THEORY OF COMPUTING, Volume 8 (8C1), 2013, pp. 8C1:1–3
www.theoryofcomputing.org

COMMENT

Comparing three proofs on the randomized communication complexity of Hamming distance

Thomas Vidick

Received February 3, 2013; Published February 26, 2013

Abstract: We compare the proofs by Chakrabarti and Regev (STOC 2011/SICOMP 2012), Sherstov (Theory of Computing 2012), and the commenter (Chicago J. Theor. Comp. Sci. 2012) of a linear lower bound on the randomized communication complexity of the gap Hamming distance problem.

Concurrently with and independently of Sherstov's work [2] published in *Theory of Computing*, the author of this comment obtained a separate proof of the same main result [3], a linear lower bound on the randomized communication complexity of the gap Hamming distance (GHD) problem (available in CJTCS).

Both Sherstov's and the author's proofs build on the work of Chakrabarti and Regev [1], who were the first to prove a linear lower bound. The three papers follow a similar path but differ in interesting ways. Although Sherstov's paper already contains a comparison, we take the opportunity of this comment

ACM Classification: F.1.3

AMS Classification: 68Q17, 68Q25

Key words and phrases: communication complexity, gap Hamming distance

DOI: 10.4086/toc.2013.v008a008c1

to briefly highlight the main differences between the three approaches. The interested reader will find additional details in either of the three papers.

As already explained by Sherstov (Section 1.2 in [2]), all three proofs have at their heart an anti-concentration statement for the inner product of a pair (x, y) of unit vectors chosen uniformly at random from subsets A, B of $\mathbb{R}^n$ of large enough measure. The proofs differ in the precise anti-concentration statement and how it is obtained, as well as how it is used to derive the communication lower bound on GHD.

First, the anti-concentration statements made in [1] and [3] apply to subsets of the n-dimensional Euclidean space equipped with the Gaussian measure, while in [2] it applies to subsets of the Hamming cube (itself viewed as a subset of the n-dimensional Euclidean space) equipped with the uniform measure.

A second difference arises in the way anti-concentration is proved. In [1] and [2] the key idea is to show that any set A of large enough measure, at least 2^{-cn} for some small enough constant $c > 0$, contains a linear (in n) number of almost-orthogonal vectors (Lemma 3.1 in [2]); an anti-concentration statement then follows without too much difficulty (see Lemma 3.2 in [2]). In [3] a different argument is used. The set A is represented as a positive semidefinite matrix $\mathbf{A} = \mathrm{E}_{x \in A}[xx^T]$. Intuitively, the spectrum of $\mathbf{A}$ describes how "pointy" the set A is. The anti-concentration result (Theorem 1.1 in [3]) follows from showing that provided A is large enough $\mathbf{A}$ must have a linear number of eigenvalues of constant magnitude (Claim 3.3 in [3]). The proof therefore does not require singling out specific vectors in A, which would in a way break the symmetry of the problem. Identifying such vectors is the more technically involved part of the proofs in [1], where they are obtained by applying a lemma of Raz (Lemma 3.4 in [1]), and in [2], where their existence is shown using a result of Talagrand (Theorem 5.1 in [2]).

Finally, Sherstov introduces a nice simplification to obtain the communication lower bound on GHD from the anti-concentration bound: by considering a related problem, "gap orthogonality," whose complexity is easily shown to be equivalent to that of GHD, he makes it possible to use Yao's corruption bound instead of the partition bound by Jain and Klauck used in [1, 3].

References

[1] AMIT CHAKRABARTI AND ODED REGEV: An optimal lower bound on the communication complexity of gap-Hamming-distance. *SIAM Journal on Computing*, 41(5):1299–1317, 2012. Preliminary version at STOC'11. [doi:10.1137/120861072,] 1, 2

[2] ALEXANDER A. SHERSTOV: The communication complexity of gap Hamming distance. *Theory of Computing*, 8(8):197–208, 2012. [doi:10.4086/toc.2012.v008a008] 1, 2

[3] THOMAS VIDICK: 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. *Chicago Journal of Theoretical Computer Science*, 2012(1), July 2012. Preprint: ECCC, TR11-051. [doi:10.4086/cjtcs.2012.001] 1, 2

AUTHOR

Thomas Vidick
CSAIL
MIT
vidick@csail.mit.edu
http://people.csail.mit.edu/vidick

Source files