In this document we are going to show that given two indistinguishable distribution ensembles $\{X_i\}_{i\in I}$ and $\{Y_i\}_{i\in I}$ and some function g that maps outputs of the ensembles to some real number, it is not correct to argue that if $\{X_i\}_{i\in I}$ and $\{Y_i\}_{i\in I}$ are computationally indistinguishable then the expectations of $g(\{X_i\}_{i\in I})$ and $g(\{Y_i\}_{i\in I})$ are the same.

To provide a simple counter example, let $\{X_i\}_{i\in I}$ and $\{Y_i\}_{i\in I}$ be distributions on $\{0,1\}^n$ for some natural number n. $\{X_i\}_{i\in I}$ takes uniform distribution on this set and $\{Y_i\}_{i\in I}$ be 'almost uniform' such that:

$$\Pr[\{Y_i\}_{i\in I} = y] = \begin{cases} 0, & \text{if } y = 0^n \\ 2^{-n+1}, & \text{if } y = 1^n \\ 2^{-n}. & \text{otherwise} \end{cases}$$

Clearly $\{X_i\}_{i\in I}$ and $\{Y_i\}_{i\in I}$ have identical distributions on everything except 0^n and 1^n. But the probability of observing those outputs is 2^{-n+1}, which is negligible in n. Therefore there is no PPT adversary who can distinguish $\{X_i\}_{i\in I}$ and $\{Y_i\}_{i\in I}$. Hence, computational indistinguishability with security parameter 1^n.

On the other hand, if we let $g : \{0,1\}^n \to \mathbb{R}$ be $g(x) = 2^n \cdot \mathbb{1}(x = 0^n)$, then we see that $\mathbb{E}(g(\{X_i\}_{i\in I})) = 1$ whereas $\mathbb{E}(g(\{Y_i\}_{i\in I})) = 0$, so one cannot really conclude anything about expectation of functions taking computationally indistinguishable inputs. Our goal is to bound the difference in expectation for some function and two computationally indistinguishable distribution ensembles.

Lemma 1. *Let $\{X_i\}_{i\in I}$ and $\{Y_i\}_{i\in I}$ be two computationally indistinguishable distribution ensembles with security parameter 1^n. Let Z be the union of the supports of the two distribution ensembles. We write Z as a disjoint union of Z_- and Z_+ such that*

1. $\forall z \in Z_-, |\Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z]| \leq 0$,
2. $\forall z \in Z_+, \Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z] > 0$.

Then

$$\sum_{z\in Z_+} \Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z] \leq \mathsf{negl}(n) \tag{1}$$

Proof. We proceed by contradiction and show that if equation 1 does not hold then there is an PPT adversary D to distinguish $\{X_i\}_{i\in I}$ and $\{Y_i\}_{i\in I}$ with non-negligible probability. We begin by considering an arbitrary adversary D and work out its advantage.

$$\begin{aligned}
&\mathrm{Adv}_D(n) && (2)\\
&= |\Pr[D(z) = 1 \mid z \leftarrow\!\!\$\, \{X_i\}_{i\in I}] - \Pr[D(z) = 1 \mid z \leftarrow\!\!\$\, \{Y_i\}_{i\in I}]| && (3)\\
&= \left| \sum_{\bar{z}\in Z} \Pr[D(z) = 1 \mid z \leftarrow\!\!\$\, \{X_i\}_{i\in I}, z = \bar{z}] \Pr[\{X_i\}_{i\in I} = \bar{z}] - \sum_{\bar{z}\in Z} \Pr[D(z) = 1 \mid z \leftarrow\!\!\$\, \{Y_i\}_{i\in I}, z = \bar{z}] \Pr[\{Y_i\}_{i\in I} = \bar{z}] \right| && (4)\\
&= \left| \sum_{\bar{z}\in Z} \Pr[D(z) = 1 \mid z \leftarrow\!\!\$\, \{X_i\}_{i\in I}, z = \bar{z}] \Pr[\{X_i\}_{i\in I} = \bar{z}] - \Pr[D(z) = 1 \mid z \leftarrow\!\!\$\, \{Y_i\}_{i\in I}, z = \bar{z}] \Pr[\{Y_i\}_{i\in I} = \bar{z}] \right| && (5)\\
&= \left| \sum_{\bar{z}\in Z} \Pr[D(\bar{z}) = 1] \left(\Pr[\{X_i\}_{i\in I} = \bar{z}] - \Pr[\{Y_i\}_{i\in I} = \bar{z}]\right) \right| && (6)
\end{aligned}$$

Since we know $\forall z \in Z_+, \Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z] > 0$, the part of sum for elements in Z_+ in

equation 6 are all non-negative. Therefore, we get

$$\mathrm{Adv}_D(n) \tag{7}$$

$$= \left| \sum_{\bar{z} \in Z_+} \Pr[D(\bar{z}) = 1] \left(\Pr[\{X_i\}_{i \in I} = \bar{z}] - \Pr[\{Y_i\}_{i \in I} = \bar{z}] \right) + \sum_{\bar{z} \in Z_1} \Pr[D(\bar{z}) = 1] \left(\Pr[\{X_i\}_{i \in I} = \bar{z}] - \Pr[\{Y_i\}_{i \in I} = \bar{z}] \right) \right| \tag{8}$$

$$\geq \left| \sum_{\bar{z} \in Z_+} \Pr[D(\bar{z}) = 1] \left(\Pr[\{X_i\}_{i \in I} = \bar{z}] - \Pr[\{Y_i\}_{i \in I} = \bar{z}] \right) \right| \tag{9}$$

Now we pick D such that on all outcomes in Z_1, it returns 1, the advantage becomes

$$\mathrm{Adv}_D(n) \tag{10}$$

$$\geq \left| \sum_{\bar{z} \in Z_+} \Pr[D(\bar{z}) = 1] \left(\Pr[\{X_i\}_{i \in I} = \bar{z}] - \Pr[\{Y_i\}_{i \in I} = \bar{z}] \right) \right| \tag{11}$$

$$= \left| \sum_{\bar{z} \in Z_+} \Pr[\{X_i\}_{i \in I} = \bar{z}] - \Pr[\{Y_i\}_{i \in I} = \bar{z}] \right| \tag{12}$$

$$= \sum_{\bar{z} \in Z_+} \Pr[\{X_i\}_{i \in I} = \bar{z}] - \Pr[\{Y_i\}_{i \in I} = \bar{z}] \tag{13}$$

But we begin by assuming equation 13 is non-negligible, so D differentiates $\{X_i\}_{i \in I}$ and $\{Y_i\}_{i \in I}$ with non-negligible probability. Since all D does is checking its input, the adversary is efficient. Therefore, D breaks computational indistinguishability between $\{X_i\}_{i \in I}$ and $\{Y_i\}_{i \in I}$. This gives a contradiction. Therefore, equation 1 must hold. □

Corollary 2. *Let $\{X_i\}_{i \in I}$ and $\{Y_i\}_{i \in I}$ be two computationally indistinguishable distribution ensembles with security parameter 1^n. Let Z be the union of the supports of the two distribution ensembles. We write Z as a disjoint union of Z_- and Z_+ such that*

1. $\forall z \in Z_-, |\Pr[\{X_i\}_{i \in I} = z] - \Pr[\{Y_i\}_{i \in I} = z]| \leq 0$,
2. $\forall z \in Z_+, \Pr[\{X_i\}_{i \in I} = z] - \Pr[\{Y_i\}_{i \in I} = z] > 0$.

Then

$$\sum_{z \in Z_-} \Pr[\{Y_i\}_{i \in I} = z] - \Pr[\{X_i\}_{i \in I} = z] \leq \mathsf{negl}(n) \tag{14}$$

Proof. The result follows immediately from the lemma above by symmetry. □

Theorem 3. *Let $\{X_i\}_{i \in I}$ and $\{Y_i\}_{i \in I}$ be two computationally indistinguishable distribution ensembles with security parameter 1^n. Let Z be the union of the supports of the two distribution ensembles. We write Z as a disjoint union of Z_- and Z_+ such that*

1. $\forall z \in Z_-, |\Pr[\{X_i\}_{i \in I} = z] - \Pr[\{Y_i\}_{i \in I} = z]| \leq 0$,
2. $\forall z \in Z_+, |\Pr[\{X_i\}_{i \in I} = z] - \Pr[\{Y_i\}_{i \in I} = z]| > 0$.

Then

$$\sum_{z \in Z} |\Pr[\{X_i\}_{i \in I} = z] - \Pr[\{Y_i\}_{i \in I} = z]| \leq \mathsf{negl}(n) \tag{15}$$

Proof. We get this by combining the two lemmas above. □

Theorem 4. *Let $\{X_i\}_{i\in I}$ and $\{Y_i\}_{i\in I}$ be two computationally indistinguishable distribution ensembles with security parameter 1^n. Let Z be the union of the supports of the two distribution ensembles. We write Z as a disjoint union of Z_- and Z_+ such that*

1. $\forall z \in Z_-, |\Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z]| \leq 0$,
2. $\forall z \in Z_+, |\Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z]| > 0$.

Let $g : Z \to \mathbb{R}$ be a function. Define $|g| = \max_{z\in Z}\{g(z), -g(z)\}$. Then

$$|\mathbb{E}[\{X_i\}_{i\in I}] - \mathbb{E}[\{Y_i\}_{i\in I}]| \leq |g| \cdot \mathsf{negl}(n)\,. \tag{16}$$

Proof. We prove the statement by manipulating probabilities.

$$\begin{aligned}
&|\mathbb{E}[\{X_i\}_{i\in I}] - \mathbb{E}[\{Y_i\}_{i\in I}]| \\
=&\left|\sum_{z\in Z} g(z)\left(\Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z]\right)\right| \\
=&\left|\sum_{z\in Z_+} g(z)\left(\Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z]\right)\right| + \left|\sum_{z\in Z_-} g(z)\left(\Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z]\right)\right| \\
\leq&|g| \sum_{z\in Z_-} |\Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z]| + |g| \sum_{z\in Z_+} |\Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z]| \\
=&|g| \sum_{z\in Z} |\Pr[\{X_i\}_{i\in I} = z] - \Pr[\{Y_i\}_{i\in I} = z]| \\
=&|g| \cdot \mathsf{negl}(n)
\end{aligned}$$

□