FISHER INFORMATION DETERMINES CAPACITY OF ε-SECURE STEGANOGRAPHY - PROOFS

TOMÁŠ FILLER
(TOMAS.FILLER@BINGHAMTON.EDU)
(HTTP://DDE.BINGHAMTON.EDU/FILLER/)

Abstract. In this report we describe the proofs that were omitted in the conference version of the paper [4].

1. Introduction

This report serves as an additional document supporting our work [4] dealing with steganographic capacity of imperfect stegosystems. The original paper [4] contains enough details and should be read first prior to reading this report. Here, we present more detailed discussions and proofs of the two main theorems. First, Theorem 1 justifies using the Fisher information rate I as a measure of capacity of ε-secure imperfect stegosystems, whereas Theorem 2 assures the existence of I and gives analytical formula for its calculation. We use the same Assumptions 1–3 and notation as described in the original paper. The Fisher information of parametric distribution $Q_\beta^{(n)}$ over n-element vectors at $\beta = 0$ is defined as

$$I_n(0) = E_P\left[\left(\frac{d}{d\beta}\ln Q_\beta^{(n)}(y_1^n)\Big|_{\beta=0}\right)^2\right], \tag{1.1}$$

where $P = Q_\beta\big|_{\beta=0}$. We use Iverson's notation — for logical expression x we define $[x]$ to be 1 if x is true, and 0 otherwise. Several lemmas and examples needed to prove Theorem 2 were moved to the appendix.

Theorem 1. [LAN of the LLRT] *Under Assumptions 1–3, the likelihood ratio*

$$T_{\beta_0}^{(n)}(X) = \ln\left(Q_{\beta_0}^{(n)}(X)/P^{(n)}(X)\right) \tag{1.2}$$

satisfies the local asymptotic normality (LAN), i.e., under both hypotheses and for values of β up to order β^2

$$\sqrt{n}\big(T_\beta^{(n)}/n + \beta^2 I/2\big) \xrightarrow{d} N\big(0,\beta^2 I\big) \text{ under } \mathrm{H}_0 \tag{1.3}$$

$$\sqrt{n}\big(T_\beta^{(n)}/n - \beta^2 I/2\big) \xrightarrow{d} N\big(0,\beta^2 I\big) \text{ under } \mathrm{H}_1, \tag{1.4}$$

where I is the Fisher information rate, $I = \lim_{n\to\infty}\frac{1}{n}I_n(0)$, and $\xrightarrow{d}$ is the convergence in distribution. The detection performance is thus completely described by the deflection coefficient

$$d^2 = \frac{(\sqrt{n}\beta^2 I/2 + \sqrt{n}\beta^2 I/2)^2}{\beta^2 I} = n\beta^2 I.$$

Proof. First, by simple algebra the leading term in the Taylor expansion w.r.t. β of the mean and variance of the likelihood ratio (1.2) is quadratic and consists of the Fisher information rate. This is valid under both H_0 and H_1.

The Gaussianity of the leading terms of the test statistic follows from a variant of the Central Limit Theorem (CLT) (this is discussed in the rest of the proof). The standard proof of the CLT uses a moment generating function and shows that it can be factorized and converges to the moment generating function of a Gaussian random variable for large n. Finally, by using Lévy's continuity theorem, we obtain the convergence in distribution. In our case, the assumption of independence is missing and is replaced by so called "exponential forgetting," which can be used to prove a similar result. This approach was used to prove the CLT for functions of Markov chains [1], because samples far enough can be seen as "almost" independent, which aloows us to use the approach from the i.i.d. case (see [1, §V, Theorem 7.5 on page 228] for an application of this idea.).

In our case, we use the prediction filter (see [3]) to write the statistic as a sum of terms that satisfy exponential forgetting. This type of description is classical in the theory of hidden Markov chains [2, p. 1538]. The exponential forgetting of the prediction filter and its derivatives, which are key to our approach, were shown in [3, Lemma 9]. □

Theorem 2. [Fisher information rate] *Let matrices $\mathbb{A} = (a_{ij})$ and $\mathbb{B}$, defined by matrix $\mathbb{C} = (c_{ij})$, define the cover model and the embedding algorithm under HMC model. Then, the normalized Fisher information $I_n(0)/n$ has a finite limit I as $n \to \infty$. This limit can be written as $I = \mathbf{c}^T\mathbb{F}\mathbf{c}$, where $\mathbf{c}$ is a column vector of size N^2 with elements c_{ij}. The matrix $\mathbb{F}$ of size $N^2 \times N^2$ is defined only in terms of matrix $\mathbb{A}$ (cover source); it does not depend on the embedding algorithm. The elements of matrix $\mathbb{F}$ are defined as*

$$f_{(i,j),(k,l)} = [j = l]V(i,j,k) - U(i,j,k,l), \tag{1.5}$$

where

$$V(i,j,k) = \left(\sum_{z\in\mathcal{X}} \pi_z a_{z,i}\frac{a_{z,k}}{a_{z,j}}\right)\left(\sum_{z\in\mathcal{X}} a_{i,z}\frac{a_{k,z}}{a_{j,z}}\right) \tag{1.6}$$

$$U(i,j,k,l) = \pi_i\left(a_{i,k} - a_{i,l}\frac{a_{j,k}}{a_{j,l}}\right) + \pi_k\left(a_{k,i} - a_{k,j}\frac{a_{l,i}}{a_{l,j}}\right). \tag{1.7}$$

Moreover, $|I_n(0)/n - I| \le \frac{C}{n}$ for some constant C. This constant depends only on elements of matrix $\mathbb{A}$ (does not depend on the embedding algorithm). The quadratic form $I(\mathbf{c}) = \mathbf{c}^T\mathbb{F}\mathbf{c}$ is semidefinite in general.

Proof. The main idea of the theorem, the decomposition of the sequence $I_n(0)/n$ into the quadratic form and its properties, can be obtained directly from the definition of Fisher information (1.1)

See Example 3.

$$\begin{aligned}\frac{1}{n}I_n(0) &= \frac{\ln 2}{n}\frac{\partial^2}{\partial\beta^2}d_n(\beta)\Big|_{\beta=0}\\ &= -\sum_{(i,j)}\sum_{(k,l)}\frac{\ln 2}{n\ln 2}E_P\Bigg[\underbrace{\frac{\partial^2}{\partial b_{ij}b_{kl}}\ln Q_\beta(Y_1^n)\Big|_{\mathbb{B}=\mathbb{I}}}_{\triangleq g(Y_1^n,i,j,k,l)}\Bigg]\underbrace{\left(\frac{\partial b_{ij}}{\partial\beta}\Big|_{\beta=0}\right)}_{=c_{ij}}\underbrace{\left(\frac{\partial b_{kl}}{\partial\beta}\Big|_{\beta=0}\right)}_{=c_{kl}}.\end{aligned}$$

The derivatives of the log-likelihood are evaluated at $\mathbb{B} = \mathbb{I}$ because $\mathbb{B}(\beta) = \mathbb{I} + \beta\mathbb{C}$ and $\beta = 0$. By using $Q_\beta(y_1^n) = \sum_{x_1^n \in \mathcal{X}^n} P(x_1^n) Q_\beta(y_1^n|x_1^n)$, the random variable $g(Y_1^n, i, j, k, l)$ does not depend on the embedding method. This is because the derivatives are evaluated at $\mathbb{B} = \mathbb{I}$ and thus only contain the elements of the cover source transition matrix $\mathbb{A}$.

In the rest of this proof, we show that $-\frac{1}{n} E_P[g(Y_1^n, i, j, k, l)]$ converges to $f_{(i,j),(k,l)}$, for which we find a closed form expression. By Lemma 4,

$$g(y_1^n, i, j, k, l) = L_1(y_1^n, i, j, k, l)\big|_{\mathbb{B}=\mathbb{I}} - [j = l] L_2(y_1^n, i, j, k)\big|_{\mathbb{B}=\mathbb{I}}.$$

By Lemma 6 and Lemma 7,

$$-f_{(i,j),(k,l)} = \underbrace{\lim_{n\to\infty} \frac{1}{n} E_P\Big[L_1(y_1^n, i, j, k, l)\big|_{\mathbb{B}=\mathbb{I}}\Big]}_{=U(i,j,k,l)} - [j = l] \underbrace{\lim_{n\to\infty} \frac{1}{n} E_P\Big[L_2(y_1^n, i, j, k)\big|_{\mathbb{B}=\mathbb{I}}\Big]}_{=V(i,j,k)}.$$

Thus, $f_{(i,j),(k,l)} = [j = l] V(i, j, k) - U(i, j, k, l)$.

The semidefinitness of the quadratic form can be proved by considering an arbitrary i.i.d. source. In this case, the rows of matrix $\mathbb{F}$ are identical and thus linearly dependent. □

Appendix A. Appendix

In this appendix, we present several examples and lemmas we consider useful to illustrate the techniques used in the proofs.

Example 3. [Fisher information and KL divergence] If Q_β is a parametric distribution over Y and $P = Q_\beta|_{\beta=0}$, then the KL divergence between P and Q_β can be written as

$$d(\beta) \triangleq D_{KL}\big(P||Q_\beta\big) = \sum_{y\in Y} P(y) \log_2 \frac{P(y)}{Q_\beta(y)} = \frac{1}{2\ln 2} I(0)\beta^2 + O(\beta^3),$$

where $I(0)$ is the Fisher information of Q_β at $\beta = 0$. If $Q'_\beta(y) = \frac{d}{d\beta} Q_\beta(y)$, then

$$\begin{aligned}
\ln 2 \frac{\partial^2 d(\beta)}{\partial \beta^2}\Big|_{\beta=0} &= \sum_{y\in Y} P(y) \frac{d^2}{d\beta^2} \ln \frac{P(y)}{Q_\beta(y)}\Big|_{\beta=0} \\
&= -\sum_{y\in Y} P(y) \frac{d}{d\beta} \frac{Q'_\beta(y)}{Q_\beta(y)}\Big|_{\beta=0} \\
&= -\underbrace{\sum_{y\in Y} P(y) \frac{Q''_\beta(y)}{P(y)}\Big|_{\beta=0}}_{=0} + \sum_{y\in Y} P(y) \Big(\frac{Q'_\beta(y)}{P(y)}\Big)^2\Big|_{\beta=0} \\
&= \sum_{y\in Y} P(y) \Big(\frac{d}{d\beta} \ln Q_\beta(y)\Big)^2\Big|_{\beta=0} \\
&= E_P\bigg[\Big(\frac{d}{d\beta} \ln Q_\beta(y)\Big)^2\bigg|\beta = 0\bigg] = I(0).
\end{aligned}$$

Lemma 4. *Derivatives of log-likelihood of Q_β (as a function of variables $\{b_{ij}|i,j \in \mathcal{X}\}$) can be written as*

$$\frac{\partial^2}{\partial b_{ij} b_{kl}} \ln Q_\beta(Y_1^n = y_1^n) = L_1(y_1^n, i, j, k, l) - [j = l] L_2(y_1^n, i, j, k),$$

where $i, j, k, l \in \mathcal{X}$, $y_1^n \in \mathcal{X}^n$ and $L_1(y_1^n, i, j, k, l)$ and $L_2(y_1^n, i, j, k)$ are defined in the proof.

Proof. The derivative of $\ln Q_\beta(y_1^n)$ for a fixed $y_1^n \in \mathcal{X}^n$ can be written as

$$\frac{\partial^2}{\partial b_{ij} b_{kl}} \ln Q_\beta(y_1^n) = \frac{\frac{\partial^2}{\partial b_{ij} b_{kl}} Q_\beta(y_1^n)}{Q_\beta(y_1^n)} - \frac{\frac{\partial}{\partial b_{ij}} Q_\beta(y_1^n)}{Q_\beta(y_1^n)} \frac{\frac{\partial}{\partial b_{kl}} Q_\beta(y_1^n)}{Q_\beta(y_1^n)}. \tag{A.1}$$

By the independence of embedding operations (MI embedding), $Q_\beta(y_1^n)$ can be written as

$$Q_\beta(y_1^n) = \sum_{x_1^n \in \mathcal{X}^n} P(x_1^n) \prod_{v=1}^{n} b_{x_v, y_v}. \tag{A.2}$$

Calculate the terms $\frac{\partial}{\partial b_{ij}} Q_\beta(y_1^n)$ and $\frac{\partial^2}{\partial b_{ij} b_{kl}} Q_\beta(y_1^n)$.

For a fixed $y_1^n \in \mathcal{X}^n$, equation (A.2) can be seen as a polynomial w.r.t. the fixed term b_{ij}. The derivative of such a polynomial w.r.t. a given b_{ij} can be written in the following general form (see Example 5 for more details)

$$\frac{\partial}{\partial b_{ij}} Q_\beta(y_1^n) = \sum_{t \in J(j)} S_y(t, i), \tag{A.3}$$

where $J(j) = \{1 \le t \le n | y_t = j\}$ and

$$S_y(t, i) = \sum_{x_1^n \in \mathcal{X}^n, x_t = i} P(x_1^n) \prod_{v=1, v \ne t}^{n} b_{x_v y_v}. \tag{A.4}$$

In the derivative of (A.2), it is sufficient to sum only over the products that contain b_{ij}. If the term is in the form Cb_{ij}^k for some constants k and C, then its derivative is Ckb_{ij}^{k-1}. This is achieved by summing over all elements from the set $J(j)$, fixing $x_t = i$ for each $t \in J(j)$, and putting 1 instead of b_{ij} in the product.

Similarly, we obtain a general form for $(\partial^2/\partial b_{ij} b_{kl}) Q_\beta(y_1^n)$ as

$$\begin{aligned}
\frac{\partial^2}{\partial b_{ij} b_{kl}} Q_\beta(y_1^n) &= \frac{\partial}{\partial b_{kl}} \sum_{t \in J(j)} S_y(t, i) \\
&= \sum_{t \in J(j)} \sum_{\substack{x_1^n \in \mathcal{X}^n \\ x_t = i}} P(x_1^n) \frac{\partial}{\partial b_{kl}} \prod_{\substack{v=1 \\ v \ne t}}^{n} b_{x_v, y_v} \\
&= \sum_{t \in J(j)} \sum_{\substack{x_1^n \in \mathcal{X}^n \\ x_t = i}} P(x_1^n) \sum_{t' \in J(l) \setminus \{t\}} [x_{t'} = k] \prod_{\substack{v=1 \\ v \notin \{t, t'\}}}^{n} b_{x_v, y_v} && \text{(A.5)} \\
&= \sum_{t \in J(j)} \sum_{t' \in J(l) \setminus \{t\}} \sum_{\substack{x_1^n \in \mathcal{X}^n \\ x_t = i, x_{t'} = k}} P(x_1^n) \prod_{\substack{v=1 \\ v \notin \{t, t'\}}}^{n} b_{x_v, y_v} \\
&= \sum_{t \in J(j)} \sum_{t' \in J(l) \setminus \{t\}} S_y(t, t', i, k), && \text{(A.6)}
\end{aligned}$$

where

$$S_y(t,t',i,k) = \sum_{\substack{x_1^n \in \mathcal{X}^n \\ x_t = i, x_{t'} = k}} P(x_1^n) \prod_{\substack{v=1 \\ v \notin \{t,t'\}}}^{n} b_{x_v,y_v}.$$

In (A.5), we used the fact that $(d/dx)Cx^k = Ckx^{k-1} = \sum_{v=1}^{k} Cx^{k-1}$ again. End of calculation.

We now substitute (A.3) and (A.6) into (A.1) and obtain

$$\begin{aligned}\frac{\partial^2}{\partial b_{ij} b_{kl}} \ln Q_\beta(y_1^n) &= \sum_{t \in J(j)} \sum_{t' \in J(l) \setminus \{t\}} \frac{S_y(t,t',i,k)}{Q_\beta(y_1^n)} - \sum_{t \in J(j)} \frac{S_y(t,i)}{Q_\beta(y_1^n)} \sum_{t' \in J(l)} \frac{S_y(t',k)}{Q_\beta(y_1^n)} \\ &= L_1(y_1^n,i,j,k,l) - [j=l] L_2(y_1^n,i,j,k),\end{aligned}$$

where

$$L_1(y_1^n,i,j,k,l) = \sum_{t \in J(j)} \sum_{t' \in J(l) \setminus \{t\}} \left(\frac{S_y(t,t',i,k)}{Q_\beta(y_1^n)} - \frac{S_y(t,i)}{Q_\beta(y_1^n)} \frac{S_y(t,k)}{Q_\beta(y_1^n)} \right) \tag{A.7}$$

$$L_2(y_1^n,i,j,k) = \sum_{t \in J(j)} \frac{S_y(t,i)}{Q_\beta(y_1^n)} \frac{S_y(t,k)}{Q_\beta(y_1^n)}. \tag{A.8}$$

□

Example 5. $\mathcal{X} = \{1,2\}$, $n=3$, $y_1^3 = (y_1,y_2,y_3) = (2,2,1)$

$$\begin{aligned}Q(y_1^3) &= \sum_{x_1^3 \in \mathcal{X}^3} Q(y_1^3 | x_1^3) P(x_1^3) \\ &= \Big(P(1,1,1)b_{1,1} + P(1,1,2)b_{2,1} \Big) b_{1,2}^2 + \Big(P(1,2,1)b_{2,2}b_{1,1} + \\ &\quad + P(1,2,2)b_{2,2}b_{2,1} + P(2,1,1)b_{2,2}b_{1,1} + P(2,1,2)b_{2,2}b_{2,1} \Big) b_{1,2} + \\ &\quad + \Big(P(2,2,1)b_{2,2}b_{2,2}b_{1,1} + P(2,2,2)b_{2,2}b_{2,2}b_{2,1} \Big)\end{aligned}$$

If $x = b_{1,2}$, then the previous result can be represented as $Ax^2 + Bx + C$. The partial derivative of $Q(y_1^3)$ w.r.t. $b_{1,2}$ accepts the following form

$$\begin{aligned}\frac{\partial Q(y_1^3)}{\partial b_{1,2}} &= 2P(1,1,1)b_{1,2}b_{1,1} + 2P(1,1,2)b_{1,2}b_{2,1} \\ &\quad + P(1,2,1)b_{2,2}b_{1,1} + P(1,2,2)b_{2,2}b_{2,1} + P(2,1,1)b_{2,2}b_{1,1} + P(2,1,2)b_{2,2}b_{2,1} \\ &= P(1,1,1)b_{1,2}b_{1,1} + P(1,1,2)b_{1,2}b_{2,1} + P(1,2,1)b_{2,2}b_{1,1} + P(1,2,2)b_{2,2}b_{2,1} \\ &\quad + P(1,1,1)b_{1,2}b_{1,1} + P(1,1,2)b_{1,2}b_{2,1} + P(2,1,1)b_{2,2}b_{1,1} + P(2,1,2)b_{2,2}b_{2,1},\end{aligned}$$

where in the last step we sum all terms for $x_1 = 1$ and $x_2 = 1$. We do not need to sum the terms with $x_1^2 = (2,2)$, because they are zero after the derivation (they do not contain $b_{1,2}$). This can be written in a general form as

$$\frac{\partial}{\partial b_{1,2}} Q(y_1^3) = \sum_{t \in J(2)} S_y(t,1),$$

where $J(2) = \{1,2\}$ (the set of indices t such that $y_t = 2$) and $S_y(t,i)$ is defined by (A.4) and

$$S_y(1,1) = P(1,1,1)b_{1,2}b_{1,1} + P(1,1,2)b_{1,2}b_{2,1} + P(1,2,1)b_{2,2}b_{1,1} + P(1,2,2)b_{2,2}b_{2,1},$$

$$S_y(2,1) = P(1,1,1)b_{1,2}b_{1,1} + P(1,1,2)b_{1,2}b_{2,1} + P(2,1,1)b_{2,2}b_{1,1} + P(2,1,2)b_{2,2}b_{2,1}.$$

The second derivative, e.g., if $i = 1$, $j = 2$, $k = 1$, $l = 1$

$$\frac{\partial^2 Q(y_1^3)}{\partial b_{1,2} b_{1,1}} = 2\Big(P(1,1,1)b_{1,2} + P(1,2,1)b_{2,2}\Big),$$

can be derived in a similar manner and written in a general form as in (A.6), where $J(j) = \{1,2\}$, $J(l) = \{3\}$, and

$$\begin{aligned} S_y(1,3,1,1) &= P(1,1,1)b_{1,2} + P(1,2,1)b_{2,2} \\ S_y(2,3,1,1) &= P(1,1,1)b_{1,2} + P(1,2,1)b_{2,2}. \end{aligned}$$

Lemma 6. *Let $L_1(y_1^n, i, j, k, l)$ be function of $y_1^n \in \mathcal{X}^n$ and let matrix $\mathbb{B} = (b_{ij})$ be defined by* (A.7)*. Then, for all $i, j, k, l \in \mathcal{X}$ the following limit exists*

$$\lim_{n\to\infty} \frac{1}{n} E_P\big[L_1(Y_1^n, i, j, k, l)\big|_{\mathbb{B}=\mathbb{I}}\big]$$

and is equal to $U(i,j,k,l)$ as defined by (1.7)*. The series converges to the limit with rate $1/n$.*

Proof. First, we show some properties of the terms in (A.7). Assuming $|t - t'| > 1$, by $\mathbb{B} = \mathbb{I}$, $y_t = j$, and $y_{t'} = l$ (remember $t \in J(j)$ and $t' \in J(l)$), we have

$$\begin{aligned} &\frac{S_y(t,t',i,k)}{Q_\beta(y_1^n)} - \frac{S_y(t,i)}{Q_\beta(y_1^n)} \frac{S_y(t',k)}{Q_\beta(y_1^n)}\Big|_{\mathbb{B}=\mathbb{I}} = \\ &= \frac{a_{y_{t-1},i} a_{i,y_{t+1}}}{a_{y_{t-1},j} a_{j,y_{t+1}}} \frac{a_{y_{t'-1},k} a_{k,y_{t'+1}}}{a_{y_{t'-1},l} a_{l,y_{t'+1}}} - \frac{a_{y_{t-1},i} a_{i,y_{t+1}}}{a_{y_{t-1},j} a_{j,y_{t+1}}} \frac{a_{y_{t'-1},k} a_{k,y_{t'+1}}}{a_{y_{t'-1},l} a_{l,y_{t'+1}}} = 0. \end{aligned}$$

This means that the only non-zero terms in (A.7) can be the terms for $|t - t'| = 1$. If $t = t' - 1$, $t \notin \{1, n-1\}$, then for $t \in J(j)$ and $t' \in J(l)$

$$\begin{aligned} \frac{S_y(t,t',i,k)}{Q_\beta(y_1^n)} - \frac{S_y(t,i)}{Q_\beta(y_1^n)} \frac{S_y(t',k)}{Q_\beta(y_1^n)}\Big|_{\mathbb{B}=\mathbb{I}} &= \frac{a_{y_{t-1},i}}{a_{y_{t-1},j}} \left(\frac{a_{i,k}}{a_{j,l}} - \frac{a_{i,y_{t+1}}}{a_{j,y_{t+1}}} \frac{a_{y_{t'-1},k}}{a_{y_{t'-1},l}} \right) \frac{a_{k,y_{t'+1}}}{a_{l,y_{t'+1}}} \\ &= \frac{a_{y_{t-1},i}}{a_{y_{t-1},j}} \left(\frac{a_{i,k}}{a_{j,l}} - \frac{a_{i,l}}{a_{j,l}} \frac{a_{j,k}}{a_{j,l}} \right) \frac{a_{k,y_{t+2}}}{a_{l,y_{t+2}}}, \end{aligned}$$

because $y_{t'-1} = y_t = j$, and $y_{t+1} = y_{t'} = l$. If $t = t' + 1$, $t \notin \{2, n\}$, then

$$\begin{aligned} \frac{S_y(t,t',i,k)}{Q_\beta(y_1^n)} - \frac{S_y(t,i)}{Q_\beta(y_1^n)} \frac{S_y(t',k)}{Q_\beta(y_1^n)}\Big|_{\mathbb{B}=\mathbb{I}} &= \frac{S_y(t',t,k,i)}{Q_{\beta=0}(y_1^n)} - \frac{S_y(t',k)}{Q_{\beta=0}(y_1^n)} \frac{S_y(t,i)}{Q_{\beta=0}(y_1^n)} \\ &= \frac{a_{y_{t-2},k}}{a_{y_{t-2},l}} \left(\frac{a_{k,i}}{a_{l,j}} - \frac{a_{k,j}}{a_{l,j}} \frac{a_{l,i}}{a_{l,j}} \right) \frac{a_{i,y_{t+1}}}{a_{j,y_{t+1}}}. \end{aligned}$$

By using both results, we can write

$$\begin{aligned} &\frac{1}{n} \sum_{y_1^n \in \mathcal{X}^n} P(y_1^n) L_1(Y_1^n, i, j, k, l)\big|_{\mathbb{B}=\mathbb{I}} = \\ &= \frac{1}{n} \sum_{t=2}^{n-2} \sum_{y_1^n \in \mathcal{X}^n} P(y_1^n) \left(\big[y_t^{t+1} = (j,l)\big] \frac{a_{y_{t-1},i}}{a_{y_{t-1},j}} \left(\frac{a_{i,k}}{a_{j,l}} - \frac{a_{i,l}}{a_{j,l}} \frac{a_{j,k}}{a_{j,l}} \right) \frac{a_{k,y_{t+2}}}{a_{l,y_{t+2}}} \right) + \\ &+ \frac{1}{n} \sum_{t=3}^{n-1} \sum_{y_1^n \in \mathcal{X}^n} P(y_1^n) \left(\big[y_{t-1}^{t} = (l,j)\big] \frac{a_{y_{t-2},k}}{a_{y_{t-2},l}} \left(\frac{a_{k,i}}{a_{l,j}} - \frac{a_{k,j}}{a_{l,j}} \frac{a_{l,i}}{a_{l,j}} \right) \frac{a_{i,y_{t+1}}}{a_{j,y_{t+1}}} \right) + g_n, \end{aligned}$$

where g_n is the sum for $(t,t') \in \{(1,2),(2,1),(n-1,n),(n,n-1)\}$. The series g_n can be sandwiched by $0 \le g_n \le C\frac{1}{n}$ for some constant C and thus $\lim_{n\to\infty} g_n = 0$

with rate $O(1/n)$. This constant depends only on elements of matrix $\mathbb{A}$. We can continue and write

$$
\begin{aligned}
&\frac{1}{n}\sum_{y_1^n\in\mathcal{X}^n} P(y_1^n)L_1(Y_1^n,i,j,k,l)\big|_{\mathbb{B}=\mathbb{I}} - g_n \\
&= \frac{1}{n}\sum_{t=2}^{n-2}\sum_{z_1,z_2\in\mathcal{X}} \frac{a_{z_1,i}}{a_{z_1,j}}\left(\frac{a_{i,k}}{a_{j,l}} - \frac{a_{i,l}}{a_{j,l}}\frac{a_{j,k}}{a_{j,l}}\right)\frac{a_{k,z_2}}{a_{l,z_2}}\underbrace{P\big(y_{t-1}^{t+2}=(z_1,j,l,z_2)\big)}_{\pi_{z_1}a_{z_1,j}a_{j,l}a_{l,z_2}} + \\
&\quad + \frac{1}{n}\sum_{t=3}^{n-1}\sum_{z_2,z_1\in\mathcal{X}} \frac{a_{z_2,k}}{a_{z_2,l}}\left(\frac{a_{k,i}}{a_{l,j}} - \frac{a_{k,j}}{a_{l,j}}\frac{a_{l,i}}{a_{l,j}}\right)\frac{a_{i,z_1}}{a_{j,z_1}}\underbrace{P\big(y_{t-2}^{t+1}=(z_2,l,j,z_1)\big)}_{\pi_{z_2}a_{z_2,l}a_{l,j}a_{j,z_1}} \\
&= \frac{n-3}{n}\sum_{z_1,z_2\in\mathcal{X}}\left\{\pi_{z_1}a_{z_1,i}\left(a_{i,k} - a_{i,l}\frac{a_{j,k}}{a_{j,l}}\right)a_{k,z_2} + \pi_{z_2}a_{z_2,k}\left(a_{k,i} - a_{k,j}\frac{a_{l,i}}{a_{l,j}}\right)a_{i,z_1}\right\} \\
&= \frac{n-3}{n}\left\{\left(a_{i,k} - a_{i,l}\frac{a_{j,k}}{a_{j,l}}\right)\sum_{z_1\in\mathcal{X}}\pi_{z_1}a_{z_1,i} + \left(a_{k,i} - a_{k,j}\frac{a_{l,i}}{a_{l,j}}\right)\sum_{z_2\in\mathcal{X}}\pi_{z_2}a_{z_2,k}\right\} \\
&= \frac{n-3}{n}\left\{\pi_i\left(a_{i,k} - a_{i,l}\frac{a_{j,k}}{a_{j,l}}\right) + \pi_k\left(a_{k,i} - a_{k,j}\frac{a_{l,i}}{a_{l,j}}\right)\right\}.
\end{aligned}
$$

Finally, the limit for $n \to \infty$ is

$$
\begin{aligned}
U(i,j,k,l) &\triangleq \lim_{n\to\infty}\frac{1}{n}E_P\left[L_1(Y_1^n,i,j,k,l)\big|_{\mathbb{B}=\mathbb{I}}\right] \\
&= \pi_i\left(a_{i,k} - a_{i,l}\frac{a_{j,k}}{a_{j,l}}\right) + \pi_k\left(a_{k,i} - a_{k,j}\frac{a_{l,i}}{a_{l,j}}\right). \qquad \text{(A.9)}
\end{aligned}
$$

□

Lemma 7. *Let $L_2(y_1^n,i,j,k)$ be function of $y_1^n \in \mathcal{X}^n$, and matrix $\mathbb{B} = (b_{ij})$ as defined by* (A.8)*, then for all $i,j,k \in \mathcal{X}$ the following limit exists*

$$
\lim_{n\to\infty}\frac{1}{n}E_P\left[L_2(Y_1^n,i,j,k)\big|_{\mathbb{B}=\mathbb{I}}\right]
$$

and is equal to $V(i,j,k)$ as defined by (1.6)*. The series converges to the limit with rate $1/n$.*

Proof. Let $y_1^n \in \mathcal{X}^n$ be a fixed realization of random variable $Y_1^n \in \mathcal{X}^n$. By substituting $\mathbb{B} = \mathbb{I}$, we simplify the term $L_2(y_1^n,i,j,k)$

$$
\begin{aligned}
L_2(y_1^n,i,j,k)\big|_{\mathbb{B}=\mathbb{I}} &= \sum_{t\in J(j)}\frac{S_y(t,i)}{Q_\beta(y_1^n)}\frac{S_y(t,k)}{Q_\beta(y_1^n)}\Big|_{\mathbb{B}=\mathbb{I}} \\
&= \sum_{t\in J(j)}\frac{P\big((y_1^{t-1},i,y_{t+1}^n)\big)}{P(y_1^n)}\frac{P\big((y_1^{t-1},k,y_{t+1}^n)\big)}{P(y_1^n)} \\
&= \sum_{t\in J(j)}\frac{a_{y_{t-1},i}a_{i,y_{t+1}}}{a_{y_{t-1},j}a_{j,y_{t+1}}}\frac{a_{y_{t-1},k}a_{k,y_{t+1}}}{a_{y_{t-1},j}a_{j,y_{t+1}}}.
\end{aligned}
$$

Now, we can rewrite the series $\frac{1}{n}E_P\big[L_2(Y_1^n,i,j,k)\big|_{\mathbb{B}=\mathbb{I}}\big]$ to calculate the limit

$$\begin{aligned}
&\frac{1}{n}E_P\big[L_2(Y_1^n,i,j,k)\big|_{\mathbb{B}=\mathbb{I}}\big] = \frac{1}{n}\sum_{y_1^n\in\mathcal{X}^n} P(y_1^n)L_2(y_1^n,i,j,k)\big|_{\mathbb{B}=\mathbb{I}} = \\
&= \frac{1}{n}\sum_{t=1}^{n}\sum_{y_1^n\in\mathcal{X}^n} P(y_1^n)\left([t\in J_y(j)]\frac{a_{y_{t-1},i}a_{i,y_{t+1}}}{a_{y_{t-1},j}a_{j,y_{t+1}}}\frac{a_{y_{t-1},k}a_{k,y_{t+1}}}{a_{y_{t-1},j}a_{j,y_{t+1}}}\right) \\
&= \frac{1}{n}\sum_{t=2}^{n-1}\sum_{y_1^n\in\mathcal{X}^n} P(y_1^n)\left([t\in J_y(j)]\frac{a_{y_{t-1},i}a_{i,y_{t+1}}}{a_{y_{t-1},j}a_{j,y_{t+1}}}\frac{a_{y_{t-1},k}a_{k,y_{t+1}}}{a_{y_{t-1},j}a_{j,y_{t+1}}}\right)+ \\
&\quad \underbrace{+\frac{1}{n}\sum_{y_1^n\in\mathcal{X}^n} P(y_1^n)\left([1\in J_y(j)]\frac{\pi_i a_{i,y_2}}{\pi_j a_{j,y_2}}\frac{\pi_k a_{k,y_2}}{\pi_j a_{j,y_2}} + [n\in J_y(j)]\frac{a_{y_{n-1},i}}{a_{y_{n-1},j}}\frac{a_{y_{n-1},k}}{a_{y_{n-1},j}}\right)}_{\triangleq f_n} = \\
&= \frac{1}{n}\sum_{t=2}^{n-1}\sum_{z_1,z_2\in\mathcal{X}}\frac{a_{z_1,i}a_{i,z_2}}{a_{z_1,j}a_{j,z_2}}\frac{a_{z_1,k}a_{k,z_2}}{a_{z_1,j}a_{j,z_2}}\sum_{y_1^n\in\mathcal{X}^n}P(y_1^n)\Big[y_{t-1}^{t+1}=(z_1,j,z_2)\Big] + f_n \\
&= \frac{1}{n}\sum_{t=2}^{n-1}\sum_{z_1,z_2\in\mathcal{X}}\frac{a_{z_1,i}a_{i,z_2}}{a_{z_1,j}a_{j,z_2}}\frac{a_{z_1,k}a_{k,z_2}}{a_{z_1,j}a_{j,z_2}}P\Big(Y_{t-1}^{t+1}=(z_1,j,z_2)\Big) + f_n \\
&= \frac{1}{n}\sum_{t=2}^{n-1}\sum_{z_1,z_2\in\mathcal{X}}\frac{a_{z_1,i}a_{i,z_2}}{a_{z_1,j}a_{j,z_2}}\frac{a_{z_1,k}a_{k,z_2}}{a_{z_1,j}a_{j,z_2}}\pi_{z_1}a_{z_1,j}a_{j,z_2} + f_n \\
&= \frac{n-2}{n}\sum_{z_1,z_2\in\mathcal{X}}\pi_{z_1}a_{z_1,i}a_{i,z_2}\frac{a_{z_1,k}a_{k,z_2}}{a_{z_1,j}a_{j,z_2}} + f_n.
\end{aligned}$$

Finally, we can calculate the limit

$$\begin{aligned}
V(i,j,k) &\triangleq \lim_{n\to\infty}\frac{1}{n}E_P\Big[L_2(Y_1^n,i,j,k)\big|_{\mathbb{B}=\mathbb{I}}\Big] \\
&= \sum_{z_1,z_2\in\mathcal{X}}\pi_{z_1}a_{z_1,i}a_{i,z_2}\frac{a_{z_1,k}a_{k,z_2}}{a_{z_1,j}a_{j,z_2}} \\
&= \left(\sum_{z\in\mathcal{X}}\pi_z a_{z,i}\frac{a_{z,k}}{a_{z,j}}\right)\left(\sum_{z\in\mathcal{X}}a_{i,z}\frac{a_{k,z}}{a_{j,z}}\right)
\end{aligned}$$

because $0\le f_n\le C/n$ for some constant C and thus the rate of convergence is $O(1/n)$. The constant C depends only on elements of matrix $\mathbb{A}$. □

References

[1] J. L. Doob. *Stochastic processes.* Wiley, New York, 1st edition, 1953.

[2] Y. Ephraim and N. Merhav. Hidden Markov processes. *Information Theory, IEEE Transactions on*, 48(6):1518–1569, June 2002.

[3] T. Filler. Important properties of normalized KL-divergence under HMC model. Technical report, DDE Lab, SUNY Binghamton, 2008. http://dde.binghamton.edu/filler/kl-divergence-hmc.pdf.

[4] T. Filler and J. Fridrich. Fisher information determines capacity of ε-secure steganography. In *Information Hiding, 11th International Workshop*, Darmstadt, Germany, June 7–10 2009. Springer-Verlag, Berlin.