A Block-Sensitivity Lower Bound for Quantum Testing Hamming Distance

Marcos Villagra
Núcleo de Investigación y Desarrollo Tecnológico
Universidad Nacional de Asunción, Paraguay
mvillagra@pol.una.py

Abstract. The Gap-Hamming distance problem is the promise problem of deciding if the Hamming distance $h$ between two strings of length $n$ is greater than $a$ or less than $b$, where the gap $g = |a - b| \geq 1$ and $a$ and $b$ could depend on $n$. In this short note, we give a lower bound of $\Omega(\sqrt{n/g})$ on the quantum query complexity of computing the Gap-Hamming distance between two given strings of length $n$. The proof is a combinatorial argument based on block sensitivity and a reduction from a threshold function.

Keywords: quantum query complexity, gap-Hamming distance, block-sensitivity.

1 Introduction

A generalized definition of the Hamming distance is the following: given two strings $x$ and $y$, decide if the Hamming distance $h(x, y)$ is greater than $a$ or less than $b$, with the condition that $b < a$. Note that this definition gives a partial boolean function for the Hamming distance with a gap. There is a entire body of work on the computation of a particular case of this notion of Hamming distance in the decision tree and communication models known as the Gap-Hamming distance (GHD) problem, which asks to differentiate the cases $h(x, y) \leq n/2 - \sqrt{n}$ and $h(x, y) \geq n/2 + \sqrt{n}$ [8]. A lower bound on GHD implies a lower bound on the memory requirements of computing the number of distinct elements in a data stream [4]. Chakrabarti and Regev [3] give a tight lower bound of $\Omega(n)$; their proof was later improved by Vidick [7] and then by Sherstov [6]. For the Hamming distance with a gap of the form $n/2 \pm g$ for some given $g$, Chakrabarti and Regev also prove a tight lower bound of $\Omega(n^2/g^2)$. In the quantum setting, there is a communication protocol with cost $O(\sqrt{n} \log n)$ [4].

Suppose we are given oracle access to input strings $x$ and $y$. In this note, we prove a lower bound on the number of queries to a quantum oracle to compute the Gap-Hamming distance with an arbitrary gap, that is, for any given $g = a - b$.

Theorem 1. Let $x, y \in \{0, 1\}^n$ and $g = a - b$ with $0 \leq b < a \leq n$. Any quantum query algorithm for deciding if $h(x, y) \geq a$ or $h(x, y) \leq b$ with bounded-error, with the promise that one of the cases hold, makes at least $\Omega(\sqrt{n/g})$ quantum oracle queries.
The proof is a combinatorial argument based on block sensitivity. The key ingredient is a reduction from a a threshold function. A previous result of Nayak and Wu \[5\] implies a tight lower bound of \(\Omega(\sqrt{n/g})\); their proof, however, is based on the polynomial method of Beals \textit{et al.} \[1\] and it is highly involved. The proof presented here, even though it is not tight, is simpler and requires no heavy machinery from the theory of polynomials.

2 Proof of Theorem \[1\]

Let \(a, b\) be such that \(0 \leq b < a \leq n\). Define the partial boolean function \(\text{GapThr}_{a,b}\) on \(\{0,1\}^n\) as

\[
\text{GapThr}_{a,b}(x) = \begin{cases} 1 & \text{if } |x| \geq a \\ 0 & \text{if } |x| \leq b. \end{cases}
\]  

(1)

To compute \(\text{GapThr}_{a,b}\) for some input \(x\), it suffices to compute the Hamming distance between \(x\) and the all 0 string. Thus, a lower bound for Gap-Hamming distance follows from a lower bound for \(\text{GapThr}_{a,b}\).

Let \(f : \{0,1\}^n \rightarrow \{0,1\}\) be a function, \(x \in \{0,1\}^n\) and \(B \subseteq \{1,\ldots,n\}\) a set of indices called a block. Let \(x^B\) denote the string obtained from \(x\) by flipping the variables in \(B\). We say that \(f\) is \textit{sensitive} to \(B\) on \(x\) if \(f(x) \neq f(x^B)\). The block sensitivity \(bs_x(f)\) of \(f\) on \(x\) is the maximum number \(t\) for which there exist \(t\) disjoint sets of blocks \(B_1,\ldots,B_t\) such that \(f\) is sensitive to each \(B_i\) on \(x\). The block sensitivity \(bs(f)\) of \(f\) is the maximum of \(bs_x(f)\) over all \(x \in \{0,1\}^n\).

From Beals \textit{et al.} \[1\] we know that the square root of block sensitivity is a lower bound on the bounded-error quantum query complexity. Thus, Theorem \[1\] follows immediately from the lemma below.

**Lemma 2.** \(bs(\text{GapThr}_{a,b}) = \Theta(n/g)\).

**Proof.** Let \(x \in \{0,1\}^n\) be such that \(\text{GapThr}_{a,b}(x) = 0\) and suppose that \(|x| = b\). To obtain a 1-output from \(x\) we need to flip at least \(g = a - b\) bits of \(x\). Hence, we divide the \(n-b\) least significant bits of \(x\) in non-intersecting blocks, where each block flips exactly \(g\) bits. The number of blocks is \(\left\lceil \frac{a-b}{g} \right\rceil\), which is at most \(bs_x(\text{GapThr}_{a,b})\). To see that \(\left\lceil \frac{a-b}{g} \right\rceil\) is the maximum number of such non-intersecting blocks, consider what happens when the size of a block is different from \(g\). If the size of a block is less than \(g\), then we cannot obtain a 1-output from \(x\); if the size of a block is greater than \(g\), then the number of blocks decreases. Thus, we have that \(bs_x(\text{GapThr}_{a,b}) = \left\lceil \frac{a-b}{g} \right\rceil\).

For any \(x'\) with \(|x'| < b\), we need to flip \(a-b\) bits plus \(b - |x'|\) bits. Using our argument of the previous paragraph, the size of each block is thus \(g + b - |x'|\), and hence, \(bs_{x'}(\text{GapThr}_{a,b}) = \left\lceil \frac{n-|x'|}{g+b-|x'|} \right\rceil\). Note that \(bs_{x'}(\text{GapThr}_{a,b}) \leq bs_x(\text{GapThr}_{a,b})\).

For the case when \(\text{GapThr}_{a,b}(x) = 1\) and \(|x| = a\), to obtain a 0-output from \(x\) we need to flip at least \(g\) bits of \(x\). Hence the same argument applies, and thus, \(bs_x(\text{GapThr}_{a,b}) = \left\lceil \frac{a-b}{g} \right\rceil\).
Taking the maximum between the cases when $|x| = b$ and $|x| = a$, we have that $bs(GapThr_{a,b}) = \max\{ (n-b)/g, (n-a)/g \} = \Theta(n/g)$. □

References

1. Beals, R., Buhrman, H., Cleve, R., Mosca, M., de Wolf, R.: Quantum lower bounds by polynomials. Journal of the ACM 48(4) (2001)

2. Buhrman, H., Cleve, R., Wigderson, A.: Quantum vs. classical communication and computation. In: Proceedings of the 30th annual ACM Symposium on Theory of Computing (STOC). pp. 63–68. ACM Press, New York, New York, USA (1998)

3. Chakrabarti, A., Regev, O.: An optimal lower bound on the communication complexity of gap-hamming-distance. In: Proceedings of the 43rd ACM Symposium on Theory of Computing (STOC) (2011)

4. Indyk, P., Woodruff, D.: Tight lower bounds for the distinct elements problem. In: Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (FOCS). pp. 283–288. IEEE Computer. Soc (2003)

5. Nayak, A., Wu, F.: The quantum query complexity of approximating the median and related statistics. In: Proceedings of the 31st annual ACM symposium on Theory of computing (STOC). pp. 384–393. ACM (1999)

6. Sherstov, A.: The Communication Complexity of Gap Hamming Distance. In: Electronic Colloquium on Computational Complexity, Report TR11-063 (2011)

7. Vidick, T.: A concentration inequality for the overlap of a vector on a large set, with application to the communication complexity of the gap-hamming-distance problem. In: Electronic Colloquium on Computational Complexity, Report TR11-051 (2010), \texttt{http://eccc.hpi-web.de/report/2011/051/}

8. Woodruff, D.: Efficient and Private Distance Approximation in the Communication and Streaming Models. Ph.D. thesis, MIT (2007)