1.
$$\Pr(X=Y) = \sum_{t\in\Omega} \Pr(X=Y=t)$$
$$\le \sum_{t\in\Omega} \min\{\Pr(X=t), \Pr(Y=t)\}$$
$$= \sum_{t\in\Omega} \min\{\mu(t), \nu(t)\}.$$

Thus,
$$\Pr(X\ne Y) = 1-\Pr(X=Y)$$
$$\ge 1-\sum_{t\in\Omega}\min\{\mu(t),\nu(t)\}$$
$$= \sum_{t\in\Omega}\mu(t) - \sum_{t\in\Omega}\min\{\mu(t),\nu(t)\}$$
$$= \sum_{t\in\Omega}(\mu(t)-\min\{\mu(t),\nu(t)\})$$
$$= \sum_{\substack{t\in\Omega\\ \mu(t)>\nu(t)}} (\mu(t)-\nu(t))$$
$$= \max_{A\subseteq\Omega} |\mu(A)-\nu(A)|$$
$$= D_{TV}(\mu,\nu).$$

If $\Pr(X\ne Y) = D_{TV}(\mu,\nu)$, we have
$$\Pr(X=Y=t) = \min\{\mu(t),\nu(t)\}.$$

Give an explicit construction of ω below:
$$\omega_{i,j} = \begin{cases} \min\{\mu(i),\nu(i)\}, & i=j \\ k\max\{0,\nu(i)-\mu(i)\}\max\{0,\mu(j)-\nu(j)\}, & i\ne j,\ k \text{ is a constant}. \end{cases}$$

First, if there exist $k\in\mathbb{R}$ such that ω is a coupling, $\Pr_{(X,Y)\sim\omega}(X\ne Y) = D_{TV}(\mu,\nu)$.

Then, compute k.
$$\begin{cases} \sum_j \omega_{ij} = \nu(i) & \forall i\in[|\Omega|] \\ \sum_i \omega_{ij} = \mu(j) & \forall j\in[|\Omega|] \end{cases}$$

Consider one of these equations.
$$\nu(i) = \sum_j \omega_{ij} = \sum_{j\ne i} k\max\{0,\nu(i)-\mu(i)\}\max\{0,\mu(j)-\nu(j)\} + \min\{\mu(i),\nu(i)\} = k\max\{0,\nu(i)-\mu(i)\}\sum_{j\ne i}\max\{0,\mu(j)-\nu(j)\} + \min\{\mu(i),\nu(i)\}$$
$$= k\max\{0,\nu(i)-\mu(i)\}\left(D_{TV}(\mu,\nu)-\max\{0,\mu(j)-\nu(j)\}\right) + \min\{\mu(i),\nu(i)\}.$$

If $\nu(i)-\mu(i) > 0$
$$\nu(i) = k(\nu(i)-\mu(i))D_{TV}(\mu,\nu)+\mu(i)$$
$$\Rightarrow k = \frac{1}{D_{TV}(\mu,\nu)}$$

If $\nu(i)-\mu(i) < 0$
$$\nu(i) = \nu(i)\ \checkmark.$$

So $k = \frac{1}{D_{TV}(\mu,\nu)}$ is a solution to all $2|\Omega|$ equations.

Thus,
$$\omega_{i,j} = \begin{cases} \min\{\mu(i),\nu(i)\}, & i=j \\ \frac{1}{D_{TV}(\mu,\nu)}\max\{0,\nu(i)-\mu(i)\}\max\{0,\mu(j)-\nu(j)\}, & i\ne j \end{cases}$$
is a distribution ~~that~~ meets the conditions.

Not use coupling :

For a transition matrix P,

2. ① $\left.\begin{array}{l}\mu_t^T = \mu_0^T P^t \\ \mu_{t+1}^T = \mu_0^T P^{t+1}\end{array}\right\} \Rightarrow \mu_{t+1}^T = \mu_t^T P$ $\quad \begin{cases} \sum_j P_{ij} = 1 \quad \sum_i P_{ij} = 1. \ \forall i, j. \\ \forall i, \forall j, P_{ij} \geq 0 \end{cases}$

$$\pi^T = \pi^T P$$

② For a vector $X = (x_1, x_2, \cdots, x_n)$, $\|X\|_{F_1} \triangleq \sum_{i=1}^{n} |x_i|$.

$$\Delta(t) = D_{TV}(\mu_t, \pi) = \frac{1}{2} \sum_{i=1}^{|\Omega|} |\mu_t(i) - \pi(i)| = \frac{1}{2} \|\mu_t - \pi\|_{F_1}.$$

Similarly, $\Delta(t+1) = \frac{1}{2} \|\mu_{t+1} - \pi\|_{F_1} = \frac{1}{2} \|P^T \mu_t - \pi\|_{F_1}$

$$= \frac{1}{2} \|P^T \mu_t - P^T \pi\|_{F_1} = \frac{1}{2} \|P^T (\mu_t - \pi)\|_{F_1}$$

③ $X_t \triangleq \mu_t - \pi$, $\quad \sum_i X_{t_i} = \sum_i \mu_t(i) - \sum_i \pi(i) = 1 - 1 = 0$

Thus,

$$\Delta(t) = \frac{1}{2} \|X_t\|_{F_1} = \frac{1}{2} \sum_{i=1}^{|\Omega|} |X_{t_i}|$$

$$\Delta(t+1) = \frac{1}{2} \|P^T X_t\|_{F_1} = \frac{1}{2} \sum_{i=1}^{|\Omega|} |P_{1i} X_{t_1} + P_{2i} X_{t_2} + \cdots + P_{|\Omega| i} X_{t_{|\Omega|}}|$$

$$\leq \frac{1}{2} \sum_{i=1}^{|\Omega|} \left(P_{1i} |X_{t_1}| + \cdots + P_{|\Omega| i} |X_{t_{|\Omega|}}| \right)$$

$$= \frac{1}{2} \sum_{i=1}^{|\Omega|} \left(\left(\sum_{j=1}^{|\Omega|} P_{ji} \right) |X_{t_i}| \right)$$

$$= \frac{1}{2} \sum_{i=1}^{|\Omega|} |X_{t_i}| = \Delta(t)$$

Use coupling:

Assume $X_t \sim \mu_t$, $Y_t \sim \pi$, ω is a coupling of $\cancel{X_t}$ and Y_t.

Then $\Delta(t) = D_{TV}(\mu_t, \pi) \le Pr_{(X_t, Y_t)\sim\omega}(X_t \ne Y_t)$

Similarly, $\Delta(t+1) \le Pr(X_{t+1} \ne Y_{t+1})$, if $X_{t+1} \sim \mu_{t+1}$, $Y_{t+1} \sim \pi$.

And in fact as what we have proved in Problem 1.

there exist an ω such that $\Delta(t) = Pr_{(X_t, Y_t)\sim\omega}(X_t \ne Y_t)$.

Consider this ω. Let ω' be a coupling such that:

① ω' is a coupling of X_{t+1} and Y_{t+1};

② If $X_t = Y_t$, then X_{t+1} and Y_{t+1} are evolved from X_t and Y_t synchronously.

which assures that if $X_t = Y_t$, $X_{t+1} = Y_{t+1}$;

③ If $X_t \ne Y_t$, then X_{t+1} and Y_{t+1} are evolved from X_t and Y_t respectively.

With ①②③, we have $Pr(X_t = Y_t) \le Pr(X_{t+1} = Y_{t+1})$;

$$\Longleftrightarrow Pr(X_t \ne Y_t) \ge Pr(X_{t+1} \ne Y_{t+1}).$$

Then $\Delta(t) = Pr(X_t \ne Y_t) \ge Pr(X_{t+1} \ne Y_{t+1}) \ge \Delta(t+1)$.

3、 $E(HH) = 1 + \frac{1}{2}E(HH|H) + \frac{1}{2}E(HH|T)$

$E(HH|H) = \frac{1}{2} \times 1 + \frac{1}{2} \times (1 + E(HH|HT))$

$= \frac{1}{2} \times 1 + \frac{1}{2} \times (1 + E(HH|T))$

$= 1 + \frac{1}{2}E(HH|T)$

$E(HH|T) = \frac{1}{2}(1 + E(HH|H)) + \frac{1}{2}(1 + E(HH|T))$

$= \frac{1}{2}\left(2 + \frac{1}{2}E(HH|T)\right) + \frac{1}{2}(1 + E(HH|T))$

$= \frac{3}{2} + \frac{3}{4}E(HH|T)$

$\Rightarrow E(HH|T) = 6$

$\Rightarrow E(HH|H) = 4$

$\Rightarrow E(HH) = 6$

$E(HT) = 1 + \frac{1}{2}E(HT|H) + \frac{1}{2}E(HT|T)$

$E(HT|H) = \frac{1}{2} \times (1 + E(HT|H)) + \frac{1}{2} \times 1$

$\Rightarrow E(HT|H) = 2$

$E(HT|T) = \frac{1}{2} \times (1 + E(HT|H)) + \frac{1}{2}(1 + E(HT|T))$

$= 2 + \frac{1}{2}E(HT|T)$

$\Rightarrow E(HT|T) = 4$

$\Rightarrow E(HT) = 4$

So $E(T_1) = 6$, $E(T_2) = 4$.

We can say that HT will be seen faster from the probability point of view.

(1) Firstly, consider the case when $k=2$.

$$Pr(z_2''=c) = \sum_{z_1''} Pr(z_1''=a)\, Pr(z_2''=c \mid z_1''=a)$$

$(z_0', z_1') \in \omega_1$
$z_1' = z_1''$
$(z_1'', z_2'') \in \omega_2$

$$= \sum_{z_1'} Pr(z_1'=a)\, Pr(z_2''=c \mid z_1'=a)$$

$$= \sum_{z_1'} \left(\sum_{z_0'} Pr(z_0'=b)\, Pr(z_1'=a \mid z_0'=b) \right) Pr(z_2''=c \mid z_1'=a)$$

$$\overset{\text{fix } z_0'}{=} \sum_{z_0'} Pr(z_0'=b) \left(\sum_{z_1'} Pr(z_1'=a \mid z_0'=b)\, Pr(z_2''=c \mid z_1'=a, z_0'=b) \right)$$

$$= \sum_{z_0'} Pr(z_0'=b) \left(\sum_{z_1'} \frac{Pr(\cancel{z_1'=a, z_0'=b})}{Pr(z_0'=b)} \, \frac{Pr(z_1'=a, z_0'=b, z_2''=c)}{Pr(\cancel{z_1'=a, z_0'=b})} \right)$$

$$= \sum_{z_0'} Pr(z_0'=b) \sum_{z_1'} Pr(z_1'=a, z_2''=c \mid z_0'=b)$$

$$= \sum_{z_0'} Pr(z_0'=b)\, Pr(z_2''=c \mid z_0'=b)$$

$$= Pr(z_2''=c)$$, i.e. (z_0, z_2) is a legal coupling of (X_t, Y_t).

$(z_0', z_2'') \in \omega$

For all $k \in \mathbb{N}^+$, we can prove similarly that if (z_0, z_{k-1}) is a legal coupling, (z_0, z_k) is a legal coupling.

By induction, (z_0, z_k) is a legal coupling of (X_t, Y_t) for all $k \in \mathbb{N}^+$.

(2) For $d(X_t, Y_t) = 1$, $\mathbb{E}(d(X_{t+1}, Y_{t+1}) \mid (X_t, Y_t)) \le (1-\alpha)\, d(X_t, Y_t)$

Compute expectations on both sides.

$$\mathbb{E}\big(\mathbb{E}(d(X_{t+1}, Y_{t+1}) \mid (X_t, Y_t))\big) \le (1-\alpha)\, \mathbb{E}(d(X_t, Y_t))$$

$\Rightarrow \mathbb{E}(d(X_{t+1}, Y_{t+1})) \le 1-\alpha$ for those $d(X_t, Y_t) = 1$.

For $d(Z_i, Z_{i+1}) = 1$ for $i = 0, 1, \cdots, k-1$, $\mathbb{E}(d(Z_i, Z_{i+1})) \le 1-\alpha$

$$d(Z_0, Z_k) \le d(Z_0, Z_1) + d(Z_1, Z_2) + \cdots + d(Z_{k-1}, Z_k)$$

Compute expectations on both sides.

$$\mathbb{E}(d(Z_0, Z_k)) \le \mathbb{E}\big(d(Z_0, Z_1) + d(Z_1, Z_2) + \cdots + d(Z_{k-1}, Z_k)\big) = \sum_{i=0}^{k-1} \mathbb{E}(d(Z_i, Z_{i+1})) \le (1-\alpha) k.$$

(3) Consider the probabilities of two types of moves.

For good moves, w.p. $\frac{d(X_t,Y_t)}{n}$, $X_t(v) \neq Y_t(v)$, and all choices of color c can make it a good move, since we have added these improper colorings to Ω.

$$\Pr(d(X_{t+1},Y_{t+1}) = d(X_t,Y_t)-1) = \Pr((v,c) \text{ is a good move}) \geq \frac{d(X_t,Y_t)}{n}$$

For bad moves, there exists a neighbor w of v that its color is different in two colorings, and in one coloring w is of color c.

$$\Pr(d(X_{t+1},Y_{t+1}) = d(X_t,Y_t)+1) = \Pr((v,c) \text{ is a bad move}) \leq \frac{\Delta d(X_t,Y_t)}{n}\frac{2}{q}$$

$$\mathbb{E}(d(X_{t+1},Y_{t+1}) \mid (X_t,Y_t)) = d(X_t,Y_t) + \Pr(d(X_{t+1},Y_{t+1}) = d(X_t,Y_t)+1) - \Pr(d(X_{t+1},Y_{t+1}) = d(X_t,Y_t)-1)$$

$$\leq d(X_t,Y_t)\left(1+\frac{2\Delta}{nq}-\frac{1}{n}\right) = d(X_t,Y_t)\left(1-\frac{q-2\Delta}{nq}\right)$$

In the case $q > 2\Delta$; $\mathbb{E}(d(X_{t+1},Y_{t+1}) \mid (X_t,Y_t)) \leq d(X_t,Y_t)\left(1-\frac{1}{nq}\right)$

If we want $D_{TV} \leq \left(1-\frac{1}{nq}\right)^t n \leq \varepsilon$.

we have $\tau_{mix}(\varepsilon) \leq nq\log\frac{n}{\varepsilon} = O(n\log n)$