Testing of Boolean functions

Pascal Huber

May 8, 2014

1 Motivation

2 Reminder

First we like to recall some basic results that we have already covered in the previous course sessions and that we will need in our investigation of linearity testing.

We consider functions from the hypercube $\Omega_n \coloneqq \{-1,1\}^n$ into the real numbers (we denote elements of $\{-1,1\}^n$ by boldfaced variables, e.g. $\boldsymbol{x}$). A function $f : \Omega_n \to \mathbb{R}$ is called *Boolean function* if it takes values only in $\{-1,1\}$.

Parity functions We recall an important type of Boolean functions. Let $[n] \coloneqq \{1,2,\ldots,n\}$ and let $S \subseteq [n]$. The *parity function over* S is defined by

$$\chi_S(\boldsymbol{x}) = \boldsymbol{x}_S \coloneqq \prod_{i\in S} x_i,$$

where we use the convention: $\chi_\emptyset(\boldsymbol{x}) = \prod_{i\in\emptyset} x_i = 1$ for all $\boldsymbol{x} \in \{-1,1\}^n$.

Great notational switch Since we will treat linearity testing of Boolean functions it is in some cases more suitable to replace the setting above by an additive version (this is sometimes called the *"great notational switch"*).
For this we replace the bit-set $\{-1,1\}$ by $\{0,1\}$ (substituting $1 \rightsquigarrow 0$ and $-1 \rightsquigarrow 1$). Moreover the multiplication in $\{-1,1\}$ is replaced by the addition in $\mathbb{F}_2$. In this setting a Boolean function has the form

$$f : \{0,1\}^n \to \{0,1\}.$$

Note that the parity functions are in this setting just the functions of the form $\chi_S(\boldsymbol{x}) = \sum_{i\in S} x_i$ for some $S \subseteq [n]$.

Unless stated otherwise we will use the multiplicative setting, (i.e. $\Omega_n = \{-1,1\}^n$).

Correlation of Boolean functions The measurable space $(\Omega_n, \mathcal{P}(\Omega_n))$ is equipped with the uniform probability measure

$$\mathbf{P}_{x\in\{-1,1\}^n} \coloneqq \left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1\right)^{\otimes n}.$$

Furthermore we define the corresponding L^2-inner product (called *correlation*) for two functions $f, g : \Omega_n \to \mathbb{R}$ on this space by

$$\langle f, g \rangle \coloneqq \mathbf{E}_{\boldsymbol{x} \in \{-1,1\}^n} [f(\boldsymbol{x}) g(\boldsymbol{x})] = \frac{1}{2^n} \sum_{\boldsymbol{x} \in \{-1,1\}^n} f(\boldsymbol{x}) g(\boldsymbol{x}).$$

Note that if f and g are Boolean functions then one always gets $\langle f, g \rangle \in [-1, 1]$ and $\|f\| \coloneqq \sqrt{\langle f, f \rangle} = 1$.

Results from Fourier analysis of Boolean functions Before introducing the notions of testability of Boolean functions we recall the most important results from Fourier analysis of Boolean functions.

- The parity functions $(\chi_S)_{S \subseteq [n]}$ form an orthonormal basis (w.r.t. the correlation) on the space of all functions from Ω_n to the real numbers, in particular
$$\langle \chi_S, \chi_T \rangle = \begin{cases} 1 & S = T \\ 0 & \text{else} \end{cases}$$
- Consequently every Boolean function f has a unique representation of the form
$$f = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S. \tag{1}$$
The coefficients $\hat{f}(S)$ are called *Fourier coefficients* and (1) is called the *Fourier expansion* of f.
- *Plancharel's theorem:* Let $f, g : \Omega_n \to \mathbb{R}$. Then $\langle f, g \rangle = \sum_{S \subseteq [n]} \hat{f}(S) \hat{g}(S)$.
- *Parseval's theorem:* Let $f : \Omega_n \to \mathbb{R}$. Then $\langle f, f \rangle = \sum_{S \subseteq [n]} \hat{f}(S)^2$. Especially, if f is Boolean one has the identity $\sum_{S \subseteq [n]} \hat{f}(S)^2 = 1$.

3 Linearity of Boolean functions and basic notions of property testing

3.1 Linearity of Boolean functions

In order to analyse linearity testing of Boolean functions (see section 4) it is a good start to define what it means for a Boolean function to be linear.

Remark. For simplicity we use in this subsection exclusively the additive notation introduced in section 2 ($\{-1, 1\}$ replaced by $\{0, 1\}$). Of course the analogous results hold also for the multiplicative setting.

Definition 1 (Linearity of Boolean functions)**.** A Boolean function $f : \{0,1\}^n \to \{0,1\}$ is called *linear* iff one of the following statements hold:

(i) For all $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$ one has $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$.

(ii) f is a parity function, i.e. there exists $S \subseteq [n]$ such that for all $\boldsymbol{x} \in \{0,1\}^n$ one has $f(\boldsymbol{x}) = \sum_{i \in S} x_i$.

Remark. In multiplicative notation Definition 1 reads as follows:

(i) For all $\boldsymbol{x}, \boldsymbol{y} \in \{-1,1\}^n$ one has $f(\boldsymbol{x} \circ \boldsymbol{y}) = f(\boldsymbol{x}) \cdot f(\boldsymbol{y})$, where $\boldsymbol{x} \circ \boldsymbol{y} \coloneqq (x_1 y_1, \ldots, x_n y_n)$.

(ii) There exists $S \subseteq [n]$ such that for all $\boldsymbol{x} \in \{-1,1\}^n$ one has $f(\boldsymbol{x}) = \prod_{i \in S} x_i$.

In order to make sure that Definition 1 is well defined we have to show that both statements of the definition are equivalent:
(ii) $\Rightarrow$ (i): is straightforward since

$$\chi_S(\boldsymbol{x} + \boldsymbol{y}) = \sum_{i \in S}(x_i + y_i) = \sum_{i \in S} x_i + \sum_{i \in S} y_i = \chi_S(\boldsymbol{x}) + \chi_S(\boldsymbol{y}).$$

(i) $\Rightarrow$ (ii): uses the representation $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$ where $\boldsymbol{e}_i \coloneqq (0, \ldots, 0, 1, 0, \ldots, 0)$ with the non-zero bit at the i-th position. Then by iterating (i) one gets

$$f(x) = f\left(\sum_{i=1}^{n} x_i \boldsymbol{e}_i\right) = \sum_{i=1}^{n} x_i f(\boldsymbol{e}_i) = \sum_{i \in S} x_i,$$

where $S \coloneqq \{i \in [n] : f(\boldsymbol{e}_i) = 1\}$.

Thus we see that the linear Boolean functions are exactly the parity functions.

3.2 Approximate linearity and basic notions of property testing

In the following we introduce the basic notions of property testing using the example of linearity testing.

Approximate linearity The requirement to be linear on $\{0,1\}^n$ is rather expensive to check (one has to do all 2^n queries) and hence for many applications too restrictive. That's why one considers a more relaxed concept of linearity:

A Boolean function $f : \{0,1\}^n \to \{0,1\}$ is called *approximate linear* iff

(i') for "most" pairs $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$ one has $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$.

(ii') there exist $S \subseteq [n]$ s.t. for "most" $\boldsymbol{x} \in \{0,1\}^n$ one has $f(\boldsymbol{x}) = \chi_S(\boldsymbol{x})$.

Analogously the multiplicative statements can be defined.

(We will soon define more precisely what is meant by "most".)

As before we would like both definitions to be equivalent and one easily sees by copying the above argument that (ii') indeed implies (i'). The converse implication on the other hand is much less obvious.

A main result of this course will be that in fact (i') implies (ii'). In order to show this we will use the results of the Fourier analysis extensively. But before dwelling on the proof we want to put the problem in the more general context of property testing. Therefore we will need some definitions.

Remark. From now on we will only use the multiplicative notation.

Definition 2 (Property of Boolean functions)**.** A *property of Boolean functions* is a subset $\mathcal{P}$ of the set of all Boolean functions. We say that a Boolean function f *has property* $\mathcal{P}$ if $f \in \mathcal{P}$.

Definition 3 (ε-far and ε-close)**.** (i) Two Boolean functions f, g are ε-close if they agree on a $(1-\varepsilon)$-fraction of $\{-1,1\}^n$, i.e.

$$\mathbf{P}_{\boldsymbol{x}\in\{-1,1\}^n}\left[f(\boldsymbol{x}) = g(\boldsymbol{x})\right] \geq (1-\varepsilon).$$

Otherwise they are *ε-far*.

(ii) A Boolean function f is *ε-close* to having property $\mathcal{P}$ if there exists some function $g \in \mathcal{P}$ such that f and g are ε-close.

Remark. Note that by definition of the correlation, two Boolean functions are ε-close if and only if

$$\mathbf{E}_{\boldsymbol{x}\in\{-1,1\}^n}\left[f(\boldsymbol{x})g(\boldsymbol{x})\right] \geq (1-2\varepsilon).$$

In our case we are interested in the property of being linear, i.e. $\mathcal{P}_{lin} \coloneqq \{\chi_S : S \subseteq [n]\}$ and with Definition 3 (ii') can be reformulated as being ε-close to $\mathcal{P}_{lin}$.

How can (i') be understood in this context? It is favorable to interpret it as a *property test*:

Definition 4 (BLR linearity test (Blum, Luby, Rubinfeld))**.** Given blackbox access to a Boolean function f do the following steps:

1. Pick $\boldsymbol{x}$ and $\boldsymbol{y}$ independently and uniformly at random from $\{-1,1\}^n$.
2. Query f on $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{x} \circ \boldsymbol{y}$.
3. "Accept" iff $f(\boldsymbol{x})f(\boldsymbol{y})f(\boldsymbol{x} \circ \boldsymbol{y}) = 1$.

Using this definition (i') can be understood in the sense that the probability of BLR accepting the function f is large, more precisely

$$\mathbf{P}_{x,y\in\{-1,1\}^n}\left[\mathsf{BLR}(f) \textit{ accepts }\right] \geq (1-\varepsilon).$$

(Here $\mathbf{P}_{x,y\in\{-1,1\}^n}$ denotes the product measure on $\{-1,1\}^n \times \{-1,1\}^n$.)

If we can prove that (i') implies (ii') then we have shown that for the linearity property there exists a querying algorithm making only 3 queries such that whenever f is ε-far from being linear then the algorithm accepts f with probability smaller than $1-\varepsilon$.
The existence of such a querying algorithm can also be shown for other properties and motivates the following definition:

Definition 5 (Locally testable property)**.** A property $\mathcal{P}$ of Boolean functions is called *locally testable* if there exists a randomized querying algorithm $\mathcal{T}$ making at most $\mathcal{O}(1)$ queries such that:

(i) If $f \in \mathcal{P}$ then $\mathbf{P}[\mathcal{T}(f) \textit{ accepts}] = 1$.

(ii) If f is ε-far from having property $\mathcal{P}$ then $\mathbf{P}[\mathcal{T}(f)\ \textit{accepts}] \leq 1 - \Omega(\varepsilon)$.

A more general definition of testability which relates the closeness to the property in question with the number of queries is due to Rubinfeld and Sudan:

Definition 6 (Testable property)**.** A property $\mathcal{P}$ of Boolean functions is *testable with $q(\varepsilon)$ queries* if there exists a randomized algorithm $\mathcal{T}$ (which gets ε as input) such that for all $\varepsilon > 0$ it makes $q(\varepsilon)$ random queries and satisfies:

(i) If $f \in \mathcal{P}$ then $\mathbf{P}[\mathcal{T}(f)\ \textit{accepts}] \geq \frac{2}{3}$.

(ii) If f is ε-far from $\mathcal{P}$ then $\mathbf{P}[\mathcal{T}(f)\ \textit{accepts}] \leq \frac{1}{3}$.

Remark. The choice of the bounds $\frac{2}{3}$ and $\frac{1}{3}$ is somewhat arbitrary and can be boosted to $1 - \delta$ and δ respectively with a few more queries.

The aim of today's course will be to show that the property $\mathcal{P}_{lin}$ of being linear is

- locally testable (this is the implication (i') $\Rightarrow$ (ii'))
- testable with $\mathcal{O}\left(\frac{1}{\varepsilon}\right)$ queries.

4 Linearity testing of Boolean functions

4.1 Testability of linearity

We will prove the testability of linearity in three steps:

1. Express the "acceptance probability" of the BLR-test for an arbitrary Boolean function f in terms of its Fourier coefficients.
2. Prove local testability of linearity using this representation.
3. Prove testability by executing the BLR-test multiple times.

Lemma 1. *Let $f : \{-1,1\}^n \to \{-1,1\}$ be a Boolean function. Then the following holds:*

$$\mathbf{P}_{\boldsymbol{x},\boldsymbol{y}\in\{-1,1\}^n}[\mathit{BLR}(f)\ \textit{accepts}] = \frac{1}{2} + \frac{1}{2}\sum_{S\subseteq[n]} \hat{f}(S)^3. \tag{2}$$

Proof. The proof consists primarly in writing the probability term in a suitable way and using the Fourier expansion of f.
We write the indicator function $\mathbb{1}_{\{\mathsf{BLR}(f)\ \textit{accepts}\}}$ in the following way for $\boldsymbol{x}, \boldsymbol{y} \in \{-1,1\}^n$

$$\mathbb{1}_{\{\mathsf{BLR}(f)\ \textit{accepts}\}}(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{2} + \frac{1}{2}f(\boldsymbol{x})f(\boldsymbol{y})f(\boldsymbol{x}\circ\boldsymbol{y}).$$

Hence we can express the left-hand side of (2) by

$$\mathbf{P}_{\boldsymbol{x},\boldsymbol{y}\in\{-1,1\}^n}[\mathsf{BLR}(f)\ \textit{accepts}] = \frac{1}{2} + \frac{1}{2}\mathbf{E}_{\boldsymbol{x},\boldsymbol{y}}[f(\boldsymbol{x})f(\boldsymbol{y})f(\boldsymbol{x}\circ\boldsymbol{y})].$$

Using the Fourier expansion of f we find by linearity of the expectation:

$$\begin{aligned}
&\mathbf{P}_{\boldsymbol{x},\boldsymbol{y}\in\{-1,1\}^n}[\mathsf{BLR}(f)\ \textit{accepts}] \\
&\quad= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U\subseteq[n]} \left(\hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{\boldsymbol{x},\boldsymbol{y}}[\chi_S(\boldsymbol{x})\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{x}\circ\boldsymbol{y})]\right) \qquad (3) \\
&\quad= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U\subseteq[n]} \left(\hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{\boldsymbol{x},\boldsymbol{y}}[\chi_S(\boldsymbol{x})\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{x})\chi_U(\boldsymbol{y})]\right), \qquad (4)
\end{aligned}$$

where we used the linearity of parity functions.
Since with respect to the product measure $\mathbf{P}_{\boldsymbol{x},\boldsymbol{y}}$, two functions $f(\boldsymbol{x}), f(\boldsymbol{y})$ are independent we finally get

$$\begin{aligned}
&\mathbf{P}_{\boldsymbol{x},\boldsymbol{y}\in\{-1,1\}^n}[\mathsf{BLR}(f)\ \textit{accepts}] \\
&\quad= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U\subseteq[n]} \left(\hat{f}(S)\hat{f}(T)\hat{f}(U)\cdot\mathbf{E}_{\boldsymbol{x}}[\chi_S(\boldsymbol{x})\chi_U(\boldsymbol{x})]\mathbf{E}_{\boldsymbol{y}}[\chi_T(\boldsymbol{y})\chi_U(\boldsymbol{y})]\right) \qquad (5) \\
&\quad= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U\subseteq[n]} \left(\hat{f}(S)\hat{f}(T)\hat{f}(U)\cdot\langle\chi_S,\chi_U\rangle\cdot\langle\chi_T,\chi_U\rangle\right) \qquad (6) \\
&\quad= \frac{1}{2} + \frac{1}{2} \sum_{S\subseteq[n]} \hat{f}(S)^3, \qquad (7)
\end{aligned}$$

since the parity functions are a orthonormal basis of all Boolean functions which means in particular

$$\langle\chi_S,\chi_U\rangle\cdot\langle\chi_T,\chi_U\rangle = \begin{cases} 1 & \text{if } S=U=T, \\ 0 & \text{else.} \end{cases}$$

□

Theorem 1 (Local testability of linearity)**.** *The property of being linear is locally testable using the BLR-test and*

$$\mathbf{P}[\textit{BLR}(f)\ \textit{accepts}] < 1-\varepsilon,$$

for every f which is ε-far from being linear.

Proof. If $f=\chi_S$ for some $S\subseteq[n]$ then the BLR-test accepts with probability one by Definition 1 and the subsequent discussion.
Let now f be ε-far from being linear and assume for the sake of contradiction

$$\mathbf{P}_{\boldsymbol{x},\boldsymbol{y}}[\mathsf{BLR}(f)\ \textit{accepts}] \geq (1-\varepsilon).$$

Then using Lemma 1 we have $1-\varepsilon \leq \frac{1}{2}+\frac{1}{2}\sum_{S\subseteq[n]}\hat{f}(S)^3$ and consequently

$$1-2\varepsilon \leq \sum_{S\subseteq[n]}\hat{f}(S)^3 \leq \left(\max_{S\subseteq[n]}\hat{f}(S)\right)\cdot\sum_{S\subseteq[n]}\hat{f}(S)^2 = \left(\max_{S\subseteq[n]}\hat{f}(S)\right), \qquad (8)$$

where we used Parseval's theorem for the last equality.
Hence there exists a subset $T\subseteq[n]$ such that

$$1-2\varepsilon \leq \mathbf{E}_{\boldsymbol{x}\in\{-1,1\}^n}\left[f(x)\chi_T(x)\right].$$

By the remark after Definition 3 this implies that f is ε-close to $\chi_T$which is a contradiction to our assumption that f is ε-far from being linear. □

Theorem 2 (Testability of linearity)**.** *The property of being linear is testable with* $\mathcal{O}\left(\frac{1}{\varepsilon}\right)$ *queries.*

Proof. We define the randomized query algorithm $\mathcal{T}(f, \varepsilon)$ taking f and ε as arguments as follows:

- Run $\mathsf{BLR}(f)$ $2/\varepsilon$ times, independently.
- Accept iff every $\mathsf{BLR}(f)$ accepts.

Note that if f is linear then $\mathcal{T}(f)$ accepts with probability one, since $\mathsf{BLR}(f)$ accepts in every case.
Assume now that f is ε-far from being linear. Since the BLR queries are independent from each other, we get from Theorem 1

$$\mathbf{P}\left[\mathcal{T}(f) \text{ accepts}\right] < (1-\varepsilon)^{2/\varepsilon} \tag{9}$$

$$= \left(\left(1 - \frac{1}{(1/\varepsilon)}\right)^{1/\varepsilon}\right)^2. \tag{10}$$

Since the last term converges for $\varepsilon \to 0$ to e^{-2} we get for ε small enough (e.g. $\varepsilon \leq \frac{1}{2}$)

$$\mathbf{P}\left[\mathcal{T}(f) \text{ accepts}\right] < \frac{1}{3}.$$

□

4.2 Local decodability of linearity

Theorem 2 gives us a possibility to determine with high probability if a given Boolean function is ε-close to a parity function.
Still the question remains how we can guess the right parity function using a minimum number of queries? The following result tells us that there is an algorithm making only two queries and predicts with high probability the right parity function; this property is called *decodability*.

Theorem 3 (Local docadiblity of linearity)**.** *The property* $\mathcal{P}_{lin}$ *of being a parity function is* locally docodable *with 2 queries, i.e. there exists a randomized 2-query algorithm* $\mathcal{T}$ *having access to a Boolean function* $f : \{-1,1\}^n \to \{-1,1\}$ *and taking strings* $\boldsymbol{x} \in \{-1,1\}^n$ *as input such that:*
If f *is* ε*-close to a parity function* χ_S*, then for any (!)* $\boldsymbol{x} \in \{-1,1\}^n$ *one has*

$$\mathbf{P}\left[\mathcal{T}(\boldsymbol{x}) = \boldsymbol{x}_S\right] \geq 1 - 2\varepsilon. \tag{11}$$

Remark. Note that the probability in (11) refers to the internal randomness of the algorithm $\mathcal{T}(\boldsymbol{x})$ for an arbitrary but fixed $\boldsymbol{x} \in \{-1,1\}^n$.
In particular one can repeat the test multiple times for the same $\boldsymbol{x}$ to boost the probability in (11).
This is very different from the statement that for *almost every* $\boldsymbol{x} \in \{-1,1\}^n$ the algorithm computes the correct value (this would be denoted by $\mathbf{P}_{\boldsymbol{x}\in\{-1,1\}^n}[\mathcal{T}(\boldsymbol{x}) = \chi_S(\boldsymbol{x})] \geq 1 - 2\varepsilon$ and is much more trivial (just query f on $\boldsymbol{x}$).

Proof. Let $\boldsymbol{x} \in \{-1,1\}^n$ be fixed. We define $\mathcal{T}$ as follows: For the given $\boldsymbol{x}$

- Pick $\boldsymbol{y} \in \{-1,1\}^n$ uniformly at random.
- Return $\mathcal{T}(\boldsymbol{x}) \coloneqq f(\boldsymbol{y})f(\boldsymbol{x} \circ \boldsymbol{y})$.

We make the following observation: Since $\boldsymbol{y}$ is uniformly distributed and f is ε-close to χ_S one has

$$\mathbf{P}_{\boldsymbol{y}\in\{-1,1\}^n}[f(\boldsymbol{y}) = \chi_S(\boldsymbol{y})] \geq 1 - \varepsilon. \tag{12}$$

Similary $\boldsymbol{x} \circ \boldsymbol{y}$ is again uniformly distributed (but not independent of $\boldsymbol{y}$) and

$$\mathbf{P}_{\boldsymbol{y}\in\{-1,1\}^n}[f(\boldsymbol{x} \circ \boldsymbol{y}) = \chi_S(\boldsymbol{x} \circ \boldsymbol{y})] \geq 1 - \varepsilon. \tag{13}$$

Using the union bound and (12) and (13) we find

$$\begin{aligned}
&\mathbf{P}_{\boldsymbol{y}}[\{f(\boldsymbol{x} \circ \boldsymbol{y}) = \chi_S(\boldsymbol{x} \circ \boldsymbol{y})\} \cap \{f(\boldsymbol{y}) = \chi_S(\boldsymbol{y})\}] && (14)\\
&\quad = \mathbf{P}_{\boldsymbol{y}}[f(\boldsymbol{x} \circ \boldsymbol{y}) = \chi_S(\boldsymbol{x} \circ \boldsymbol{y})] + \mathbf{P}_{\boldsymbol{y}}[f(\boldsymbol{y}) = \chi_S(\boldsymbol{y})] && (15)\\
&\quad - \mathbf{P}_{\boldsymbol{y}}[\{f(\boldsymbol{x} \circ \boldsymbol{y}) = \chi_S(\boldsymbol{x} \circ \boldsymbol{y})\} \cup \{f(\boldsymbol{y}) = \chi_S(\boldsymbol{y})\}] && (16)\\
&\quad \geq (1 - \varepsilon) + (1 - \varepsilon) - 1 && (17)\\
&\quad = 1 - 2\varepsilon. && (18)
\end{aligned}$$

But in this case that booth equalities hold we have

$$\mathcal{T}(\boldsymbol{x}) = f(\boldsymbol{y})f(\boldsymbol{x} \circ \boldsymbol{y}) = \boldsymbol{y}_S\boldsymbol{x}_S\boldsymbol{y}_S = \boldsymbol{x}_S,$$

hence

$$\mathbf{P}_{\boldsymbol{y}}[\mathcal{T}(\boldsymbol{x}) = \boldsymbol{x}_S] \geq 1 - 2\varepsilon.$$

□