponds to permuting columns according to the permutation σ and left-multiplying by σ corresponds to permuting rows according to σ.

Next, we define two types of symmetry that our query problem could exhibit.

Definition 3.4 (Function symmetry)**.** *A permutation $\sigma : D \to D$ is said to be* function symmetric with respect to a function $F : D \to E$ *if* $F(\sigma(x)) = F(\sigma(y)) \iff F(x) = F(y)$, $\forall x, y \in D$.

Definition 3.5 (Query symmetry)**.** *A permutation $\sigma : D \to D$ is said to be* query symmetric with respect to the set of queries $\mathcal{R}$ *if* $\{\sigma \Delta_q \sigma^T \mid q \in \mathcal{R}\} = \{\Delta_q \mid q \in \mathcal{R}\}$.

We now define a group that we use for the rest of this section and in the appendices.

Definition 3.6. *Let G denote the group generated by bit-flip permutations, i.e., $G = \{\sigma_y \mid y \in \{-1,1\}^n\}$ where $\sigma_y : x \mapsto x \oplus y$.*

As a first step, we prove that G exhibits function symmetry with the parity function, and query symmetry with any substring-query set $\mathcal{Q}$.

Lemma 3.7. *Let $n \in \mathbb{N}$ and $\mathcal{Q} \subseteq 2^{[n]}$. Then, every $\sigma \in G$ is function-symmetric w.r.t. $\oplus$ and query-symmetric w.r.t. $\mathcal{Q}$.*

Proof. Recall that $\forall x \in \{-1,1\}^n$, $|x \oplus y|$ is even $\iff \oplus(x) = \oplus(y)$. We use this fact to show that every $\sigma_z \in G$ is function symmetric with respect to the parity function.

$$|\sigma_z(x) \oplus \sigma_z(y)| = |(x \oplus z) \oplus (y \oplus z)| = |x \oplus y|.$$

Therefore,

$$\oplus(\sigma_z(x)) = \oplus(\sigma_z(y)) \iff |\sigma_z(x) \oplus \sigma_z(y)| \text{ is even} \iff |x \oplus y| \text{ is even} \iff \oplus(x) = \oplus(y).$$

[3]Here, we use the convention that $0/0 = 0$.

Let $S \in \mathcal{Q}$, $b \in \{-1,1\}^S$, σ_z be a permutation in G and let $\sigma_{z|S} : x_S \mapsto x_S \oplus z_S$ denote the permutation restricted to indices in S. Then,

$$\mathbb{I}[x_S = b] = \mathbb{I}[(\sigma_z(x))_S = \sigma_{z|S}(b)].$$

By the definition of our query set, the latter is a valid query on input $\sigma_z(x)$. Since this is true for all $S \in \mathcal{Q}$, all $b \in \{-1,1\}^S$ and all $\sigma_z \in G$, this implies

$$\{\Delta_{S,b} \mid S \in \mathcal{Q}, b \in \{-1,1\}^S\} = \{\sigma_z \Delta_{S,b} \sigma_z^T \mid S \in \mathcal{Q}, b \in \{-1,1\}^S\}.$$

This shows that every in $\sigma_z \in G$ is query symmetric with respect to $\mathcal{Q}$. □

Lemma 3.8. *Let n be a positive integer and let $\mathcal{R}$ be a query set. Let G be the group generated by bit flip permutations. If every permutation $\sigma \in G$ is query symmetric with respect to $\mathcal{R}$, then there exists an $f : \{-1,1\}^n \to \mathbb{R}$ such that*

- *$f(x) = 0$ for all x of even Hamming weight, and*
- *$M_{f \circ \mathsf{XOR}}$ is an optimal solution for the adversary bound $\mathrm{ADV}^{\mathcal{R},\pm}(\oplus)$.*

Proof. Let $D = \{-1,1\}^n$. Let (β, Γ) be an optimal solution to $\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(\oplus)$. In particular,

$$\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(\oplus) = \sum_{x,y \in \{-1,1\}^n} \Gamma[x,y].$$

For the next part of the proof, we show that $(\sigma\beta, \sigma\Gamma\sigma^T)$ is also a valid and optimal solution for all $\sigma \in G$. We organize this part of the proof by bullets, one for each constraint in the alternate formulation of the primal adversary bound.

- Entrywise non-negativity of $\sigma\beta$ is clear since this holds true for β, and hence any permutation of β as well. Note that $\sigma\Gamma\sigma^T$ is simply reordering the rows and columns of Γ by the same permutation σ. We observe below that $\sigma\Gamma\sigma^T$ is symmetric for any $\sigma \in G$.

$$(\sigma\Gamma\sigma^T)^T = (\sigma^T)^T \Gamma^T \sigma^T = \sigma\Gamma^T\sigma^T = \sigma\Gamma\sigma^T. \qquad (\Gamma = \Gamma^T)$$

- Since every $\sigma \in G$ in function symmetric with respect to $\oplus$ (see Theorem 3.7), we have $\oplus(x) = \oplus(y) \iff \oplus(\sigma(x)) = \oplus(\sigma(y))$. Therefore, $\sigma\Gamma\sigma^T[x,y] = \Gamma[\sigma(x), \sigma(y)] = 0$.

- Now we show that $\forall \sigma \in G, \quad \sigma\mathsf{diag}(\beta)\sigma^T - \sigma\Gamma\sigma^T \circ \Delta_q \succeq 0$. It is easy to see that every $\sigma \in G$ is its own inverse, and hence $\sigma^{-1} \in G$. Since every $\sigma \in G$ is query symmetric with respect to $\mathcal{R}$, we have $\sigma^{-1}\Delta_q(\sigma^{-1})^T = \Delta_{q'}$ for some $q' \in \mathcal{R}$ and hence $\Delta_q = \sigma\Delta_{q'}\sigma^T$. Let $|v\rangle \in \mathbb{R}^D$.

$$\begin{aligned}
\langle v| \left(\sigma\mathsf{diag}(\beta)\sigma^T - \sigma\Gamma\sigma^T \circ \Delta_q\right) |v\rangle &= \langle v| \left(\sigma\mathsf{diag}(\beta)\sigma^T - \sigma\Gamma\sigma^T \circ \sigma\Delta_{q'}\sigma^T\right) |v\rangle, \\
&= \langle v| \sigma(\mathsf{diag}(\beta) - \Gamma \circ \Delta_{q'})\sigma^T |v\rangle, \\
&= \left\langle v' \right| (\mathsf{diag}(\beta) - \Gamma \circ \Delta_{q'}) \left|v'\right\rangle, \qquad (|v'\rangle = \sigma^T |v\rangle \in \mathbb{R}^D) \\
&\geq 0,
\end{aligned}$$

  where the last line uses $\beta - \Gamma \circ \Delta_q \succeq 0, \forall q \in \mathcal{R}$, which holds true since (β, Γ) is an optimal (and hence valid) solution to $\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F)$.

- There are two cases to consider. From the function symmetry, we find that $\{\oplus^{-1}(1), \oplus^{-1}(-1)\} = \{\sigma(\oplus^{-1}(1)), \sigma(\oplus^{-1}(-1))\}$. That is, either σ maps all 1-inputs to 1-inputs and -1-inputs to -1-inputs, or vice versa. In either case, we obtain that

$$\sum_{x \in \oplus^{-1}(1)} (\sigma\beta)[x] = \sum_{x \in \oplus^{-1}(1)} \beta[\sigma^{-1}(x)] = \sum_{x \in \sigma(\oplus^{-1}(1))} \beta[x] = \frac{1}{2},$$

  and similarly when we take the summation over $x \in \oplus^{-1}(-1)$.

Thus, $(\sigma\beta, \sigma\Gamma\sigma^T)$ is a valid solution for all $\sigma \in G$. We now show its optimality. Since the objective value in the alternate formulation is defined as the sum of entries of Γ, both (β, Γ) and $(\sigma\beta, \sigma\Gamma\sigma^T)$ have the same objective value, formally argued below. Let $|\bar{1}\rangle$ denote the all-1 vector of length 2^n. Then

$$\begin{aligned}\sum_{x,y \in \{-1,1\}^n} \Gamma[x,y] &= \langle\bar{1}|\,\Gamma\,|\bar{1}\rangle = \langle\bar{1}|\,\sigma^T\sigma\Gamma\sigma^T\sigma\,|\bar{1}\rangle\,, \\ &= \langle\bar{1}|\,\left(\sigma\Gamma\sigma^T\right)|\bar{1}\rangle = \sum_{x,y \in \{-1,1\}^n} (\sigma\Gamma\sigma^T)[x,y].\end{aligned}$$

Next, we let (β', Γ') be obtained by averaging over all permutations in G as shown below.

$$(\beta', \Gamma') = \frac{1}{|G|} \sum_{\sigma \in G} (\sigma\beta, \sigma\Gamma\sigma^T). \tag{5}$$

Since $\overline{\text{ADV}}^{\mathcal{R},\pm}(\oplus)$ is a convex optimization program and (β', Γ') is a linear combination of feasible solutions to $\overline{\text{ADV}}^{\mathcal{R},\pm}(\oplus)$, we obtain that (β', Γ') is also a feasible solution. Moreover, since for all $\sigma \in G$, all the objective values of the feasible solutions $(\sigma\beta, \sigma\Gamma\sigma^T)$ are equal as argued above, the objective value of (β', Γ') is the same as that of (β, Γ). We conclude that (β', Γ') is also an optimal solution.

Finally, we observe from Theorem 3.3 that an optimal solution to $\text{ADV}^{\mathcal{R},\pm}(\oplus)$ can be written as

$$M[x,y] = \frac{\Gamma'[x,y]}{\sqrt{\beta'[x]\beta'[y]}}.$$

Since β' and Γ' are invariant under σ by construction, we observe for all $\sigma \in G$ that

$$M[\sigma(x), \sigma(y)] = \frac{\Gamma'[\sigma(x), \sigma(y)]}{\sqrt{\beta'[\sigma(x)]\beta'[\sigma(y)]}} = \frac{\Gamma'[x,y]}{\sqrt{\beta[x]\beta[y]}},$$

and so in particular,

$$M[x,y] = M[\sigma_y(x), \sigma_y(y)] = M[\sigma_y(x), 1^n] = M[x \oplus y, 1^n].$$

Thus, if we define $f : \{-1,1\}^n \to \mathbb{R}$ as $f(x) = M[x, 1^n]$, then we find that $M = M_{f \circ \mathsf{XOR}}$ is an optimal solution to $\text{ADV}^{\mathcal{R},\pm}(\oplus)$. Finally, for any $x \in \{-1,1\}^n$, if $|x|$ is even, then $\oplus(x) = \oplus(1^n)$, and so $f(x) = M[x, 1^n] = 0$. □

Corollary 3.9. *Let $n \in \mathbb{N}$ and $\mathcal{Q} \subseteq 2^{[n]}$. Then*

$$\mathrm{ADV}^{\mathcal{Q},\pm}(\oplus) = \max_{\substack{f:\{-1,1\}^n \to \mathbb{R} \\ |z| \text{ even} \implies f(z)=0}} \frac{\|M_{f\circ\mathsf{XOR}}\|}{\max\limits_{\substack{S\in\mathcal{Q}, \\ b\in\{-1,1\}^S}} \|M_{f\circ\mathsf{XOR}} \circ \Delta_{S,b}\|}.$$

Proof. We first show that $M_{f\circ\mathsf{XOR}}$ is a feasible solution to $\mathrm{ADV}^{\mathcal{Q},\pm}(\oplus)$, for all $f : \{-1,1\}^n \to \mathbb{R}$ satisfying $f(x) = 0$ when $|x|$ is even. The matrix $M_{f\circ\mathsf{XOR}}$ is symmetric since $M[x,y] = f(x \oplus y) = f(y \oplus x) = M[y,x]$ for all $x, y \in \{-1,1\}^n$. We know that $|x \oplus y|$ is even if both $|x|, |y|$ are even or if both $|x|, |y|$ are odd. Therefore,

$$\oplus(x) = \oplus(y) \implies |x \oplus y| \text{ is even} \implies M_{f\circ\mathsf{XOR}}[x,y] = 0.$$

Theorem 3.8 (which is applicable since Theorem 3.7 guarantees all elements of G to be query symmetric w.r.t. $\mathcal{Q}$ and function symmetric w.r.t. $\oplus$) yields the corollary. □

3.3 Simplification by Fourier Analysis

In this subsection we show that using the structure we derived in the last subsection, we can simplify the expression on the right-hand side of Theorem 3.9 to a concise analytic optimization problem over the Boolean hypercube. Lemma 2.5 says that the numerator equals $2^n \max_{S\subseteq[n]} |\widehat{f}(S)|$. For the denominator, we require the following lemma.

Lemma 3.10. *For all $f : \{-1,1\}^n \to \mathbb{R}$, and all $S \subseteq [n], b \in \{-1,1\}^S$, we have*

$$M_{f\circ\mathsf{XOR}} \circ \Delta_{S,b} = \begin{array}{c|c|c} & y_S = b & y_S \neq b \\ \hline x_S = b & 0 & \cdots \; M_{b,y_S} \; \cdots \\ \hline x_S \neq b & \begin{matrix}\vdots \\ M_{x_S,b} \\ \vdots\end{matrix} & 0 \end{array}, \tag{6}$$

where for all x_S, y_S, $M_{x_S,y_S} := M_{f_{\bar{S}|x_S\oplus y_S}\circ\mathsf{XOR}}$.

Proof. We verify the equality entry-wise. To that end, let $x, y \in \{-1,1\}^n$. If $\Delta_{S,b}[x,y] = 0$, then the left-hand side is 0. Moreover, we have $\mathbb{I}[x_S = b] = \mathbb{I}[y_S = b]$, i.e., either $x_S = b$ and $y_S = b$, or $x_S \neq b$ and $y_S \neq b$. In both cases, we also have a 0-entry on the right-hand side as well.

On the other hand, suppose that $\Delta_{S,b}[x,y] = 1$. Then, we know that either $x_S = b$ or $y_S = b$, but not both. Then, on the left-hand side of the equation, we have $M_{f\circ\mathsf{XOR}}[x,y] = f(x \oplus y)$, and on the right-hand side we have

$$M_{f_{\bar{S}|x_S\oplus y_S}\circ\mathsf{XOR}}[x_{\bar{S}}, y_{\bar{S}}] = f_{\bar{S}|x_S\oplus y_S}(x_{\bar{S}} \oplus y_{\bar{S}}) = f(1_S x_{\bar{S}} \oplus 1_S y_{\bar{S}} \oplus 1_{\bar{S}} x_S \oplus 1_{\bar{S}} y_S) = f(x \oplus y).$$

In the above equation, $1_S x_{\bar{S}}$ denotes the string that equals 1_S on S and x on $\bar{S}$. The strings $1_S y_{\bar{S}}$, $1_{\bar{S}} x_S$ and $1_{\bar{S}} y_S$ are defined analogously. □

Now, we turn to the computation of the spectral norm of this matrix.

Lemma 3.11. *For all $f : \{-1,1\}^n \to \mathbb{R}$, and all $S \subseteq [n]$ and $b \in \{-1,1\}^S$, we have*

$$\|M_{f\circ\mathsf{XOR}} \circ \Delta_{S,b}\| = \left\| \begin{array}{c|c|c} & y_S = b & y_S \neq b \\ \hline x_S = b & 0 & \cdots \; D_{b,y_S} \; \cdots \\ \hline x_S \neq b & \begin{matrix}\vdots \\ D_{x_S,b} \\ \vdots\end{matrix} & 0 \end{array} \right\|,$$

where $D_{x_S,y_S} := D_{2^{n-|S|}\widehat{f_{\bar{S}|x_S\oplus y_S}}}$. The rows and columns of the matrix on the RHS are indexed by $x, y \in \{-1,1\}^n$.

Proof. It suffices to show that the matrices on the left- and right-hand side are related by a similarity transform, since they leave the spectrum, and hence spectral norm, invariant. To that end, let $H_{\bar{S}}$ be the Hadamard transform over $\bar{S}$, and let $X = H_{\bar{S}} \otimes I_S$, where $I_S \in \mathbb{R}^{\{-1,1\}^S \times \{-1,1\}^S}$ is the identity matrix. We observe that $X^{-1} = H_{\bar{S}}^{-1} \otimes I_S$, and so combining Theorem 3.10 and Theorem 2.6, we observe that

$$X^{-1}(M_{f\circ\mathsf{XOR}} \circ \Delta_{S,b})X = \left[\begin{array}{c|c|c} & y_S = b & y_S \neq b \\ \hline x_S = b & 0 & \cdots \; H_{\bar{S}}^{-1}H_{\bar{S}}D_{b,y_S}H_{\bar{S}}^{-1}H_{\bar{S}} \; \cdots \\ \hline x_S \neq b & \begin{matrix}\vdots \\ H_{\bar{S}}^{-1}H_{\bar{S}}D_{x_S,b}H_{\bar{S}}^{-1}H_{\bar{S}} \\ \vdots\end{matrix} & 0 \end{array} \right],$$

which equals the matrix on the RHS of the equation in the lemma statement. □

The following lemma rewrites the spectral norm further. Specifically, using the structure of the matrix in the RHS of Theorem 3.11 we are able to rewrite the matrix as a direct sum of $2^{|\bar{S}|}$ many matrices, each of which is $2^{|S|}$-dimensional, with rows and columns indexed by strings in $\{-1,1\}^S$.

Lemma 3.12. *For all $f : \{-1,1\}^n \to \mathbb{R}$, and all $S \subseteq [n]$ and $b \in \{-1,1\}^S$, we have*

$$\|M_{f\circ\mathsf{XOR}} \circ \Delta_{S,b}\| = 2^{n-|S|} \max_{z\in\{-1,1\}^{\bar{S}}} \left\| \begin{array}{c|c|c} & y = b & y \neq b \\ \hline x = b & 0 & \cdots \; \widehat{f_{\bar{S}|b\oplus y}}(z) \; \cdots \\ \hline x \neq b & \begin{matrix}\vdots \\ \widehat{f_{\bar{S}|x\oplus b}}(z) \\ \vdots\end{matrix} & 0 \end{array} \right\|,$$

where the row and column labels range over $x, y \in \{-1,1\}^S$.

Proof. We start from the matrix on the RHS in Theorem 3.11. We observe that its $[x,y]$'th entry is only non-zero when $x_{\bar{S}} = y_{\bar{S}}$, since each of the blocks D_{x_S,y_S} is diagonal. Thus, by rearranging rows and columns of this matrix by the same permutations (which doesn't affect spectral norm), we obtain a block-diagonal matrix, with every block labeled by $z \in \{-1,1\}^{\bar{S}}$. Hence, its operator norm

is the maximum operator norm of each of these blocks. The block corresponding to $z \in \{-1,1\}^{\bar{S}}$ is precisely the matrix that appears in the RHS of the lemma statement, which concludes the proof. □

Now, we can prove the required bound on the operator norm of the denominator in Theorem 3.9.

Theorem 3.13. *For all $f : \{-1,1\}^n \to \mathbb{R}, S \subseteq [n]$ and all $b \in \{-1,1\}^S$, we have*

$$\|M_{f\circ\mathsf{XOR}} \circ \Delta_{S,b}\| = 2^n \max_{x\in\{-1,1\}^{\bar{S}}} \sqrt{\mathbf{Var}(\hat{f}_{S|x})}.$$

Proof. For $x \in \{-1,1\}^S$ and $y \in \{-1,1\}^{\bar{S}}$, let xy denote the bit string that equals x on S and y on $\bar{S}$. For the rest of this proof we identify a bitstring with its characteristic set. Specifically, this means for every $f : \{-1,1\}^n \to \mathbb{R}$, we use the notation $\hat{f}(x) := \hat{f}(\{j \in [n] : x_j = -1\})$. The spectral norm of a symmetric matrix with a single non-zero row is simply the norm of that row. We can thus combine Theorem 3.12 and Theorem 2.13 to obtain that

$$\begin{aligned}
\|M_{f\circ\mathsf{XOR}} \circ \Delta_{S,b}\| &= 2^{n-|S|} \max_{x\in\{-1,1\}^{\bar{S}}} \left\| \left(\widehat{f_{\bar{S}|y\oplus b}}(x)\right)_{y\in\{-1,1\}^S\setminus\{b\}} \right\|_2 \\
&= 2^{n-|S|} \max_{x\in\{-1,1\}^{\bar{S}}} \sqrt{\sum_{y\in\{-1,1\}^S\setminus\{b\}} \widehat{f_{\bar{S}|y\oplus b}}(x)^2} \\
&= 2^{n-|S|} \max_{x\in\{-1,1\}^{\bar{S}}} \sqrt{\sum_{y\in\{-1,1\}^S\setminus\{1^S\}} \widehat{f_{\bar{S}|y}}(x)^2} \qquad \text{by variable substitution} \\
&= 2^{n-|S|} \max_{x\in\{-1,1\}^{\bar{S}}} \sqrt{\sum_{y\in\{-1,1\}^S} \widehat{f_{\bar{S}|y}}(x)^2 - \widehat{f_{\bar{S}|1^S}}(x)^2} \\
&= 2^{n-|S|} \max_{x\in\{-1,1\}^{\bar{S}}} \sqrt{2^{|S|}\mathbb{E}_{y\in\{-1,1\}^S}[\widehat{f_{\bar{S}|y}}(x)^2] - \widehat{f_{\bar{S}|1^S}}(x)^2} \\
&= 2^{n-|S|} \max_{x\in\{-1,1\}^{\bar{S}}} \sqrt{2^{|S|}\sum_{y\in\{-1,1\}^S} \widehat{f}(xy)^2 - \left(\sum_{y\in\{-1,1\}^S} \widehat{f}(xy)\right)^2} \\
&\qquad \text{By Corollaries 2.9 and 2.10} \\
&= 2^{n-|S|} \max_{x\in\{-1,1\}^{\bar{S}}} \sqrt{2^{2|S|}\mathbb{E}[\hat{f}_{S|x}^2] - 2^{2|S|}\mathbb{E}[\hat{f}_{S|x}]^2} \\
&= 2^n \max_{x\in\{-1,1\}^{\bar{S}}} \sqrt{\mathbf{Var}(\hat{f}_{S|x})}.
\end{aligned}$$

□

Observe that the RHS in the theorem above is independent of b. Using this derived expression for the denominator, we are able to now state and prove our main theorem of this section. Recall the definition of $\mathsf{val}_{\mathcal{Q}}$ from Equation (1).

Theorem 1.1. *Let n be a positive integer and $\mathcal{Q} \subseteq 2^{[n]}$. Then $\mathsf{Q}^{\mathcal{Q}}(\oplus) = \Theta(\mathsf{val}_{\mathcal{Q}})$.*

Proof. By Theorem 3.9, we have

$$\mathrm{ADV}^{\mathcal{Q},\pm}(\oplus) = \max_{\substack{f:\{-1,1\}^n\to\mathbb{R},\\ |z| \text{ even} \implies f(z)=0}} \frac{\|M_{f\circ\mathsf{XOR}}\|}{\max\limits_{\substack{S\in\mathcal{Q},\\ b\in\{-1,1\}^S}} \|M_{f\circ\mathsf{XOR}}\circ\Delta_{S,b}\|}.$$

We can use Lemma 2.5 and Theorem 3.13 to rewrite the numerator and denominator, respectively. This implies that

$$\mathrm{ADV}^{\mathcal{Q},\pm}(\oplus) = \max_{\substack{f:\{-1,1\}^n\to\mathbb{R},\\ |z| \text{ even} \implies f(z)=0}} \frac{2^n\|\hat{f}\|_\infty}{2^n \max\limits_{S\in\mathcal{Q}} \max\limits_{b\in\{-1,1\}^{\bar{S}}} \sqrt{\mathbf{Var}(\widehat{f}_{S|b})}}.$$

Lemma 2.11 says that the condition '$|z|$ even $\implies f(z)=0$' is equivalent to the condition 'for all $S\subseteq[n]$, we have $\widehat{f}(S) = -\widehat{f}(\bar{S})$' (equivalently, $\widehat{f}$ is an odd function, identifying sets with their characteristic vectors). The RHS above is thus equal to

$$\max_{\substack{\widehat{f}:2^{[n]}\to\mathbb{R},\\ \widehat{f} \text{ is an odd function}}} \frac{2^n\|\hat{f}\|_\infty}{2^n \max\limits_{S\in\mathcal{Q}} \max\limits_{b\in\{-1,1\}^{\bar{S}}} \sqrt{\mathbf{Var}(\widehat{f}_{S|b})}}.$$

Syntactically replacing $\widehat{f}$ by f everywhere and using Theorem 2.3, we obtain

$$\mathrm{Q}^{\mathcal{Q}}(\oplus) = \Theta(\mathrm{ADV}^{\mathcal{Q},\pm}(\oplus)) = \Theta\left(\max_{\substack{f:\{-1,1\}^n\to\mathbb{R},\\ f \text{ is an odd function}}} \frac{\|f\|_\infty}{\max\limits_{\substack{S\in\mathcal{Q},\\ b\in\{-1,1\}^{\bar{S}}}} \sqrt{\mathbf{Var}(f_{S|b})}}\right). \qquad \square$$

Using Observation 3.1 and Lemma 3.2, we obtain the following immediate corollary about the quantum $\mathcal{Q}$-substring quantum query complexity of learning an input string.

Corollary 3.14. *Let $n\in\mathbb{N}$ and $\mathcal{Q}\subseteq 2^{[n]}$. Then, $\mathrm{Q}^{\mathcal{Q}}(\mathsf{rec}) = \Omega(\mathsf{val}_{\mathcal{Q}})$. If $\mathcal{Q}$ is also downward closed, then $\mathrm{Q}^{\mathcal{Q}}(\mathsf{rec}) = \Theta(\mathsf{val}_{\mathcal{Q}})$.*

4 Applications of the Framework

In this section we use the framework developed in the last section to analyze $\mathsf{val}_{\mathcal{Q}}$, and thereby the quantum $\mathcal{Q}$-substring query complexity of computing parity, for various interesting collections $\mathcal{Q}$.

4.1 Notation

We define some notation that we will use for the rest of this section.

For a collection of sets $\mathcal{Q}\subseteq 2^{[n]}$, we use the notation $\mathsf{val}_{\mathcal{Q}}$ as defined in Theorem 1.1. That is,

$$\mathsf{val}_{\mathcal{Q}} := \max_{\substack{f:\{-1,1\}^n\to\mathbb{R}\\ f \text{ is an odd function}}} \frac{\|f\|_\infty}{\max\limits_{\substack{S\in\mathcal{Q},\\ b\in\{-1,1\}^{\bar{S}}}} \sqrt{\mathbf{Var}(f_{S|b})}}. \tag{7}$$

Recall the following interpretation of $\mathsf{val}_{\mathcal{Q}}$: it is the maximum, over all odd functions $f : \{-1,1\}^n \to \mathbb{R}$, ratio of the maximum absolute value of f to the maximum standard deviation of f in a subcube. The subcubes under consideration here are precisely those where the set of free variables must be equal to an allowed substring query set (i.e., a set in $\mathcal{Q}$). The fixed variables can be fixed to any value. We organize the rest of this section into various subsections, one for each type of $\mathcal{Q}$ that we consider.

4.2 Contains All Singletons

In the case that the allowed substring queries include all singletons, there is an easy quantum query upper bound of n: simply query all of the variables one at a time. As a sanity check, we first show how this can be recovered by our framework.

Proposition 4.1. *Let $\mathcal{Q} \supseteq \{\{i\} : i \in [n]\}$. Then $\mathsf{val}_{\mathcal{Q}} \leq n$.*

Proof. Since f is odd, there is always an x^* such that $f(x^*) = \|f\|_\infty > 0$. Let $M = \max_x f(x)$ be attained at x^*. As f is odd, $f(-x^*) = -M$.

Consider a shortest path from x^* to $-x^*$ in the hypercube, this has length n. Along this path the total value of f decreases by $2M$, thus there must exist an edge with a drop of at least $2M/n$. That edge corresponds to flipping one coordinate, say $i \in [n]$. This means there exists an $z \in \{-1,1\}^n$ such that $|f(z) - f(z \oplus e_i)| \geq 2M/n$. Choose the subcube that fixes all variables but i to be consistent with z, and leave the i'th variable free. Note that this is a valid subcube to consider as $\mathcal{Q}$ contains all singletons. For this choice of subcube we have the variance (using the formula $\mathbf{Var}(g) = \frac{1}{2}\mathbb{E}[(g(x) - g(y))^2]$) to be $\frac{(f(z)-f(z\oplus e_i))^2}{4} \geq \frac{4M^2}{4n^2}$. Thus $\mathsf{val}_{\mathcal{Q}} \leq \frac{M}{\sqrt{4M^2/4n^2}} = n$. $\square$

4.3 Sets of Bounded Size

The setting $\mathcal{Q} = \left\{S \subseteq 2^{[n]} : |S| \leq k\right\}$ corresponds to the model where the allowed substrings to be queried are bounded in size.

In particular, $k = n$ corresponds to the usual search with wildcards setting, where an algorithm is allowed to query OR's of arbitrary sets of literals.

Theorem 4.2. *Let $1 \leq k \leq n$ be an integer, and let $\mathcal{Q} = \left\{S \subseteq 2^{[n]} : |S| \leq k\right\}$. Then,*

$$\frac{n}{\sqrt{k}} \leq \mathsf{val}_{\mathcal{Q}} \leq \frac{2n}{\sqrt{k}}.$$

We show the lower bound using the (appropriately shifted) Hamming weight function. We prove the upper bound by induction on n. We first state a convenient induction hypothesis, then show how it implies the upper bound, and finally prove the induction hypothesis.

Lemma 4.3. *For all integers $n > 0$ and all functions $f : \{-1,1\}^n \to \mathbb{R}$ with $\mathbb{E}[f(x)] = 0$, there exists a $S \subseteq [n]$ and $b \in \{-1,1\}^{\bar{S}}$ where:*

- $\mathbf{Var}(f_{S|b}) \geq \frac{\|f\|_\infty^2}{4n}$, *and*
- $\left|\widehat{f_{S|b}}(\{i\})\right| \geq \frac{\|f\|_\infty}{2n}$ *for all $i \in S$.*

In other words, there exists a subcube where the restricted function has large variance, as well as large degree-1 Fourier coefficients. We feel this is an independently interesting result in Fourier analysis. We first show how Theorem 4.2 follows from Lemma 4.3 (note that Lemma 4.3 gives no guarantee on $|S|$, which we require for proving Theorem 4.2).

Proof of Theorem 4.2. We first show the lower bound on $\mathsf{val}_{\mathcal{Q}}$, and then the upper bound.

- **Lower bound:** Consider the function $f : \{-1,1\}^n \to \mathbb{R}$ defined by $f(x) = |x| - (n/2)$. It is clear that f is an odd function since $f(-x) = |-x| - (n/2) = (n - |x|) - (n/2) = n/2 - |x| = -f(x)$. Let $m \leq k$. For any subcube that fixes all but a collection of m variables, it is easy to see that the random variable obtained by restricting f to this subcube and picking a uniformly random function value in the restricted subcube is simply a constant additive shift of the binomial random variable $\mathrm{Bin}(m, 1/2)$. Since such shifts do not affect variance, the variance in this subcube equals $m/4 \leq k/4$, since the only subcubes under consideration are those where the number of free variables is at most k. Since $\max |f(x)| = n/2$, this function f witnesses
$$\mathsf{val}_{\mathcal{Q}} \geq \frac{n/2}{\sqrt{k/4}} = \frac{n}{\sqrt{k}}.$$

- **Upper bound:** Let $M := \|f\|_\infty = \max_{x \in \{-1,1\}^n} |f(x)|$. Let $S \subseteq [n]$ and $b \in \{-1,1\}^{\bar{S}}$ be obtained from Lemma 4.3 (which is applicable since f is odd, and hence balanced).

  If $|S| \leq k$, then this is a valid subcube that can be chosen in the denominator of $\mathsf{val}_{\mathcal{Q}}$, and we get $\mathsf{val}_{\mathcal{Q}} \leq \frac{M}{\sqrt{M^2/4n}} = 2\sqrt{n} \leq \frac{2n}{\sqrt{k}}$, which proves the theorem. So for the rest of the proof we may assume that $|S| > k$. For convenience, define $g := f_{S|b} : \{-1,1\}^S \to \mathbb{R}$.

  Let T be any subset of S of size k, and define $\bar{T} := S \backslash T$ (i.e., the complement of T with respect to the universe S). For an arbitrary $x \in \{-1,1\}^{\bar{T}}$, consider the function $g_{T|x} : \{-1,1\}^T \to \mathbb{R}$ obtained by restricting g further by setting the variables in $\bar{T}$ to x. We have for each $i \in T \subseteq S$,
$$\begin{aligned} \mathbb{E}_x[\widehat{g_{T|x}}(\{i\})^2] &= \sum_{A \subseteq \bar{T}} \widehat{g}(A \cup \{i\})^2 && \text{by Corollary 2.10} \\ &\geq \widehat{g}(\emptyset \cup \{i\})^2 = \widehat{g}(\{i\})^2 \\ &\geq \frac{M^2}{4n^2}, \end{aligned}$$
  where the last line holds by Lemma 4.3. By linearity of expectation, and since $|T| = k$, we have
$$\mathbb{E}_x\left[\sum_{i \in T} \widehat{g_{T|x}}(\{i\})^2\right] = \sum_{i \in T} \mathbb{E}_x\left[\widehat{g_{T|x}}(\{i\})^2\right] \geq \frac{M^2 k}{4n^2}.$$
  Therefore there exists a $x^* \in \{-1,1\}^{\bar{T}}$ such that $\sum_{i \in T} \left(\widehat{g_{T|x^*}}(\{i\})^2\right) \geq \frac{M^2 k}{4n^2}$. Since $\mathbf{Var}(h) = \sum_{S \neq \emptyset} \widehat{h}(S)^2$ for all functions h by Lemma 2.12, we have
$$\mathbf{Var}(g_{T|x^*}) = \sum_{S \subseteq T, S \neq \emptyset} \widehat{g_{T|x^*}}(S)^2 \geq \sum_{i \in T} \widehat{g_{T|x^*}}(\{i\})^2 \geq \frac{M^2 k}{4n^2}.$$

Note that $g_{T|x^*}$ is the restriction of f to the subcube T, by setting the variables in $\bar{S}$ and $\bar{T}$ to b and x^*, respectively. This is a valid subcube for the denominator of $\mathsf{val}_{\mathcal{Q}}$ (see Equation (7)), and so this subcube witnesses

$$\mathsf{val}_{\mathcal{Q}} \leq \frac{M}{\sqrt{\frac{M^2 k}{4n^2}}} = \frac{2n}{\sqrt{k}},$$

which completes the proof. □

It only remains to prove our induction hypothesis, Lemma 4.3, which we do now.

Proof of Theorem 4.3. We prove the lemma by induction on n. Let $\|f\|_\infty = M$ and assume that $|f(x^*)| = M$.

- **Base case**: $n = 1$. Thus we have $f(-1) = M$ and $f(1) = -M$ (or the opposite, in which case essentially the same proof works). We have the variance over the whole cube as $\mathbf{Var}(f) = M^2 > \frac{M^2}{4}$. We also have $\widehat{f}(\{1\}) =$ either M or $-M$. In either case, $|\widehat{f}(\{1\})| = M \geq \frac{M}{2}$. This proves the base case.

- **Assumption for inductive step**: We assume the induction hypothesis to be true for $n = \ell$, i.e., for every balanced function f on ℓ variables, there exists a $S \subseteq [\ell]$ and $b \in \{-1,1\}^{\bar{S}}$ such that $\mathbf{Var}(f_{S|b}) \geq \frac{M^2}{4\ell}$, and $\left|\widehat{f_{S|b}}(\{i\})\right| \geq \frac{M}{2\ell}$ for all $i \in S$.

- **Inductive step**: For the inductive step, we assume that the hypothesis holds for $n = \ell$ (see above) and show it is true for $n = \ell + 1$. Define $\mu_{i,b} := \mathbb{E}_{x:x_i=b}[f(x)]$ for $b \in \{-1,1\}$. We consider two cases.

  - **Case 1:** $\exists i \in [\ell+1]$ such that $|\mu_{i,x_i^*} - \mu_{i,-x_i^*}| \leq \frac{M}{\ell+1}$. Since we know f to be balanced, we have $\mu_{i,x_i^*} + \mu_{i,-x_i^*} = 0$, and hence by the triangle inequality

$$|\mu_{i,x_i^*}| = \left|\frac{\mu_{i,x_i^*} - \mu_{i,-x_i^*}}{2} + \frac{\mu_{i,x_i^*} + \mu_{i,-x_i^*}}{2}\right| \leq \frac{|\mu_{i,x_i^*} - \mu_{i,-x_i^*}|}{2} + \frac{|\mu_{i,x_i^*} + \mu_{i,-x_i^*}|}{2} \leq \frac{M}{2(\ell+1)}.$$

    Define $f_i := f|_{x_i=x_i^*} - \mu_{i,x_i^*}$. Note that we have $\mathbb{E}[f_i] = 0$ and $\|f_i\|_\infty \geq M - \frac{M}{2(\ell+1)}$, where we used that the maximum absolute value of f is attained in the domain of f_i by construction (i.e., x^* is in the domain of f_i). Applying the induction hypothesis to f_i, we get a $S \subseteq [\ell+1]$ and $b \in \{-1,1\}^{\bar{S}}$ (with $i \in \bar{S}$ and $b_i = x_i^*$) such that:

    * We write $b' := b|_{\bar{S}\setminus\{i\}}$, i.e., the bit string b without the bit at position i, and obtain

$$\mathbf{Var}((f_i)_{S|b'}) \geq \frac{\left(M\left(1 - \frac{1}{2(\ell+1)}\right)\right)^2}{4\ell} \geq \frac{\left(M\left(\sqrt{1 - \frac{1}{\ell+1}}\right)\right)^2}{4\ell} = \frac{M^2}{4\ell} \cdot \frac{\ell}{\ell+1} = \frac{M^2}{4(\ell+1)}.$$

      Here we are using $1 - \frac{1}{2\alpha} \geq \sqrt{1 - \frac{1}{\alpha}}$ for all $\alpha > 1$, which can easily be verified by squaring both sides and comparing terms. Since $f_{S|b} = (f_i)_{S|b'} + \mu_{i,x_i^*}$ and variance is unaffected by a constant additive shift, this implies $\mathbf{Var}(f_{S|b}) \geq \frac{M^2}{4(\ell+1)}$ as well.

* First note that for any function $g : \{-1,1\}^n \to \mathbb{R}$ and any constant c, all Fourier coefficients of g are equal to the corresponding Fourier coefficients of $g-c$, except for the empty coefficient which gets subtracted by c. Thus, we again write $b' := b|_{\bar{S}\setminus\{i\}}$, to find that for all $j \in S$,

$$|\widehat{f_{S|b}}(\{j\})| = |\widehat{(f_i)_{S|b'}}(\{j\})| \geq \frac{M\left(1-\frac{1}{2(\ell+1)}\right)}{2\ell} = \frac{M\cdot\frac{2\ell+1}{2(\ell+1)}}{2\ell} > \frac{M\cdot 2\ell}{2(\ell+1)}\cdot\frac{1}{2\ell} = \frac{M}{2(\ell+1)}.$$

– **Case 2:** For all $i \in [\ell+1]$, we have $|\mu_{i,x_i^*} - \mu_{i,-x_i^*}| > \frac{M}{\ell+1}$. In that case, we choose $S = [\ell+1]$ and b becomes the empty string, so that $f_{S|b} = f$. Then, we find

* By definition of Fourier coefficients, we have for all $i \in [\ell+1]$,

$$|\widehat{f}(\{i\})| = |\mathbb{E}_x[f(x)\cdot x_i]| = \left|\frac{x_i^*}{2}(\mu_{i,x_i^*} - \mu_{i,-x_i^*})\right| > \frac{M}{2(\ell+1)},$$

which implies $|\widehat{f}(\{i\})| > \frac{M}{2(\ell+1)}$ for all $i \in [\ell+1]$.

* By Lemma 2.12, we have the variance of the whole subcube as

$$\mathbf{Var}(f) = \sum_{S\neq\emptyset}\widehat{f}(S)^2 \geq \sum_{i\in[\ell+1]}\widehat{f}(\{i\})^2 > (\ell+1)\frac{M^2}{4(\ell+1)^2} = \frac{M^2}{4(\ell+1)}. \qquad \square$$

4.4 Contiguous Blocks

The setting $\mathcal{Q} = \{[i,j] : i \leq j \in [n]\}$ corresponds to the model where the allowed substrings to be queried form contiguous blocks. While we don't phrase it as such, our results in this subsection also hold when the notion of contiguity allows for wraparound.

The following is our main result of this subsection.

Theorem 4.4. *Let n be a positive integer and let $\mathcal{Q} = \{[i,j] : i \leq j \in [n]\}$. Then there exists a universal constant c such that*

$$\frac{cn}{\log n} \leq \mathsf{val}_{\mathcal{Q}} \leq n.$$

Proof. Since $\mathcal{Q}$ contains all singletons, Proposition 4.1 implies the upper bound. For the lower bound, define $f : \{-1,1\}^{mk} \to \mathbb{R}$ as $f(z^{(1)},\ldots,z^{(m)}) = \sum_{i=1}^m g(z^{(i)})$. Here each $z^{(i)} \in \{-1,1\}^k$, and

$$g(z^{(i)}) = \begin{cases} 1, & z^{(i)} = 1^k, \\ -1, & z^{(i)} = (-1)^k, \\ 0, & \text{otherwise.} \end{cases}$$

First note that f is odd and $\max|f(x)| = m = n/k$. We show that every valid subcube in the denominator of the RHS of the expression of $\mathsf{val}_{\mathcal{Q}}$ (see Equation (7)) has small variance, which proves the theorem.

For any subcube (S,b) where S is a contiguous block, since a contiguous block can partially intersect at most 2 blocks (let the free variables in these blocks be denoted by strings $y^{(1)}$ and $y^{(2)}$), we can assume that the restriction of f to the subcube (S,b) has the following form:

$$f_{S|b}(y^{(1)},y^{(2)},z^{(1)},z^{(2)},\ldots z^{(t)}) = h_1(y^{(1)}) + h_2(y^{(2)}) + \sum_{i=1}^t g(z^{(i)}) + c, \tag{8}$$

where $h_1, h_2 \in \{r_1, r_{-1}, r_0\}$ depending on the values of the set variables in the corresponding blocks, where for all $b \in \{-1, 0, 1\}$,

$$r_b(y) := \begin{cases} b, & \text{if } y \text{ is the all-}b \text{ string,} \\ 0, & \text{otherwise.} \end{cases}$$

In particular, r_0 is the constant 0 function. Since each block in the restricted function contains disjoint variables, the variance of the sum on the RHS of Equation (8) equals the sum of variances of the individual terms, which equals

$$\begin{aligned} &= \mathbf{Var}(h_1) + \mathbf{Var}(h_2) + \sum_{i=1}^{t} \mathbf{Var}(g) + 0, \\ &= O(1) + t \cdot \mathbb{E}[g^2], \\ &= O(1) + m \cdot \frac{1}{2^{k-1}}, \end{aligned}$$

where the second line follows since since h_i is a $\{-1, 0, 1\}$-valued function, and each g is a balanced function, and the last line follows since $t \leq m$ and g^2 takes value 1 at precisely two points and 0 everywhere else.

If we set n to be the number of variables for f, which equals mk, then the above variance equals $O(1) + \frac{n/k}{2^{k-1}}$. Thus we have

$$\mathsf{val}_{\mathcal{Q}} = \frac{\|f\|_\infty}{\sqrt{\max\limits_{\substack{S \in \mathcal{Q}, \\ b \in \{-1,1\}^{\bar{S}}}} \mathbf{Var}(f_{S|b})}} = \frac{n/k}{\sqrt{O(1) + \frac{n/k}{2^{k-1}}}} = \Omega\left(\min\left\{\frac{n}{k}, \sqrt{\frac{n}{k}} \cdot 2^{\frac{k-1}{2}}\right\}\right).$$

Setting k to maximize this quantity asymptotically, we get $\frac{n}{k} = 2^{k-1}$, which implies $n = k \cdot 2^{k-1}$, implying $k = \log n - \log\log n + O(1)$, and hence $\mathsf{val}_{\mathcal{Q}} = \Omega(n/\log n)$, proving the theorem. □

4.5 Prefixes

In this subsection we deal with the setting where $\mathcal{Q} = \{[1, i] : i \in [n]\}$, that is, the allowed substrings to be queried are only prefixes of the input string.

Theorem 4.5. *Let n be a positive integer and let $\mathcal{Q} = \{[1, i] : i \in [n]\}$. Then,*

$$\frac{n}{\sqrt{8}} \leq \mathsf{val}_{\mathcal{Q}} \leq n.$$

We first define decision lists.

Definition 4.6 (Decision lists)**.** *A decision list is a sequence $D = ((L_1, a_1), (L_2, a_2), \ldots, (L_k, a_k), b)$, where each $a_i, b \in \mathbb{R}$ and each L_i is either a variable or the negation (i.e. complement) of a variable. The decision list computes a function $f : \{-1, 1\}^n \to \mathbb{R}$ as follows. If $L_1(x) = -1$, then $f(x) = a_1$; elseif $L_2(x) = -1$, then $f(x) = a_2$, elseif $\ldots$, elseif $L_k(x) = -1$, then $f(x) = a_k$, else $f(x) = b$.*

Proof of Theorem 4.5. We first show the upper bound and then the lower bound.

- The upper bound follows immediately from the observation that the values of all the variables (and hence their parity) can be learned by n (classical) prefix queries and the characterization given in Theorem 1.1. Below we recover the same bound by analyzing $\mathsf{val}_{\mathcal{Q}}$ directly. To this end, let $f : \{-1,1\}^n \to \mathbb{R}$ be any odd function, and say its maximum value M is attained at $x^* = (x_1^*, \ldots, x_n^*) \in \{-1,1\}^n$. For $i \in [n]$ define subcube C_i to be the one that sets variables $x_i, \ldots, x_n$ to $x_i^*, \ldots, x_n^*$ respectively and define $C_{n+1} := \{-1,1\}^n$. Note that the set of variables set in each of these subcubes is the complement of a prefix, and hence each of these subcubes is a valid subcube to consider in the denominator of the expression for $\mathsf{val}_{\mathcal{Q}}$. Since f is odd and hence balanced, $\mathbb{E}_{x \in C_{n+1}}[f(x)] = 0$. We also have that $\mathbb{E}_{x \in C_1}[f(x)] = f(x^*) = M$. Thus there exists an $i \in [n]$ such that $\mathbb{E}_{x \in C_{i+1}}[f(x)] - \mathbb{E}_{x \in C_i}[f(x)] \leq -M/n$. We will show that $\mathbf{Var}(f_{C_{i+1}})$ is large.

  Note that $C_i = C_{i+1} \cap \{x : x_i = x_i^*\}$. Let $C_i' := C_{i+1} \cap \{x : x_i = -x_i^*\}$. Thus $\mathbb{E}_{x \in C_{i+1}}[f(x)] = \frac{1}{2}\left(\mathbb{E}_{x \in C_i'}[f(x)] + \mathbb{E}_{x \in C_i}[f(x)]\right)$. Hence,

$$\begin{aligned} -\frac{M}{n} &\geq \mathbb{E}_{x \in C_{i+1}}[f(x)] - \mathbb{E}_{x \in C_i}[f(x)] \\ &= \frac{1}{2} \cdot \left(\mathbb{E}_{x \in C_i'}[f(x)] - \mathbb{E}_{x \in C_i}[f(x)]\right) \\ &\geq -|\mathbb{E}_{x \in C_{i+1}}[x_i \cdot f(x)]| \\ &= -|\widehat{f_{C_{i+1}}}(\{i\})|, \end{aligned}$$

  implying $|\widehat{f_{C_{i+1}}}(\{i\})| \geq M/n$. This in turn implies via Lemma 2.12 that $\mathbf{Var}(f_{C_{i+1}}) \geq |\widehat{f_{C_{i+1}}}(\{i\})|^2 \geq M^2/n^2$. We thus have that $\mathsf{val}_{\mathcal{Q}} \leq \frac{M}{\sqrt{M^2/n^2}} = n$.

- For the lower bound, we define a function such that every valid subcube has small variance. Define $g : \{-1,1\}^{n-1} \to \mathbb{R}$ to be the decision list $((x_{n-1}, 1), (x_{n-2}, 2), \ldots, (x_2, n-2), (x_1, n-1), n)$. Then define $f : \{-1,1\}^n \to \mathbb{R}$ as follows:

$$f(x_1, \ldots, x_{n-1}, y) = \begin{cases} g(x_1, \ldots, x_{n-1}), & \text{if } y = 1, \\ -g(-x_1, \ldots, -x_{n-1}), & \text{if } y = -1. \end{cases}$$

  Note that f is odd by construction. Towards bounding val_Q from below, we first note that $\|f\|_\infty = n$. Next, we bound the variance of f in each subcube that fixes every variable in the complement of a prefix. We show that the variance of f in each such subcube is $O(1)$. This will show that $\mathsf{val}_{\mathcal{Q}} \geq \frac{n}{\sqrt{O(1)}} = \Omega(n)$.

  Let $S \in \mathcal{Q}$ and $b \in \{-1,1\}^{\bar{S}}$. We split our analysis into three cases.

  - **Case 1:** $S = [1, n]$. Since f is odd and hence balanced, we have that

$$\begin{aligned} \mathbf{Var}(f) &= \mathbb{E}_{x_1, \ldots, x_{n-1}, y}[f(x_1, \ldots, x_{n-1}, y)^2] \\ &= \frac{1}{2} \cdot \mathbb{E}_{x_1, \ldots, x_{n-1}}[f(x_1, \ldots, x_{n-1}, 1)^2] + \frac{1}{2} \cdot \mathbb{E}_{x_1, \ldots, x_{n-1}}[f(x_1, \ldots, x_{n-1}, -1)^2] \\ &= \frac{1}{2} \cdot \mathbb{E}_{x_1, \ldots, x_{n-1}}[g(x_1, \ldots, x_{n-1})^2] + \frac{1}{2} \cdot \mathbb{E}_{x_1, \ldots, x_{n-1}}[(-g(-x_1, \ldots, -x_{n-1}))^2] \\ &= \mathbb{E}_{x_1, \ldots, x_{n-1}}[g(x_1, \ldots, x_{n-1})^2] \end{aligned}$$

$$= \sum_{i=1}^{n-1} \frac{1}{2^i} \cdot i^2 + \frac{1}{2^{n-1}} \cdot n^2 = \sum_{i=1}^{n} \frac{i^2}{2^i} + \frac{n^2}{2^n} \leq \sum_{i=1}^{\infty} \frac{i^2}{2^i} + \max_{n \geq 1} \frac{n^2}{2^n} < 6 + 2 = 8.$$

- **Case 2:** $S = [1, n-1]$. In this case, the restricted function is either $g(x_1, \ldots, x_{n-1})$ or $-g(-x_1, \ldots, -x_{n-1})$ depending on the setting of y. In either case, the variance of the restricted function is bounded above by $\mathbb{E}_{x_1,\ldots,x_{n-1}}[g(x_1, \ldots, x_{n-1})^2]$ which is at most 8 by the calculations in the previous case.
- **Case 3:** $S = [1, i]$ for some $i \in \{1, \ldots, n-2\}$ and $b = (b_{i+1}, \ldots, b_{n-1}, b_0) \in \{-1, 1\}^{n-i}$. Thus the subcube is obtained by setting $x_j = b_j$ for $j \in \{i+1, \ldots, n-1\}$ and $y = b_0$.
  - **Case 3a:** $b_0 = 1$. In this case, if $b_j = -1$ for some $j \neq 0$, then $f_{S|b}$ is a constant function, and hence has variance 0. So, we assume that $b_j = 1$ for all $j \in \{i+1, \ldots, n-1\}$. Then $f_{S|b}(x_1, \ldots, x_i)$ is the function given by the decision list $((x_i, n-i), (x_{i-1}, n-i+1), \ldots, (x_2, n-2), (x_1, n-1), n)$.
    In order to bound $\mathbf{Var}(f_{S|b})$ we use the following formulation of variance which is easy to verify from the definition of variance.

$$\mathbf{Var}(f_{S|b}) = \frac{1}{2} \cdot \mathbb{E}_{x,y}[(f_{S|b}(x) - f_{S|b}(y))^2].$$

    Note that for $|f_{S|b}(x) - f_{S|b}(y)|$ to be at least k, at least one of $f_{S|b}(x)$ and $f_{S|b}(y)$ must be at least $n - i + k$. For $f_{\bar{S}|b}(x)$ ($f_{\bar{S}|b}(y)$, resp.) to be at least $n - i + k$, all x_j (all y_j, resp.) for $j \in \{i, i-1, \ldots, i-k-1\}$ must equal 1, an event whose probability is $1/2^k$. We thus have that

$$\mathbf{Var}(f_{S|b}) = \frac{1}{2} \cdot \mathbb{E}_{x,y}[(f_{S|b}(x) - f_{S|b}(y))^2] = \frac{1}{2} \sum_{k=0}^{n} k^2 \mathbb{P}_{x,y}\left[|f_{S|b}(x) - f_{S|b}(y)| = k\right]$$

$$\leq \frac{1}{2} \cdot \sum_{k=1}^{i-1} 2 \cdot \frac{1}{2^k} \cdot k^2 < 6.$$

  - **Case 3b:** $b_0 = -1$. In this case, if $b_j = 1$ for some $j \neq 0$, then $f_{S|b}$ is a constant function, and hence has variance 0. So, we assume that $b_j = -1$ for all $j \in \{i+1, \ldots, n-1\}$. Then $f_{S|b}(x_1, \ldots, x_{i-1})$ is the function given by the decision list $((-x_i, -(n-i)), (-x_{i-1}, -(n-i+1)), \ldots, (-x_2, -(n-2)), (-x_1, -(n-1)), -n)$. By a similar analysis as in the previous case, it follows that $\mathbf{Var}(f_{S|b}) < 6$.

The theorem now follows from the definition of $\mathsf{val}_{\mathcal{Q}}$ in Equation (7). □

We remark that Theorem 4.5 and Theorem 1.1 imply a lower bound of $\Omega(n)$ for computing parity using prefix queries. We have not tried to optimize the constants in Theorem 4.5. In fact, we show a tighter bound of approximately $n/\sqrt{2}$ in the appendix using the adversary bound directly (see Theorem B.1) for the function that computes the last input bit x_n (which can only be easier that computing the parity of all the bits).

4.6 Only Full Set

Finally, we consider the setting where $\mathcal{Q} = \{[n]\}$. That is, the only allowed substrings must contain all indices. The task of learning $x \in \{-1, 1\}^n$ is equivalent to search over a space of 2^n

elements, where one element is promised to be marked.[4] Our bound of $\mathsf{val}_{\mathcal{Q}} = 2^{(n-1)/2}$ in this case recovers the now well-known lower bound of $\Omega(\sqrt{N})$ of searching for a marked item among N elements [BBHT98]. The upper bound of $O(\sqrt{N})$ for search follows from [Gro96].

Theorem 4.7. *Let n be a positive integer and let $\mathcal{Q} = \{[n]\}$. Then,*

$$\mathsf{val}_{\mathcal{Q}} = 2^{(n-1)/2}.$$

Proof. We first show the lower bound, and then the upper bound. Since $\mathcal{Q} = \{[n]\}$, the denominator of the expression for $\mathsf{val}_{\mathcal{Q}}$ (Equation (7)) is the standard deviation of f on the whole cube.

- For the lower bound, consider $f : \{-1,1\}^n \to \mathbb{R}$ defined by $f(1^n) = 1, f(-1^n) = -1$, and $f(x) = 0$ for all other $x \in \{-1,1\}^n$. We have the standard deviation of f on the whole cube equal to $\sqrt{\mathbf{Var}(f)} = \sqrt{\mathbb{E}[f^2] - \mathbb{E}[f]^2} = \sqrt{\mathbb{E}[f^2]} = \sqrt{\frac{2}{2^n}} = 2^{-(n-1)/2}$, and hence f witnesses $\mathsf{val}_{\mathcal{Q}} \geq 2^{(n-1)/2}$.
- For the upper bound let $f : \{-1,1\}^n \to \mathbb{R}$ be such that f is odd, which in particular means $\mathbb{E}[f] = 0$. The standard deviation over the whole cube equals $\sqrt{\mathbf{Var}(f)} = \sqrt{\mathbb{E}[f^2] - \mathbb{E}[f]^2} = \sqrt{\mathbb{E}[f^2]} \geq \sqrt{\frac{2}{2^n} \max |f(x)|^2}$. Here the last inequality holds because f is odd: if $x^* = \arg\max_x |f(x)|$, then since $f(x) = -f(-x)$, we also have $-x^* = \arg\max_x |f(x)|$. Thus, for all odd functions f we have $\mathsf{val}_{\mathcal{Q}} \leq \frac{\max |f(x)|}{\sqrt{\frac{2}{2^n} \max |f(x)|^2}} = 2^{(n-1)/2}$. □

5 Conclusions

This work introduces a general framework for analyzing query algorithms for learning input strings in models that extend the search with wildcards setting [AM14]. As applications of our framework we are able to recover existing upper bounds [AM14, Bel15] in the search-with-wildcards setting and also show new bounds in generalized search-with-wildcards settings. Interestingly, our upper bounds in this paper are all shown (without explicitly resorting to SDP duality) using the *primal* version of the negative-weight adversary bound [HLŠ07], which is a maximization SDP, typically used to show *lower bounds*. To the best of our knowledge, ours is the first work to use the primal adversary bound to establish novel upper bounds. Our work makes progress towards answering an open question of [BBGK18], who asked if one could devise algorithms for learning input strings or computing some function of the input string, assuming non-standard query access to the input. It would be interesting to see if our results can be used to show any classical lower bounds using the framework developed in [BBGK18].

An interesting related direction is that of the AND/OR-query complexity of computing Boolean functions rather than learning an input string. It is not hard to show that the query model in the search-with-wildcards setting is equivalent to allowing unit-cost query access to arbitrary ORs and arbitrary ANDs of inputs. It was recently shown [CDM+23] that randomness does not give a superpolynomial advantage over determinism for computing total Boolean functions in this model. It would be interesting to see if ideas from our work could be generalized to show that quantumness also does not provide a superpolynomial advantage over randomness/determinism. Our framework

[4] Note that the query set $\mathcal{Q}$ is not downward closed (see Theorem 3.2) and therefore $\mathsf{val}_{\mathcal{Q}}$ only gives us a lower bound on $\mathsf{Q}^{\mathcal{Q}}(\mathsf{rec})$.

relies on the optimizing matrix for the primal adversary bound being the communication matrix of an XOR function, which crucially uses symmetries of Parity (the function being computed), so it is not clear how to generalize our framework in this direction easily.

Acknowledgments

N.S.M. gratefully acknowledges support from the London Mathematical Society through a Scheme 7 grant, which enabled preliminary work that sparked this project. The authors gratefully acknowledge OpenAI's ChatGPT (GPT-5), which assisted in developing several crucial ideas presented in this work. The authors bear full responsibility for any mistakes in the work.

A.C. is supported by a Simons-CIQC postdoctoral fellowship through NSF QLCI Grant No. 2016245.

References

[AM14] Andris Ambainis and Ashley Montanaro. Quantum algorithms for search with wildcards and combinatorial group testing. *Quantum Inf. Comput.*, 14(5-6):439–453, 2014.

[BBGK18] Shalev Ben-David, Adam Bouland, Ankit Garg, and Robin Kothari. Classical lower bounds from quantum upper bounds. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 339–349. IEEE, 2018.

[BBHT98] Michel Boyer, Gilles Brassard, Peter Høyer, and Alain Tapp. Tight bounds on quantum searching. *Fortschritte der Physik: Progress of Physics*, 46(4-5):493–505, 1998.

[BC02] Anna Bernasconi and Bruno Codenotti. Spectral analysis of boolean functions as a graph eigenvalue problem. *IEEE transactions on computers*, 48(3):345–351, 2002.

[Bel15] Aleksandrs Belovs. Quantum algorithms for learning symmetric juntas via the adversary bound. *Comput. Complex.*, 24(2):255–293, 2015.

[BV97] Ethan Bernstein and Umesh V. Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26(5):1411–1473, 1997. Earlier version in STOC'93.

[CDM+23] Arkadev Chattopadhyay, Yogesh Dahiya, Nikhil S. Mande, Jaikumar Radhakrishnan, and Swagato Sanyal. Randomized versus deterministic decision tree size. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing (STOC)*, pages 867–880. ACM, 2023.

[Cor23] Arjan Cornelissen. *Quantum multivariate estimation and span program algorithms*. University of Amsterdam, 2023.

[CWE15] Mark JP Chaisson, Richard K Wilson, and Evan E Eichler. Genetic variation and the de novo assembly of human genomes. *Nature Reviews Genetics*, 16(11):627–640, 2015.

[FGGS99] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. Bound on the number of functions that can be distinguished with k quantum queries. *Physical Review A*, 60(6):4331, 1999.

[Gro96] Lov K Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM Symposium on Theory of Computing (STOC)*, pages 212–219, 1996.

[HLŠ07] Peter Høyer, Troy Lee, and Robert Špalek. Negative weights make adversaries stronger. In *Proceedings of the thirty-ninth annual ACM Symposium on Theory of Computing (STOC)*, pages 526–535, 2007.

[LD24] Heng Li and Richard Durbin. Genome assembly in the telomere-to-telomere era. *Nature Reviews Genetics*, 25(9):658–670, 2024.

[Man18] Nikhil Shekhar Mande. *Communication complexity of XOR functions*. PhD thesis, PhD thesis, Tata Institute of Fundamental Research Mumbai, 2018.

[O'D14] Ryan O'Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014.

[Rei09] Ben W. Reichardt. Span programs and quantum query complexity: the general adversary bound is nearly tight for every Boolean function. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science—FOCS*, pages 544–551. IEEE Computer Soc., Los Alamitos, CA, 2009.

[Rei11] Ben W. Reichardt. Reflections for quantum query algorithms. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 560–569. SIAM, Philadelphia, PA, 2011.

[RSC+10] Gordon Robertson, Jacqueline Schein, Readman Chiu, Richard Corbett, Matthew Field, Shaun D Jackman, Karen Mungall, Sam Lee, Hisanaga Mark Okada, Jenny Q Qian, et al. De novo assembly and analysis of RNA-seq data. *Nature methods*, 7(11):909–912, 2010.

[SS95] Steven S Skiena and Gopalakrishnan Sundaram. Reconstructing strings from substrings. *Journal of Computational Biology*, 2(2):333–353, 1995.

[vD98] Wim van Dam. Quantum oracle interrogation: Getting all information for almost half the price. In *Proceedings 39th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 362–367. IEEE, 1998.

[XZL23] Yongzhen Xu, Shihao Zhang, and Lvzhou Li. Quantum algorithm for learning secret strings and its experimental demonstration. *Physica A: Statistical Mechanics and its Applications*, 609:128372, 2023.

A Proof of Theorem 3.3

We first observe that for any Boolean function F on $D \subseteq \{-1,1\}^n$, the optimal adversary matrix Γ for $\mathrm{ADV}^{\mathcal{Q},\pm}(F)$ attains its norm without needing absolute values, i.e., its spectral norm equals its largest eigenvalue.

Lemma A.1. *Let $D \subseteq \{-1,1\}^n$ and $F : D \to \{-1,1\}$ be a Boolean function. Let Γ be an optimal solution to $\mathrm{ADV}^{\mathcal{Q},\pm}(F)$. Then $\|\Gamma\| = \lambda_{\max}(\Gamma)$.*

Proof. To prove the lemma, it is sufficient to show that if λ is an eigenvalue of Γ, then so is $-\lambda$. We first group the elements in D as $F^{-1}(1)$ and $F^{-1}(-1)$. Grouping the indices of Γ, it takes the following form.

$$\Gamma = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}. \tag{9}$$

Let the rank of the matrix A be r. Then, the rank of the matrix Γ is $2r$. Furthermore, let $A = \sum_{j=1}^{r} s_j |u_j\rangle \langle v_j|$ be the singular value decomposition of A. Now, for every $j \in [r]$, we find

$$\Gamma(|v_j\rangle \oplus \pm |u_j\rangle) = (\pm A^T |u_j\rangle) \oplus A |v_j\rangle = s_j(\pm |v_j\rangle \oplus |u_j\rangle) = \pm s_j(|v_j\rangle \oplus \pm |u_j\rangle),$$

and so we have found all $2r$ eigenvalues of Γ. It follows that the spectrum is $\{s_j, -s_j : j \in [r]\}$, which is symmetric around 0. It follows that $\|\Gamma\| = \lambda_{\max}(\Gamma)$. □

We require the following fact that is easy to verify.

Fact A.2. *For a diagonal matrix A, and matrices B, C, we have $A(B \circ C)A = (ABA) \circ C$.*

For the sake of clarity, we restate Theorem 3.3.

Theorem 3.3. *Let $D \subseteq \{-1,1\}^n$ and $F : D \to \{-1,1\}$ be a Boolean function. Then $\mathrm{ADV}^{\mathcal{R},\pm}(F) = \overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F)$. Moreover, if (β, Γ) is an optimal solution for $\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F)$, then Γ', defined as*[5]

$$\Gamma'[x,y] = \frac{\Gamma[x,y]}{\sqrt{\beta[x]\beta[y]}},$$

is an optimal solution for $\mathrm{ADV}^{\mathcal{R},\pm}(F)$.

Proof. We split our proof into two parts. First, we show how to convert a feasible solution Γ' of $\mathrm{ADV}^{\mathcal{R},\pm}(F)$ into a feasible solution (β, Γ) of $\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F)$ with the same objective value. Next, we show that if (β, Γ) is an optimal solution for $\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F)$, then Γ' is an optimal solution for $\mathrm{ADV}^{\mathcal{R},\pm}(F)$.

Standard formulation to alternate formulation Let M be an optimal solution to $\mathrm{ADV}^{\mathcal{R},\pm}(F)$. Then there exists a unit vector $|v\rangle$ such that $M |v\rangle = \lambda_{\max}(M) |v\rangle$, and so

$$\begin{aligned} \|M\| &= |\lambda_{\max}(M)|, && (10) \\ &= \lambda_{\max}(M), && \text{(by Theorem A.1)} \\ &= \langle v| M |v\rangle, && (M |v\rangle = \lambda_{\max}(M) |v\rangle) \end{aligned}$$

If an entry of $|v\rangle$ has a negative sign, we can switch this to a positive sign and instead multiply the corresponding row and column of M with a negative sign. Define a diagonal matrix $A \in \{-1,0,1\}^{D \times D}$ as follows.

$$A[x,x] = \begin{cases} -1, & v[x] < 0, \\ 1, & v[x] \geq 0. \end{cases}$$

[5] Here, we use the convention that $0/0 = 0$.

We write $|v_{\mathsf{abs}}\rangle$ for the vector containing the entry-wise absolute value of $|v\rangle$. Then $A|v_{\mathsf{abs}}\rangle = |v\rangle$, and

$$\langle v| M |v\rangle = \langle v_{\mathsf{abs}}| AMA |v_{\mathsf{abs}}\rangle = \sum_{x,y} |v[x]|A[x,x]M[x,y]A[y,y]|v[y]|. \tag{11}$$

Since M is an optimal (and hence feasible) solution to $\mathrm{ADV}^{\mathcal{R},\pm}(F)$, $F(x) = F(y) \implies M[x,y] = 0$, which implies, since AMA is simply flipping the signs of certain rows and columns of M, that $F(x) = F(y) \implies AMA[x,y] = 0$, which means AMA is a feasible solution to $\mathrm{ADV}^{\mathcal{R},\pm}(F)$ as well. By Equation (11), this new solution is also optimal, and hence we can assume without loss of generality that all entries of $|v\rangle$ are non-negative. For the rest of the proof we assume that M is an optimal solution to $\mathrm{ADV}^{\mathcal{R},\pm}(F)$ whose principal eigenvector v has only non-negative entries.

Now, we define a vector β such that $\beta[x] = v[x]^2$. Then, Equation (11) implies

$$\langle v| M |v\rangle = \sum_{x,y\in D} \sqrt{\beta[x]}M[x,y]\sqrt{\beta[y]}. \tag{12}$$

Define the quantity $s = \sum_{x\in F^{-1}(1)} \beta[x]$. By the definition of M, the contribution to the sum $\sum_{x,y\in D} \sqrt{\beta[x]}M[x,y]\sqrt{\beta[y]}$ from (x,y) with $F(x) = F(y)$ is 0. This means every non-zero summand must correspond to $x \in F^{-1}(1), y \in F^{-1}(-1)$ (or vice versa). Since the objective value considered is non-zero, we must have $s > 0$. Note that since $\|v\| = 1$, this means $\sum_x \beta[x] = 1$, which implies $s \leq 1$ (a similar argument to the above shows $s < 1$). Consider vector $\beta' \in \mathbb{R}^D$ whose entries are defined as follows.

$$\beta'[x] = \begin{cases} \frac{1}{2s}\beta[x], & x \in F^{-1}(1) \\ \frac{1}{2(1-s)}\beta[x], & x \in F^{-1}(-1) \end{cases}$$

We have

$$\begin{aligned} &\sum_{x,y\in D} \sqrt{\beta'[x]}M[x,y]\sqrt{\beta'[y]}, && (13)\\ &= \sum_{x,y\in D} \frac{1}{2\sqrt{s(1-s)}}\sqrt{\beta[x]}M[x,y]\sqrt{\beta[y]}, && (F(x) = F(y) \implies M[x,y] = 0)\\ &\geq \sum_{x,y\in D} \sqrt{\beta[x]}M[x,y]\sqrt{\beta[y]}. && (s \in (0,1)) \end{aligned}$$

This choice of β' guarantees that

$$\sum_{x\in F^{-1}(1)} \beta'[x] = \sum_{x\in F^{-1}(-1)} \beta'[x] = \frac{1}{2}. \tag{14}$$

Define the diagonal matrix $B \in \mathbb{R}^{D\times D}$ such that $B[x,x] = \sqrt{\beta'[x]}$. Next, we define a new matrix $\Gamma = BMB$.

$$\sum_{x,y\in D} \Gamma[x,y] = \sum_{x,y\in D} \sqrt{\beta'[x]}M[x,y]\sqrt{\beta'[y]}. \tag{15}$$

Observe that both M and Δ_q are symmetric matrices. Therefore, $M \circ \Delta_q$ is also a symmetric matrix. Due to this, $\|M \circ \Delta_q\| = |\lambda_{\max}(M \circ \Delta_q)|$. Then

$$\|M \circ \Delta_q\| \leq 1 \implies |\lambda_{\max}(M \circ \Delta_q)| \leq 1,$$

which implies all eigenvalues of $M \circ \Delta_q$ are in $[-1, 1]$. Note that any vector in $\mathbb{R}^D$ is an eigenvector of I_D with eigenvalue 1. Then $I_D - M \circ \Delta_q$ has the same eigenvectors as $M \circ \Delta_q$ with eigenvalues of the form $1 - \lambda$ where $\lambda \in [-1, 1]$. Therefore, all eigenvalues of $I_D - M \circ \Delta_q$ are non-negative and hence $I_D - M \circ \Delta_q \succeq 0$. This implies, since B is diagonal with all non-negative entries,

$$\begin{aligned} B(I_D - M \circ \Delta_q)B &\succeq 0, && (16) \\ BB - BMB \circ \Delta_q &\succeq 0, && \text{by Theorem A.2} \\ \mathsf{diag}(\beta') - \Gamma \circ \Delta_q &\succeq 0. \end{aligned}$$

Note that (β', Γ) has the following properties.

- $\Gamma = BMB$ is symmetric since B is diagonal and M is symmetric, β' is entry-wise non-negative by construction.
- $\Gamma[x, y] = 0$ if $F(x) = F(y)$ since $\Gamma[x, y] = \sqrt{\beta'[x]} M[x, y] \sqrt{\beta'[y]}$ and $M[x, y] = 0$ if $F(x) = F(y)$.
- $\mathsf{diag}(\beta) - \Gamma \circ \Delta_q \succeq 0$ for all $q \in \mathcal{R}$ due to Equation (16)
- $\sum_{x \in F^{-1}(1)} \beta'[x] = \sum_{x \in F^{-1}(-1)} \beta'[x] = \frac{1}{2}$ due to Equation (14)

Thus, (β', Γ) is a feasible solution to $\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F)$, and the objective value attained is

$$\begin{aligned} \sum_{x,y \in D} \Gamma[x, y] &= \sum_{x,y \in D} \sqrt{\beta'[x]} M[x, y] \sqrt{\beta'[y]} && \text{(by Equation (15))} \\ &\geq \sum_{x,y \in D} \sqrt{\beta[x]} M[x, y] \sqrt{\beta[y]} && \text{(by Equation (13))} \\ &= \langle v | \, M \, | v \rangle && \text{(by Equation (12))} \\ &= \|M\|. && \text{(by Equation (10))} \end{aligned}$$

This shows that $\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F) \geq \mathrm{ADV}^{\mathcal{R},\pm}(F)$.

Alternate formulation to standard formulation In the other direction, let (β, Γ) be an optimal solution to $\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F)$. Define $\Gamma' = B\Gamma B$ where the matrix B is a diagonal matrix defined as follows.

$$B[x, x] = \begin{cases} \frac{1}{\sqrt{\beta[x]}}, & \beta[x] \neq 0 \\ 0, & \text{otherwise} \end{cases} \qquad (17)$$

Suppose that $\Gamma[x,y] > 0$. Then, $F(x) \neq F(y)$, which in particular means that $x \neq y$. Let q be a query that distinguishes between x and y. Since $\mathsf{diag}(\beta) - \Gamma \circ \Delta_q \succeq 0$, so are all its submatrices. In particular, the 2×2 submatrix of indexed by x and y is also PSD, and it looks as follows;

$$\begin{bmatrix} \beta[x] & -\Gamma[x,y] \\ -\Gamma[x,y] & \beta[y] \end{bmatrix} \succeq 0.$$

For a 2×2 matrix to be PSD, the determinant must be non-negative, which implies that $\beta[x]\beta[y] - \Gamma[x,y]^2 \geq 0$. Since $\Gamma[x,y] \neq 0$ by assumption, we find that $\beta[x]\beta[y] \geq \Gamma[x,y]^2 > 0$, and so we find that $\beta[x] > 0$ and $\beta[y] > 0$. By the contrapositive, if $\beta[x] = 0$, then the entire row and column in Γ labeled by x is 0. As such, we can rewrite the objective value in $\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F)$ as

$$\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F) = \sum_{x,y \in D} \Gamma[x,y] = \sum_{\substack{x,y \in D \\ \beta[x]>0 \wedge \beta[y]>0}} \Gamma[x,y].$$

Now, since (β, Γ) is a feasible solution, $\mathsf{diag}(\beta) - \Gamma \circ \Delta_q \succeq 0$. Then

$$\begin{aligned} B(\mathsf{diag}(\beta) - \Gamma \circ \Delta_q)B \succeq 0, \qquad & (18) \\ B\mathsf{diag}(\beta)B - B\Gamma B \circ \Delta_q \succeq 0, \qquad & \text{by Theorem A.2} \\ I_D - \Gamma' \circ \Delta_q \succeq 0. \qquad & \end{aligned}$$

This implies that eigenvalues of $I_D - \Gamma' \circ \Delta_q$ are non-negative. Since $I_D - \Gamma' \circ \Delta_q$ and $\Gamma' \circ \Delta_q$ have the same eigenvectors, eigenvalues of $\Gamma' \circ \Delta_q$ are at most 1. Furthermore, Γ' and $\Gamma' \circ \Delta_q$ are of the form

$$\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}.$$

This observation allows us to use the same argument as in the proof of Theorem A.1, implying that for all eigenvalues λ, $-\lambda$ is also an eigenvalue. Combining these two observations, we get $\|\Gamma' \circ \Delta_q\| \leq 1$.

Observe that Γ' has the following properties.

- $\Gamma' = B\Gamma B$ is symmetric since B is diagonal and Γ is symmetric.
- Since 0-entries in Γ continue to be 0-entries in Γ', this means $F(x) = F(y) \implies \Gamma'(x,y) = 0$.
- $\|\Gamma' \circ \Delta_q\| \leq 1$ for all $q \in \mathcal{R}$ as shown above.

Let the vector $|v\rangle$ be defined as $v[x] = \sqrt{\beta[x]}$. Then

$$\begin{aligned} \mathrm{ADV}^{\mathcal{R},\pm}(F) \geq \left\|\Gamma'\right\| \geq \langle v| \Gamma' |v\rangle &= \sum_{x,y} \sqrt{\beta[x]}\Gamma'[x,y]\sqrt{\beta[y]} \\ &= \sum_{\substack{x,y \\ \beta[x]>0 \wedge \beta[y]>0}} \sqrt{\beta[x]}\frac{1}{\sqrt{\beta[x]}}\Gamma[x,y]\frac{1}{\sqrt{\beta[y]}}\sqrt{\beta[y]} = \sum_{\substack{x,y \\ \beta[x]>0 \wedge \beta[y]>0}} \Gamma[x,y] = \overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F). \end{aligned}$$

Moreover, from an optimal solution (β, Γ) for $\overline{\mathrm{ADV}}^{\mathcal{R},\pm}(F)$, we obtained a solution Γ' to $\mathrm{ADV}^{\mathcal{R},\pm}(F)$ defined by $\Gamma'[x,y] = \frac{\Gamma[x,y]}{\sqrt{\beta[x]\beta[y]}}$ with at least as large objective value (and hence equal, by the first part of this proof). This concludes the proof. □

B Prefix-Query Complexity of Dictator Function

In this section we study the query complexity of the Dictator function $\mathsf{Dict}_n : \{-1,1\}^n \to \{-1,1\}$ defined by $\mathsf{Dict}_n(x) = x_n$ in the substring query model with $\mathcal{Q} = \{[1,i] \mid i \in [n]\}$ as the set of all prefixes.

Lemma B.1. *Let n be a positive integer. Let $\mathcal{Q} = \{[1,i] \mid i \in [n]\}$. Then,*

$$\mathrm{ADV}^{\mathcal{Q},\pm}(\mathsf{Dict}_n) \geq 2 + \frac{n-3}{\sqrt{2}}.$$

Proof. We prove a lower bound for $\mathrm{ADV}^{\mathcal{Q},\pm}(\mathsf{Dict}_n)$ by exhibiting a feasible solution, denoted by Γ, to the standard formulation of the primal adversary bound for the Dict_n function and by analyzing the quantity

$$\frac{\|\Gamma\|}{\max\limits_{\substack{S\in\mathcal{Q},\\ b\in\{-1,1\}^S}} \|\Gamma \circ \Delta_{S,b}\|}. \tag{19}$$

Defining matrix Γ Let $\mathsf{Index} : \{-1,1\}^n \to [n] \cup \{0\}$ be a function defined as $\mathsf{Index}(x) = 0$ if $x_n = 1$, and $\mathsf{Index}(x) = \min\{i \in [n] : x_i = -1\}$ otherwise.

Let $f : \{-1,1\}^n \to \mathbb{R}$ be a function defined as follows.

$$f(x) = \begin{cases} 0, & x_n = 1 \\ \frac{1}{2}, & \mathsf{Index}(x) \in \{n-1, n\} \\ \frac{1}{2^{d-1}\sqrt{2}}, & \mathsf{Index}(x) = n-d \quad \forall d \in [2, n-2] \\ \frac{1}{2^{n-2}}, & \mathsf{Index}(x) = 1. \end{cases}$$

Let $\Gamma \in \mathbb{R}^{2^n \times 2^n}$ be the matrix defined as $\Gamma = M_{f\circ\mathsf{XOR}}$.

First observe that this is a valid matrix to consider for the adversary bound: it is symmetric, and $\mathsf{Dict}_n(x) = \mathsf{Dict}_n(y) \implies (x \oplus y)_n = 1 \implies f(x \oplus y) = 0$. To compute the quantity in Equation (19), we first reorder the rows and similarly the columns of Γ in a way that Γ can be partitioned into blocks containing all strings with $x_n = 1$ appear first, followed by all strings with $x_n = -1$. Clearly, with such an ordering the two diagonal blocks will contain all zeroes. Furthermore, the remaining two off-diagonal blocks will be transpose of each other because Γ is a symmetric matrix by construction. Consequently, Γ is of the form

$$\Gamma = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}. \tag{20}$$

Further ordering is simply lexicographic in the reverse of the strings. Equivalently this ordering is done by the value of the function $\mathsf{Index}(\cdot)$: the bit-string x appears before the bit-string y if $\mathsf{Index}(x) > \mathsf{Index}(y)$, and ties are broken by similarly comparing the values of Index on the substrings of x and y obtained by deleting the first $\mathsf{Index}(x)$ bits.

Lower bounding numerator of Equation (19) Let $x \in \{-1,1\}^n$ be an arbitrary string. The x'th row of Γ contains 2^{n-1} entries with value 0 (corresponding to y with $x_n = y_n$), 2^{n-2} entries with value $\frac{1}{2^{n-2}}$ (corresponding to all remaining y with $x_1 \neq y_1$), 2^{d-1} entries of $\frac{1}{2^{d-1}\sqrt{2}}$ $\forall d \in [2, n-2]$ (corresponding to all y with $x_n \neq y_n$ and $\mathsf{Index}(x \oplus y) = n - d$) and 2 entries of values $\frac{1}{2}$ (corresponding to the input strings $(1^{n-2}, -1, -1)$ and $(1^{n-1}, -1)$). Let $|\bar{1}\rangle$ denote the all-1 vector of length 2^n. Then,

$$\frac{1}{2^n} \langle \bar{1}| \Gamma |\bar{1}\rangle = \frac{1}{2^n} \left[2^n \left(\frac{1}{2} + \frac{1}{2} + \sum_{d=2}^{n-2} \frac{2^{d-1}}{2^{d-1}\sqrt{2}} + \frac{2^{n-2}}{2^{n-2}} \right) \right] = 2 + \frac{n-3}{\sqrt{2}}.$$

Since $\|\Gamma\| \geq \frac{1}{2^n} \langle \bar{1}| \Gamma |\bar{1}\rangle$, we get $\|\Gamma\| \geq 2 + \frac{n-3}{\sqrt{2}}$.

Computing denominator of Equation (19) Recall that G is the bit-flip group (see Theorem 3.6). Since Γ is a communication matrix of a XOR function, $\forall \sigma \in G, \Gamma = \sigma \Gamma \sigma^T$. Consider an arbitrary query $q = ([k], b)$ where $b \in \{-1,1\}^k$. Let $\sigma := \sigma_{b1^{n-k}}$ That is, σ is the permutation in G such that $\sigma(b1^{n-k}) = 1^n$.

$$\begin{aligned} \sigma(\Gamma \circ \Delta_q)\sigma^T &= (\sigma \Gamma \sigma^T) \circ (\sigma \Delta_q \sigma^T) = \Gamma \circ \Delta_{1^k}, \\ \implies \|\Gamma \circ \Delta_q\| &= \left\|\sigma(\Gamma \circ \Delta_q)\sigma^T\right\| = \|\Gamma \circ \Delta_{1^k}\|. \end{aligned}$$

Therefore, it is sufficient to bound the denominator for queries of the form $q = ([k], 1^k), \forall k \in [n]$.

Case $q = (\{1\}, 1)$ The matrix $\Gamma \circ \Delta_q$ is the following block matrix where each block is of size 2^{n-2}.

$$\Gamma \circ \Delta_q = \frac{1}{2^{n-2}} \left(\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}_{4\times 4} \otimes J_{2^{n-2}\times 2^{n-2}} \right). \tag{21}$$

Clearly, this is a rank 2 matrix with eigenvectors $(v, \bar{0}, \bar{0}, v)^T$ and $(v, \bar{0}, \bar{0}, -v)^T$ where $\bar{0}$ represents the all zero column vector of length 2^{n-2} and v is the all one column vector of length 2^{n-2} multiplied with $\frac{1}{2^{n-2}}$. The corresponding eigenvalues are $+1$ and -1 respectively.

Case $q = ([n], 1^n)$ The matrix $\Gamma \circ \Delta_q$ is the following.

$$\Gamma \circ \Delta_q = \frac{1}{2^{n-2}} \begin{bmatrix} 0 & \cdots & 0 & A \\ \vdots & \ddots & \vdots & 0_{1\times 2^{n-1}} \\ 0 & \cdots & 0 & 0 \\ A^T & 0_{2^{n-1}\times 1} & 0 & 0_{2^{n-1}\times 2^{n-1}} \end{bmatrix}, \tag{22}$$

where A is the row vector defined as

$$A = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}, \underbrace{\frac{1}{4\sqrt{2}}, \ldots, \frac{1}{4\sqrt{2}}}_{2^2}, \ldots, \underbrace{\frac{1}{2^{n-3}\sqrt{2}}, \ldots, \frac{1}{2^{n-3}\sqrt{2}}}_{2^{n-3}}, \underbrace{\frac{1}{2^{n-2}}, \ldots, \frac{1}{2^{n-2}}}_{2^{n-2}}).$$

Once again, $\Gamma \circ \Delta_q$ is a rank 2 matrix. The eigenvectors are the following with the appropriate normalization factors.

$$(\pm 1, 0, \ldots, 0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}, \underbrace{\frac{1}{4\sqrt{2}}, \ldots, \frac{1}{4\sqrt{2}}}_{2^2}, \ldots, \underbrace{\frac{1}{2^{n-3}\sqrt{2}}, \ldots, \frac{1}{2^{n-3}\sqrt{2}}}_{2^{n-3}}, \underbrace{\frac{1}{2^{n-2}}, \ldots, \frac{1}{2^{n-2}}}_{2^{n-2}}). \quad (23)$$

The corresponding eigenvalues are ± 1 respectively.

Case $q = ([k], 1^k), k \in [2, n-1]$ The corresponding matrix $\Gamma \circ \Delta_q$ is of the form $\begin{bmatrix} 0 & B \\ B & 0 \end{bmatrix}$ where B is a $2^{n-1} \times 2^{n-1}$ symmetric matrix defined below. The eigenvectors of $\Gamma \circ \Delta_q$ are $(v, v)^T$ and $(v, -v)^T$ where v is any eigenvector of B and the eigenvalues are $\pm\lambda$ where λ is the eigenvalue of B corresponding to eigenvector v.

$$B = \begin{bmatrix} 0 & A_0 & A_1 & \ldots & A_{k-2} & A_{k-1} \\ A_0^T & & & & & \\ A_1^T & & & & & \\ \vdots & & & 0 & & \\ A_{k-2}^T & & & & & \\ A_{k-1}^T & & & & & \end{bmatrix}, \quad (24)$$

$$A_i = \frac{1}{2^{n-k-1+i}\sqrt{2}}[1]_{(2^{n-k-1})\times(2^{n-k-1+i})}, \qquad (\forall i \in \{0, 1, \ldots, k-2\})$$

$$A_{k-1} = \frac{1}{2^{n-2}}[1]_{(2^{n-k-1})\times(2^{n-2})}.$$

The matrix B is clearly a rank 2 matrix since each A_i is a constant matrix, and each A_i has the same height. The two eigenvectors of the matrix are the following with the appropriate normalization factor. The corresponding eigenvalues are ± 1. Therefore, the eigenvalues of $\Gamma \circ \Delta_q$ are also ± 1.

$$(\underbrace{\pm 1, \ldots, \pm 1}_{2^{n-k-1}}, \underbrace{\frac{1}{\sqrt{2}}, \ldots, \frac{1}{\sqrt{2}}}_{2^{n-k-1}}, \underbrace{\frac{1}{2\sqrt{2}}, \ldots, \frac{1}{2\sqrt{2}}}_{2^{n-k}}, \ldots, \underbrace{\frac{1}{2^{k-2}\sqrt{2}}, \ldots, \frac{1}{2^{k-2}\sqrt{2}}}_{2^{n-3}}, \underbrace{\frac{1}{2^{k-1}\sqrt{2}}, \ldots, \frac{1}{2^{k-1}\sqrt{2}}}_{2^{n-2}})^T.$$

For every prefix query q, the quantity $\|\Gamma \circ \Delta_q\|$ is always 1. Therefore, $\max_q \|\Gamma \circ \Delta_q\| = 1$.

Adversary bound Combining the lower bound for the numerator with the denominator value, we get the following.

$$\mathrm{ADV}^{\mathcal{Q},\pm}(\mathsf{Dict_n}) = \frac{\|\Gamma\|}{\max\limits_{\substack{S \in \mathcal{Q}, \\ b \in \{-1,1\}^S}} \|\Gamma \circ \Delta_{S,b}\|} \geq 2 + \frac{n-3}{\sqrt{2}}. \qquad \square$$

Theorem 2.3 and the naive upper bound of query prefixes of increasing size, learning 1 bit in each step, immediately yields the following corollary.

Corollary B.2. *Let n be a positive integer. Let $\mathcal{Q} = \{[1, i] \mid i \in [n]\}$. Then,*

$$\mathsf{Q}^{\mathcal{Q}}(\mathsf{Dict}_n) = \Theta(n).$$

Remark B.3. *We observe that when $\mathcal{Q} = \{[1, i] \mid i \in [n]\}$, computing parity is at least as hard as computing the last bit, more formally, $\mathsf{Q}^{\mathcal{Q}}(\oplus) \geq (1/2)\mathsf{Q}^{\mathcal{Q}}(\mathsf{Dict}_n)$: Given an algorithm for computing parity, first run this algorithm on the input string x and get output b, say. Next run this algorithm on the input string with the last bit deleted (note that prefix queries to this input are valid prefix queries to the original input) and get output b', say. It is easy to see that $x_n = b \oplus b'$.*