Reduction From Public Coin HVSZK To Statistical Closeness

Personal Notes

Ari Biswas
University Of Warwick
aribiswas3@gmail.com

1 The Goal

We have some promise problem $\Pi = (\Pi^{\mathsf{Yes}}, \Pi^{\mathsf{No}})$ and a public coin HVSZK proof system (P, V) for Π with constant completeness and soundness error. Let Sim denote the simulator for the proof system described above. Given $x \in \Pi^{\mathsf{Yes}} \cup \Pi^{\mathsf{No}}$, Fix $\kappa = |x|$, we want to use the output of the simulator $\mathsf{Sim}(1^\kappa, x)$[1] to come up with two distributions X and Y such that

$$x \in \Pi^{\mathsf{Yes}} \implies d_{\mathtt{TV}}(X, Y) \leq \frac{1}{3}$$

$$x \in \Pi^{\mathsf{No}} \implies d_{\mathtt{TV}}(X, Y) > \frac{2}{3}$$

So when we write $x \mapsto (X, Y)$ we mean that an HVSZK proof for Π Karp reduces to $\overline{SD}$. Throughout this document $\kappa = |x|$ is the security parameter, and $v(\kappa)$ denotes the number of messages the prescribed verifier sends to the prover, where v is some polynomial.

1.1 Simulation As A Virtual Interaction

For honest verifier zero-knowledge proofs, simulation can be viewed as an interaction between the virtual verifier and a virtual prover. The simulator is responsible for constructing the virtual verifier and virtual prover. The virtual prover strategy (also called simulation-based prover) $\mathsf{P}_{\mathsf{Sim}}$ is constructed as follows:

Given an $x \in \Pi^{\mathsf{Yes}} \cup \Pi^{\mathsf{No}}$, and a conversation history γ (consisting of $2i$ messages exchanged between the virtual prover and virtual verifier), $\mathsf{P}_{\mathsf{Sim}}$ responds as follows:

- If $\mathsf{Sim}(x)$ outputs transcripts beginning with γ with probability 0, then $\mathsf{P}_{\mathsf{Sim}}$ outputs $\perp$.
- Otherwise, $\mathsf{P}_{\mathsf{Sim}}$ responds with β with probability p_β where

  $$p_\beta = \Pr[\mathsf{Sim}(x)_{2i+1} = (\gamma, \beta) | \mathsf{Sim}(x)_{2i} = \gamma]$$

  In simpler words, the virtual prover is simply the simulator pretending to be a real-world prover (but without any of the computation resources of a real-world prover).

From the definition of HVSZK we get that if $x \in \Pi^{\mathsf{Yes}}$, we must have both

[1] Henceforth, I'll often drop the security parameter from the notation to make it cleaner and write $\mathsf{Sim}(1^\kappa, x)$ as just $\mathsf{Sim}(x)$.

1. The virtual verifier must accept with high probability (given by completeness).
2. The virtual verifier "behaves like" the real verifier (see below for a formal definition).

However, the definition of HVSZK does not say anything explicitly about the quality of simulation for $x \in \Pi^{\mathsf{No}}$. Despite this, it is possible to show that at least one of the following must be true for No instances:

1. The virtual verifier must accept with low probability.
2. The virtual verifier must behave very differently from the real verifier.

Note that if both of these statements are false, we have a virtual prover that convinces the virtual verifier with high probability. This virtual verifier is very close to the real verifier. Thus, if I swapped the virtual and real verifier, we would break soundness as we have found a cheating prover strategy $\mathsf{P}_{\mathsf{Sim}}$ that gets the verifier to accept. Note that we have not yet technically defined what "behaves" as the real verifier means, but we do so in the next section. In summary, this is the game

Given a simulator Sim for an HVSZK proof system for a promise problem Π and an instance $x \in \Pi$, our goal is to construct distributions X and Y that enable us to distinguish between Yes and No using the statistical difference problem.

1.2 The Behaviour Of The Virtual Verifier and The Real Verifier

Next, we formalise what it means for a virtual verifier to behave like a real verifier. Let $v(\kappa)$ denote the number of messages the verifier sends to a real prover. For $j \leq 2v(\kappa) + 2$, we refer to a tuple of strings $(m_1, \ldots, m_j)$ as a conversation transcript if the even indexed messages correspond to the messages sent by the verifier and the odd indices of messages represent messages sent by the prover. For $i = 1, \ldots, v(\kappa)$ define

$X_i(x)$: Let $\gamma \xleftarrow{\$} \mathsf{Sim}(x)$ and let γ_{2_i} denote the first $2i$ messages exchanged. Set $X_i(x) = \gamma_{2i}$.

$Y_i(x)$: Let $\gamma \xleftarrow{\$} \mathsf{Sim}(x)$ and let γ_{2_i-1} denote the first $2i-1$ messages exchanged. Let l_i represent a polynomial which computes the number of public coins $l_i(x, \gamma_{2_i-1})$[a] the real verifier sends in message round i. Let $r \xleftarrow{\$} \{0,1\}^{l_i}$. Then set $Y_i(x) = (\gamma_{2i-1}, r)$.

[a] We abuse notation to use l_i to denote both the polynomial and the polynomial evaluation i.e. $l_i = l_i(x, \gamma_{2i-1})$.

In X_i, the i'th message is computed per the virtual verifier strategy defined by the Simulator. In Y_i, the i'th verifier message is chosen uniformly and independently of the history, like the real verifier would have done when interacting with any arbitrary prover. Define $\delta_i(x)$ as

$$\delta_i(x) := SD(X_i(x), Y_i(x)) \tag{1}$$

and it quantifies how much the virtual verifier differs from the real verifier in computing the i'th message. If $\delta_i(x)$ is small for all $i = 1, \ldots, v(\kappa)$, then it implies that the virtual verifier behaves like the real verifier. If this is the case and the virtual verifier accepts with high probability, then $x \in \Pi^{\mathsf{Yes}}$. On the other hand, if the virtual prover accepts with low probability $x \in \Pi^{\mathsf{No}}$.

1.3 Key Technical Lemmas

We need a few technical lemmas before we can prove our final Karp Reduction in Section 2. Lemma 1 upper bounds the difference between the real and virtual verifier in terms of the simulator deviation. If $x \in \Pi^{\mathsf{Yes}}$, then the simulation deviation is small, and hence the virtual verifier is close to the real verifier.

Lemma 1 Let $\mu = d_{\mathtt{TV}}\Big(\mathsf{Sim}(x), \mathrm{view}[(\mathsf{P}, \mathsf{V})(x)]\Big)$ be the simulator deviation from the view of the real verifier in the HVSZK proof system (P, V) and $v(\kappa)$ be the number of messages the real verifier sends to the prover P. For all $i = 1, \dots, v(\kappa)$ we have

$$\delta_i(x) \leq 2\mu \tag{2}$$

Proof. Dropping x from the notation we have

$$\begin{aligned} \delta_i &= d_{\mathtt{TV}}(X_i, Y_i) &(3)\\ &\leq d_{\mathtt{TV}}(X_i, (\mathsf{P}, \mathsf{V})_{2_i}) + d_{\mathtt{TV}}((\mathsf{P}, \mathsf{V})_{2i}, Y_i) &(4) \end{aligned}$$

The last inequality comes from the triangle inequality. Note that X_i is the same distribution as Sim_{2_i}, and thus

$$d_{\mathtt{TV}}(X_i, (\mathsf{P}, \mathsf{V})_{2_i}) = d_{\mathtt{TV}}(\mathsf{Sim}_{2_i}, (\mathsf{P}, \mathsf{V})_{2_i}) \leq d_{\mathtt{TV}}(\mathsf{Sim}, (\mathsf{P}, \mathsf{V})) \tag{5}$$

On the other hand Y_i is obtained from Sim_{2i-1} and then concatenating l_i random coins (mimicking what the verifier would do for its i'th turn). If we did the same thing to $(\mathsf{P}, \mathsf{V})_{2i-1}$ we get $(\mathsf{P}, \mathsf{V})_{2_i}$. Based on the fact below, as we are applying the same random function on both distributions, we have

$$d_{\mathtt{TV}}((\mathsf{P}, \mathsf{V})_{2i}, Y_i) \leq d_{\mathtt{TV}}((\mathsf{P}, \mathsf{V})_{2i-1}, \mathsf{Sim}_{2i-1}) \leq d_{\mathtt{TV}}(\mathsf{Sim}, (\mathsf{P}, \mathsf{V})) \tag{6}$$

Combining (5) and (6) gives us our proof. □

Fact 1 For any two distributions A and B on $\mathcal{X}$, any randomised procedure f on $\mathcal{X}$, if we have $d_{\mathtt{TV}}(f(A), f(B)) \leq d_{\mathtt{TV}}(A, B)$

In the Yes instances, Lemma 1 guarantees that the virtual and real behaviour is close (as μ is small). This also guarantees that the virtual verifier accepts with high probability (otherwise, it would not be similar to the real verifier, which by the completeness property, accepts with high probability). Thus at this point, it might be tempting to declare $\bar{X} = (X_1(x), \otimes \dots \otimes X_{v(\kappa)}(x)$ and $\bar{Y} = (Y_1(x) \otimes \dots, \otimes Y_{v(\kappa)}(x)$ and claim we are done. As $x \in \Pi^{\mathsf{Yes}}$ implies, that $\sum_{i=1}^{v(\kappa)} \delta_i(x) = d_{\mathtt{TV}}(\bar{X}, \bar{Y}) \leq 2v(\kappa)\mu$ which is small (we will formalise how small later).

However, this definition of $\bar{X}$ and $\bar{Y}$ does not quite handle the No instances as efficiently. The next lemma relates closeness between the virtual and real verifier and the virtual verifier's acceptance probability.

Lemma 2 Let p be the probability that $\mathsf{Sim}(x)$ outputs an accepting transcript, i.e. the virtual verifier accepts. Let $q = \max_{\widetilde{\mathsf{P}}} \mathsf{out}[(\widetilde{\mathsf{P}}, \mathsf{V}) = \text{accept}]$

$$p - q \leq d_{\texttt{TV}}\Big(\text{view}[(\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})(x)], \mathsf{Sim}(x)\Big) \leq \sum_{i=1}^{v(\kappa)} \delta_i(x) \tag{7}$$

Proof. We will prove this by induction. The base case for $j = 0$, by the definition of δ_i we get the distributions are the same. Note for any j, we obtain $(\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+1}$ by applying the same procedure to $(\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j}$ as we do by when we go from Sim_{2j} to Sim_{2j+1}. Thus by above fact

$$d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+1}, \mathsf{Sim}_{2j+1}) \leq d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j}, \mathsf{Sim}_{2j}) \tag{8}$$

Induction step. Assume $d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j}, \mathsf{Sim}_{2j}) \leq \sum_{i=0}^{j} \delta_i$.

$$\begin{aligned} d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+2}, \mathsf{Sim}_{2j+2}) &= d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+2}, X_{j+1}) \quad (9) \\ &\leq d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+2}, Y_{j+1}) + d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+2}, Y_{j+1}) \quad (10) \end{aligned}$$

$(\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+2}$ obtained from $(\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+1}$ by applying the same random process as how get Y_{j+1} from Sim_{2j+1}. Thus,

$$\begin{aligned} d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+2}, \mathsf{Sim}_{2j+2}) &\leq d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+1}, \mathsf{Sim}_{2j+1}) + d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+2}, Y_{j+1}) \quad (11) \\ &\leq d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j}, \mathsf{Sim}_{2j}) + d_{\texttt{TV}}((\mathsf{P}_{\mathsf{Sim}}, \mathsf{V})_{2j+2}, Y_{j+1}) \quad (12) \\ &\leq \sum_{i=0}^{j} \delta_i + \delta_j \quad (13) \end{aligned}$$

□

If $x \in \Pi^{\mathsf{No}}$, then $q \leq \frac{1}{3}$ (from soundness). Here p is the probability with which the virtual verifier accepts. Thus, if p is high, it implies that $p - q$ is large; therefore, the virtual and real verifiers are not close, i.e. $d_{\texttt{TV}}(\bar{X}, \bar{Y})$ is large as we desire. But what happens when p is small? Now we have no guarantees about distance between $\bar{X}$ and $\bar{Y}$ as we desire. So clearly, the current definitions of the distributions is not enough. Thus we define $\bar{X}' = \bar{X} \otimes X_0$ and $\bar{Y}' = \bar{Y} \otimes Y_0$, where

X_0: Output 1 always.

Y_0: Run the simulator $216 \log(v(\kappa))$ times and output the decision made by most virtual verifiers.

Now if $x \in \Pi^{\mathsf{Yes}}$, we want $d_{\texttt{TV}}(X_0, Y_0)$ to be small i.e. we want to upper bound the probability that $\Pr[Y_0 = 0]$. Let $w = 216 \log(v(\kappa))$ and $A_1, \ldots, A_w$ be the independent accept/reject outputs of the simulator $\mathsf{Sim}(x)$. We have

$$d_{\texttt{TV}}(X_0, Y_0) = \Pr[Y_0 = 0] = \Pr[\texttt{majority}(A_1, \ldots, A_w) = 0] = \Pr\left[\frac{1}{w}\sum_{j=1}^{w} A_j \leq \frac{1}{2}\right]$$

When $x \in \Pi^{\mathsf{Yes}}$, from completeness and the definition of simulator deviation in zero knowledge, the simulator must output accept with probability $\frac{2}{3} - \mu$ i.e. for any $j \in [w]$, $\Pr[A_j = 1] = p \geq \frac{2}{3} - \mu \geq \frac{7}{12}$. The maximum deviation between X_0 and Y_0 occurs when $p = \frac{7}{12}$, and thus we can re-write

$$\begin{aligned}
\Pr\left[\frac{1}{w}\sum_{j=1}^{w} A_j \leq \frac{1}{2}\right] &= \Pr\left[\frac{1}{w}\sum_{j=1}^{w} A_j \leq \frac{7}{12} - \frac{1}{12}\right] && (14)\\
&\leq \mathsf{exp}\left(-2 \cdot w \cdot \frac{1}{12^2}\right) && (15)\\
&\leq \frac{1}{(12v)^3} && (16)
\end{aligned}$$

Equation (15) comes from the Chernoff bound described in Lemma 5.

2 Reducing To Statistical Closeness

Finally, we are ready to prove the final reduction.

Theorem 1 Let Sim be the simulator for an HVSZK proof system (P, V) for a promise problem Π. Let $v = O(\mathtt{poly}(\kappa))$ denote the number of messages the honest verifier sends to the prescribed prover, and $\mu \leq \frac{1}{4(12v)^3}$[a] denote the simulator deviation. Fix $s = 4(12v)^2$. Then there exists an algorithm that constructs distributions $\otimes^s \bar{X}'$ and $\otimes^s \bar{Y}'$, where $\bar{X}'$ and $\bar{Y}'$ are constructed in polynomial time as defined above, such that

$$x \in \Pi^{\mathsf{Yes}} \implies d_{\mathtt{TV}}(\otimes^s \bar{X}', \otimes^s \bar{Y}') \leq \frac{1}{3} \qquad (17)$$

$$x \in \Pi^{\mathsf{No}} \implies d_{\mathtt{TV}}(\otimes^s \bar{X}', \otimes^s \bar{Y}') > \frac{2}{3} \qquad (18)$$

[a] Note this automatically follows from the definition of zero knowledge which requires the deviation to typically be negligible.

Proof. We begin by proving the statement described in Equation (17). Assume $x \in \Pi^{\mathsf{Yes}}$, then

$$\begin{aligned}
d_{\mathtt{TV}}(\bar{X}', \bar{Y}') &\leq d_{\mathtt{TV}}(X_0, Y_0) + \sum_{i=1}^{v} \delta_i(x) && (19)\\
&\leq d_{\mathtt{TV}}(X_0, Y_0) + 2\mu && (20)\\
&\leq d_{\mathtt{TV}}(X_0, Y_0) + \frac{v}{2(12v)} && (21)\\
&\leq \frac{1}{(12v)^3} + \frac{v}{2(12v)} && (22)\\
&\leq \frac{1}{12(12v)^2} && (23)
\end{aligned}$$

Equation (19) comes from lemma 3. Equation (20) comes from lemma 1. Equation (21) comes from the assumption about simulator deviation, and Equation (22) comes from the Chernoff bound described in Equation (15).

Therefore if $x \in \Pi^{\mathsf{Yes}}$, then we have

$$\begin{aligned} d_{\mathtt{TV}}(\otimes^s \bar{X}', \otimes^s \bar{Y}') &\leq s \cdot d_{\mathtt{TV}}(\bar{X}', \bar{Y}') && (24)\\ &\leq 4(12v)^2 \frac{1}{12(12v)^2} && (25)\\ &= \frac{1}{3} && (26) \end{aligned}$$

Here Equation (24) comes from the Direct Product Lemma 3.

Now for the second case, where $x \in \Pi^{\mathsf{No}}$, where we want to show that $d_{\mathtt{TV}}(\otimes^s \bar{X}', \otimes^s \bar{Y}')$ is large. First we show that $d_{\mathtt{TV}}(\bar{X}', \bar{Y}') \geq \frac{1}{12v}$ and to do this, it suffices to show that there exists a single index i in $\{0, \ldots, v\}$ such that $d_{\mathtt{TV}}(X_i, Y_i) \geq \frac{1}{12v}$.

Assume that simulator acceptance probability p is low, i.e. $p \leq \frac{5}{12}$, then by the Chernoff bound again,

$$\begin{aligned} \Pr[Y_0 = 1] &= \Pr\left[\frac{1}{w}\sum_{j=1}^{w} A_j \geq \frac{1}{2}\right] && (27)\\ &= \Pr\left[\frac{1}{w}\sum_{j=1}^{w} A_j \geq \frac{5}{12} + \frac{1}{12}\right] && (28)\\ &\leq \mathsf{exp}\left(-2 \cdot w \cdot \frac{1}{12^2}\right) && (29)\\ &< \frac{1}{2} && (30) \end{aligned}$$

Thus if the simulator acceptance probability is low i.e. $p \leq \frac{5}{12}$ then $d_{\mathtt{TV}}(X_0, Y_0) \geq \frac{1}{2} > \frac{1}{12v}$.

Now assume that $p \geq \frac{5}{12}$, then by Lemma 2 we have $\frac{5}{12} - \frac{1}{3} = \frac{1}{12} \leq \sum_{i=1}^{v} d_{\mathtt{TV}}(X_i, Y_i)$. Thus,I there must be at least one $j \in \{1, \ldots, v\}$ such that $d_{\mathtt{TV}}(X_j, Y_j) \geq \frac{1}{12v}$.

Therefore if $x \in \Pi^{\mathsf{No}}$ then we have

$$\begin{aligned} d_{\mathtt{TV}}(\otimes^s \bar{X}', \otimes^s \bar{Y}') &\geq 1 - \mathsf{exp}\left(-s \cdot \frac{1}{2(12v)^2}\right) && (31)\\ &= 1 - \mathsf{exp}\left(-4.(12v)^2 \cdot \frac{1}{2(12v)^2}\right) && (32)\\ &\geq \frac{2}{3} && (33) \end{aligned}$$

Here equation (32) comes from the Direct Product Lemma 3.

□

3 Deferred Proofs

Fact 2 Given two vectors $\vec{X}$ and $\vec{Y}$, $\vec{X} \otimes \vec{Y}$ represents their tensor product (outer prodct).

$$||\vec{X} \otimes \vec{Y}||_1 = ||\vec{X}||_1 \cdot ||\vec{Y}||_1$$

Fact 3 Let $X = (X_0, X_1)$ be a joint distribution where X_0 and X_1 are independent. Similarly, define $Y = (Y_0, Y_1)$. Then we have

$$d_{\texttt{TV}}(X, Y) = d_{\texttt{TV}}(X_0, Y_0) + d_{\texttt{TV}}(X_1, Y_1) \tag{34}$$

Lemma 3 (Direct Product Lemma) Let X and Y be distributions such that $d_{\texttt{TV}}(X, Y) = \delta$. Then for all $k \in \mathbb{N}$, we have

$$1 - 2\texttt{exp}(-\frac{k\delta^2}{2}) \leq d_{\texttt{TV}}(\otimes^k X, \otimes^k Y) \leq k\delta$$

Proof. The upper bound follows directly from Lemma 3 by replacing X_0 and X_1 as i.i.d samples from X, likewise with Y.

Let $\delta = d_{\texttt{TV}}(X, Y)$ and define S^* as $\Pr[X \in S^*] - \Pr[Y \in S^*] = \delta$. Let $p = \Pr[Y \in S^*]$, then $\Pr[X \in S^*] = p + \delta$. Thus, here we have two Bernoulli random variables with mean p and $p + \delta$. By the Chernoff Bound (additive version in Lemma 5):

1. The probability that at most $k(p + \frac{\delta}{2})$ fraction of k samples like in S is equivalent to

$$\begin{aligned}
\Pr[\sum_{i \in [k]} X_i \leq k(p + \frac{\delta}{2})] &= \Pr[(\frac{1}{k} \sum_{i \in [k]} X_i) - (p + \delta) \leq -\frac{\delta}{2}] && (35) \\
&\leq \mathsf{exp}(-\frac{k}{2}\delta^2) && (36)
\end{aligned}$$

2. Similarly, the probability that at least $k(p + \frac{\delta}{2})$ fraction of k samples lie in S is at most $\mathsf{exp}(-\frac{k}{2}\delta^2)$. (This is just the other tail of the Chernoff bound, writing it out explicitly like above makes it pretty obvious).

$$\begin{aligned}
\Pr[\sum_{i \in [k]} Y_i \geq k(p + \frac{\delta}{2})] &= \Pr[(\frac{1}{k} \sum_{i \in [k]} Y_i) - p \geq \frac{\delta}{2}] && (37) \\
&\leq \mathsf{exp}(-\frac{k}{2}\delta^2) && (38)
\end{aligned}$$

Define S' as a set of all k-tuples such that:

$$S' = \left\{(z_1, \ldots, z_k) \in (\{0,1\}^n)^k : \text{at least } k(p + \frac{\delta}{2}) \text{ fraction of samples, lie in } S^*\right\}$$

Thus we have,

$$\begin{aligned}
d_{\texttt{TV}}(\otimes^k X, \otimes^k Y) &\geq \Pr[\otimes^k X \in S'] - \Pr[\otimes^k Y \in S'] && (39) \\
&\geq 1 - 2\mathsf{exp}(-\frac{k}{2}\delta^2) && (40)
\end{aligned}$$

Equation (39) comes from the definition of statistical difference. Equation (40) come from $\Pr[\otimes^k Y \in S'] \leq \mathsf{exp}(-\frac{k}{2}\delta^2)$ as that is the definition of item 2. And $\Pr[\otimes^k X \in S'] \geq 1 - \mathsf{exp}(-\frac{k}{2}\delta^2)$ from the definition given in item (1).

□

Lemma 4 (XOR Lemma) Given a pair of distributions $X_0, X_1 \in \Delta(\Omega_\mathsf{n})$ and $k \in \mathbb{N}$ as input, there exists a polynomial time computable function that outputs a pair of distributions (Y_0, Y_1) such that

$$d_{\mathtt{TV}}(Y_0, Y_1) = d_{\mathtt{TV}}(X_0, X_1)^k$$

Specifically, Y_0, Y_1 are defined as follows:

Y_0: Uniformly select $(b_1, \ldots, b_k) \in \{0,1\}^k$ such that $\oplus_{i=1}^k b_i = 0$, and set $Y_0 = (X_{b_1} \otimes \cdots \otimes X_{b_k})$.

Y_1: Uniformly select $(b_1, \ldots, b_k) \in \{0,1\}^k$ such that $\oplus_{i=1}^k b_i = 1$, and set $Y_1 = (X_{b_1} \otimes \cdots \otimes X_{b_k})$.

Proof. Prove the base case with $k = 2$. Then the rest follows from induction. Define a_i such that all $i \in [k]$, we have $a_i \in \{0,1\}$.

Let $X_0^{(1)}, X_1^{(1)}, X_0^{(2)}, X_1^{(2)}$ be random variables such that

$Z_0 = X_{a_1}^{(1)} \otimes X_{a_2}^{(2)}$ such that $a_1 \oplus a_2 = 0$

$Z_1 = X_{a_1}^{(1)} \otimes X_{a_2}^{(2)}$ such that $a_1 \oplus a_2 = 1$

$$\begin{aligned} d_{\mathtt{TV}}(Z_0, Z_1) &= \frac{1}{2}|Z_0 - Z_1| \\ &= \frac{1}{2}\left|\frac{1}{2}[(X_0^{(1)} \otimes X_0^{(2)}) + (X_1^{(1)} \otimes X_1^{(2)})] - \frac{1}{2}[(X_1^{(1)} \otimes X_0^{(2)}) + (X_0^{(1)} \otimes X_1^{(2)})]\right| \\ &= \frac{1}{4}\left|(X_0^{(1)} - X_1^{(1)}) \otimes (X_0^{(2)} - X_1^{(2)})\right| \\ &= \frac{1}{2}||X_0^{(1)} - X_1^{(1)}||_1 \cdot \frac{1}{2}||X_0^{(2)} - X_1^{(2)}||_1 \\ &= d_{\mathtt{TV}}(X_0^{(1)}, X_1^{(1)})\ d_{\mathtt{TV}}(X_0^{(2)}, X_1^{(2)}) \end{aligned}$$

Induction assumption for $k > 2$, we have random variables $\{X_0^{(i)}\}_{i \in [k]}$ and $\{X_1^{(i)}\}_{i \in [k]}$ and that

$$d_{\mathtt{TV}}(Z_0, Z_1) = d_{\mathtt{TV}}(X_0^{(1)}, X_1^{(1)}) \ldots d_{\mathtt{TV}}(X_0^{(k)}, X_1^{(k)})$$

When we add variables $X_0^{(k+1)}$ and $X_1^{(k+1)}$, and computing the statistical distance between Z_0 and Z_1 reduces to the base case again as exactly half of the values will have parity 0 and a half will have parity 1. Thus the lemma holds by induction.

□

Lemma 5 (Chernoff Bound) Let $X_1, \ldots, X_n$ be n independent bernoulli random variables with $\Pr[X_i = 1] = p$ for all $i \in [n]$ for some $p > 1/2$. Then for any $\delta > 0$, we have

$$\Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_i \leq p + \delta\right] \leq \mathsf{exp}(-2n\delta^2) \tag{41}$$

$$\Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_i \leq p - \delta\right] \leq \mathsf{exp}(-2n\delta^2) \tag{42}$$

Proof. Look it up in any probability textbook. □

References