Electronic Colloquium on Computational Complexity, Report No. 38 (2011)

On the query complexity for Showing Dense Model

Jiapeng Zhang

Shanghai Jiao Tong University, Shanghai 200030, China
jpeng.zhang@gmail.com

Abstract. A theorem of Green, Tao, and Ziegler can be stated as follows: if R is a pseudorandom distribution, and D is a dense distribution of R, then D can be modeled as a distribution M which is dense in uniform distribution such that D and M are indistinguishable. The reduction involved in the proof has exponential loss in the distinguishing probability. Reingold et al give a new proof of the theorem with polynomial loss in the distinguishing probability. In this paper, we are focus on query complexity for showing dense model, and then give a optimal bound of the query complexity. We also follow the connection between Impagliazzo's Hardcore Theorem and Tao's Regularity lemma, and obtain a proof of L_2-norm version Hardcore Theorem via Regularity lemma.

Keywords: pseudorandomness, regularity lemma, query complexity, on-line learning algorithm

ISSN 1433-8092

1 Introduction

Green and Tao[GT] have proved that the primes contains arbitrarily long arithmetic progression. To prove this theorem, a key result is the following Dense Model Theorem,

Theorem 1 (informal). *Let R be a pseudorandom set of integers and D be a subset of R with constant density in R. Then there is a set M that has constant density in the integers and is indistinguishable from D.*

Tao and Ziegler[TZ] have proved such a result in board generality. It not only consider the pseudorandom set of integers, but also consider other domains, like $\{0,1\}^n$. Roughly speaking, they indicates that if R is a pseudorandom distribution on X, then every δ-dense distribution D in R is indistinguishable from some distribution which is $\delta/2$-dense in the uniform distribution on X, where X is an arbitrary finite universal. This result seems applicable for both complexity theory and cryptography. However, the reduction implicit in their has exponential loss in the distinguishing probability, making it inapplicable.

Reingold, Trevisan, Tulsiani and Vadhan[$RTTV$] have introduced the Dense Model Theorem into complexity-theoretic. Means in that paper, a quantitatively improved characterization was obtained using an argument based on Nisan's proof of the Impagliazzo's Hardcore Theorem[Imp], i.e., in their proof, the reduction has polynomial loss in the distinguishing probability.

It seems that Dense Model Theorem is dual with Hardcore Theorem which indicates that if f is a δ-hard function, then it is extreme hard in a δ-dense measure. Trevisan, Tulsiani and Vadhan[TTV] give a decomposition theorem that show strong connections between Hardcore Theorem, Dense Model Theorem and Weak Graph Regularity lemma of Frieze and Kannan[FK].

Similar as Hardcore Theorem, we will consider the query complexity of the reduction which showing dense model. In this paper, we will provide a different reduction to prove Dense Model Theorem, in which has query complexity better than $[RTTV], [TTV]$. Our reduction is inspired by $[BHK]$'s proof of Hardcore Theorem. And in further, we will prove that, the query complexity of our reduction has touched the optimal bound with constant factor(in black-box reduction). The optimal bound is same as the optimal bound in query complexity of hard-core set constructions$[BHK], [KS], [LTW]$, and also is same as the optimal bound in query complexity of reductions which showing hardness amplification$[SV], [Imp]$. We also interesting in the connections between Tao's[$Tao1$] arithmetic version of regularity lemma and Hardcore Theorem, and give a proof for a L_2-norm version of Hardcore Theorem.

Tao[$Tao1, Tao2, Tao3$] has developed series of regularity lemmas. All of them are structure theorems in different perspectives, i.e, arithmetic-theoretic perspective, information-theoretic perspective, graph-theoretic perspective, and so on. In tuition, all of this theorems are relative, and in this viewpoint, both Hardcore Theorem and Dense Model Theorem are special perspectives of the structure theorems.

1.1 Dense Model Theorem

Let us first recall some definitions in complexity theory. We have a finite universal X, for example $\{0,1\}^s$, then we will always consider the distributions, measures on X.

A *measure* on the set X is a function $M : X \to [0,1]$. We let $|M| = \sum_{x\in X} M(x)$ denote the absolute size of M and $\mu(M) = |M|/|X|$ denote its density(relative size). The distribution D_M induced by M is defined by $D_M(x) = M(x)/|M|$.

Let S be a subset of X, we always treat it as a measure on X, i.e., $S(x)$ is equal to 1 when $x \in S$, and equal to 0 otherwise. We use D_S to denote the uniform distribution over S. In particular, we use D_X to denote uniform distribution over X.

We say that a measure M(or a set S) is δ-dense if $\mu(M) \geq \delta$(or $\mu(S) \geq \delta$, respectively). And we say that a distribution D is δ-dense in a distribution R if $\Pr[D = x] \leq \frac{1}{\delta}\Pr[R = x]$, for all $x \in X$. In particular, D is δ-dense in the uniform distribution if and only if D is induced by some δ-dense measure.

Let $\mathcal{F} = \{g_1, g_2, \cdots, g_k\}$ be a finite collection of bounded functions $g_i : X \to [0,1]$. Let $I = \{i_1, \cdots, i_q\}$ be a subset of $[k]$, let g_I denote the function such that $g_I(x) = (g_{i_1}(x), \cdots, g_{i_q}(x))$.

We say that a distribution R on X is ϵ-pseudorandom for $\mathcal{F}$ if for every function $f \in \mathcal{F}$ we have that

$$\left|\mathbb{E}[f(D_X)] - \mathbb{E}[f(R)]\right| \leq \epsilon,$$

i.e., $\mathcal{F}$ can't distinguish R from uniform distribution.

In the paper, we will always consider the parameters, i.e., ϵ, δ, q and so on, are functions of $|X|$, and we use $f = O(g)$ to denote the quantity bounded by $c \cdot g$, where c is a constant.

Definition 1. *Let X be a finite universal. We say that a distribution D on X has $(\delta, \epsilon, \mathcal{F})$-model if for some distribution D_1 that is δ-dense in the uniform distribution, it has*

$$\left|\mathbb{E}[f(D_1)] - \mathbb{E}[f(D)]\right| \leq \epsilon,$$

for all $f \in \mathcal{F}$.

Roughly speaking, D has a $(\delta, \epsilon, \mathcal{F})$-model means that D looks like a δ-dense distribution in the uniform distribution.

Definition 2. *Let X be a finite universal. A black-box (q, ϵ, δ, a)-reduction showing dense model for $\mathcal{F}$ is an oracle algorithm $Dec^{(\cdot)}(\cdot,\cdot) : X \times \{0,1\}^a \to \{0,1\}$. It is required that*

(i) black-box: *there is a function C such that $C(x, g_I(x), \alpha) = Dec^{\mathcal{F},I}(x, \alpha)$ for each $I \subseteq [|\mathcal{F}|]$ with $|I| = q$ and $\alpha \in \{0,1\}^a$;*

(ii) showing dense model: *for every distribution D on X which* doesn't *have $(\delta, \epsilon, \mathcal{F})$-model, there exists a string $\alpha \in \{0,1\}^a$ and a subset $I \subseteq [\mathcal{F}]$ with $|I| = q$ such that for every distribution R with D δ-dense in R, the function $f(x) = Dec^{\mathcal{F},I}(x, \alpha)$ distinguishes R and D_X, i.e.,*

$$\left|\mathbb{E}[f(D_X)] - \mathbb{E}[f(R)]\right| \geq c\epsilon\delta,$$

where c is a universal constant, for example 0.01.

We call q the query complexity of the reduction.

Remark 1. One may has another definition for the reduction Dec where the non-uniform advices I and α may depend on R, i.e. they define the reduction Dec that showing dense model with non-uniform on D and R. In our definition, the reductions have non-uniform advices which only depend on D.

Dense Model Theorem mainly indicates that the reduction Dec exists. Notice that the distinguishing probability $c\epsilon\delta$ can not be much better, i.e., with constant factor. For example, let $S_1 \subseteq S_2 \subseteq X$ with $|S_1| = \delta(1-\epsilon)|X|$, $|S_2| = \delta|X|$, and let g be the character function for S_1. Then D_{S_1} doesn't have $(\delta, \epsilon, \{g\})$-model, and there is no function can distinguish $R = \delta D_{S_1} + (1-\delta)D_{X\setminus S_2}$ from D_X better than $\epsilon\delta$.

In $[RTTV]$'s proof for Dense Model Theorem, they provided a reduction Dec with query complexity $q = O(\log(\frac{1}{\epsilon\delta})\frac{1}{\epsilon^2})$, i.e., a $(O(\log(\frac{1}{\epsilon\delta})\frac{1}{\epsilon^2}), \epsilon, \delta, poly(1/\epsilon, 1/\delta))$-reduction which showing dense model for arbitrary finite $\mathcal{F}$.

Inspired by $[BHK]$, we provide a $(O(\log(\frac{1}{\delta})\frac{1}{\epsilon^2}), \epsilon, \delta, O(\log(\frac{1}{\delta})\frac{1}{\epsilon^2}))$-reduction which showing dense model for arbitrary finite $\mathcal{F}$, and in further, we will prove that $q = O(\log(\frac{1}{\delta})\frac{1}{\epsilon^2})$ is optimal, and it's the same as Hardcore Lemma.

1.2 Tao's regularity lemma, and a L_2-norm version of Hardcore Theorem

In Tao's arithmetic perspective of regularity lemma[$Tao1$], it is studying lies in some real finite-dimension Hilbert space. Let H be a real Hilbert space, $S \subseteq H$ is a finite collection of ″basic structured″vectors with bounded length, i.e., $\|v\| \leq 1$ for all $v \in S$.

Then, given $f \in H$, we say that f is (M, K)-structured for some $M, K > 0$ if it has a decomposition

$$f = \sum_{1 \leq i \leq M} c_i v_i$$

with $v_i \in S$ and $c_i \in [-K, K]$ for all $1 \leq i \leq M$. We say that f is ϵ-pseudorandom for some $\epsilon > 0$ if for all $v \in S$, we have $|\langle f, v\rangle| \leq \epsilon$.

Remark 2. In (M, K)-structured, we notice that S is correspond to $\mathcal{F}$ in Dense Model Theorem, and M, K are corresponding to query complexity and the non-uniform advise α respectively.

Tao's regularity lemma shows that it often has a *dichotomy* between structure and pseudorandomness.

Theorem 2. [*Tao*1] *Let H, S be as above. Let $f \in H$ be such that $\|f\| \leq 1$, and let $0 < \epsilon \leq 1$. Then there exists a decomposition*

$$f = f_{str} + f_{psd}$$

such that f_{str} is $(1/\epsilon^2, 1/\epsilon)$-structured, f_{psd} is ϵ-pseudorandom.

On the other hand, in Hardcore Theorem, it's always considering the hardness of functions. In this paper, we consider **Boolean Circuits which output 1 or -1**. Let $f : X \to \{-1, 1\}$ be a function, and C is a Boolean circuit, the advantage of C on computing f is defined as

$$Adv_C(f) := \mathbb{E}[C(D_X)f(D_X)] = \sum_x C(x)f(x)/|X|,$$

i.e., if $C(x) = f(x)$, it will contribute $1/|X|$, and will contribute $-1/|X|$ otherwise. And we say that $Adv_s(f) \leq \epsilon$ if $Adv_C(f) \leq \epsilon$ for every circuit C with size s.

Let M be a measure on X, we define $Adv_C^M(f) := \mathbb{E}[C(D_M)f(D_M)]$. We call f ϵ-hard-core on M for size s, if $Adv_s^M(f) \leq \epsilon$. Hardcore Theorem mainly indicates that, every mildly hard function f has a ϵ-hard-core. Formally,

Theorem 3 (Hardcore Theorem). [*Imp*], [*BHK*], [*KS*] *Let $0 < \delta, \epsilon < 1$ be parameters, and let $f : X \to \{-1, 1\}$ be a function with $Adv_s(f) \leq 1 - 2\delta$. Then there is a measure M with $\mu(M) \geq c\delta$ so that $Adv_{s'}^M(f) \leq \epsilon$, where $s' = O(s\epsilon^2/\log(1/\delta))$ and c is an universal constant.*

In our result, we will give a proof for L_2-version of Hardcore Theorem via regularity lemma.

2 Black-Box Construction of Dense Model Distribution via Bregman Projections

In this section, we will prove the following Dense Model Theorem,

Theorem 4 (Dense Model Theorem). *Let X be a finite universe, $\mathcal{F}$ a collection of bounded functions $f : X \to [0, 1]$. Let $0 < \epsilon, \delta < 1$ be parameters, D a distribution over X. Suppose for every distribution D_δ that is δ-dense in D_X there is a function $g \in \mathcal{F}$ such that*

$$\left|\mathbb{E}[g(D_\delta)] - \mathbb{E}[g(D)]\right| \geq \epsilon,$$

i.e., D doesn't have $(\delta, \epsilon, \mathcal{F})$-model. Then there are functions $g_1, \ldots, g_T \in \mathcal{F}$, and parameters $a_1, \cdots, a_T \in \{-1, +1\}$ with $T = O((1/\epsilon^2) \cdot \log(1/\delta))$, and $t_0 \in [-T, T] \cap \mathbb{Z}$ such that if we define $h : X \to \{0, 1\}$ by

$$h(x) = 1 \Leftrightarrow \sum_i a_i g_i(x) \geq t_0,$$

then for every distribution R with D δ-dense in R,

$$\left|\mathbb{E}[h(D_X)] - \mathbb{E}[h(R)]\right| \geq \Omega(\epsilon\delta).$$

Remark 3. There are two parts of non-uniform advises and one part of oracle advices above, i.e., the parameters $(a_i)_{i\in[T]}$ and the threshold t_0 are the non-uniform advises, and $(g_i)_{i\in[T]}$ are the oracle advices respectively.

We can encode the non-uniform advice by a string $\alpha \in \{0,1\}^{2T}$, thus we have a $(O(\log(\frac{1}{\delta})\frac{1}{\epsilon^2}), \epsilon, \delta, O(\log(\frac{1}{\delta})\frac{1}{\epsilon^2}))$-reduction which showing dense model for arbitrary $\mathcal{F}$.

2.1 Preparations

In $[BHK]$, they provided an algorithm based on the technology as Freund and Schapire's$[FS]$ well known *AdaBoost* algorithm. And our algorithm is similar as $[BHK]$'s algorithm.

Let X be a finite set, M and N are measures on X. The *Kullback-Leibler divergence* between M and N is defined as

$$D(M\|N) = \sum_{x\in X} M(x)\log\left(\frac{M(x)}{N(x)}\right) + N(x) - M(x).$$

In further, let $\Gamma \subseteq \mathbb{R}^{|X|}$ be a non-empty closed convex set of measures. Then the *Bregman projection* of N onto Γ is defined as the measure $P_\Gamma N \in \Gamma$ such that

$$D(P_\Gamma N\|N) \leq D(M\|N)$$

for all $M \in \Gamma$, i.e., with minimized distance.

The definition above is well-defined, since one can show that for every N, the minimized $P_\Gamma N$ exists and is unique via the following theorem.$[CZ]$

Theorem 5 (Bregman). *Let N, M be measures such that $M \in \Gamma$. Then,*

$$D(M\|P_\Gamma N) + D(P_\Gamma N\|N) \leq D(M\|N).$$

Let $\Gamma_\delta := \{M|\mu(M) \geq \delta\}$, i.e., Γ_δ are the δ-dense measures. We will denote the Bregman projection onto the set Γ_δ by P_δ. One can show that, for every measure N with support at least $\delta|X|$ and $\mu(N) < \delta$, then $\mu(P_\delta N) = \delta$.

Lemma 1. $[BHK]$ *Let N be a measure with support at least $\delta|X|$ and let $c \geq 1$ be the smallest constant such that the measure $M^* = \min(1, c\cdot N)$ has density δ. Then, $P_\delta N = M^*$.*

Then we will consider the standard model of online algorithm. Let penalty be the vectors $\mathbf{m} = (m_x)_{x\in X}$ with $m_x \in [0,1]$ for each $x \in X$. Let M be a measure on X, we set the loss function $L(M,\mathbf{m}) = \sum_{x\in X} M(x)m_x$. Similar as $[BHK]$, we have the following lemma, and the proof is omitted here.

Lemma 2. *Let Γ be a closed convex set of measures. Let $M^{(1)} \in \Gamma$ be an arbitrary initial measure, and let $\mathbf{m}^{(t)}, t \in [T]$ be arbitrary penalties. We define $N^{(t+1)}$ be the measure with that $N^{(t+1)}(x) = (1-\epsilon/4)^{m_x^{(t)}} M^{(t)}(x)$, and let $M^{(t+1)} := P_\Gamma N^{(t+1)}$. Then for every measure $M \in \Gamma$, we have*

$$\sum_{t=1}^{T} L(M^{(t)}, \mathbf{m}^{(t)}) \leq \left(1+\frac{\epsilon}{4}\right) \sum_{t=1}^{T} L(M, \mathbf{m}^{(t)}) + 4 \cdot \frac{D(M\|M^{(1)})}{\epsilon}.$$

2.2 Proof of the Dense Model Theorem

In this section, we will prove the Dense Model Theorem via the *On-line Learning* algorithm.

Proof. To prove the theorem, we will iterate the following processes for $T = \frac{16}{\epsilon^2} \log \frac{1}{\delta}$ rounds, and in each round, we make sure that $M^{(t)}$ with support at least $\delta|X|$ and $\mu(M^{(t)}) = \delta$.

- Step 0. Let $t = 1$, and let $M^{(1)}$ be the initial measure that is δ at every point. Note that $\mu(M^{(1)}) = \delta$.
- Step 1. Since $M^{(t)} \in \Gamma_\delta$, $D_{M^{(t)}} = M^{(t)}/|M^{(t)}|$ is a δ-dense distribution in D_X, then by the assumption of D, we have a function $g_t \in \mathcal{F}$ such that

$$\left|\mathbb{E}[g_t(D_{M^{(t)}})] - \mathbb{E}[g_t(D)]\right| \geq \epsilon.$$

- Step 2. There are possible cases in this step.
  Case 1. $\mathbb{E}[g_t(D_{M^{(t)}})] - \mathbb{E}[g_t(D)] \geq \epsilon$, then set $a_t = 1$, and define $\mathbf{m}^{(t)}$ by putting $m_x^{(t)} := g_t(x)$;
  Case 2. $\mathbb{E}[g_t(D_{M^{(t)}})] - \mathbb{E}[g_t(D)] \leq -\epsilon$, then set $a_t = -1$, and define $\mathbf{m}^{(t)}$ by putting $m_x^{(t)} := 1 - g_t(x)$.

- Step 3. Define $N^{(t+1)}$ by setting $N^{(t+1)}(x) := (1-\epsilon/4)^{m_x^{(t)}} M^{(t)}(x)$, and let $M^{(t+1)} := P_\delta N^{(t+1)}$.
- Step 4. Set $t := t+1$, and return to Step 1.

Define $k(x) := \sum_{t=1}^{T} a_t g_t(x)$, one may hope that k learns D well, i.e., k distinguishes every δ-dense subset S from D,

Claim. Let S be an arbitrary subset of X with $|S| = \delta|X|$. Then,

$$\left(1+\frac{\epsilon}{4}\right) \mathbb{E}[k(D_S)] \geq \mathbb{E}[k(D)] + \frac{\epsilon}{2} T.$$

Proof. By the construction of g_t and a_t, we have that

$$\sum_{t=1}^{T} \mathbb{E}[a_t g_t(D_{M^{(t)}})] - \sum_{t=1}^{T} \mathbb{E}[a_t g_t(D)] \geq \epsilon T \tag{1}$$

Also, apply Lemma 2 with $M = U_S$, it has

$$\sum_{t=1}^{T}\sum_{x\in X} M^{(t)}(x)m^{(t)}(x) \le \left(1+\frac{\epsilon}{4}\right)\sum_{t=1}^{T}\sum_{x\in X} S(x)m^{(t)}(x) + 4\cdot\frac{D(U_S\|M^{(1)})}{\epsilon}.$$

Then by definitions, $\mathbb{E}[a_t g_t(D_{M^{(t)}})] = \sum_x a_t g_t(x) M^{(t)}(x)/|M|$, thus

$$\sum_{t=1}^{T}\mathbb{E}[a_t g_t(D_{M^{(t)}})] \le \left(1+\frac{\epsilon}{4}\right)\sum_{t=1}^{T}\mathbb{E}[a_t g_t(D_S)] + \frac{\epsilon}{4}T + 4\cdot\frac{D(U_S\|M^{(1)})}{\epsilon\delta|X|}. \quad (2)$$

Also,

$$D(U_S\|M^{(1)}) = \sum_{x\in S}\log(1/M^{(1)}(x)) + |M^{(1)}| - |U_S| = \delta|X|\log(1/\delta).$$

Combined with Eqs. (1) and (2),

$$\epsilon T \le \left(1+\frac{\epsilon}{4}\right)\sum_{t=1}^{T}\mathbb{E}[a_t g_t(D_S)] + \frac{\epsilon}{4}T + 4\cdot\frac{1}{\epsilon}\log\frac{1}{\delta} - \sum_{t=1}^{T}\mathbb{E}[a_t g_t(D)].$$

The claim then follows since $T = \frac{16}{\epsilon^2}\log\frac{1}{\delta}$. □

Note that $\left|\mathbb{E}[k(D_S)]\right| \le T$, thus we have

$$\mathbb{E}[k(D_S)] \ge \mathbb{E}[k(D)] + \frac{\epsilon}{4}T.$$

We then show that D and D_S can be distinguished via a Boolean function, i.e., find the threshold t_0.

Lemma 3. *[RTTV]Let $F : X \to [0, 2T]$ be a bounded function, let D_Z and D_W be distributions such that $\mathbb{E}[F(D_Z)] \ge \mathbb{E}[F(D_W)] + (\epsilon/4)T$. Then there exists $t \in [0, 2T]$ such that*

$$\Pr[F(D_W) \ge t - \frac{\epsilon}{16}T] + \frac{\epsilon}{16} \le \Pr[F(D_Z) \ge t].$$

Applying this lemma with $F = k + T$, we have that, for each $S \subseteq X$ with $|S| = \delta|X|$, there is a $t_S \in [-T, T]$ such that

$$\Pr[k(D) \ge t_S - \frac{\epsilon}{16}T] + \frac{\epsilon}{16} \le \Pr[k(D_S) \ge t_S].$$

Let S be the set consisting of $\delta|X|$ elements of X with the smallest value of $k(x)$, and let t_0 be a integer with $t_0 \in [t_S - (\epsilon/16)T, t_S]$,($t_0$ exists since $T > 16/\epsilon$). Thus,

$$\Pr[k(D) \ge t_0] + \frac{\epsilon}{16} \le \Pr[k(D_S) \ge t_0].$$

Denote $r := \Pr[k(D_S) \geq t_0]$. Since $\epsilon > 0$, we have that $r > 0$, i.e. there is a $x \in S$ such that $k(x) \geq t_0$, then by the definition of S, we have that

$$\Pr[k(D_X) \geq t_0] = 1 - \delta(1 - r).$$

On the other hand, let R be a distribution on X with D δ-dense in R, then we have

$$\Pr[k(R) < t_0] \geq \delta \Pr[k(D) < t_0] \geq \delta(1 + \frac{\epsilon}{16} - r),$$

then,

$$\Pr[k(R) \geq t_0] \leq 1 - \delta(1 - r) - \delta\frac{\epsilon}{16}.$$

Thus if we define $h : X \to \{0, 1\}$ by

$$h(x) = 1 \Leftrightarrow k(x) \geq t_0,$$

the theorem then follows. □

3 Lower Bound on the Query Complexity in Black-Box Constructions

In this section, we will give a lower bound on the query complexity of the reductions which showing dense model. Our proof is inspired by [*LTW*]. W.l.o.g, we will assume the finite universal $X = \{0, 1\}^s$. Formally, we will prove the following theorem.

Theorem 6. *Suppose* $2^{-c_* \cdot s} \leq \delta, \epsilon \leq c_*$ *with* $\epsilon = \delta^{O(1)}$, $a \leq 2^{c_* \cdot s}$ *and* $q = o(\frac{1}{\epsilon^2}\log(\frac{1}{\delta}))$. *Then for every* $\omega(\frac{1}{\epsilon^2}\log(\frac{1}{\delta})) \leq k \leq 2^{2^{c_* \cdot s}}$, *there exists a collection of boolean functions* $\mathcal{F} : \{0, 1\}^s \to \{0, 1\}$ *with* $|\mathcal{F}| = k$ *such that, there doesn't exist* $(q, 0.25\epsilon, \delta, a)$*-reduction which showing dense model for* $\mathcal{F}$.

The constant c_* *here is a small universal constant, for example,* $c_* = 0.0001$.

Remark 4. The assumption $a \leq 2^{c_* \cdot s}$ is reasonable. For example, if $a = 2^s$, we can encode arbitrary boolean function $C : \{0, 1\}^s \to \{0, 1\}$ by an advice $\alpha \in \{0, 1\}^a$, and then the reductions will be trivial.

Remark 5. The error parameter $e = 0.25\epsilon$ above is not critical. Our proof is also applicable in the case $e = \epsilon$, we set $e = 0.25\epsilon$ just for easier notations.

3.1 Preparations

We prove this theorem by probability method. We will consider the following probability space.

Probability space. The probability space will consists of independent identically distributed random variables $(V(x))_{x \in \{0,1\}^s}$ and $(P_i(x))_{i \in [k], x \in \{0,1\}^s}$, where for each $x \in \{0, 1\}^s$, $V(x) = 1$ with probability $\delta/2$, $V(x) = 0$ with probability $1 - \delta/2$, and for each $i \in [k]$, $x \in \{0, 1\}^s$, $P_i(x) = 1$ with probability $\frac{1-\epsilon}{2}$,

$P_i(x) = 0$ with probability $\frac{1+\epsilon}{2}$. We define a random measure $W(x) = V(x)$ and k random functions $g_i(x) = V(x) \oplus P_i(x)$, for each $i \in [k]$.

First, we will need the following bound on binomial distribution. Let $Z_1, \cdots, Z_n$ be *i.i.d.* binary random variables, with successful probability p, i.e. $E[Z_i] = p$ for $i \in [n]$. Define $Z := \sum_{i\in[n]} Z_i$, and let

$$F(k; n, p) := \Pr(Z \leq k) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i},$$

be the cumulative distribution function. Then,

Lemma 4. *Suppose parameters* $\epsilon, \delta \leq 0.01$, $q = o(\frac{1}{\epsilon^2} log(\frac{1}{\delta}))$, $t = O(1)$ *and* $c_1 > 0$ *with* $c_1 = \Omega(1)$, *then the following holds:*

$$F(k; q, \frac{1+\epsilon}{2}) + c_1\delta^t \geq \delta F(k; q, \frac{1-\epsilon}{2}),$$

for all $0 \leq k \leq q-1$.

Proof. We will represent $F(k; n, p)$ in terms of the regularized incomplete beta function$[PTVF]$ as follows:

$$\begin{aligned} F(k; n, p) &= \Pr(Z \leq k) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \\ &= I_{1-p}(n-k, k+1) = (n-k)\binom{n}{k} \int_0^{1-p} t^{n-k-1}(1-t)^k dt. \end{aligned}$$

Thus, we only need to prove

$$(q-k)\binom{q}{k} \int_0^{\frac{1-\epsilon}{2}} t^{q-k-1}(1-t)^k dt + c_1\delta^t \geq \delta(q-k)\binom{q}{k} \int_0^{\frac{1+\epsilon}{2}} t^{q-k-1}(1-t)^k dt, .$$

We will only consider the case that $k \leq \frac{1-\epsilon}{2}q$, since the other case is similar. Let $k = \frac{1-r\epsilon}{2}q$, $r \in [1, +\infty)$. Suppose, for the sake of contradiction, the inequality is failed. Then,

$$(1-\delta) \int_{\frac{1-3\epsilon}{2}}^{\frac{1-\epsilon}{2}} t^{q-k-1}(1-t)^k dt \leq \delta \int_{\frac{1-\epsilon}{2}}^{\frac{1+\epsilon}{2}} t^{q-k-1}(1-t)^k dt.$$

By derivative, $t^{q-k-1}(1-t)^k$ is monotonically increasing in $t \in [0, \frac{1+r\epsilon}{2}]$, thus

$$\begin{aligned} &(1-\delta)\epsilon \left(\frac{1-3\epsilon}{2}\right)^{q-k-1} \left(\frac{1+3\epsilon}{2}\right)^{k} \\ &\leq (1-\delta) \int_{\frac{1-3\epsilon}{2}}^{\frac{1-\epsilon}{2}} t^{q-k-1}(1-t)^k dt \\ &\leq \delta \int_{\frac{1-\epsilon}{2}}^{\frac{1+\epsilon}{2}} t^{q-k-1}(1-t)^k dt \\ &\leq \delta\epsilon \left(\frac{1+\epsilon}{2}\right)^{q-k-1} \left(\frac{1-\epsilon}{2}\right)^{k}, \end{aligned}$$

i.e., we obtain that

$$\left(1-\frac{4\epsilon}{1+\epsilon}\right)^{q-k-1}\cdot\left(1+\frac{4\epsilon}{1-\epsilon}\right)^{k}\leq\frac{\delta}{1-\delta}.$$

By the fact that $(1+\frac{1}{n})^n\leq e\leq(1+\frac{1}{n})^{n+1}$ and $(1-\frac{1}{n})^n\leq\frac{1}{e}\leq(1-\frac{1}{n})^{n-1}$, it follows

$$\exp(-c_t\epsilon^2 rq)=\exp(-c_t\epsilon(q-2k))\leq\frac{\delta}{1-\delta},\tag{3}$$

for some $c_t=O(1)$. Thus $r=\omega(1)$, since $q=o(\frac{1}{\epsilon^2}\log(\frac{1}{\delta}))$.
Similarly, it has that

$$\left(\frac{1+s\epsilon}{1+(s-1)\epsilon}\right)^{q-k-1}\left(\frac{1-s\epsilon}{1-(s-1)\epsilon}\right)^{k}\geq\exp(\epsilon^2 q(r-s)),$$

for all $0\leq s\leq r$. Applying repeatedly for $s=2,\cdots,r-2$ yields

$$\left(\frac{1+(r-2)\epsilon}{1+\epsilon}\right)^{q-k-1}\left(\frac{1-(r-2)\epsilon}{1-\epsilon}\right)^{k}\geq\exp(\frac{r^2-r}{2}\epsilon^2 q).\tag{4}$$

Combining Eqs. (3), (4) and the fact that $t^{q-k-1}(1-t)^k$ is monotonically increasing in $t\in[0,\frac{1+r\epsilon}{2}]$, we get that

$$\begin{aligned}
&\int_{\frac{1-\epsilon}{2}}^{\frac{1+\epsilon}{2}}t^{q-k-1}(1-t)^k dt\\
&\leq\epsilon\left(\frac{1+\epsilon}{2}\right)^{q-k-1}\left(\frac{1-\epsilon}{2}\right)^{k}\\
&\leq\epsilon\left(\frac{1+r\epsilon}{2}-\epsilon\right)^{q-k-1}\left(\frac{1-r\epsilon}{2}+\epsilon\right)^{k}\cdot\exp(-\frac{r^2-r}{2}\epsilon^2 q)\\
&\leq\delta^{c_u r}\int_{\frac{1+r\epsilon}{2}-\epsilon}^{\frac{1+r\epsilon}{2}}t^{q-k-1}(1-t)^k dt,
\end{aligned}$$

for some $c_u=\Omega(1)$. On the other hand, by the property of regularized incomplete beta function,

$$(q-k)\binom{q}{k}\int_0^{\frac{1+r\epsilon}{2}}t^{q-k-1}(1-t)^k dt=F(k;q,\frac{1-r\epsilon}{2})\leq 1,$$

thus

$$\begin{aligned}
&(q-k)\binom{q}{k}\int_{\frac{1-\epsilon}{2}}^{\frac{1+\epsilon}{2}}t^{q-k-1}(1-t)^k dt\\
&\leq\delta^{c_u r}(q-k)\binom{q}{k}\int_{\frac{1+r\epsilon}{2}-\epsilon}^{\frac{1+r\epsilon}{2}}t^{q-k-1}(1-t)^k dt\\
&\leq\delta^{c_u r}\leq c\delta^t,
\end{aligned}$$

where the last inequality comes from that $r = \omega(1)$. Thus we have derived a contradiction, the claim then follows. □

3.2 Proof of Lower Bound

Let $Dec^{(\cdot)}(\cdot,\cdot)$ be a oracle algorithm. To show that Dec is not a (q,ϵ,δ,a)-reduction which showing dense model, we need the following lemmas.

Lemma 5. *Suppose* $k := |\mathcal{F}| = \omega(\frac{1}{\epsilon^2}\log(\frac{1}{\delta}))$. *Then,*

$$\Pr_{V,P}[D_W \text{ has a } (\delta, 0.25\epsilon, \mathcal{F})\text{-model}] = o(1).$$

Proof. Let $\mathcal{W} := \{W \subseteq \{0,1\}^s : \big||W| - 0.5\delta 2^s\big| \leq 0.001\epsilon\delta 2^s\}$, then by a simple application of Chernoff bound, we have that

$$\Pr_V[W \notin \mathcal{W}] = 2^{-\Omega(\epsilon^2\delta^2 2^s)} = o(1),$$

thus, by the conditional probability, it suffice to prove that for every $W' \in \mathcal{W}$,

$$\Pr_{V,P}[D_W \text{ has a } (\delta, 0.25\epsilon, \mathcal{F})\text{-model}|W = W'] = o(1).$$

For easier notations, we write it as

$$\Pr_P[D_W \text{ has a } (\delta, 0.25\epsilon, \mathcal{F})\text{-model}] = o(1). \tag{5}$$

Let $\mathcal{S} = \{S \subseteq \{0,1\}^s : \delta 2^s \leq |S| \leq (1+0.001\epsilon)\delta 2^s\}$. Similar as $[Imp]$, we will first prove the following claim.

Claim. Let $W \in \mathcal{W}$, and let M be a measure with $|M| \geq \delta 2^s$ such that $\max_{g\in\mathcal{F}}(\big|\mathbb{E}[g(D_M)] - \mathbb{E}[g(D_W)]\big|) \leq 0.25\epsilon$, then there is a $S \in \mathcal{S}$ such that

$$\max_{g\in\mathcal{F}}(\big|\mathbb{E}[g(D_S)] - \mathbb{E}[g(D_W)]\big|) \leq 0.4\epsilon.$$

Proof. We will assume that $|M| = \delta 2^s$, and otherwise, we can set $M'(x) = \delta 2^s M(x)/|M|$. Define $R(x) := \delta 2^s W(x)/|W|$, it has that $R(x) \leq 2.5$ since $W \in \mathcal{W}$. Let $g \in \mathcal{F}$, and pick S by placing $x \in S$ with probability $M(x)$. By the assumptions,

$$\big|\sum_x g(x)(M(x) - R(x))\big| \leq 0.25\epsilon|M| = 0.25\epsilon\delta 2^s,$$

thus

$$\big|\mathbb{E}_S[\sum_x g(x)(S(x) - R(x))]\big| \leq 0.25\epsilon\delta 2^s,$$

Note that $\sum_x g(x)(S(x)-R(x))$ is the sum of 2^s independent random variables that are in $[-2.5,1]$. Hence by Hoeffding's inequality[Hoe],

$$\Pr_S\big(\big|\sum_x g(x)(S(x)-R(x))\big| \geq 0.3\epsilon\delta 2^s\big) \leq 2^{-c\epsilon^2\delta^2 2^s},$$

for some small constant c, for example $c=0.01$. Thus, the probability that there is such a $g\in\mathcal{F}$ at most $|\mathcal{F}|2^{-c\epsilon^2\delta^2 2^s} \leq \frac{1}{4}$ since $|\mathcal{F}|=k\leq 2^{2^{c_*s}}$ and $\epsilon,\delta\geq 2^{-c_*\cdot s}$.

On the other hand, it has that $\Pr_S[S\in\mathcal{S}]\geq 1/3$ since $\mathbb{E}_S[|S|]=\delta 2^s$. Then, we have a $S\in\mathcal{S}$ with $\big|\sum_x g(x)(S(x)-R(x))\big|\leq 0.3\epsilon\delta 2^s$ for all $g\in\mathcal{F}$. Thus

$$\begin{aligned}
&\max_{g\in\mathcal{F}}\big(\big|\mathbb{E}[g(D_S)]-\mathbb{E}[g(D_W)]\big|\big)\\
=&\max_{g\in\mathcal{F}}\big|\sum_x g(x)\left(\frac{S(x)}{|S|}-\frac{R(x)}{|R|}\right)\big|\\
\leq& 0.4\epsilon,
\end{aligned}$$

the claim then follows. □

Thus, it remains to prove that

$$\Pr_P(\exists S\in\mathcal{S}, \max_{g\in\mathcal{F}}\big(\big|\mathbb{E}[g(D_S)]-\mathbb{E}[g(D_W)]\big|\big)\leq 0.4\epsilon) = o(1),$$

for each $W\in\mathcal{W}$.

Let $S\in\mathcal{S}$, $W=\mathcal{W}$, and assume that $S\cup W=\{x_1,\cdots,x_r\}$. Notice that $r\leq 2\delta 2^s$. Let $(Z_{i,j})_{i\in[k],j\in[r]}$ be the random variables such that $Z_{i,j}=\delta 2^s g_i(x_j)(\frac{W(x_j)}{|W|}-\frac{S(x_j)}{|S|})$. Clearly, $Z_{i,j}$ are $i.i.d$ with $Z_{i,j}\in[-1,2.1]$, and in further, by the fact that $g_i(x_j)=W(x_j)\oplus P_i(x_j)$, it has $\mathbb{E}[\sum_j Z_{i,j}]\geq 0.45\epsilon\delta 2^s$ for each $i\in[k]$. Then by Hoeffding's inequality,

$$\Pr_P\big(\sum_{i,j} Z_{i,j}\leq 0.4\epsilon\delta k 2^s\big) = 2^{-\Omega(k\epsilon^2\delta 2^s)}.$$

We first note that, conditioned on $\sum_{i,j}Z_{i,j}>0.4\epsilon\delta k 2^s$, there is a $i_0\in[k]$ such that $\sum_j Z_{i_0,j}>0.4\epsilon\delta 2^s$. Means

$$\big|E[g_{i_0}(D_W)]-E[g_{i_0}(D_S)]\big|>0.4\epsilon.$$

Hence,

$$\Pr_P(\max_{g\in\mathcal{F}}\big(\big|\mathbb{E}[g(D_{U_S})]-\mathbb{E}[g(D_W)]\big|\big)\leq 0.4\epsilon) = 2^{-\Omega(k\epsilon^2\delta 2^s)}.$$

On the other hand, by Stirling's formula, $|\mathcal{S}|=\sum_l\binom{2^s}{l}=2^{O(\log(\frac{1}{\delta})\delta 2^s)}$, then by union bound,

$$\Pr_P(\exists S\in\mathcal{S}, \max_{g\in\mathcal{F}}\big(\big|\mathbb{E}[g(D_S)]-\mathbb{E}[g(D_W)]\big|\big)\leq 0.4\epsilon) = o(1),$$

since $k=\omega(\log(\frac{1}{\delta})\frac{1}{\epsilon^2})$. The claim then follows. □

Let $\mathcal{S}' = \{S \subseteq \{0,1\}^s : \mu(S) = \delta\}$. Next, we show that a black-box algorithm Dec is unlikely to approximate W well. Formally, we have the following lemma.

Lemma 6. *Let c be a constant, $k := |\mathcal{F}|$. Consider the probability space, let E be the event that,* there exist non-uniform advice $\alpha \in \{0,1\}^a$ and $I \subseteq [k]$ with $|I| = q$, such that for all $S \in \mathcal{S}'$, it has

$$\left|\mathbb{E}[Dec^{\mathcal{F},I}(D_W, \alpha)] - \mathbb{E}[Dec^{\mathcal{F},I}(D_S, \alpha)]\right| > c\epsilon.$$

Then for $q = o(\frac{1}{\epsilon^2}\log(\frac{1}{\delta}))$, $\Pr_{V,P}[E] = o(1)$.

Proof. We first notice soma basic facts. Suppose there are $S_1, S_2 \in \mathcal{S}'$ with

$$\mathbb{E}[Dec^{\mathcal{F},I}(D_W, \alpha)] - \mathbb{E}[Dec^{\mathcal{F},I}(D_{S_1}, \alpha)] > c\epsilon,$$
$$\mathbb{E}[Dec^{\mathcal{F},I}(D_W, \alpha)] - \mathbb{E}[Dec^{\mathcal{F},I}(D_{S_2}, \alpha)] < -c\epsilon,$$

then there is a subset $S_3 \in \mathcal{S}'$ such that $S_3 \subseteq S_1 \cup S_2$ and

$$\left|\mathbb{E}[Dec^{\mathcal{F},I}(D_W, \alpha)] - \mathbb{E}[Dec^{\mathcal{F},I}(D_{S_3}, \alpha)]\right| \leq c\epsilon.$$

Thus, let $E_1 := \{\exists I, \alpha, \forall S \in \mathcal{S}', \mathbb{E}[Dec^{\mathcal{F},I}(D_S, \alpha)] - \mathbb{E}[Dec^{\mathcal{F},I}(D_W, \alpha)] > c\epsilon\}$ and $E_2 := \{\exists I, \alpha, \forall S \in \mathcal{S}', \mathbb{E}[Dec^{\mathcal{F},I}(D_S, \alpha)] - \mathbb{E}[Dec^{\mathcal{F},I}(D_W, \alpha)] < -c\epsilon\}$, it has

$$E = E_1 \cup E_2.$$

Thus, it suffice to prove that $\Pr_{V,P}[E_1] = o(1)$ and $\Pr_{V,P}[E_2] = o(1)$. We will only show $\Pr_{V,P}[E_1] = o(1)$ since the analyses of E_1 and E_2 are similar.

Consider any subset $I \subseteq [k]$ with $|I| = q$ and $\alpha \in \{0,1\}^a$. Let $C : \{0,1\}^s \times \{0,1\}^q \to \{0,1\}$ be the function such that $C(x, g_I(x)) = Dec^{\mathcal{F},I}(x, \alpha)$. Note that C is well defined since $\mathcal{F}$ are boolean functions and Dec is black-box.

For every $x \in \{0,1\}^s$, let $p_1(x) := \Pr_{V,P}[C(x, g_I(x)) = 0 | W(x) = 0]$ and $p_2(x) := \Pr_{V,P}[C(x, g_I(x)) = 1 | W(x) = 1]$. We first prove that

$$\frac{(1 - 0.5c\epsilon)p_1(x)}{\delta} + (1 - 0.5c\epsilon)p_2(x) \geq 1 - c\epsilon. \tag{6}$$

Define $C_x^{-1}(0) := \{y \in \{0,1\}^q : C(x,y) = 0\}$, and let $r_x := |C_x^{-1}(0)|$. It suffice to consider in the case that $r_x = \sum_{i \leq k} \binom{q}{i}$ for some $0 \leq k \leq q$(the other case is straightforward in our proof). Since $W(x) = V(x)$ and $g_i(x) = V(x) \oplus P_i(x)$, it has

$$\begin{aligned} p_1(x) &= \Pr_{V,P}(C(x, g_I(x)) = 0 | W(x) = 0) = \Pr_{P}(C(x, P_I(x)) = 0) \\ &= \sum_{y \in C_x^{-1}(0)} \Pr_P(P_I = y) \\ &\geq \sum_{i \leq k} \binom{q}{i} \left(\frac{1+\epsilon}{2}\right)^i \left(\frac{1-\epsilon}{2}\right)^{q-i} = F(k; q, \frac{1+\epsilon}{2}). \end{aligned}$$

Similarly, we can prove that

$$p_2(x) \geq \sum_{i>k} \binom{q}{i} \left(\frac{1-\epsilon}{2}\right)^i \left(\frac{1+\epsilon}{2}\right)^{q-i} = 1 - F(k, q, \frac{1-\epsilon}{2})$$

Thus,

$$\begin{aligned}
&\frac{(1-0.5c\epsilon)p_1(x)}{\delta} + (1-0.5c\epsilon)p_2(x) \\
&\geq \frac{1-0.5c\epsilon}{\delta}\left(F(k;q,\frac{1+\epsilon}{2}) - \delta F(k;q,\frac{1-\epsilon}{2})\right) + 1 - 0.5c\epsilon \\
&\geq 1 - c\epsilon,
\end{aligned}$$

where the last inequality holds since $\epsilon = \delta^{O(1)}$ and with applying Lemma 4.

Let Z_x be a random variable with that

$$Z_x := 1 - C(x, g_I(x)) + W(x)C(x, g_I(x)) - W(x)(1 - C(x, g_I(x))).$$

then,

$$\mathbb{E}[Z_x] = p_1(x) + \delta p_2(x) - \frac{\delta}{2}(p_1(x) + p_2(x)).$$

Since $p_1(x) + \delta p_2(x) \geq \delta \frac{1-c\epsilon}{1-0.5c\epsilon}$, it can be shown that

$$\mathbb{E}[Z_x] \geq \delta \frac{1-c\epsilon}{1-0.5c\epsilon} - \frac{\delta}{2}$$

by cases analysis, i.e. $p_1(x) + p_2(x) \geq 1$ or $p_1(x) + p_2(x) < 1$.
Thus,

$$\Pr_{V,P}[\sum Z_x \leq \big((1-c\epsilon)(1+0.3c\epsilon)\delta - \frac{\delta}{2}\big)2^s] = 2^{-\Omega(c^2\epsilon^2\delta^2 2^s)},$$

by Hoeffding's inequality.
Also, let $\mathcal{W}' := \{W \subseteq \{0,1\}^s : \big||W| - 0.5\delta 2^s\big| \leq 0.01c\epsilon\delta 2^s\}$, we have that

$$\Pr_V[W \notin \mathcal{W}'] = 2^{-\Omega(c^2\epsilon^2\delta^2 2^s)}.$$

Let $A_1 := \{x : C(x, g_I(x)) = 0\}$, $A_2 := \{x : W(x) = 1 \wedge C(x, g_I(x)) = 1\}$, $A_3 := \{x : W(x) = 1 \wedge C(x, g_I(x)) = 0\}$ be random sets.

It can be shown that conditioned on $\sum Z_x \geq \big((1-c\epsilon)(1+0.3c\epsilon)\delta - \frac{\delta}{2}\big)2^s$ and $W \in \mathcal{W}'$, we have

$$|A_1| + 2|A_2| \geq (1-c\epsilon)(1+0.2c\epsilon)\delta 2^s,$$

then let $S \in \mathcal{S}'$ with $\mu(S) = \delta$ such that $A_1 \subseteq S$ ($S \subseteq A_1$ when $\mu(A_1) \geq \delta$), it has that

$$\begin{aligned}
&\mathbb{E}[Dec^{\mathcal{F},I}(D_S, \alpha)] - \mathbb{E}[Dec^{\mathcal{F},I}(D_W, \alpha)] \\
&= \mathbb{E}[C(D_S, g_I(D_S))] - \mathbb{E}[C(D_W, g_I(D_W))] \\
&= \frac{|S| - |A_1|}{|S|} - \frac{|A_2|}{|W|} \\
&\leq c\epsilon.
\end{aligned}$$

Thus,

$$\Pr_{V,P}[\forall S \in \mathcal{S}', \mathbb{E}[Dec^{\mathcal{F},I}(D_S,\alpha)] - \mathbb{E}[Dec^{\mathcal{F},I}(D_W,\alpha)] > c\epsilon] = 2^{-\Omega(c^2\epsilon^2\delta^2 2^s)},$$

then by union bound,

$$\Pr_{V,P}[E_1] = 2^a k^q 2^{-\Omega(c^2\epsilon^2\delta^2 2^s)} = o(1),$$

since $\epsilon, \delta \geq 2^{-c_* s}$, $a \leq 2^{c_* s}$ and $k \leq 2^{2^{c_* s}}$. The claim then follows. □

Combined Lemma 5 and Lemma 6, there exist W and $\mathcal{F}$ such that

- D_W doesn't have $(\delta, 0.25\epsilon, \mathcal{F})$-model;
- for every $\alpha \in \{0,1\}^a$ and $I \subseteq [k]$ with $|I| = q$, there is a $S_{I,\alpha} \in \mathcal{S}'$ such that $\left|\mathbb{E}[Dec^{\mathcal{F},I}(D_W,\alpha)] - \mathbb{E}[Dec^{\mathcal{F},I}(D_{S_{I,\alpha}},\alpha)]\right| > c\epsilon$.

Let $R_{I,\alpha} := \delta D_W + (1-\delta) D_{X \setminus S_{I,\alpha}}$, then it shows that Dec can't be $(q, 0.25\epsilon, \delta, a)$-reduction which showing dense model for $\mathcal{F}$.

4 Hardcore via regularity lemma

In this section, we will pay special attention that $H = \{f : X \to \mathbb{R}\}$ be a Hilbert space with inner product $\langle f, g\rangle := \mathbb{E}(f(D_X) \cdot g(D_X)) = \sum_x f(x)g(x)/|X|$, i.e., $\|f\| := \|f\|_{L_2}$.
Let $sgn : H \to H$ be a map by putting

$$sgn(f)(x) = \begin{cases} f(x) & \text{if } |f(x)| \leq 1 \\ 1 & \text{if } f(x) > 1 \\ -1 & \text{if } f(x) < -1. \end{cases}$$

Let $S \subset H$ be the structured vectors such that $\|v\| = 1$ for all $v \in S$. We define $S_1 := \{c \cdot f : |c| \leq 1, f \in S\}$, and recursively define

$$S_k := \{sgn(f_1 + cf_2) : f_1 \in S_{k-1}, |c| \leq 1, f_2 \in S\},$$

we say the vectors $f \in S_k$ has complexity k. Then similar as [$Tao1$], we have the following lemma.

Lemma 7. *Let H, S as above. Let $f \in H$ with $\|f\| \leq 1$, such that f is not $\epsilon\|f\|$-pseudorandom for some $0 < \epsilon \leq 1$. Then there exists $v \in S$ and $c \in [-1,1]$ such that $|\langle f, v\rangle| \geq \epsilon\|f\|$ and $\|f - cv\|^2 \leq \|f\|^2(1-\epsilon^2)$.*

Proof. By the definitions, we can find $v \in S$ such that $|\langle f, v\rangle| \geq \epsilon\|f\|$, and then set $c := \langle f, v\rangle/\|v\|^2$(i.e. cv is the orthogonal projection of f to v). By the Cauchy-Schwarz, we have

$$|\langle f, v\rangle|/\|v\|^2 \leq \|f\|\|v\|/\|v\|^2 \leq 1/\|v\| = 1,$$

thus $c \in [-1, 1]$.
Also, by Pythagoras' theorem, we will have

$$\|f - cv\|^2 = \|f\|^2 - (|\langle f, v\rangle|/\|v\|)^2 \leq \|f\|^2(1 - \epsilon^2),$$

we obtain the claim. □

Then, we can prove another version of regularity lemma, and prove a l_2-norm version of Hardcore Theorem.

Theorem 7. *Let H, S as above, $0 < \delta, \epsilon < 1$ be parameters, and let $t = 2\log(\frac{1}{2\delta})/\epsilon^2$. Suppose $f \in H$ such that:*
(i)$|f(x)| \leq 1$ for all $x \in X$;
(ii)for all $g \in S_t$ such that $\|g\| \geq (1-\delta/2)\|f\|$, it has $|\langle f, g\rangle| \leq (1-\delta)\|g\|\|f\|$.
there exists a decomposition $f = f_{str} + f_{psd}$ such that $f_{str} \in S_t$ and f_{psd} is $\epsilon\|f_{psd}\|$-pseudorandom with $\|f_{psd}\| \geq \delta\|f\|/2$.

Proof. To prove this theorem, we will repeat the following processes.

- Step 0. Initialise $f_{0,str} := 0$, and $f_{0,psd} := f$.
- Step 1. If $f_{i,psd}$ is $\epsilon\|f_{i,psd}\|$-pseudorandom then STOP, and set $f_{str} = f_{i,str}$, $f_{psd} = f_{i,psd}$. Otherwise, by the Lemma 7, it has $v \in S$ and $c \in [-1, 1]$ such that $\|f_{i,psd} - cv\|^2 \leq (1 - \epsilon^2)\|f_{i,psd}\|^2$.
- Step 2. Let $f_{i+1,str} = sgn(f_{i,str} + cv), f_{i+1,psd} = f - f_{i+1,str}$. And back to Step 1 with $f_{i+1,str}$ and $f_{i+1,psd}$.

We will first prove that $\|f_{i,psd}\| \geq \delta\|f\|/2$, for every $i < t$. For the sake of contradiction, we assume $\|f_{i,psd}\| < \delta\|f\|/2$. Notice that $f = f_{i,str} + f_{i,psd}$, then by triangle inequality, it has

$$\|f_{i,str}\| > (1 - \delta/2)\|f\|,$$

thus

$$\begin{aligned}\langle f, f_{i,str}\rangle &= \langle f_{i,str}, f_{i,str}\rangle + \langle f_{i,psd}, f_{i,str}\rangle \\ &\geq \|f_{i,str}\|^2 - \|f_{i,str}\|\|f_{i,psd}\| \\ &> (1 - \delta)\|f_{i,str}\|\|f\|.\end{aligned}$$

On the other hand, by the assumptions, we have $\langle f, f_{i,str}\rangle \leq (1 - \delta)\|f_{i,str}\|\|f\|$ since $f_{i,str} \in S_t$, which is a contradiction.

Then we prove that the process will halt before t steps, it suffice to prove that $f_{t-1,str} < \delta\|f\|/2$. By the construction, we have $f_{i,psd} = f - sgn(f_{i-1,str} + cv)$, i.e.,

$$f_{i,psd}(x) = \begin{cases} f(x) - f_{i-1,str}(x) - cv(x) & \text{if } |f_{i-1,str}(x) + cv(x)| \leq 1 \\ f(x) - 1 & \text{if } f_{i-1,str}(x) + cv(x) > 1 \\ f(x) + 1 & \text{if } f_{i-1,str}(x) + cv(x) < -1 \end{cases}$$

In each of the case, we have $|f_{i,psd}(x)| \leq |f(x) - f_{i-1,str}(x) - cv(x)|$ since $|f(x)| \leq 1$. Then

$$\|f_{i,psd}\| \leq \|f - f_{i-1,str} - cv\| = \|f_{i-1,psd} - cv\| \leq (1 - \epsilon^2)\|f_{i-1,psd}\|.$$

Applying repeatedly for $i = 1, \cdots, t-1$ yields

$$\|f_{t-1,psd}\|^2 \leq \|f_{0,psd}\|^2(1 - \epsilon^2)^{t-1} = \|f\|^2(1 - \epsilon^2)^{t-1} < \delta\|f\|/2.$$

The claim then follows. □

Set the structured set $S \subseteq \{g : X \to \{-1, 1\}\}$ be the functions which can be computed by some circuits with size s', it has the following corollary, and the proof is omitted here.

Corollary 1. *Let $0 < \epsilon, \delta < 1$ be parameters, and $f : X \to \{-1, 1\}$ be a function with $Adv_s(f) \leq 1 - \delta$. Then there is a measure M with $\|M\| \geq c\delta$ such that $Adv^M_{s'}(f) \leq \epsilon\frac{\|M\|}{\mu(M)}$, where $s' = O(s\epsilon^2/\log(1/\delta))$ and c is an universal constant.*

Remark 6. In fact, we have $M(x) = f_{psd}(x) \cdot f(x) \geq 0$ since $|f_{str}(x)| \leq 1 = |f(x)|$. And we have decomposed $f = f_{psd} + f_{str}$, informally,one part is easy to compute, and the other one is hard.

Remark 7. In our result, we get a hardcore measure M with size $\|M\| \geq c\delta$ and $\left(\delta\frac{\|M\|}{\mu(M)}\right)$-hardness. In fact, we have that $\mu(M) = \|M\|_{L_1} \leq \|M\|_{L_2} = \|M\|$, thus our result is weaker than the classic one. There is an open problem here, is there a essential gap between L_1-norm and L_2-norm. Based on $[GLR]$ and $[Kas]$, we conjecture that there are no huge gaps between them in general case.

References

[AS00] Noga Alon and Joel Spencer. The probabilistic method, 2nd edn. Wiley-Interscience, 2000.

[AS] Sergei Artemenko and Ronen Shaltiel: Lower bounds on the query complexity of non-uniform and adaptive reductions showing hardness amplification. In preprint, ECCC-TR11-016.

[BHK] Boaz Barak, Moritz Hardt and Satyen Kale: The Uniform Hardcore Lemma via Approximate Bregman Projections. In Proceeding of ACM-SIAM Symposium on Discrete Algorithms, page 1193-1200, 2009.

[BSW] Boaz Barak, Ronen Shaltiel and Avi Wigderson: Computational analogues of entropy. In Proceeding of RANDOM, page 200-215,2003.

[CZ] Yair Censor and Stavros A.Zenios: Parallel Optimization – Theory, Algorithms, and Applications. Oxford University Press, 1997.

[FK] Alan M.Freize and Ravi Kannan. Quick approximation to matrices and applications. Combinatorica, 19(2):175-220,1999.

[FS] Yoav Freund and Robert E.Schapire. A decision-theoretic generalization of on-line learing and an application to boosting. Journal of Computer and System Sciences. 55(1):119-139, August 1997.

[GLR] Venkatesan Guruswami, James R.Lee and Alexander Razborov. Almost Euclidean subspace of L_1^n expander coders. In Proceeding of SODA, 2008.

[GT] Ben Green and Terence Tao. The primes contains arbitrarily long arithmetic progression. In Annals of Mathematics, 2004.

[Hoe] Wassily Hoeffding. Probability inequalities for sums of bounded random variables, in Journal of the American Statistical Association, 58(301):13-30, March 1963.

[Imp] Russel Impagliazzo: Hard-core distributions for somewhat hard problems. In *FOCS*, page 528-545,1995.

[Kas] B.S.Kashin. Diameters of some finite-demensional sets and classes of smooth functions. Izv. Akad. Nauk. SSSR, 41(2), 1977.

[KS] Adam R.Klivans and Rocco A.Servedio: Boosting and hard-core sets. Machine Learing, 53(3):217-238,2003.

[LTW] Chi-Jen Lu, Shi-Chum Tsai and Hsin-Lung Wu: On the complexity of hard-core set constructions. In Automata, Languages and Programming, 34th International Colloquium, page 183-194, 2007.

[PTVF] William H.Press, Saul A.Teukolsky, William T. Vetterling and Brian P.Flannery: Numberical Recipes in C – The Art of Scientific Computing. Cambridge University Press, 1992.

[RTTV] Omer Reingold, Luca Trevisan, Madhur Tulsiani and Salil Vadhan: Dense subsets of pseudorandom sets. In *FOCS*, 2008.

[SV] Ronen Shaltiel and Emanuele Viola: Hardness Amplification Proofs Require Majority. In Proceeding of *STOC*, 2008.

[Tao1] Terence Tao: Structure and randomness in combinatorics. In *FOCS*, page 3-18,2007.

[Tao2] Terence Tao: Szemeredi's regularity lemma revisited. In Contrib.Discrete Math, 1(1):8-28, 2006.

[Tao3] Terence Tao: A variant of the hypergraph removal lemma. In J.Combin.Thy.A 113:1257-1280, 2006.

[TTV] Luca Trevisan, Madhur Tulsiani and Salil Vadhan: Regularity, Boosting, and Efficiently Simulating Every High-Entropy Distribution. In IEEE Computational Complexity Conference, 2009.

[TZ] Terence Tao and Tamar Ziegler: The primes contains arbitrary long polynomial progressions. arXiv:math/0610050,2006.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de