On Tight Uncertainty Principle for Biased Compressions

Alyazeed Basyoni

February 2020

1 Introduction

In this work, we aim to prove a tighter version of the Uncertainty Principle for communication complexity of a biased compression operator $\mathcal{C} \in \mathbb{B}(\alpha)$ for $\alpha < 1$. This is the the operators $\mathcal{C} : \mathbb{R}^d \to \mathbb{R}^d$ such that for any $x \in \mathbb{R}^d$, one has $\mathbb{E}\left[\|\mathcal{C}(x) - x\|^2\right] \leq \alpha \|x\|^2$. The goal of this work is to analyze the difficulty of compressing operators in $\mathbb{B}(\alpha)$. In particular, we uncover a lower bound on the number of bits needed to communicate $\mathcal{C}(x)$, where $\mathcal{C} \in \mathbb{B}(\alpha)$. In addition, we show the tightness of this lower bound by designing an operator requiring very few bits of excess communication in expectation and with very high probability. We start off with the following lemma.

Lemma 1.1 *If $\alpha < 1$ and $\mathcal{C} \in \mathbb{B}(\alpha)$ then $\mathcal{C}$ assumes infinitely many values.*

Proof: Otherwise, assume the values are $c_1, c_2, ..., c_N$, one can choose x such that $\|x\| < \frac{\min_i \|c_i\|}{(\sqrt{\alpha}+1)}$ or $\|x\| > \frac{\max_i \|c_i\|}{(1-\sqrt{\alpha})}$ to ensure that for all i, one has $\|c_i - x\| > \sqrt{\alpha} \|x\|$, which is against that $\mathcal{C} \in \mathbb{B}(\alpha)$.

From this, we can see that one needs to have an additional constrain on space where $\mathcal{C}$ is acting. Since it is easy to communicate $\|x\|$ up to reasonable precision in finitely many bits, we can assume that $\|x\| = 1$. In this case, we redefine $\mathbb{B}(\alpha)$ as the set of biased compression operators $\mathcal{C} : \mathbb{R}^d \to \mathbb{R}^d$ such that: $\mathbb{E}\left[\|\mathcal{C}(x) - x\|^2\right] \leq \alpha$ for all $x \in \mathbb{R}^d, \|x\| = 1$. Under this constrain, it is possible to ensure that $\mathcal{C}$ takes finitely many values, and hence can be communicated with finitely many bits. This leads us to our second and third lemmas.

Lemma 1.2 *Assume that $\mathcal{C}$ takes only finitely many values $c_1, ..., c_N$, then the spheres centered at $c_1, ..., c_N$ with radii $\sqrt{\alpha}$ cover the boundary of the unit sphere completely.*

Proof: Otherwise, let x be such that $\|x\| = 1$ and x is not covered by those spheres. Then $\|c_i - x\| > \sqrt{\alpha}$ which is against that $\mathcal{C} \in \mathbb{B}(\alpha)$.

Lemma 1.3 *Let $c \in \mathbb{R}^d$, the ball centered at c with radius $\sqrt{\alpha}$ covers at most $P(\alpha)$ fraction of the surface of the unit sphere, where:*
$P(\alpha) = Pr(x_1 \geq \sqrt{1 - \alpha})$, where $(x_1, ..., x_d)$ is chosen uniformly at random on the unit sphere.

Proof: Without loss of generality, assume that $c = e_1$. In order to maximize the covered fraction of the unit sphere, one needs to maximize $\sup(\|v_1 - v_2\|)$ where $v_!, v_2$ are on the unit sphere and within a radius of $\sqrt{\alpha}$ of c. Clearly, this quantity is bounded above by $2\sqrt{\alpha}$, and equality is achieve when c is the midpoint, in which case it is to check that $\|c\| = \sqrt{1 - \alpha}$, or that $c = \sqrt{1 - \alpha}e_1$. In this case, the fraction of the surface covered by the ball centered at c with radius $\sqrt{\alpha}$ is $P(\alpha) = Pr(x_1 >= \sqrt{1 - \alpha})$ with $(x_!, ..., x_d)$ chosen uniformly on the surface of the unit sphere.

From all the above results, one can finally arrive at the following uncertainty principle:

Lemma 1.4 *Let $\mathcal{C} \in \mathbb{B}(\alpha)$. If $\mathcal{C}(x)$ can always be communicated in b bits, then one has that $2^b P(\alpha) \geq 1$*

Proof: If one can encode $\mathcal{C}(x)$ with b bits, that means $\mathcal{C}(x)$ can assume at most 2^b values. Nonetheless, since the balls centered at these values with radii $\sqrt{\alpha}$ cover the surface of the unit sphere, and each one covers at most a small fraction of $P(\alpha)$, we must have $2^b P(\alpha) \geq 1$, as desired.

Next, we show that this bound is tight.

Lemma 1.5 *There exists a compression operator $\mathcal{C} \in \mathbb{B}(\alpha)$ along with a communication protocol such that the expected number of communicated bits b satisfies the following:*
1) $\mathbb{E}[b] \leq 1 + \log \frac{1}{P(\alpha)}$
2) $\mathbb{E}[2^b] \leq \frac{2}{P(\alpha)}$
3) $Pr(b \geq K + 1 + \log \frac{1}{P(\alpha)}) \leq 2^{-k}$ for any $K \in \mathbb{N}$

Proof: Let $x \in \mathbb{R}^d$ such that $\|x\| = 1$. Let the compressing node do the following: keep generating point $c \in \mathbb{R}^d$ such that $\|c\| = \sqrt{1 - \alpha}$, uniformly at random, until $\|c - x\| \leq \sqrt{\alpha}$. Let τ be the stopping time for this experiment. The compressing node communicates τ to the server. The server starts from the same random seed as the compressing node (the random seed should be shared ahead of time), runs the same experiment τ times to arrive at the same c, which is, by definition, within $\sqrt{\alpha}$ of x. Since each ball centered at c with radius $\sqrt{\alpha}$ covers exactly $P(\alpha)$ fraction of the surface of the unit sphere, the probability that a random c contains x is $P(\alpha)$. Hence, the expected time to get c close enough to x is $\mathbb{E}[\tau] = \frac{1}{P(\alpha)}$. Since $\tau \leq 2^b \leq 2\tau$, we get $\mathbb{E}[2^b] \leq \frac{2}{P(\alpha)}$. Moreover, since log is a concave function, we get: $\mathbb{E}[b] \leq \mathbb{E}[\log 2\tau] \leq \log \mathbb{E}[2\tau] = \log \frac{2}{P(\alpha)} = 1 + \log \frac{1}{P(\alpha)}$. Finally, the probability of sending K or more additional

bits, $Pr(b \geq K + 1 + log(\frac{1}{P(\alpha)})) = Pr(2^b \geq \frac{2^{K+1}}{P(\alpha)}) \leq Pr(2\tau \geq \frac{2^{K+1}}{P(\alpha)}) \leq Pr(\tau \geq \frac{2^K}{P(\alpha)}) = Pr(\tau \geq 2^K \mathbb{E}[\tau])$ which, by Markov inequality, is bounded above by 2^{-k}, as desired.