Solutions 4

1. a) By the definition of coupling, we need to prove that

$$\sum_{j\in K_m} \mathbb{P}(X_n = i, Y_n = j) = \mu_i^{(n)}, \qquad \sum_{i\in K_m} \mathbb{P}(X_n = i, Y_n = j) = \nu_j^{(n)}.$$

We will prove it by induction. For the base case $n = 0$ we have by definition $\mathbb{P}(X_0 = i, Y_0 = j) = \delta_{ii_0}\delta_{jj_0}$ which implies

$$\sum_{j\in K_m} \mathbb{P}(X_0 = i, Y_0 = j) = \delta_{ii_0} = \mu_i^{(0)}, \qquad \sum_{i\in K_m} \mathbb{P}(X_0 = i, Y_0 = j) = \delta_{jj_0} = \nu_j^{(0)}.$$

Now we assume the equalities for n and prove them for $n + 1$. We do it explicitly for the first equality (the second is similar).

$$\begin{aligned}
\sum_{j\in K_m} \mathbb{P}(X_{n+1} = i, Y_{n+1} = j) &= \sum_{j,k,l\in K_m} \mathbb{P}(X_{n+1} = i, Y_{n+1} = j | X_n = k, Y_n = l)\mathbb{P}(X_n = k, Y_n = l) \\
&= \sum_{k\neq l, k\neq i, j\neq l} \frac{1}{(m-1)^2}\mathbb{P}(X_n = k, Y_n = l) + \sum_{k=l,k\neq i, j=i} \frac{1}{(m-1)}\mathbb{P}(X_n = k, Y_n = l) \\
&= \frac{m-1}{(m-1)^2} \sum_{k:k\neq i} \sum_{l:l\neq k} \mathbb{P}(X_n = k, Y_n = l) + \frac{1}{m-1} \sum_{k:k\neq i} \sum_{l:l=k} \mathbb{P}(X_n = k, Y_n = l) \\
&= \frac{1}{m-1} \sum_{k:k\neq i} \sum_{l\in K_m} \mathbb{P}(X_n = k, Y_n = l) \\
&= \frac{1}{m-1} \sum_{k:k\neq i} \mu_k^{(n)} \\
&= \mu_i^{(n+1)}.
\end{aligned}$$

To get the third line we perform the sums over j. To get the fifth line we use the induction hypothesis. To get the last line we remark

$$\mu_i^{(n+1)} = \sum_{k\in K_m} \mathbb{P}(X_{n+1} = i | X_n = k)\mu_k^{(n)} = \frac{1}{m-1} \sum_{k\neq i} \mu_k^{(n)}.$$

b) For $n \geq 1$,

$$\begin{aligned}
\mathbb{P}(T > n) &= \mathbb{P}(X_n \neq Y_n, X_{n-1} \neq Y_{n-1}, \ldots, X_1 \neq Y_1) \\
&= \sum_{i\neq j} \mathbb{P}(X_n \neq Y_n, X_{n-1} = i, Y_{n-1} = j, X_{n-2} \neq Y_{n-2}, \ldots, X_1 \neq Y_1) \\
&= \sum_{i\neq j} \mathbb{P}(X_n \neq Y_n \,|\, X_{n-1} = i, Y_{n-1} = j, X_{n-2} \neq Y_{n-2}, \ldots, X_1 \neq Y_1) \\
&\qquad \cdot \mathbb{P}(X_{n-1} = i, Y_{n-1} = j, X_{n-2} \neq Y_{n-2}, \ldots, X_1 \neq Y_1).
\end{aligned}$$

By the Markov property, we have

$$\begin{aligned}
&\mathbb{P}(X_n \neq Y_n \,|\, X_{n-1} = i, Y_{n-1} = j, X_{n-2} \neq Y_{n-2}, \dots, X_1 \neq Y_1) \\
&= \mathbb{P}(X_n \neq Y_n | X_{n-1} = i, Y_{n-1} = j) = 1 - \mathbb{P}(X_n = Y_n | X_{n-1} = i, Y_{n-1} = j) \\
&= 1 - \sum_{\substack{\ell \in K_m \\ \ell \neq i \\ \ell \neq j}} \mathbb{P}(X_n = \ell, Y_n = \ell \mid X_{n-1} = i, Y_{n-1} = j) = 1 - \frac{m-2}{(m-1)^2}.
\end{aligned}$$

So

$$\begin{aligned}
\mathbb{P}(T > n) &= \sum_{i \neq j} \left(1 - \frac{m-2}{(m-1)^2}\right) \mathbb{P}(X_{n-1} = i, Y_{n-1} = j, X_{n-2} \neq Y_{n-2}, \dots, X_1 \neq Y_1) \\
&= \left(1 - \frac{m-2}{(m-1)^2}\right) \mathbb{P}(X_{n-1} \neq Y_{n-1}, X_{n-2} \neq Y_{n-2}, \dots, X_1 \neq Y_1) \\
&= \left(1 - \frac{m-2}{(m-1)^2}\right) \mathbb{P}(T > n-1).
\end{aligned}$$

The result follows by simple induction. Furthermore,

$$\begin{aligned}
\mathbb{P}(T = n) = \mathbb{P}(T > n-1) - \mathbb{P}(T > n) &= \left(1 - \frac{m-2}{(m-1)^2}\right)^{n-1} \left(1 - 1 + \frac{m-2}{(m-1)^2}\right) \\
&= \frac{m-2}{(m-1)^2} \left(1 - \frac{m-2}{(m-1)^2}\right)^{n-1}.
\end{aligned}$$

NB: In order to explain the above detailed equations, let us mention that one cannot write directly

$$\mathbb{P}(X_n \neq Y_n \,|\, X_{n-1} \neq Y_{n-1}, \dots, X_1 \neq Y_1) = \mathbb{P}(X_n \neq Y_n \,|\, X_{n-1} \neq Y_{n-1}) \tag{1}$$

"by the Markov property". The Markov property requires indeed that the state at time $n-1$ (which is the present, here) is *known*. It happens that equality (1) holds in the present case because of the symmetry of the Markov chain under study, but this equality is not true in general.

c) Using the first definition of total variation distance, we obtain

$$\begin{aligned}
\|\mu^{(n)} - \nu^{(n)}\|_{\mathrm{TV}} &= \frac{1}{2} \sum_j |\mathbb{P}(X_n = j) - \mathbb{P}(Y_n = j)| \\
&= \frac{1}{2} \sum_{j \in K_m} |\mathbb{P}(X_n = j | T > n) - \mathbb{P}(Y_n = j | T > n)| \, \mathbb{P}(T > n) \\
&+ \frac{1}{2} \sum_{j \in K_m} |\mathbb{P}(X_n = j | T \leq n) - \mathbb{P}(Y_n = j | T \leq n)| \, \mathbb{P}(T \leq n).
\end{aligned}$$

Note that the last sum is zero, since the walks have coalesced. Therefore,

$$\begin{aligned}
\|\mu^{(n)} - \nu^{(n)}\|_{\mathrm{TV}} &= \frac{1}{2} \sum_{j \in K_m} |\mathbb{P}(X_n = j | T > n) - \mathbb{P}(Y_n = j | T > n)| \, \mathbb{P}(T > n) \\
&\leq \frac{1}{2} \sum_{j \in K_m} (\mathbb{P}(X_n = j | T > n) + \mathbb{P}(Y_n = j | T > n)) \, \mathbb{P}(T > n) \\
&= \mathbb{P}(T > n) = \left(1 - \frac{m-2}{(m-1)^2}\right)^n \leq e^{-\frac{n(m-2)}{(m-1)^2}}.
\end{aligned}$$

d) Let P be the matrix of transition probabilities of the random walk on the complete graph. Then, the elements on the principal diagonal of P are 0, and the remaining elements are all equal to $1/(m-1)$. Therefore, it is easy to check that $\pi = (1/m, \cdots, 1/m)$ satisfies the equation $\pi = \pi P$ and, consequently, it is a stationary distribution.

By the corollary to Lemma 1 seen in class (cf. lecture notes), we have that

$$\|\mu^{(n)} - \pi\|_{\mathrm{TV}} \leq \mathbb{P}(T > n) \leq e^{-\frac{n(m-2)}{(m-1)^2}},$$

which gives us the desired claim.

e) In the (very) particular case $m = 2$, X_n and Y_n never coalesce and the coupled chain is periodic of period 2. Indeed, $X_0 = i_0$ and $Y_0 = j_0$ with $i_0 \neq j_0$. Since $m = 2$, i_0 and j_0 are the only possible states. Therefore, $X_k = i_0$ and $Y_k = j_0$, when k is even; $X_k = j_0$ and $Y_k = i_0$, when k is odd.

2. a) $\mathbf{1} \leq \mathbf{2}$: Consider $A = \{j \in S : \mu_j \geq \nu_j\}$. Then,

$$\begin{aligned}\sum_{j\in S} |\mu_j - \nu_j| &= \sum_{j\in A}(\mu_j - \nu_j) + \sum_{j\in A^c}(\nu_j - \mu_j) = \sum_{j\in A}(\mu_j - \nu_j) + \sum_{j\in S}(\nu_j - \mu_j) - \sum_{j\in A}(\nu_j - \mu_j) \\ &= 2\left(\sum_{j\in A}(\mu_j - \nu_j)\right) = 2(\mu(A) - \nu(A)),\end{aligned}$$

as $\sum_{j\in S} \mu_j = \sum_{j\in S} \nu_j = 1$. Therefore,

$$\max_{A\subset S} |\mu(A) - \nu(A)| \geq \frac{1}{2}\sum_{j\in S} |\mu_j - \nu_j|.$$

$\mathbf{2} \leq \mathbf{3}$: For every $A \subset S$, define

$$\phi_j = \begin{cases} +1 & \text{if } j \in A \\ -1 & \text{if } j \in A^c \end{cases}$$

Then

$$\mu(\phi) - \nu(\phi) = \sum_{j\in A}(\mu_j - \nu_j) - \sum_{j\in A^c}(\mu_j - \nu_j) = 2\,(\mu(A) - \nu(A)),$$

by the same argument as above. Hence,

$$\frac{1}{2}\max_{\phi:S\to[-1,+1]} |\mu(\phi) - \nu(\phi)| \geq \max_{A\subset S} |\mu(A) - \nu(A)|$$

$\mathbf{3} \leq \mathbf{1}$: Simply observe that for every $\phi : S \to [-1, +1]$, we have

$$|\mu(\phi) - \nu(\phi)| = \left|\sum_{j\in S}(\mu_j - \nu_j)\,\phi_j\right| \leq \sum_{j\in S} |\mu_j - \nu_j|\,|\phi_j| \leq \sum_{j\in S} |\mu_j - \nu_j|.$$

Therefore,

$$\frac{1}{2}\max_{\phi:S\to[-1,+1]} |\mu(\phi) - \nu(\phi)| \leq \frac{1}{2}\sum_{j\in S} |\mu_j - \nu_j|.$$

b) Let us consider the definition $||\mu - \nu||_{\mathrm{TV}} = \frac{1}{2}\sum_{i\in S}|\mu_i - \nu_i|$. Then, $||\mu - \nu||_{\mathrm{TV}} \geq 0$, because it is the sum of absolute values, and,

$$\sum_{i\in S}|\mu_i - \nu_i| = 0 \qquad \Longleftrightarrow \qquad \mu_i = \nu_i \quad \forall i \in S \qquad \Longleftrightarrow \qquad \mu = \nu,$$

i.e., $||\mu - \nu||_{\mathrm{TV}} = 0$ if and only if $\mu = \nu$.

In addition,

$$||\mu - \nu||_{\mathrm{TV}} = \frac{1}{2}\sum_{i\in S}|\mu_i - \nu_i| = \frac{1}{2}\sum_{i\in S}|\nu_i - \mu_i| = ||\nu - \mu||_{\mathrm{TV}},$$

i.e., $||\mu - \nu||_{\mathrm{TV}}$ is symmetric.

Furthermore, for any τ s.t. $\sum_{i\in S}\tau_i = 1$,

$$||\mu-\nu||_{\mathrm{TV}} = \frac{1}{2}\sum_{i\in S}|\mu_i-\nu_i| = \frac{1}{2}\sum_{i\in S}|\mu_i-\tau_i+\tau_i-\nu_i| \leq \frac{1}{2}\sum_{i\in S}|\mu_i-\tau_i|+\frac{1}{2}\sum_{i\in S}|\tau_i-\nu_i| = ||\mu-\tau||_{\mathrm{TV}}+||\tau-\nu||_{\mathrm{TV}},$$

i.e., the triangle inequality is satisfied. As a result, $||\mu - \nu||_{\mathrm{TV}}$ is a distance.