Solutions 8

1. The only place that could be possible problematic is when we find the polynomial $Q(x, y)$ which vanishes at all points (x_i, y_i). But this is not a problem at all, as long as the vectors containing the values of the monomials are still linearly independent. Indeed, it is clearly possible to present the given set of points as union of sets of points S_i having all their first coordinates distinct, so that all situations are covered. Each set S_i has its own polynomial $Q_i(X, Y)$, the product of which is a valid polynomial for our purpose (after suppression of useless factors).

2. The first observation is that X itself has minimal entropy $k - 1$. To see this, observe that, $\mathbb{P}[Y = a] \leq 2^{-k}$ for all values of $a \in \Omega$, since Y has minimal entropy k. Moreover,

$$|\mathbb{P}[X \in A] - \mathbb{P}[Y \in A]| \leq 2^{-k}$$

for all $A \subseteq \Omega$, since X is 2^{-k}-close to Y. Therefore, if we take A to be the singleton set $\{a\}$, we obtain

$$\mathbb{P}[X = a] - \mathbb{P}[Y = a] \leq 2^{-k} \Rightarrow \mathbb{P}[X = a] \leq \mathbb{P}[Y = a] + 2^{-k} \leq 2^{-(k-1)},$$

thereby proving that X has minimal entropy $k-1$. But the similarities between X and Z end there. For instance if we let Z to be a flat distribution on a set of size exactly 2^{k-1}, and X is any distribution having minimal entropy at least k, then the statistical distance (or total variation distance) between X and Z is at least $\frac{1}{2}$. To prove our claim, take A to be the support of Z. Then $\mathbb{P}[Z \in A] = 1$, while $\mathbb{P}[X \in A] \leq |A|2^{-k} = \frac{1}{2}$. Hence

$$d_{TV}(X, Z) \geq |\mathbb{P}[Z \in A] - \mathbb{P}[X \in A]| \geq \frac{1}{2}.$$

3. First, we observe that every entry of $H(U_n)$ is distributed uniformly on $\{0, 1\}$, except the first one which is constant and equals to zero. In fact, since

$$H(x)_y = \langle x, y \rangle \pmod 2,$$

hence $H(x)_y = \sum_{i:y_i \neq 0} x_i \pmod 2$. Thus we have that $\mathbb{P}[H(U_n)_y = 0] = \mathbb{P}[H(U_n)_y = 1] = \frac{1}{2}$ for all $y \neq 0$, because $H(U_n)_y$ is the sum of independent uniformly distributed $0, 1$-random variables.

Bitte wenden!

In order to prove that all the entries in $H(U_n)$ are pairwise independent, we just need to prove that for all $y_1 \neq y_j$ we have $\mathbb{P}[H(U_n)_{y_1} = H(U_n)_{y_2}] = \frac{1}{2}$. But

$$H(x)_{y_1} = H(x)_{y_2} \iff H(x)_{y'} = 0,$$

where $y' = y_1 - y_2 \pmod 2$. Therefore, clearly we have

$$\mathbb{P}[H(U_n)_{y_1} = H(U_n)_{y_2}] = \frac{1}{2},$$

finishing the proof.