1 Proof of ACMMD expression in the synthetic setting

Recall that we have

$$\text{ACMMD}^2 = \text{MMD}(\mathbb{P}_X\mathbb{P}_|, \mathbb{P}_X Q_|)^2 \quad = \int k(p,p')\left(T_{11} + T_{22} - 2T_{12}\right) \mathrm{d}\mathbb{P}_X(p)\mathrm{d}\mathbb{P}_X(p')$$

where

$$T_{12} = \int k_y(y,y')p(y|p)q(y'|p')\mathrm{d}(y)\mathrm{d}(y')$$

and T_{22} and T_{12} are defined similarly. For a sequence y, we define the the function len given by $\text{len}(y) \coloneqq \min\{i \in \mathbb{N} | y_i = \text{STOP}\}$, which intuitively returns the length of the sequence.

Computing T_{ij} As we will see, a lot of the computations are agnostic to whether we are computing T_{11}, T_{22} or T_{12}. Note that the kernel on Y writes as a product

$$k_y(y,y') = e^{-\lambda_y d_H(y,y')} = e^{-\lambda_y \sum_{i=0}^{\infty} \delta(y_i \neq y_i')} = \prod_{i=0}^{\infty} e^{-\lambda_y \delta(y_i \neq y_i')} = \prod_{i=0}^{\max(\text{len}(y), \text{len}(y'))} e^{-\lambda_y \delta(y_i \neq y_i')}$$

let us define the following events

$$\begin{aligned} F(m) &\coloneqq \Big\{\min(\text{len}(y), \text{len}(y')) = m\Big\} \\ G(m, \delta m) &\coloneqq \Big\{\max(\text{len}(y), \text{len}(y')) = m + \delta m\Big\} \end{aligned}$$

which we further break down as

$$\begin{aligned} F_1(m) &= \{\text{len}(y) = m\} \cap \{\text{len}(y') > m\} \\ F_2(m) &= \{\text{len}(y) > m\} \cap \{\text{len}(y') = m\} \\ F_3(m) &= \{\text{len}(y) = m\} \cap \{\text{len}(y') = m\} \\ \implies F(m) &= F_1(m) \cup F_2(m) \cup F_3(m) \end{aligned}$$

For which the following probabilities hold:

$$\begin{aligned} P(F_1(m)) &= P(\text{len}(y) = m) \times P(\text{len}(y') > m) = ((2p)^m \times (1-2p)) \times (2p')^{m+1} \\ P(F_2(m)) &= P(\text{len}(y') = m) \times P(\text{len}(y) > m) = ((2p')^m \times (1-2p')) \times (2p)^{m+1} \\ P(F_3(m)) &= P(\text{len}(y') = m) \times P(\text{len}(y) = m) = ((2p')^m \times (1-2p')) \times ((2p)^m \times (1-2p)) \\ P(G(m,\delta m)|F_1(m)) &= P(\text{len}(y') = m + \delta m | \text{len}(y) = m, \text{len}(y') > m) = (2p')^{\delta m - 1} \times (1-2p')\delta_{(\delta m \geq 1)} \\ P(G(m,\delta m)|F_2(m)) &= P(\text{len}(y) = m + \delta m | \text{len}(y') = m, \text{len}(y) > m) = (2p)^{\delta m - 1} \times (1-2p)\delta_{(\delta m \geq 1)} \\ P(G(m,\delta m)|F_3(m)) &= \delta(\delta m = 0) \end{aligned}$$

Let us note

$$E(m, \delta m, i) \coloneqq F_i(m) \cap G(m, \delta m)$$

We have that $E(m, \delta_m, i) \cap E(m', \delta m', j) = \emptyset$ if $(m, \delta m, i) \neq (m', \delta m', j)$.

$$\Omega = \bigcup_{m=0}^{+\infty} \bigcup_{i=1}^{3} \bigcup_{\delta m=0}^{+\infty} E(m, \delta m)$$

Using the law of total probability, we have that Thus, using the law of total probability:

$$\begin{aligned}
T_{ij}(p,p') &= \sum_{m=0}^{+\infty}\sum_{i=1}^{3}\sum_{\delta m=0}^{+\infty} \mathbb{P}(E_i(m,\delta m,i))\mathbb{E}(e^{-\lambda d_H(y,y')}|E(m,\delta m,i)) \\
&= \sum_{m=0}^{+\infty}\sum_{i=1}^{3}\sum_{\delta m=0}^{+\infty} \mathbb{P}(F_i(m)\cap G(m,\delta m))\mathbb{E}(e^{-\lambda d_H(y,y')}|E(m,\delta m,i)) \\
&= \sum_{m=0}^{+\infty}\sum_{i=1}^{3} \mathbb{P}(F_i(m)) \sum_{\delta m=0}^{+\infty} P(G(m,\delta m)|F_i(m))\mathbb{E}(e^{-\lambda d_H(y,y')}|E(m,\delta m,i)) \\
&= \sum_{m=0}^{+\infty}\sum_{i=1}^{3} \mathbb{P}(F_i(m))\mathbb{E}(e^{-\lambda d_H(y_{:m},y'_{:m})}|F_i(m)) \\
&\times \sum_{\delta m=0}^{+\infty} P(G(m,\delta m)|F_i(m))\mathbb{E}(e^{-\lambda d_H(y_{m+1:m+\max(\delta m,1)},y'_{m+1:m+\max(\delta m,1)})}|E(m,\delta m,i),p,p') \\
&= \sum_{m=0}^{+\infty}\sum_{i=1}^{3} \mathbb{P}(F_i(m))\mathbb{E}(e^{-\lambda d_H(y_{:m},y'_{:m})}|F_i(m)) \times \sum_{\delta m=0}^{+\infty} P(G(m,\delta m)|F_i(m))e^{-\lambda(\max(0,\delta m-1)+\delta(m>0))} \\
&= \sum_{m=0}^{+\infty}\sum_{i=1}^{3} \mathbb{P}(F_i(m))\mathbb{E}(e^{-\lambda d_H(y_{:m},y'_{:m})}|F_i(m)) \left(\prod_{i=1}^{\max(m-1,1)} \mathbb{E}(e^{-\lambda\delta(y_i\neq y'_i)}|F_i(m))\right)^{\delta(m\geq 2)} \\
&\times \sum_{\delta m=0}^{+\infty} P(G(m,\delta m)|F_i(m))e^{-\lambda(\max(0,\delta m-1)+\delta(m>0))}
\end{aligned}$$

where we break down the factorized hamming distance over the sequence into the sum of the hamming distances over each coordinate, and made use of the fact that

$$d_H(y_{m:m+\delta m}, y'_{m:m+\delta m}) = \max(0,\delta m-1)+\delta(m>0)$$

conditioned on $F_i(m)$ and $G(m,\delta m)$. The disjunction of cases is necessary in order to not count the term 0^{th} term twice in the event when $m=0$. This representation is convenient since whenever $m \geq 2$, for any $1 \leq i \leq m-1$,

$$P(\delta(y_i,y'_i)=1|F_i(m)) = \frac{(pp')+(pp')}{(p+p)\times(p'+p')} = \frac{1}{2} = P(\delta(y_i,y'_i)=0|F_i(m))$$

meaning we have

$$\begin{aligned}
T_{ij}(p,p') &= \sum_{m=0}^{+\infty}\sum_{i=1}^{3} \mathbb{E}(e^{-\lambda\delta(y_0\neq y'_0)}|F_i(m))\mathbb{P}(F_i(m)) \left(\frac{1+e^{-\lambda}}{2}\right)^{\max(m-1,0)} \\
&\times \sum_{\delta m=0}^{+\infty} P(G(m,\delta m)|F_i(m))e^{-\lambda(\max(0,\delta m-1)+\delta(m>0))}
\end{aligned}$$

Inserting the relevant event probabilities into the expression for T_{ij}, we have

$$
\begin{aligned}
T_{ij}(p,p') &= \sum_{m=0}^{+\infty} \left(\frac{1+e^{-\lambda}}{2}\right)^{\max(m-1,0)} \\
&\times \Bigg(\mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_1(m))(2p)^m \times (1-2p)(2p')^{m+1} \times (1-2p')e^{-\lambda\delta(m>0)} \sum_{\delta m=1}^{+\infty} e^{-\lambda(\delta m-1)}(2p')^{\delta m-1} \\
&+ \mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_2(m))(2p')^m \times (1-2p')(2p)^{m+1} \times (1-2p)e^{-\lambda\delta(m>0)} \sum_{\delta m=1}^{+\infty} e^{-\lambda(\delta m-1)}(2p)^{\delta m-1} \\
&+ \mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_3(m))(2p')^m \times (1-2p')(2p)^m(1-2p) \Bigg) \\
&= \sum_{m=0}^{+\infty} \left(\frac{1+e^{-\lambda}}{2}\right)^{\max(m-1,0)} \\
&\times \Bigg(\mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_1(m))(2p)^m \times (1-2p)(2p')^{m+1} \times (1-2p')e^{-\lambda\delta(m>0)} \sum_{\delta m=0}^{+\infty} e^{-\lambda\delta m}(2p')^{\delta m} \\
&+ \mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_2(m))(2p')^m \times (1-2p')(2p)^{m+1} \times (1-2p)e^{-\lambda\delta(m>0)} \sum_{\delta m=0}^{+\infty} e^{-\lambda\delta m}(2p)^{\delta m} \\
&+ \mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_3(m))(2p')^m \times (1-2p')(2p)^m(1-2p) \Bigg) \\
&= \sum_{m=0}^{+\infty} \left(\frac{1+e^{-\lambda}}{2}\right)^{\max(m-1,0)} \\
&\times \Bigg(\mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_1(m))(2p)^m \times (1-2p)(2p')^{m+1} \times (1-2p') \times \frac{e^{-\lambda\delta(m>0)}}{1-2p'e^{-\lambda}} \\
&+ \mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_2(m))(2p')^m \times (1-2p')(2p)^{m+1} \times (1-2p) \times \frac{e^{-\lambda\delta(m>0)}}{1-2pe^{-\lambda}} \\
&+ \mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_3(m))(2p')^m \times (1-2p')(2p)^m(1-2p) \Bigg)
\end{aligned}
$$

Now, some simplification arise when $m \geq 1$. Indeed, in that case, $\mathbb{E}(e^{-\lambda\delta(y_0,y_0')}|F_i(m))$ is independent of i. Noting $T_{ij}^1(p,p')$ the sum of the terms for $m \geq 1$, we thus have

$$
\begin{aligned}
T_{ij}^1(p,p') &= \sum_{m=1}^{+\infty} \mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F(m)) \left(\frac{1+e^{-\lambda}}{2}\right)^{m-1} \\
&\times \Bigg((2p)^m \times (1-2p)(2p')^{m+1} \times (1-2p') \times \frac{e^{-\lambda}}{1-2p'e^{-\lambda}} \\
&+ (2p')^m \times (1-2p')(2p)^{m+1} \times (1-2p) \times \frac{e^{-\lambda}}{1-2pe^{-\lambda}} \\
&+ (2p')^m \times (1-2p')(2p)^m(1-2p) \Bigg)
\end{aligned}
$$

Noting A_{ij} the term $\mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F(m))$, which is constant for all $m \geq 1$

$$\begin{aligned}
T^1_{ij}(p,p') &= A_{ij} \sum_{m=1}^{+\infty} \left(\frac{1+e^{-\lambda}}{2}\right)^{m-1} \\
&\times \Bigg((2p)^m \times (1-2p)(2p')^{m+1} \times (1-2p') \times \frac{e^{-\lambda}}{1-2p'e^{-\lambda}} \\
&+ (2p')^m \times (1-2p')(2p)^{m+1} \times (1-2p) \times \frac{e^{-\lambda}}{1-2pe^{-\lambda}} \\
&+ (2p')^m \times (1-2p')(2p)^m(1-2p) \Bigg) \\
&= A_{ij}(1-2p)(1-2p')4pp'(\frac{2p'e^{-\lambda}}{1-2p'e^{-\lambda}} + \frac{2pe^{-\lambda}}{1-2pe^{-\lambda}} + 1) \sum_{m=0}^{+\infty} (4pp'(1+e^{-\lambda})/2)^m \\
&= A_{ij} \times \frac{(1-2p)(1-2p')4pp'}{1-4pp'(1+e^{-\lambda})/2} \left(\frac{2p'e^{-\lambda}}{1-2p'e^{-\lambda}} + \frac{2pe^{-\lambda}}{1-2pe^{-\lambda}} + 1\right) \\
&= C \times A_{ij}
\end{aligned}$$

where

$$C(p,p') = \frac{(1-2p)(1-2p')4pp'}{1-4pp'(1+e^{-\lambda})/2} \left(\frac{2p'e^{-\lambda}}{1-2p'e^{-\lambda}} + \frac{2pe^{-\lambda}}{1-2pe^{-\lambda}} + 1\right)$$

is a constant does does not depend on i, j. We compute the $m = 0$ sum, noted $T^0_{ij}(p,p')$. We have

$$\begin{aligned}
T^0_{ij}(p,p') = \Bigg(&\mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_1(0)) \times (1-2p)(2p') \times (1-2p') \times \frac{1}{1-2p'e^{-\lambda}} \\
&+ \mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_2(0)) \times (1-2p')(2p) \times (1-2p) \times \frac{1}{1-2pe^{-\lambda}} \\
&+ \mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_3(0)) \times (1-2p')(1-2p) \Bigg)
\end{aligned}$$

And we need to compute the terms $\mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_i(0))$ indivdually.

i=1, i=2 For $i = 1$, we must have $y_0 \neq y_0'$, since $y_0 = \text{STOP}$, and $\text{len}(y') > 0$. Thus, $\mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_1(0)) = e^{-\lambda}$. Similarly, $\mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_2(0)) = e^{-\lambda}$.

i=3 In that case, we must have $y_0 = y_0' = \text{STOP}$, since $\text{len}(y) = \text{len}(y') = 0$. Thus, $\mathbb{E}(e^{-\lambda\delta(y_0\neq y_0')}|F_3(0)) = 1$.
Putting this together, we have

$$T^0_{ij}(p,p') = (1-2p)(1-2p') \left(\frac{2p'e^{-\lambda}}{1-2p'e^{-\lambda}} + \frac{2pe^{-\lambda}}{1-2pe^{-\lambda}} + 1\right)$$

With that notation, we have:

$$
\begin{aligned}
\mathrm{ACMMD}(\mathbb{P}_|, \mathbb{Q}_|) &= \int k(p,p')C(p,p')(A_{11} + A_{22} - 2A_{12})\mathrm{d}\mathbb{P}_X(p)\mathrm{d}\mathbb{P}_X(p') \\
&+ \int k(p,p')(T^0_{11}(p,p') + T^0_{22}(p,p') - 2T^0_{12}(p,p'))\mathrm{d}\mathbb{P}_X(p)\mathrm{d}\mathbb{P}_X(p') \\
&= \int k(p,p')C(p,p')(A_{11} + A_{22} - 2A_{12})\mathrm{d}\mathbb{P}_X(p)\mathrm{d}\mathbb{P}_X(p')
\end{aligned}
$$

since T^0_{ij} does not depend on i, j. We can narrow the variation down even further: by noting $p^A_{ij} = P(\delta(y_i \neq y'_i) = 0|F(m))$ (resp $p^B_{ij} = P(\delta(y_i \neq y'_i) = 0|F(0))$), since $\mathbb{E}(e^{-\lambda\epsilon}) = p(\epsilon = 0)(1-e^{-\lambda})+e^{-\lambda}$ if ϵ is a Bernoulli random variable,

$$
\mathrm{ACMMD}(\mathbb{P}_|, \mathbb{Q}_|) = \int k(p,p')C(p,p')(1 - e^{-\lambda})(p^A_{11} + p^A_{22} - 2p^A_{12})\mathrm{d}\mathbb{P}_X(p)\mathrm{d}\mathbb{P}_X(p')
$$

We now compute the probabilities p^A_{ij} for $i, j \in \{1, 2\}$. In every case, such p^A_{ij} can be written as:

$$
p^A_{ij} = \frac{P(y_0 = y'_0 = A) + P(y_0 = y'_0 = B)}{P(\{y_0 \in \{A, B\}\} \cap \{y'_0 \in \{A, B\}\})} \quad \frac{P(y_0 = y'_0 = A) + P(y_0 = y'_0 = B)}{4pp'}
$$

and we have

$$
\begin{aligned}
p^A_{11} &= \frac{pp' + pp'}{4pp'} = \frac{1}{2} \\
p^A_{22} &= \frac{(p + \Delta p)(p' + \Delta p) + (p - \Delta p)(p' - \Delta p)}{4pp'} = \frac{2pp' + 2\Delta p^2}{4pp'} \\
p^A_{12} &= \frac{(p)(p' + \Delta p) + (p)(p' - \Delta p)}{4pp'} = \frac{1}{2} \\
\implies p^A_{11} + p^A_{22} - 2p^A_{12} &= \frac{2pp' + 2\Delta p^2}{4pp'} - \frac{1}{2} = \frac{2\Delta p^2}{4pp'}
\end{aligned}
$$

Putting it together We thus have

$$
\mathrm{ACMMD}(\mathbb{P}_|, \mathbb{Q}_|) = \int C(p,p')k(p,p')(1 - e^{-\lambda})\frac{2\Delta p^2}{4pp'}\mathrm{d}\mathbb{P}_X(p)\mathrm{d}\mathbb{P}_X(p')
$$

Recalling that

$$
C(p,p') = \frac{(1 - 2p)(1 - 2p')4pp'}{1 - 4pp'(1 + e^{-\lambda})/2}\left(\frac{2p'e^{-\lambda}}{1 - 2p'e^{-\lambda}} + \frac{2pe^{-\lambda}}{1 - 2pe^{-\lambda}} + 1\right)
$$

yields the desired result.