6.842 Randomness and Computation April 27, 2022

Homework 5

Yuchong Pan *MIT ID: 911346847*

1. *Collaborators and sources:* none.

Proof. Let $f : \{\pm 1\}^n \to \{\pm 1\}$. Let $\varepsilon \in (0, 1/2)$. Then

$$\begin{aligned}
NS_\varepsilon(f) &= \mathop{\mathbb{P}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[f(x) \neq f\left(N_\varepsilon(x)\right) \right] = \mathop{\mathbb{P}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[f(x) f\left(N_\varepsilon(x)\right) = -1 \right] \\
&= \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[\frac{1}{2} - \frac{1}{2} f(x) f\left(N_\varepsilon(x)\right) \right] = \frac{1}{2} - \frac{1}{2} \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[f(x) f\left(N_\varepsilon(x)\right) \right] \\
&= \frac{1}{2} - \frac{1}{2} \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[\left(\sum_{S \subset [n]} \hat{f}(S) \chi_S(x) \right) \left(\sum_{T \subset [n]} \hat{f}(T) \chi_T\left(N_\varepsilon(x)\right) \right) \right] \\
&= \frac{1}{2} - \frac{1}{2} \sum_{S, T \subset [n]} \hat{f}(S) \hat{f}(T) \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[\chi_S(x) \chi_T\left(N_\varepsilon(x)\right) \right].
\end{aligned}$$

For all $x \in \{\pm 1\}^n$ and $i \in [n]$, we denote by x_i and $N_\varepsilon(x)_i$ the i^{th} coordinates of x and $N_\varepsilon(x)$, respectively. For all $S \subset [n]$,

$$\begin{aligned}
\mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[\chi_S(x) \chi_S\left(N_\varepsilon(x)\right) \right] &= \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[\left(\prod_{i \in S} x_i \right) \left(\prod_{i \in S} N_\varepsilon(x)_i \right) \right] = \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[\prod_{i \in S} x_i N_\varepsilon(x)_i \right] \\
&= \prod_{i \in S} \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[x_i N_\varepsilon(x)_i \right] = \left(\varepsilon \cdot (-1) + (1 - \varepsilon) \cdot 1 \right)^{|S|} \qquad (1) \\
&= (1 - 2\varepsilon)^{|S|}.
\end{aligned}$$

Note that (1) is due to the independence of each bit in $N_\varepsilon(x)$ and the fact that each bit of x uniformly chosen from $\{\pm 1\}^n$ is uniform in $\{\pm 1\}$. For all $S, T \subset [n]$ with $S \neq T$,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[\chi_S(x) \chi_T\left(N_\varepsilon(x)\right) \right] \\
&= \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[\left(\prod_{i \in S} x_i \right) \left(\prod_{i \in T} N_\varepsilon(x)_i \right) \right] \\
&= \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[\left(\prod_{i \in S \cap T} x_i N_\varepsilon(x)_i \right) \left(\prod_{i \in S \setminus T} x_i \right) \left(\prod_{i \in T \setminus S} N_\varepsilon(x)_i \right) \right] \\
&= \left(\prod_{i \in S \cap T} \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[x_i N_\varepsilon(x)_i \right] \right) \left(\prod_{i \in S \setminus T} \mathop{\mathbb{E}}_{x \in \{\pm 1\}^n} \left[x_i \right] \right) \left(\prod_{i \in T \setminus S} \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} \left[N_\varepsilon(x)_i \right] \right). \qquad (2)
\end{aligned}$$

Note that (2) is again due to the independence of each bit in $N_\varepsilon(x)$. For $S, T \subset [n]$ with $S \neq T$, either $S \setminus T \neq \emptyset$ or $T \setminus S \neq \emptyset$. Note that each bit of x uniformly chosen from $\{\pm 1\}^n$ is uniform in $\{\pm 1\}$. Therefore, if $S \setminus T \neq \emptyset$,

$$\prod_{i \in S \setminus T} \mathop{\mathbb{E}}_{x \in \{\pm 1\}^n} [x_i] = \left(\mathop{\mathbb{E}}_{b \in \{\pm 1\}} [b] \right)^{|S \setminus T|} = 0^{|S \setminus T|} = 0.$$

Moreover, if $T \setminus S \neq \emptyset$,

$$\prod_{i \in T \setminus S} \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} [N_\varepsilon(x)_i] = \left(\frac{1}{2}(\varepsilon(-1) + (1-\varepsilon) \cdot 1) + \frac{1}{2}(\varepsilon \cdot 1 + (1-\varepsilon)(-1)) \right)^{|T \setminus S|} = 0^{|T \setminus S|} = 0.$$

Therefore, for all $S, T \subset [n]$ with $S \neq T$,

$$\mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} [\chi_S(x) \chi_T(N_\varepsilon(x))] = 0.$$

It follows that

$$NS_\varepsilon(f) = \frac{1}{2} - \frac{1}{2} \sum_{S, T \subset [n]} \hat{f}(S) \hat{f}(T) \mathop{\mathbb{E}}_{\substack{x \in \{\pm 1\}^n \\ N_\varepsilon}} [\chi_S(x) \chi_T(N_\varepsilon(x))] = \frac{1}{2} - \frac{1}{2} \sum_{S \subset [n]} \hat{f}(S)^2 (1 - 2\varepsilon)^{|S|}.$$

This completes the proof. $\square$

2. (a) *Collaborators and sources:* none.

Proof. Let $f : \{\pm 1\}^n \to \{\pm 1\}$ be monotone. Let $i \in [n]$. WLOG, assume $i = 1$. Then

$$\begin{aligned}
\hat{f}(\{1\}) &= \langle f, \chi_{\{1\}} \rangle \\
&= \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} f(x) \chi_{\{1\}}(x) \\
&= \frac{1}{2^n} \sum_{x=(x_1,\ldots,x_n) \in \{\pm 1\}^n} f(x) x_1 \\
&= \frac{1}{2^n} \left(\sum_{\substack{x=(x_1,\ldots,x_n) \in \{\pm 1\}^n \\ x_1 = 1}} f(x) \cdot 1 + \sum_{\substack{x=(x_1,\ldots,x_n) \in \{\pm 1\}^n \\ x_1 = -1}} f(x) \cdot (-1) \right) \\
&= \frac{1}{2^n} \left(\sum_{x' \in \{\pm 1\}^{n-1}} f\left(1, x'\right) - \sum_{x' \in \{\pm 1\}^{n-1}} f\left(-1, x'\right) \right) \\
&= \frac{1}{2^n} \sum_{x' \in \{\pm 1\}^{n-1}} \left(f\left(1, x'\right) - f\left(-1, x'\right) \right) \\
&= \frac{1}{2^n} \sum_{\substack{x' \in \{\pm 1\}^{n-1} \\ f(1,x') \neq f(-1,x')}} \left(f\left(1, x'\right) - f\left(-1, x'\right) \right).
\end{aligned}$$

Since f is monotone, then $f(1, x') \geq f(-1, x')$ for all $x' \in \{\pm 1\}^{n-1}$. Hence, for all $x' \in \{\pm 1\}^{n-1}$, if $f(1, x') \neq f(-1, x')$, then $f(1, x') = 1$ and $f(-1, x') = -1$, so $f(1, x') - f(-1, x') = 1 - (-1) = 2$. Therefore,

$$\begin{aligned}
\hat{f}(\{1\}) &= \frac{1}{2^n} \sum_{\substack{x' \in \{\pm 1\}^{n-1} \\ f(1,x') \neq f(-1,x')}} 2 \\
&= \frac{1}{2^n} \cdot 2 \left| \left\{ x' \in \{\pm 1\}^{n-1} : f\left(1, x'\right) \neq f\left(-1, x'\right) \right\} \right| \\
&= \frac{1}{2^{n-1}} \left| \left\{ x' \in \{\pm 1\}^{n-1} : f\left(1, x'\right) \neq f\left(-1, x'\right) \right\} \right|.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\mathit{Inf}_1(f) &= \mathop{\mathbb{P}}_{x \in \{\pm 1\}^n} \left[f(x) \neq f\left(x^{\oplus 1}\right) \right] \\
&= \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} \mathbb{1}\left[f(x) \neq f\left(x^{\oplus 1}\right) \right] \\
&= \frac{1}{2^n} \cdot 2 \sum_{x' \in \{\pm 1\}^{n-1}} \mathbb{1}\left[f\left(1, x'\right) \neq f\left(-1, x'\right) \right] \\
&= \frac{1}{2^{n-1}} \left| \left\{ x' \in \{\pm 1\}^{n-1} : f\left(1, x'\right) \neq f\left(-1, x'\right) \right\} \right| \\
&= \hat{f}(\{1\}).
\end{aligned}$$

This completes the proof. □

(b) *Collaborators and sources:* none.

Proof. Let $n \in \mathbb{N}$ be odd. Let $f : \{\pm 1\}^n \to \{\pm 1\}$ be the majority function, i.e., $f(x) = \operatorname{sign}(\sum_{i=1}^n x_i)$ for all $x = (x_1, \ldots, x_n) \in \{\pm 1\}^n$. First, we show that f is monotone. Let $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \{\pm 1\}^n$ be such that $x_i \leq y_i$ for all $i \in [n]$. Then $\sum_{i=1}^n x_i \leq \sum_{i=1}^n y_i$, so $f(x) = \operatorname{sign}(\sum_{i=1}^n x_i) \leq \operatorname{sign}(\sum_{i=1}^n y_i) = f(y)$. This proves that f is monotone.

Second, let $g : \{\pm 1\}^n \to \{\pm 1\}$ be monotone. Then

$$\begin{aligned}
\mathit{Inf}(g) &= \sum_{i=1}^n \mathit{Inf}_i(g) \\
&= \sum_{i=1}^n \hat{g}(\{i\}) && \text{(part (a))} \\
&= \sum_{i=1}^n \langle g, \chi_{\{i\}} \rangle \\
&= \sum_{i=1}^n \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} g(x)\chi_{\{i\}}(x) \\
&= \frac{1}{2^n} \sum_{x=(x_1,\ldots,x_n)\in\{\pm 1\}^n} g(x) \sum_{i=1}^n x_i \\
&\leq \left| \frac{1}{2^n} \sum_{x=(x_1,\ldots,x_n)\in\{\pm 1\}^n} g(x) \sum_{i=1}^n x_i \right| \\
&\leq \frac{1}{2^n} \sum_{x=(x_1,\ldots,x_n)\in\{\pm 1\}^n} |g(x)| \left| \sum_{i=1}^n x_i \right| && \text{(triangle inequality)} \\
&= \frac{1}{2^n} \sum_{x=(x_1,\ldots,x_n)\in\{\pm 1\}^n} \left| \sum_{i=1}^n x_i \right|. && \text{(since } g(x) \in \{\pm 1\} \text{ for all } x \in \{\pm 1\}^n\text{)}
\end{aligned}$$

Third, since f is monotone,

$$\mathit{Inf}(f) = \frac{1}{2^n} \sum_{x=(x_1,\ldots,x_n)\in\{\pm 1\}^n} f(x) \sum_{i=1}^n x_i.$$

Since n is odd, then $\sum_{i=1}^n x_i \neq 0$. If $\sum_{i=1}^n x_i < 0$, then $f(x) = \operatorname{sign}(\sum_{i=1}^n x_i) < 0$, so

$$\mathit{Inf}(f) = \frac{1}{2^n} \sum_{x=(x_1,\ldots,x_n)\in\{\pm 1\}^n} |f(x)| \left| \sum_{i=1}^n x_i \right| = \frac{1}{2^n} \sum_{x=(x_1,\ldots,x_n)\in\{\pm 1\}^n} \left| \sum_{i=1}^n x_i \right|, \tag{3}$$

since $f(x) \in \{\pm 1\}$ for all $x \in \{\pm 1\}^n$. Otherwise, $\sum_{i=1}^n x_i > 0$, so $f(x) = \operatorname{sign}(\sum_{i=1}^n x_i) > 0$, implying that (3) holds. Hence, (3) holds in both cases. It follows that $\mathit{Inf}(g) \leq \mathit{Inf}(f)$ for any monotone $g : \{\pm 1\}^n \to \{\pm 1\}$, completing the proof. $\square$

3. (a) *Collaborators and sources:* none.

Proof. Let $f : \{\pm 1\}^n \to \{\pm 1\}$. Let $\varepsilon > 0$. We show that the statement holds with $C = 1$. Suppose for the sake of contradiction that $\sum_{S\subset[n],|S|\geq \mathit{Inf}(f)/\varepsilon} \hat{f}(S)^2 > C\varepsilon = \varepsilon$. For each $i \in [n]$, let $g_i : \{\pm 1\}^n \to \{0, \pm 1\}$ be defined by

$$\begin{aligned} g_i(x) &= \frac{f(x) - f\left(x^{\oplus i}\right)}{2} = \frac{1}{2}\left(\sum_{S\subset[n]} \hat{f}(S)\chi_S(x) - \sum_{S\subset[n]} \hat{f}(S)\chi_S\left(x^{\oplus i}\right)\right) \\ &= \frac{1}{2}\sum_{S\subset[n]} \hat{f}(S)\left(\chi_S(x) - \chi_S\left(x^{\oplus i}\right)\right). \end{aligned}$$

Then $g_i(x)^2 = \mathbb{1}[f(x) \neq f(x^{\oplus i})]$ for all $i \in [n]$ and $x \in \{\pm 1\}^n$. Fix $i \in [n]$, $x = (x_1, \ldots, x_n) \in \{\pm 1\}^n$ and $S \subset [n]$. If $i \in S$, then

$$\begin{aligned} \chi_S(x) - \chi_S\left(x^{\oplus i}\right) &= \prod_{j\in S} x_j - (-x_i)\prod_{j\in S\setminus\{i\}} x_j = x_i \prod_{j\in S\setminus\{i\}} x_j - (-x_i)\prod_{j\in S\setminus\{i\}} x_j \\ &= (x_i - (-x_i))\prod_{j\in S\setminus\{i\}} x_j = 2x_i \prod_{j\in S\setminus\{i\}} x_j = 2\prod_{j\in S} x_j = 2\chi_S(x). \end{aligned}$$

If $i \notin S$, then

$$\chi_S(x) - \chi_S\left(x^{\oplus i}\right) = \prod_{j\in S} x_j - \prod_{j\in S} x_j = 0.$$

Hence, for all $i \in [n]$ and $x \in \{\pm 1\}^n$,

$$g_i(x) = \frac{1}{2}\sum_{\substack{S\subset[n]\\ i\in S}} \hat{f}(S)\cdot 2\chi_S(x) = \frac{1}{2}\cdot 2\sum_{\substack{S\subset[n]\\ i\in S}} \hat{f}(S)\chi_S(x) = \sum_{\substack{S\subset[n]\\ i\in S}} \hat{f}(S)\chi_S(x).$$

For all $i \in [n]$, by the orthonormality of the Fourier basis $\{\chi_S : S \subset [n]\}$,

$$\begin{aligned} \mathit{Inf}_i(f) &= \mathop{\mathbb{P}}_{x\in\{\pm 1\}^n}\left[f(x) \neq f\left(x^{\oplus i}\right)\right] = \mathop{\mathbb{E}}_{x\in\{\pm 1\}^n}\left[\mathbb{1}\left[f(x) \neq f\left(x^{\oplus i}\right)\right]\right] = \mathop{\mathbb{E}}_{x\in\{\pm 1\}^n}\left[g_i(x)^2\right] \\ &= \mathop{\mathbb{E}}_{x\in\{\pm 1\}^n}\left[\left(\sum_{\substack{S\subset[n]\\ i\in S}} \hat{f}(S)\chi_S(x)\right)^2\right] = \mathop{\mathbb{E}}_{x\in\{\pm 1\}^n}\left[\sum_{\substack{S\subset[n]\\ i\in S}}\sum_{\substack{T\subset[n]\\ i\in T}} \hat{f}(S)\hat{f}(T)\chi_S(x)\chi_T(x)\right] \\ &= \mathop{\mathbb{E}}_{x\in\{\pm 1\}^n}\left[\sum_{\substack{S\subset[n]\\ i\in S}} \hat{f}(S)^2\right] = \sum_{\substack{S\subset[n]\\ i\in S}} \hat{f}(S)^2. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathit{Inf}(f) &= \sum_{i=1}^n \mathit{Inf}_i(f) = \sum_{i=1}^n \sum_{\substack{S\subset[n]\\ i\in S}} \hat{f}(S)^2 = \sum_{S\subset[n]}\sum_{i\in S} \hat{f}(S)^2 = \sum_{S\subset[n]} |S|\hat{f}(S)^2 \\ &\geq \sum_{\substack{S\subset[n]\\ |S|\geq \frac{\mathit{Inf}(f)}{\varepsilon}}} |S|\hat{f}(S)^2 \geq \frac{\mathit{Inf}(f)}{\varepsilon}\sum_{\substack{S\subset[n]\\ |S|\geq \frac{\mathit{Inf}(f)}{\varepsilon}}} \hat{f}(S)^2 > \frac{\mathit{Inf}(f)}{\varepsilon}\cdot\varepsilon = \mathit{Inf}(f), \end{aligned}$$

a contradiction. This completes the proof. $\square$

(b) *Collaborators and sources:* Guanghao Ye.

Proof. Let $f : \{\pm 1\}^n \to \{\pm 1\}$ be monotone. Then

$$\begin{aligned}
\mathit{Inf}(f) &= \sum_{i=1}^{n} \mathit{Inf}_i(f) \\
&= \sum_{i=1}^{n} \hat{f}(\{i\}) && \text{(Problem 2 part (a))} \\
&\le \sqrt{n \sum_{i=1}^{n} \hat{f}(\{i\})^2} && \text{(Cauchy-Schwarz inequality)} \\
&\le \sqrt{n \sum_{S \subset [n]} \hat{f}(S)^2} \\
&= \sqrt{n \cdot 1} && \text{(Boolean Parseval's identity)} \\
&= \sqrt{n}.
\end{aligned}$$

By part (a), there exists an absolute constant C such that for all $\varepsilon > 0$,

$$\sum_{\substack{S \subset [n] \\ |S| \ge \frac{\sqrt{n}}{\varepsilon}}} \hat{f}(S)^2 \le \sum_{\substack{S \subset [n] \\ |S| \ge \frac{\mathit{Inf}(f)}{\varepsilon}}} \hat{f}(S)^2 \le C\varepsilon.$$

Hence, any monotone Boolean function has Fourier concentration $\alpha(\varepsilon, n) = C\sqrt{n}/\varepsilon$. The low degree algorithm gives a uniform distribution learning algorithm $\mathcal{A}$ for the class of monotone Boolean functions with sample complexity

$$\begin{aligned}
O\left(\frac{n^{\alpha(\varepsilon,n)}}{\varepsilon} \log \frac{n^{\alpha(\varepsilon,n)}}{\delta}\right) &= O\left(\frac{n^{\alpha(\varepsilon,n)}}{\varepsilon} \cdot \alpha(\varepsilon, n) \log n\right) \\
&= O\left(\frac{n^{\frac{C\sqrt{n}}{\varepsilon}}}{\varepsilon} \cdot \frac{C\sqrt{n}}{\varepsilon} \log n\right) \\
&= O\left(\frac{n^{\frac{C\sqrt{n}}{\varepsilon} + \frac{1}{2}}}{\varepsilon^2} \log n\right).
\end{aligned}$$

Since $1/\varepsilon^2 \le (n^{\sqrt{n}})^{1/\varepsilon}$ for sufficiently large n and sufficiently small ε, then the sample complexity of $\mathcal{A}$ is

$$O\left(n^{\frac{C\sqrt{n}}{\varepsilon} + \frac{1}{2}} \log n\right) \le n^{\Theta\left(\frac{\sqrt{n}}{\varepsilon}\right)} = \left(2^{\log_2 n}\right)^{\Theta\left(\frac{\sqrt{n}}{\varepsilon}\right)} = 2^{(\log_2 n)\Theta\left(\frac{\sqrt{n}}{\varepsilon}\right)} = 2^{\Theta\left(\frac{\sqrt{n}}{\varepsilon} \log n\right)} = 2^{\widetilde{O}\left(\frac{\sqrt{n}}{\varepsilon}\right)}.$$

This completes the proof. $\square$

4. (a) *Collaborators and sources:* none.

Proof. Let $X = (X_1, \dots, X_n) \in \{\pm 1\}^n$ be an (ε, k)-wise independent random vector for some $\varepsilon \in (0,1)$ and $k \in [n]$. Let $S \subset [n]$ be such that $0 < |S| \leq k$. By straightforward calculations (see, e.g., the proof of Problem 1 part (a) in Homework 2), for all $\ell \in [n]$,

$$\mathop{\mathbb{P}}_{(W_1,\dots,W_\ell)\sim \mathsf{Unif}\{\pm 1\}^\ell}\left[\prod_{i=1}^{\ell} W_\ell = 1\right] = \frac{1}{2}.$$

Since X is (ε, k)-wise independent and since $0 < |S| \leq k$,

$$\left|\mathop{\mathbb{P}}_{X}\left[\prod_{i\in S} X_i = 1\right] - \frac{1}{2}\right| = \left|\mathop{\mathbb{P}}_{X}\left[\prod_{i\in S} X_i = 1\right] - \mathop{\mathbb{P}}_{(W_1,\dots,W_\ell)\sim \mathsf{Unif}\{\pm 1\}^\ell}\left[\prod_{i=1}^{\ell} W_\ell = 1\right]\right| \leq \varepsilon.$$

WLOG, assume that

$$\mathop{\mathbb{P}}_{X}\left[\prod_{i\in S} X_i = 1\right] = \frac{1+\varepsilon_0}{2},$$

for some $\varepsilon_0 \in [0, 2\varepsilon]$ (the case $\mathbb{P}_X[\prod_{i\in S} X_i = 1] = (1-\varepsilon_0)/2$ for some $\varepsilon_0 \in [0, 2\varepsilon]$ is symmetric). Let $\lambda = 1/(1+\varepsilon_0) \in (0, 1]$. Let $Y = (Y_1, \dots, Y_n) \in \{\pm 1\}^n$ be a random vector defined as follows:

(i) With probability λ, let $Y = X$.
(ii) With probability $1 - \lambda$, let Y be uniform over

$$\mathcal{W} := \left\{(x_1, \dots, x_n) \in \{\pm 1\}^n : \prod_{i\in S} x_i = -1\right\}.$$

Then

$$\mathop{\mathbb{P}}_{Y}\left[\prod_{i\in S} Y_i = 1\right] = \lambda \mathop{\mathbb{P}}_{X}\left[\prod_{i\in S} X_i = 1\right] + (1-\lambda)\cdot 0 = \frac{1}{1+\varepsilon_0}\cdot\frac{1+\varepsilon_0}{2} = \frac{1}{2}.$$

For all $\mathcal{T} \subset \{\pm 1\}^n$,

$$\begin{aligned}
\mathop{\mathbb{P}}_{X}[X\in\mathcal{T}] - \mathop{\mathbb{P}}_{Y}[Y\in\mathcal{T}] &= \mathop{\mathbb{P}}_{X}[X\in\mathcal{T}] - \left(\lambda \mathop{\mathbb{P}}_{X}[X\in\mathcal{T}] + (1-\lambda)\mathop{\mathbb{P}}_{W\sim \mathsf{Unif}\,\mathcal{W}}[W\in\mathcal{T}]\right) \\
&= (1-\lambda)\left(\mathop{\mathbb{P}}_{X}[X\in\mathcal{T}] - \mathop{\mathbb{P}}_{W\sim \mathsf{Unif}\,\mathcal{W}}[W\in\mathcal{T}]\right) \\
&\leq (1-\lambda)(1-0) = 1 - \frac{1}{1+\varepsilon_0} = \frac{\varepsilon_0}{1+\varepsilon_0} \\
&\leq \varepsilon_0 \leq 2\varepsilon.
\end{aligned}$$

Therefore,

$$\Delta(X,Y) = \max_{\mathcal{T}\subset\{\pm 1\}^n}\left(\mathop{\mathbb{P}}_{X}[X\in\mathcal{T}] - \mathop{\mathbb{P}}_{Y}[Y\in\mathcal{T}]\right) \leq 2\varepsilon.$$

This completes the proof. ∎