Various Views on the Trapdoor Channel and an Upper Bound on its Capacity

Tobias Lutz

Abstract

The problem of maximizing the n-letter mutual information of the trapdoor channel is considered. It is shown that $\frac{1}{2}\log_2\left(\frac{5}{2}\right) \approx 0.6610$ bits per use is an upper bound on the capacity of the trapdoor channel. This upper bound, which is the tightest upper bound known, proves that feedback increases the capacity. In the second part of the paper, two novel views are presented on the trapdoor channel. First, by deriving the underlying iterated function system (IFS), it is shown that the trapdoor channel with input blocks of length n can be regarded as the nth element of a sequence of shapes approximating a fractal. Second, an algorithm is presented that fully characterizes the trapdoor channel and resembles the recursion of generating all permutations of a given string.

Index Terms

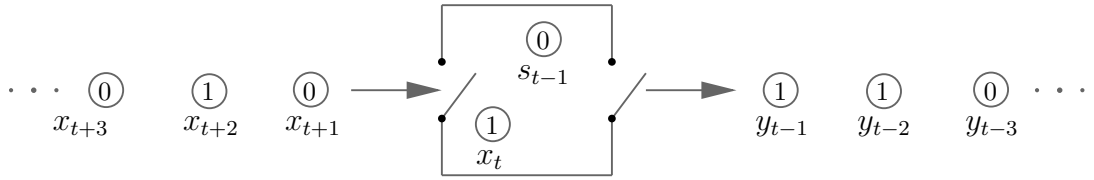
Trapdoor channel, Lagrange multipliers, convex optimization, iterated function systems, fractals, channels with memory, recursions, permutations.

I. Introduction and Channel Model

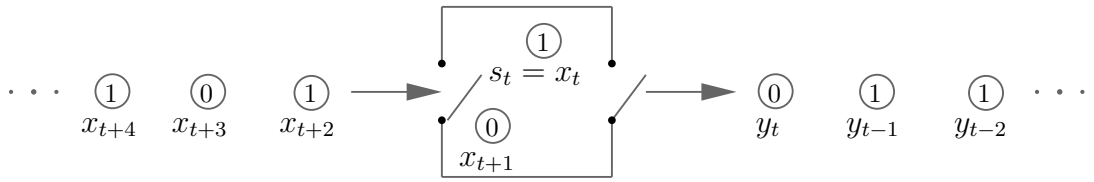
The trapdoor channel was introduced by David Blackwell in 1961 [1] and is used by Robert Ash both as a book cover and as an introductory example for channels with memory [2]. The mapping of channel inputs to channel outputs can be described as follows. Consider a box that

Submitted for publication on September 17, 2014. This paper was presented in part at the IEEE Int. Symp. Inf. Theory in Honolulu, HI, USA, 2014.

Tobias Lutz is with the Lehrstuhl für Nachrichtentechnik, Technische Universität München, D-80290 München, Germany (e-mail: tobi.lutz@tum.de).



(a) The trapdoor channel at time t.



(b) The trapdoor channel at time $t+1$. Here $y_t = s_{t-1}$.

Fig. 1: The trapdoor channel.

contains a ball that is labeled $s_0 \in \{0, 1\}$, where the index 0 refers to time 0. Both the sender and the receiver know the initial ball. In time slot 1, the sender places a new ball labeled $x_1 \in \{0, 1\}$ in the box. In the same time slot, the receiver chooses one of the two balls s_0 or x_1 at random while the other ball remains in the box. The chosen ball is interpreted as channel output y_1 at time $t = 1$ while the remaining ball becomes the channel state s_1. The same procedure is applied in every future channel use. In time slot 2, for instance, the sender places a new ball $x_2 \in \{0, 1\}$ in the box and the corresponding channel output y_2 is either x_2 or s_1. The transmission process is visualized in Fig. 1. Fig. 1a shows the trapdoor channel at time t when the sender places ball x_t in the box. In the same time slot, the receiver chooses randomly one of the two balls x_t or s_{t-1} as channel output, in the figure the ball labeled with s_{t-1}. Consequently, the upcoming channel state s_t becomes x_t (see Fig. 1b). At time $t+1$ the sender places a new ball x_{t+1} in the box and the receiver draws y_{t+1} from s_t and x_{t+1}. Table I depicts the probability of an output y_t given an input x_t and state s_{t-1}.

Despite the simplicity of the trapdoor channel, deriving its capacity seems challenging and is an open problem. One feature that makes the problem cumbersome is that the distribution of the output symbols may depend on events happening arbitrarily far back in the past since each ball has a positive probability to remain in the channel over any finite number of channel uses. Instead of maximizing $I(X;Y)$ one rather has to consider the multi-letter mutual information, i.e., $\limsup_{n\to\infty} I(\mathbf{X}^n; \mathbf{Y}^n)$.

TABLE I: Transition Probabilities of the Trapdoor Channel

x_t	s_{t-1}	$P_{Y_t\|X_tS_{t-1}}(y_t=0\|x_t,s_{t-1})$	$P_{Y_t\|X_tS_{t-1}}(y_t=1\|x_t,s_{t-1})$
0	0	1	0
0	1	0.5	0.5
1	0	0.5	0.5
1	1	0	1

Let $\mathbf{P}_{n|s_0}$ denote the matrix of conditional probabilities of output sequences of length n given input sequences of length n where the initial state equals s_0. The following ordering of the entries of $\mathbf{P}_{n|s_0}$ is assumed. Row indices represent input sequences and column indices represent output sequences. To be more precise, the $(i,j)^{\text{th}}$ entry of $\left[\mathbf{P}_{n|s_0}\right]$, indicated as $\left[\mathbf{P}_{n|s_0}\right]_{i,j}$, is the conditional probability of the binary output sequence corresponding to the integer $j-1$ given the binary input sequence corresponding the the integer $i-1$, $1 \leq i,j \leq 2^n$. For instance, if $n=3$, then $\left[\mathbf{P}_{3|s_0}\right]_{5,3}$ denotes the conditional probability that the channel input $x_1,x_2,x_3 = 1,0,0$ will be mapped to the channel output $y_1,y_2,y_3 = 0,1,0$. It was shown in [3] that $\mathbf{P}_{n|s_0}$, $s_0 \in \{0,1\}$, satisfies the recursion laws

$$\mathbf{P}_{n+1|0} = \begin{bmatrix} \mathbf{P}_{n|0} & \mathbf{0} \\ \frac{1}{2}\mathbf{P}_{n|1} & \frac{1}{2}\mathbf{P}_{n|0} \end{bmatrix} \tag{1}$$

$$\mathbf{P}_{n+1|1} = \begin{bmatrix} \frac{1}{2}\mathbf{P}_{n|1} & \frac{1}{2}\mathbf{P}_{n|0} \\ \mathbf{0} & \mathbf{P}_{n|1} \end{bmatrix}, \tag{2}$$

where the initial matrices are given by $\mathbf{P}_{0|0} = \mathbf{P}_{0|1} = 1$. Ahlswede and Kaspi [4] derived the *zero-error capacity* of the trapdoor channel, which equals 0.5 b/u. Permuter et al. [5] considered the trapdoor channel under the additional assumption of having a unit delay feedback link available from the receiver to the sender. They established that the capacity of the trapdoor channel with feedback is equal to the logarithm of the golden ratio.

In this paper, we consider the problem of maximizing the n-letter mutual information of the trapdoor channel for any $n \in \mathbb{N}$. We relax the problem by permitting distributions that are not probability distributions. The resulting optimization problem is convex but the feasible set is larger than the probability simplex. Using the method of Lagrange multipliers via a theorem presented in [2], we find explicit solutions for any $n \in \mathbb{N}$. It is then shown that $\frac{1}{2}\log_2\left(\frac{5}{2}\right) \approx$

0.6610 b/u is an upper bound on the capacity of the trapdoor channel. Specifically, the same absolute maximum $\frac{1}{2}\log_2\left(\frac{5}{2}\right) \approx 0.6610$ b/u results for all trapdoor channels which process input blocks of even length. And the sequence of absolute maxima corresponding to trapdoor channels which process inputs of odd lengths converges to $\frac{1}{2}\log_2\left(\frac{5}{2}\right)$ b/u from below as the block length increases. Unfortunately, the absolute maxima of our relaxed optimization are attained outside the probability simplex. Otherwise we would have established the capacity. Nevertheless, $\frac{1}{2}\log_2\left(\frac{5}{2}\right) \approx 0.6610$ b/u is, to the best of our knowledge, the tightest capacity upper bound. Moreover, this bound is less than the feedback capacity of the trapdoor channel proving that feedback increases the capacity. In the second part of the paper, we propose two different views on the trapdoor channel. Based on the underlying stochastic matrices (1) and (2), the trapdoor channel can be described geometrically as a fractal or algorithmically as a recursive procedure.

The organization of the paper is as follows. Section II presents the derivation of the upper bound. In particular, the problem is set up and a useful result from the literature is reviewed. Two recursions are then developed for the trapdoor channel based on which the main result is derived. Section III interprets the trapdoor channel as a fractal and derives the underlying iterated function system (IFS). To be more precise, we introduce the mathematical background of fractals and, in particular, the notion of an IFS in Subsection III-A. In Subsection III-B, the IFS corresponding to the trapdoor channel is derived. In Section IV, we study the trapdoor channel as a recursive procedure. The paper is concluded with Section V.

A. Notation

The notation is as follows. The symbols $\mathbb{N}_0$ and $\mathbb{N}$ refer to the natural numbers with and without 0, respectively. The input corresponding to the ith row of $\mathbf{P}_{n|s_0}$ is denoted as $\mathbf{x}_i$. Further, $\mathbf{I}_n$ denotes the $2^n \times 2^n$ identity matrix, $\tilde{\mathbf{I}}_n$ is a $2^n \times 2^n$ matrix whose secondary diagonal entries are all equal to 1 while the remaining entries are all equal to 0, and $\mathbf{1}_n$ denotes a column vector of length 2^n consisting only of ones. The vector $\mathbf{1}_n^T$ is the transpose of $\mathbf{1}_n$. The functions $\exp_2(\cdot)$ and $\log_2(\cdot)$ indicate the exponential function to base 2 and the logarithm to base 2. If applied to a vector/matrix, $\log_2(\cdot)$ or $\exp_2(\cdot)$ of each element is taken and a vector/matrix results. The symbol $\circ$ refers to the Hadarmard product, i.e., the entry wise product of two matrices. The canonical basis vectors of $\mathbb{R}^3$ are denoted by $\mathbf{e}_x$, $\mathbf{e}_y$ and $\mathbf{e}_z$. They are assumed to be row vectors. Finally, the n-fold composition of a function, say Φ, is denoted as $\Phi^{\circ n}$.

II. A Lagrange Multiplier Approach to the Trapdoor Channel

A. Problem Formulation

We derive an upper bound on the capacity of the trapdoor channel. Specifically, for any $n \in \mathbb{N}$, we find a solution to the optimization problem

$$\begin{aligned}
\underset{P_{\mathbf{X}^n}}{\text{maximize}} \quad & \frac{1}{n} I\left(\mathbf{X}^n; \mathbf{Y}^n | s_0\right) \\
& = \frac{1}{n} \sum_{i=1}^{2^n} \sum_{j=1}^{2^n} p_i \left[\mathbf{P}_{n|s_0}\right]_{i,j} \log \frac{\left[\mathbf{P}_{n|s_0}\right]_{i,j}}{\sum_{k=1}^{2^n} p_k \left[\mathbf{P}_{n|s_0}\right]_{k,j}} \qquad (3) \\
\text{subject to} \quad & \sum_{i=1}^{2^n} p_i = 1 \qquad (4) \\
& \sum_{k=1}^{2^n} p_k \left[\mathbf{P}_{n|s_0}\right]_{k,j} \geq 0 \quad \text{for all } 1 \leq j \leq 2^n. \qquad (5)
\end{aligned}$$

Note that $P_{\mathbf{X}^n}$ is a 2^n-sequences $(p_1, \ldots, p_{2^n})$ where p_i denotes the probability of the i^{th} input sequence $\mathbf{x}_i$, i.e., the binary sequence corresponding to the integer $i-1$. Constraint (5) guarantees that the argument of the logarithm does not become negative. The feasible set, defined by (4) and (5), is convex. It includes the set of probability mass functions, but might be larger. To see this note that (5) is a weighted sum of all p_k where each weight $\left[\mathbf{P}_{n|s_0}\right]_{k,j}$ is non negative. Clearly, (4) and (5) are satisfied by probability distributions. However, there might exist "distributions" which involve negative values and sum up to one but still satisfy (5). Moreover, the objective function $n^{-1} I\left(\mathbf{X}^n; \mathbf{Y}^n | s_0\right)$ is concave on the set of "distributions" satisfying (4) and (5). Consequently, the optimization problem is convex and every solution maximizes $n^{-1} I\left(\mathbf{X}^n; \mathbf{Y}^n | s_0\right)$. In the following, we use the notation

$$C_n^{\uparrow} \overset{def}{=} \max_{P_{\mathbf{X}^n}} n^{-1} I\left(\mathbf{X}^n; \mathbf{Y}^n | s_0\right).$$

Taking the limit of the sequence $\left(C_n^{\uparrow}\right)_{n \in \mathbb{N}}$, one obtains either the capacity of the trapdoor channel or an upper bound on the capacity, depending on whether the limit is attained inside or outside the set of probability distributions. Since it does not matter whether the optimization is with respect to initial state 0 or 1 (due to symmetry reasons), we do not have to distinguish between *lower capacity* and *upper capacity* [6, Chapter 4.6]

B. Using a Result from the Literature

The reason for considering (5) and not the more natural constraints $p_k \geq 0$ for all k is that a closed form solution can be obtained by applying the method of *Lagrange multipliers* to (3) and (4). As a byproduct, (5) will be automatically satisfied. In particular, setting the partial derivatives of

$$\frac{1}{n} I\left(\mathbf{X}^n; \mathbf{Y}^n | s_0\right) + \lambda \sum_{i=1}^{2^n} p_i \tag{6}$$

with respect to each of the p_i equal to zero results in a closed form solution of the considered optimization problem.

This was done in [2, Theorem 3.3.3] for general discrete memoryless channels which are square and non singular. Note that $\mathbf{P}_{n|s_0}$ is square and non singular (see, e.g., Lemma II.2 (b)). Moreover, we assume that the channel $\mathbf{P}_{n|s_0}$ is memoryless by repeatedly using it over a large number of input blocks of length n. Consequently, $C_n^{\uparrow}$ might be an upper bound on the capacity of the channel $\mathbf{P}_{n|s_0}$. The reason is that some input blocks possibly drive the channel $\mathbf{P}_{n|s_0}$ into the opposite state $s_0 \oplus 1$, i.e., the upcoming input block sees channel $\mathbf{P}_{n|s_0 \oplus 1}$ (whose $C_n^{\uparrow}$ is equal to $C_n^{\uparrow}$ of $\mathbf{P}_{n|s_0}$ by symmetry) but not $\mathbf{P}_{n|s_0}$. However, by assuming that the channel does not change over time, the sender always knows the channel state before a new block is transmitted. Hence, $C_n^{\uparrow}$ might be an upper bound (even though it is attained on the set of probability distributions). Nevertheless, this issue can be ignored if n goes to infinity because in the asymptotic regime the channel $\mathbf{P}_{n|s_0}$ is used only once.

In summary, it is valid to apply [2, Theorem 3.3.3] which yields

$$C_n^{\uparrow} = \frac{1}{n} \log_2 \sum_{j=1}^{2^n} \exp_2 \left(- \sum_{i=1}^{2^n} \left[\mathbf{P}_{n|s_0}^{-1} \right]_{j,i} H\left(\mathbf{Y}^n | \mathbf{X}^n = \mathbf{x}_i\right) \right), \tag{7}$$

attained at

$$p_k = 2^{-C_n^{\uparrow}} d_k, \quad k = 1, \ldots, 2^n \tag{8}$$

where

$$d_k = \sum_{j=1}^{2^n} \left[\mathbf{P}_{n|s_0}^{-1} \right]_{j,k} \exp_2 \left(- \sum_{i=1}^{2^n} \left[\mathbf{P}_{n|s_0}^{-1} \right]_{j,i} H\left(\mathbf{Y}^n | \mathbf{X}^n = \mathbf{x}_i\right) \right). \tag{9}$$

Clearly, $(p_1, \ldots, p_{2^n})$ is a probability distribution only if $d_k \geq 0$. The Lagrangian (6) does not involve constraint (5). However, the proof of [2, Theorem 3.3.3] shows that

$$\sum_{k=1}^{2^n} p_k \left[\mathbf{P}_{n|s_0} \right]_{k,j} = \exp \left(\lambda - \sum_{i=1}^{M} \left[\mathbf{P}_{n|s_0}^{-1} \right]_{j,i} H\left(\mathbf{Y}^n | \mathbf{X}^n = \mathbf{x}_i\right) - 1 \right) \tag{10}$$

for all $1 \leq j \leq 2^n$. Hence, (5) is satisfied.

For computational reasons, we write (7) in matrix vector notation, which reads

$$C_n^{\uparrow} = \frac{1}{n} \log_2 \left\{ \mathbf{1}_n^T \exp_2 \left[\mathbf{P}_{n|s_0}^{-1} \left(\mathbf{P}_{n|s_0} \circ \log_2 \mathbf{P}_{n|s_0} \right) \mathbf{1}_n \right] \right\}, \tag{11}$$

where $\mathbf{1}_n$ is a column vector of length 2^n consisting only of ones while $\circ$ denotes the Hadamard product. Observe that

$$- \left(\mathbf{P}_{n|s_0} \circ \log_2 \mathbf{P}_{n|s_0} \right) \mathbf{1}_n = \left[H\left(\mathbf{Y}^n | \mathbf{X}^n = \mathbf{x}_1\right), \ldots, H\left(\mathbf{Y}^n | \mathbf{X}^n = \mathbf{x}_{2^n}\right) \right]^T. \tag{12}$$

In the remainder, we use (11) instead of (7). In particular, we find exact numerical expressions for (11) in Theorem II.9 below.

C. Useful Recursions

Definition II.1. *(a) The conditional entropy vector* $\mathbf{h}_{n|s_0}$ *of* $\mathbf{P}_{n|s_0}$*,* $s_0 \in \{0,1\}$*, is*

$$\mathbf{h}_{n|s_0} \stackrel{def}{=} \left[H(\mathbf{Y}^n | \mathbf{X}^n = \mathbf{x}_1) \quad \ldots \quad H(\mathbf{Y}^n | \mathbf{X}^n = \mathbf{x}_{2^n}) \right]^T \tag{13}$$

$$= - \left(\mathbf{P}_{n|s_0} \circ \log_2 \mathbf{P}_{n|s_0} \right) \mathbf{1}_n, \tag{14}$$

where $n \in \mathbb{N}_0$.

(b) The weighted conditional entropy vector $\boldsymbol{\omega}_{n|s_0}$ *of* $\mathbf{P}_{n|s_0}$*,* $s_0 \in \{0,1\}$*, is*

$$\boldsymbol{\omega}_{n|s_0} \stackrel{def}{=} -\mathbf{P}_{n|s_0}^{-1} \cdot \mathbf{h}_{n|s_0} \tag{15}$$

$$= \mathbf{P}_{n|s_0}^{-1} \left(\mathbf{P}_{n|s_0} \circ \log_2 \mathbf{P}_{n|s_0} \right) \mathbf{1}_n, \tag{16}$$

where $n \in \mathbb{N}_0$.

The following three lemmas provide tools that we need in order to prove recursions for $\mathbf{h}_{n|s_0}$ and $\boldsymbol{\omega}_{n|s_0}$, as stated in Lemma II.5 and Lemma II.6.

Lemma II.2. *(a) The trapdoor channel matrices* $\mathbf{P}_{2n+2|0}$ *and* $\mathbf{P}_{2n+2|1}$*,* $n \in \mathbb{N}_0$*, satisfy the following recursions:*

$$\mathbf{P}_{2n+2|0} = \begin{bmatrix} \mathbf{P}_{2n|0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{1}{2}\mathbf{P}_{2n|1} & \frac{1}{2}\mathbf{P}_{2n|0} & \mathbf{0} & \mathbf{0} \\ \frac{1}{4}\mathbf{P}_{2n|1} & \frac{1}{4}\mathbf{P}_{2n|0} & \frac{1}{2}\mathbf{P}_{2n|0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2}\mathbf{P}_{2n|1} & \frac{1}{4}\mathbf{P}_{2n|1} & \frac{1}{4}\mathbf{P}_{2n|0} \end{bmatrix} \tag{17}$$

$$\mathbf{P}_{2n+2|1} = \begin{bmatrix} \frac{1}{4}\mathbf{P}_{2n|1} & \frac{1}{4}\mathbf{P}_{2n|0} & \frac{1}{2}\mathbf{P}_{2n|0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2}\mathbf{P}_{2n|1} & \frac{1}{4}\mathbf{P}_{2n|1} & \frac{1}{4}\mathbf{P}_{2n|0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{2}\mathbf{P}_{2n|1} & \frac{1}{2}\mathbf{P}_{2n|0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{P}_{2n|1} \end{bmatrix}. \tag{18}$$

(b) Let $\mathbf{M}_0 \stackrel{def}{=} \mathbf{P}_{2n|0}^{-1}\mathbf{P}_{2n|1}\mathbf{P}_{2n|0}^{-1}$ *and* $\mathbf{M}_1 \stackrel{def}{=} \mathbf{P}_{2n|1}^{-1}\mathbf{P}_{2n|0}\mathbf{P}_{2n|1}^{-1}$. *The inverses of* $\mathbf{P}_{2n+2|0}$ *and* $\mathbf{P}_{2n+2|1}$, $n \in \mathbb{N}_0$, *satisfy the following recursions:*

$$\mathbf{P}_{2n+2|0}^{-1} = \begin{bmatrix} \mathbf{P}_{2n|0}^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\mathbf{M}_0 & 2\mathbf{P}_{2n|0}^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{P}_{2n|0}^{-1} & 2\mathbf{P}_{2n|0}^{-1} & \mathbf{0} \\ 2\mathbf{M}_0\mathbf{P}_{2n|1}\mathbf{P}_{2n|0}^{-1} & -3\mathbf{M}_0 & -2\mathbf{M}_0 & 4\mathbf{P}_{2n|0}^{-1} \end{bmatrix} \tag{19}$$

$$\mathbf{P}_{2n+2|1}^{-1} = \begin{bmatrix} 4\mathbf{P}_{2n|1}^{-1} & -2\mathbf{M}_1 & -3\mathbf{M}_1 & 2\mathbf{M}_1\mathbf{P}_{2n|0}\mathbf{P}_{2n|1}^{-1} \\ \mathbf{0} & 2\mathbf{P}_{2n|1}^{-1} & -\mathbf{P}_{2n|1}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 2\mathbf{P}_{2n|1}^{-1} & -\mathbf{M}_1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{P}_{2n|1}^{-1} \end{bmatrix}. \tag{20}$$

Proof: (a): Substituting $\mathbf{P}_{2n+1|0}$ and $\mathbf{P}_{2n+1|1}$ into $\mathbf{P}_{2n+2|0}$ and $\mathbf{P}_{2n+2|1}$, where the four matrices are expressed as in (1) and (2), yields (17) and (18).

(b): Two versions of the matrix inversion lemma are [7]

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{D}^{-1} \end{bmatrix} \tag{21}$$

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{D}^{-1} \end{bmatrix}. \tag{22}$$

Now divide (17) and (18) into four blocks of equal size. A twofold application of (21) and (22), first to $\mathbf{P}_{2n+2|0}$ and $\mathbf{P}_{2n+2|1}$ and, subsequently, to each of the blocks of $\mathbf{P}_{2n+2|0}$ and $\mathbf{P}_{2n+2|1}$ yields (19) and (20). ■

Lemma II.3. *Let* $\tilde{\mathbf{I}}_n$ *be the* $2^n \times 2^n$ *matrix whose secondary diagonal entries are equal to* 1 *while the remaining entries are* 0. *Let* $\mathbf{A}$ *be an arbitrary* $2^n \times 2^n$ *matrix and* $\mathbf{b}$ *an arbitrary column vector of size* 2^n. *A left and right multiplication of* $\mathbf{A}$ *with* $\tilde{\mathbf{I}}_n$ *results in a permutation of the*

elements of $\mathbf{A}$. *In particular, the element* $[\mathbf{A}]_{i,j}$ *of* $\mathbf{A}$ *is shifted to position* $(2^n+1-i, 2^n+1-j)$ *in* $\tilde{\mathbf{I}}_n\mathbf{A}\tilde{\mathbf{I}}_n$, $1 \leq i, j \leq 2^n$. *Similarly, a left multiplication of* $\mathbf{b}$ *with* $\tilde{\mathbf{I}}_n$ *turns* $\mathbf{b}$ *upside down, i.e., the* i^{th} *entry of* $\mathbf{b}$ *is shifted to the* $(2^n+1-i)^{th}$ *position in* $\tilde{\mathbf{I}}_n\mathbf{b}$, $1 \leq i \leq 2^n$. *Moreover,* $\left(\tilde{\mathbf{I}}_n\mathbf{A}\tilde{\mathbf{I}}_n\right) \circ \log_2\left(\tilde{\mathbf{I}}_n\mathbf{A}\tilde{\mathbf{I}}_n\right) = \tilde{\mathbf{I}}_n\left(\mathbf{A} \circ \log_2 \mathbf{A}\right)\tilde{\mathbf{I}}_n$.

Proof: The first two properties follow from the rules of matrix multiplication and noting that the i^{th} row and the i^{th} column of $\tilde{\mathbf{I}}_n$ has a one at position 2^n+1-i and zeros else. The final equality holds because it does not matter whether the Hadamard product and the elementwise logarithm is applied before or after permuting the elements of $\mathbf{A}$. ■

A transformation relating $\mathbf{P}_{n|0}$ to $\mathbf{P}_{n|1}$, $\mathbf{P}_{n|0}^{-1}$ to $\mathbf{P}_{n|1}^{-1}$, $\mathbf{h}_{n|0}$ to $\mathbf{h}_{n|1}$ and $\boldsymbol{\omega}_{n|0}$ to $\boldsymbol{\omega}_{n|1}$ is derived next.

Lemma II.4. *Let* $\mathbf{P}_{n|0}$ *and* $\mathbf{P}_{n|1}$ *be trapdoor channel matrices,* $n \in \mathbb{N}_0$. *We have the following identities:*

(a)

$$\mathbf{P}_{n|1} = \tilde{\mathbf{I}}_n\mathbf{P}_{n|0}\tilde{\mathbf{I}}_n \tag{23}$$

$$\mathbf{P}_{n|0} = \tilde{\mathbf{I}}_n\mathbf{P}_{n|1}\tilde{\mathbf{I}}_n \tag{24}$$

(b)

$$\mathbf{P}_{n|1}^{-1} = \tilde{\mathbf{I}}_n\mathbf{P}_{n|0}^{-1}\tilde{\mathbf{I}}_n \tag{25}$$

$$\mathbf{P}_{n|0}^{-1} = \tilde{\mathbf{I}}_n\mathbf{P}_{n|1}^{-1}\tilde{\mathbf{I}}_n \tag{26}$$

(c)

$$\mathbf{h}_{n|1} = \tilde{\mathbf{I}}_n\mathbf{h}_{n|0} \tag{27}$$

$$\mathbf{h}_{n|0} = \tilde{\mathbf{I}}_n\mathbf{h}_{n|1} \tag{28}$$

(d)

$$\boldsymbol{\omega}_{n|1} = \tilde{\mathbf{I}}_n\boldsymbol{\omega}_{n|0} \tag{29}$$

$$\boldsymbol{\omega}_{n|0} = \tilde{\mathbf{I}}_n\boldsymbol{\omega}_{n|1} \tag{30}$$

(e) The row sums of $\mathbf{P}_{n|0}^{-1}$ *and* $\mathbf{P}_{n|1}^{-1}$ *are* 1.

Proof: (a): The proof is by induction. For $n = 0$, the identities $\mathbf{P}_{0|1} = \tilde{\mathbf{I}}_0 \mathbf{P}_{0|0} \tilde{\mathbf{I}}_0$ and $\mathbf{P}_{0|0} = \tilde{I}_0 \mathbf{P}_{0|1} \tilde{\mathbf{I}}_0$ clearly hold. Now suppose that (23) and (24) are true if n is replaced by $n-1$. Then we have

$$\tilde{\mathbf{I}}_n \mathbf{P}_{n|0} \tilde{\mathbf{I}}_n = \begin{bmatrix} \mathbf{0} & \tilde{\mathbf{I}}_{n-1} \\ \tilde{\mathbf{I}}_{n-1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{n-1|0} & \mathbf{0} \\ \frac{1}{2}\mathbf{P}_{n-1|1} & \frac{1}{2}\mathbf{P}_{n-1|0} \end{bmatrix} \begin{bmatrix} 0 & \tilde{\mathbf{I}}_{n-1} \\ \tilde{\mathbf{I}}_{n-1} & 0 \end{bmatrix} \tag{31}$$

$$= \begin{bmatrix} \frac{1}{2}\tilde{\mathbf{I}}_{n-1}\mathbf{P}_{n-1|0}\tilde{\mathbf{I}}_{n-1} & \frac{1}{2}\tilde{\mathbf{I}}_{n-1}\mathbf{P}_{n-1|1}\tilde{\mathbf{I}}_{n-1} \\ \mathbf{0} & \tilde{\mathbf{I}}_{n-1}\mathbf{P}_{n-1|0}\tilde{\mathbf{I}}_{n-1} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2}\mathbf{P}_{n-1|1} & \frac{1}{2}\mathbf{P}_{n-1|0} \\ \mathbf{0} & \mathbf{P}_{n-1|1} \end{bmatrix} \tag{32}$$

$$= \mathbf{P}_{n-1|1}, \tag{33}$$

where (31) and (33) are due to the recursive expressions (1) and (2) while (32) follows from the induction hypothesis. It remains to show (24). But (24) is a direct consequence of the just proven equation and using the identity $\tilde{\mathbf{I}}_n \tilde{\mathbf{I}}_n = \mathbf{I}_n$.

(b): Follows immediately from (a) and the identity $\tilde{\mathbf{I}}_n \tilde{\mathbf{I}}_n = \mathbf{I}_n$.

(c): Starting with the definition of $\mathbf{h}_{n|1}$, we have

$$\begin{aligned} \mathbf{h}_{n|1} &= -\left(\mathbf{P}_{n|1} \circ \log_2 \mathbf{P}_{n|1}\right) \mathbf{1}_n \\ &= -\left[\left(\tilde{\mathbf{I}}_n \mathbf{P}_{n|0} \tilde{\mathbf{I}}_n\right) \circ \log_2 \left(\tilde{\mathbf{I}}_n \mathbf{P}_{n|0} \tilde{\mathbf{I}}_n\right)\right] \mathbf{1}_n && (34) \\ &= -\tilde{\mathbf{I}}_n \left(\mathbf{P}_{n|0} \circ \log_2 \mathbf{P}_{n|0}\right) \tilde{\mathbf{I}}_n \mathbf{1}_n && (35) \\ &= \tilde{\mathbf{I}}_n \mathbf{h}_{n|0}, \end{aligned}$$

where (34) and (35) hold because of (23) and Lemma II.3, respectively.

Equation (28) follows from (27) and the identity $\tilde{\mathbf{I}}_n \tilde{\mathbf{I}}_n = \mathbf{I}_n$.

(d): Starting with the definition of $\boldsymbol{\omega}_{n|1}$, we have

$$\begin{aligned} \boldsymbol{\omega}_{n|1} &= -\mathbf{P}_{n|1}^{-1} \mathbf{h}_{n|1} \\ &= -\tilde{\mathbf{I}}_n \mathbf{P}_{n|0}^{-1} \mathbf{h}_{n|0} && (36) \\ &= \tilde{\mathbf{I}}_n \boldsymbol{\omega}_{n|0}, \end{aligned}$$

where (36) follows by replacing $\mathbf{P}_{n|1}$ and $\mathbf{h}_{n|1}$ with (23) and (27), respectively, and using the identity $\tilde{\mathbf{I}}_n \tilde{\mathbf{I}}_n = \mathbf{I}_n$.

Equation (30) follows from (29) and the identity $\tilde{\mathbf{I}}_n\tilde{\mathbf{I}}_n = \mathbf{I}_n$.

(e): A standard way to compute $\mathbf{P}_{n|0}^{-1}$ is by Gauss-Jordan elimination. That is, a sequence of elementary row operations is applied to the augmented matrix $\begin{bmatrix}\mathbf{P}_{n|0} & \mathbf{I}_n\end{bmatrix}$ such that $\begin{bmatrix}\mathbf{I}_n & \mathbf{P}_{n|0}^{-1}\end{bmatrix}$ eventually results. Clearly, $\mathbf{P}_{n|0}$ and $\mathbf{I}_n$ are stochastic matrices, i.e., all row sums are equal to one. Thus, at each stage of performing elementary row operations, the row sum of the left matrix equals the row sum of the right matrix. In particular, $\mathbf{P}_{n|0}^{-1}$ has the same row sum as $\mathbf{I}_n$. The same arguments hold for $\mathbf{P}_{n|1}^{-1}$. ■

We can now state the recursive law for the conditional entropy vector.

Lemma II.5. *For $n \geq 1$, $\mathbf{h}_{2n+2|0}$ satisfies the recursion*

$$\mathbf{h}_{2n+2|0} = \begin{bmatrix} \mathbf{h}_{2n|0} \\ \frac{1}{2}\mathbf{h}_{2n|0} + \frac{1}{2}\tilde{\mathbf{I}}_{2n}\mathbf{h}_{2n|0} + \mathbf{1}_{2n} \\ \frac{3}{4}\mathbf{h}_{2n|0} + \frac{1}{4}\tilde{\mathbf{I}}_{2n}\mathbf{h}_{2n|0} + \frac{3}{2}\mathbf{1}_{2n} \\ \frac{1}{4}\mathbf{h}_{2n|0} + \frac{3}{4}\tilde{\mathbf{I}}_{2n}\mathbf{h}_{2n|0} + \frac{3}{2}\mathbf{1}_{2n} \end{bmatrix}. \tag{37}$$

The initial value for $n = 0$ is given by $\mathbf{h}_{0|0} = 0$.

Before proving Lemma II.5, we remark that in order to refer to the i^{th} subvector, $1 \leq i \leq 4$, of the conditional entropy vector $\mathbf{h}_{2n+2|0}$, i.e., the subvector composed of the $\left((i-1)\cdot 2^{2n}+1\right)^{\text{th}}$ to the $\left(i \cdot 2^{2n}\right)^{\text{th}}$ element, we use the superscript (i). For instance, $\mathbf{h}_{2n+2|0}^{(2)}$ refers to $\frac{1}{2}\mathbf{h}_{2n|0} + \frac{1}{2}\tilde{\mathbf{I}}_{2n}\mathbf{h}_{2n|0} + \mathbf{1}_{2n}$. The same notation is used for the weighted conditional entropy vector $\boldsymbol{\omega}_{2n+2|0}$.

Proof of Lemma II.5: The initial value $\mathbf{h}_{0|0}$ can be computed using $\mathbf{P}_{0|0} = 1$ in (14). To show (37), we replace $\mathbf{P}_{2n+2|0}$ in (14) with (17) and compute each of the four entries of the resulting vector. Clearly, $\mathbf{h}_{2n+2|0}^{(1)} = -\left(\mathbf{P}_{2n|0} \circ \log_2 \mathbf{P}_{2n|0}\right)\mathbf{1}_{2n} = \mathbf{h}_{2n|0}$. The three remaining terms are

$$\begin{aligned}
\mathbf{h}_{2n+2|0}^{(2)} &= \left[-\frac{1}{2}\mathbf{P}_{2n|1} \circ \log_2\left(\frac{1}{2}\mathbf{P}_{2n|1}\right) - \frac{1}{2}\mathbf{P}_{2n|0} \circ \log_2\left(\frac{1}{2}\mathbf{P}_{2n|0}\right)\right]\mathbf{1}_{2n} \\
&= \left[\frac{1}{2}\mathbf{P}_{2n|1} - \frac{1}{2}\left(\tilde{\mathbf{I}}_{2n}\mathbf{P}_{2n|0}\tilde{\mathbf{I}}_{2n}\right) \circ \log_2\left(\tilde{\mathbf{I}}_{2n}\mathbf{P}_{2n|0}\tilde{\mathbf{I}}_{2n}\right) + \frac{1}{2}\mathbf{P}_{2n|0}\right. \\
&\qquad \left. -\frac{1}{2}\mathbf{P}_{2n|0} \circ \log_2 \mathbf{P}_{2n|0}\right]\mathbf{1}_{2n} \qquad (38) \\
&= \mathbf{1}_{2n} - \frac{1}{2}\tilde{\mathbf{I}}_{2n}\left(\mathbf{P}_{2n|0} \circ \log_2 \mathbf{P}_{2n|0}\right)\mathbf{1}_{2n} + \frac{1}{2}\mathbf{h}_{2n|0} \qquad (39) \\
&= \frac{1}{2}\mathbf{h}_{2n|0} + \frac{1}{2}\tilde{\mathbf{I}}_{2n}\mathbf{h}_{2n|0} + \mathbf{1}_{2n};
\end{aligned}$$

$$\begin{aligned}
\mathbf{h}^{(3)}_{2n+2|0} &= \left[-\frac{1}{4}\mathbf{P}_{2n|1}\circ\log_2\left(\frac{1}{4}\mathbf{P}_{2n|1}\right)-\frac{1}{4}\mathbf{P}_{2n|0}\circ\log_2\left(\frac{1}{4}\mathbf{P}_{2n|0}\right)\right.\\
&\quad\left.-\frac{1}{2}\mathbf{P}_{2n|0}\circ\log_2\left(\frac{1}{2}\mathbf{P}_{2n|0}\right)\right]\mathbf{1}_{2n}\\
&= \left[\frac{1}{2}\mathbf{P}_{2n|1}-\frac{1}{4}\left(\tilde{\mathbf{I}}_{2n}\mathbf{P}_{2n|0}\tilde{\mathbf{I}}_{2n}\right)\circ\log_2\left(\tilde{\mathbf{I}}_{2n}\mathbf{P}_{2n|0}\tilde{\mathbf{I}}_{2n}\right)+\mathbf{P}_{2n|0}\right.\\
&\quad\left.-\frac{3}{4}\mathbf{P}_{2n|0}\circ\log_2\mathbf{P}_{2n|0}\right]\mathbf{1}_{2n} \qquad (40)\\
&= \frac{3}{2}\mathbf{1}_{2n}-\frac{1}{4}\tilde{\mathbf{I}}_{2n}\left(\mathbf{P}_{2n|0}\circ\log_2\mathbf{P}_{2n|0}\right)\mathbf{1}_{2n}+\frac{3}{4}\mathbf{h}_{2n|0} \qquad (41)\\
&= \frac{3}{4}\mathbf{h}_{2n|0}+\frac{1}{4}\tilde{\mathbf{I}}_{2n}\mathbf{h}_{2n|0}+\frac{3}{2}\mathbf{1}_{2n};
\end{aligned}$$

$$\begin{aligned}
\mathbf{h}^{(4)}_{2n+2|0} &= \left[-\frac{1}{2}\mathbf{P}_{2n|1}\circ\log_2\left(\frac{1}{2}\mathbf{P}_{2n|1}\right)-\frac{1}{4}\mathbf{P}_{2n|1}\circ\log_2\left(\frac{1}{4}\mathbf{P}_{2n|1}\right)\right.\\
&\quad\left.-\frac{1}{4}\mathbf{P}_{2n|0}\circ\log_2\left(\frac{1}{4}\mathbf{P}_{2n|0}\right)\right]\mathbf{1}_{2n}\\
&= \left[\mathbf{P}_{2n|1}-\frac{3}{4}\left(\tilde{\mathbf{I}}_{2n}\mathbf{P}_{2n|0}\tilde{\mathbf{I}}_{2n}\right)\circ\log_2\left(\tilde{\mathbf{I}}_{2n}\mathbf{P}_{2n|0}\tilde{\mathbf{I}}_{2n}\right)\right.\\
&\quad\left.+\frac{1}{2}\mathbf{P}_{2n|0}-\frac{1}{4}\mathbf{P}_{2n|0}\circ\log_2\mathbf{P}_{2n|0}\right]\mathbf{1}_{2n} \qquad (42)\\
&= \frac{3}{2}\mathbf{1}_{2n}-\frac{3}{4}\tilde{\mathbf{I}}_{2n}\left(\mathbf{P}_{2n|0}\circ\log_2\mathbf{P}_{2n|0}\right)\mathbf{1}_{2n}+\frac{1}{4}\mathbf{h}_{2n|0} \qquad (43)\\
&= \frac{1}{4}\mathbf{h}_{2n|0}+\frac{3}{4}\tilde{\mathbf{I}}_{2n}\mathbf{h}_{2n|0}+\frac{3}{2}\mathbf{1}_{2n}.
\end{aligned}$$

Observe that (38), (40), (42) follow from using (23) and

$$\log_2\left(\frac{1}{2^r}\mathbf{P}_{2n|s_0}\right)=\log_2\left(\frac{1}{2^r}\mathbf{1}_{2n\times 2n}\circ\mathbf{P}_{2n|s_0}\right)=-r\mathbf{1}_{2n\times 2n}+\log_2\mathbf{P}_{2n|s_0},\quad r=1,2.$$

Summing up the scaled vectors $\mathbf{P}_{2n|0}\mathbf{1}_{2n}$ and $\mathbf{P}_{2n|1}\mathbf{1}_{2n}$ in (38), (40), (42) yields the first term in (39), (41), (43). Finally, the second term in (39), (41),(43) follows because it does not matter whether the Hadamard product and the elementwise logarithm is applied before or after permuting the elements of $\mathbf{P}_{2n|0}$ (see Lemma II.3). ■

Lemma II.6. *(a) For $n\in\mathbb{N}_0$, $\boldsymbol{\omega}_{2n|0}$ satisfies the recursion*

$$\boldsymbol{\omega}_{2n+2|0}=\begin{bmatrix}\boldsymbol{\omega}_{2n|0}\\ \boldsymbol{\omega}_{2n|0}-2\cdot\mathbf{1}_{2n}\\ \boldsymbol{\omega}_{2n|0}-2\cdot\mathbf{1}_{2n}\\ \boldsymbol{\omega}_{2n|0}\end{bmatrix} \qquad (44)$$

with initial value $\boldsymbol{\omega}_{0|0} = 0$.

(b) For $n \in \mathbb{N}$, $\boldsymbol{\omega}_{2n+1|0}$ *satisfies the recursion*

$$\boldsymbol{\omega}_{2n+1|0} = \begin{bmatrix} \boldsymbol{\omega}_{2n-1|0} \\ \tilde{\mathbf{I}}_{2n-1}\boldsymbol{\omega}_{2n-1|0} \\ \boldsymbol{\omega}_{2n-1|0} - 2 \cdot \mathbf{1}_{2n-1} \\ \tilde{\mathbf{I}}_{2n-1}\boldsymbol{\omega}_{2n-1|0} - 2 \cdot \mathbf{1}_{2n-1} \end{bmatrix} \tag{45}$$

with initial value $\boldsymbol{\omega}_{1|0} = \begin{bmatrix} 0 & -2 \end{bmatrix}^T$.

Remark II.7. *The weighted conditional entropy vector* $\boldsymbol{\omega}_{n|0}$ *is a palindrome for even* $n \in \mathbb{N}_0$*, i.e., the vector reads the same backwards as forward.*

Proof of Lemma II.6: (a): We show by induction that (44) holds. The case $n = 0$ can be verified using Definition II.1(b) and noting that $\mathbf{P}_{0|0} = \mathbf{P}_{0|0}^{-1} = 1$. Now assume that (44) holds for $n-1$. In order to show (44) for n, we evaluate $\boldsymbol{\omega}_{2n+2|0}$ using (15) and replacing $\mathbf{P}_{2n+2|0}^{-1}$ and $\mathbf{h}_{2n+2|0}$ with (19) and (37). Then

$$\boldsymbol{\omega}_{2n+2|0} = \begin{bmatrix} -\mathbf{P}_{2n|0}^{-1}\mathbf{h}_{2n+2|0}^{(1)} \\ \mathbf{P}_{2n|0}^{-1}\left(\mathbf{P}_{2n|1}\mathbf{P}_{2n|0}^{-1}\mathbf{h}_{2n+2|0}^{(1)} - 2\mathbf{h}_{2n+2|0}^{(2)}\right) \\ \mathbf{P}_{2n|0}^{-1}\left(\mathbf{h}_{2n+2|0}^{(2)} - 2\mathbf{h}_{2n+2|0}^{(3)}\right) \\ \mathbf{M}_0\left(-2\mathbf{P}_{2n|1}\mathbf{P}_{2n|0}^{-1}\mathbf{h}_{2n+2|0}^{(1)} + 3\mathbf{h}_{2n+2|0}^{(2)} + 2\mathbf{h}_{2n+2|0}^{(3)}\right) - 4\mathbf{P}_{2n|0}^{-1}\mathbf{h}_{2n+2|0}^{(4)} \end{bmatrix}. \tag{46}$$

Recall from Lemma II.5 that $\mathbf{h}_{2n+2|0}^{(1)} = \mathbf{h}_{2n|0}$. Hence, by definition, the first entry of (46) is equal to $\boldsymbol{\omega}_{2n|0}$. Replacing $\mathbf{h}_{2n+2|0}^{(1)}$ and $\mathbf{h}_{2n+2|0}^{(2)}$ in (46) with the corresponding expressions from Lemma II.5, we obtain

$$\boldsymbol{\omega}_{2n+2|0}^{(2)} = \mathbf{P}_{2n|0}^{-1}\left(\mathbf{P}_{2n|1}\mathbf{P}_{2n|0}^{-1}\mathbf{h}_{2n|0} - \mathbf{h}_{2n|0} - \tilde{\mathbf{I}}_{2n}\mathbf{h}_{2n|0} - 2 \cdot \mathbf{1}_{2n}\right). \tag{47}$$

In order to simplify (47), observe that

$$-\tilde{\mathbf{I}}_{2n}\boldsymbol{\omega}_{2n|0} + \boldsymbol{\omega}_{2n|0} = 0 \tag{48}$$

since $\boldsymbol{\omega}_{2n|0}$ is a palindrome by hypothesis. Further, using (15), (26) and the relation $\tilde{\mathbf{I}}_{2n}\tilde{\mathbf{I}}_{2n} = \mathbf{I}_{2n}$, (48) reads

$$\mathbf{P}_{2n|0}^{-1} \cdot \mathbf{h}_{2n|0} - \mathbf{P}_{2n|1}^{-1}\tilde{\mathbf{I}}_{2n} \cdot \mathbf{h}_{2n|0} = 0, \tag{49}$$

which becomes, after a right multiplication with $\mathbf{P}_{2n|1}$,

$$\mathbf{P}_{2n|1}\mathbf{P}_{2n|0}^{-1}\mathbf{h}_{2n|0} - \tilde{\mathbf{I}}_{2n} \cdot \mathbf{h}_{2n|0} = 0. \tag{50}$$

Finally, using (50) in (47) as well as the definition of $\boldsymbol{\omega}_{2n|0}$ and noting that $2\mathbf{P}_{2n|0}^{-1}\mathbf{1}_{2n} = 2 \cdot \mathbf{1}_{2n}$ (since $\mathbf{P}_{2n|0}^{-1}$ is a stochastic matrix by Lemma II.4 (e)), we obtain

$$\boldsymbol{\omega}_{2n+2|0}^{(2)} = \boldsymbol{\omega}_{2n|0} - 2 \cdot \mathbf{1}_{2n}.$$

We continue with the third entry of (46). After replacing $\mathbf{h}_{2n+2|0}^{(2)}$ and $\mathbf{h}_{2n+2|0}^{(3)}$ in (46) with the corresponding expressions from Lemma II.5, it immediately follows that $\boldsymbol{\omega}_{2n+2|0}^{(3)} = \boldsymbol{\omega}_{2n|0} - 2 \cdot \mathbf{1}_{2n}$. Regarding the fourth entry in (46), we start with the first term in parentheses. Observe that

$$\begin{aligned}
&- 2\mathbf{P}_{2n|1}\mathbf{P}_{2n|0}^{-1}\mathbf{h}_{2n+2|0}^{(1)} + 3\mathbf{h}_{2n+2|0}^{(2)} + 2\mathbf{h}_{2n+2|0}^{(3)} \\
&= -2\left(\mathbf{P}_{2n|1}\mathbf{P}_{2n|0}^{-1}\mathbf{h}_{2n+2|0}^{(1)} - 2\mathbf{h}_{2n+2|0}^{(2)}\right) - \left(\mathbf{h}_{2n+2|0}^{(2)} - 2\mathbf{h}_{2n+2|0}^{(3)}\right) && (51) \\
&= -3\mathbf{P}_{2n|0}\left(\boldsymbol{\omega}_{2n|0} - 2 \cdot \mathbf{1}_{2n}\right). && (52)
\end{aligned}$$

Under consideration of the second and third entry of (46), the first parentheses of (51) equals $-2\mathbf{P}_{2n|0}\boldsymbol{\omega}_{2n+2|0}^{(2)}$ and the second parentheses $\mathbf{P}_{2n|0}\boldsymbol{\omega}_{2n+2|0}^{(3)}$. Hence, equation (52) holds since both $\boldsymbol{\omega}_{2n+2|0}^{(2)}$ and $\boldsymbol{\omega}_{2n+2|0}^{(3)}$ are equal to $\boldsymbol{\omega}_{2n|0} - 2 \cdot \mathbf{1}_{2n}$. Using (52) in the fourth entry of (46), replacing $\mathbf{h}_{2n+2|0}^{(4)}$ with the corresponding expression from Lemma II.5 and $\mathbf{M}_0$ with its definition $\mathbf{P}_{2n|0}^{-1}\mathbf{P}_{2n|1}\mathbf{P}_{2n|0}^{-1}$, we have

$$\begin{aligned}
\boldsymbol{\omega}_{2n+2|0}^{(4)} &= \mathbf{P}_{2n|0}^{-1}\left[-3\mathbf{P}_{2n|1}\left(\boldsymbol{\omega}_{2n|0} - 2 \cdot \mathbf{1}_{2n}\right) - \mathbf{h}_{2n|0} - 3\tilde{\mathbf{I}}_{2n}\mathbf{h}_{2n|0} - 6 \cdot \mathbf{1}_{2n}\right] \\
&= 3\mathbf{P}_{2n|0}^{-1}\left(-\mathbf{P}_{2n|1}\boldsymbol{\omega}_{2n|0} - \tilde{\mathbf{I}}_{2n}\mathbf{h}_{2n|0}\right) + 6 \cdot \mathbf{P}_{2n|0}^{-1}\left(\mathbf{P}_{2n|1}\mathbf{1}_{2n} - \mathbf{1}_{2n}\right) \\
&\quad - \mathbf{P}_{2n|0}^{-1}\mathbf{h}_{2n|0} \\
&= -\mathbf{P}_{2n|0}^{-1}\mathbf{h}_{2n|0} \\
&= \boldsymbol{\omega}_{2n|0}.
\end{aligned} \tag{53}$$

Observe that the first parentheses of equation (53) evaluates to 0 since it is equal to the left hand side of (50). Similarly, the second parentheses in (53) evaluates to 0 because $\mathbf{P}_{2n|1}$ is a stochastic matrix.

(b): Recall the recursions

$$\mathbf{P}_{2n+2|0} = \begin{bmatrix} \mathbf{P}_{2n+1|0} & \mathbf{0} \\ \frac{1}{2}\mathbf{P}_{2n+1|1} & \frac{1}{2}\mathbf{P}_{2n+1|0} \end{bmatrix} \tag{54}$$

$$\mathbf{P}_{2n+2|0}^{-1} = \begin{bmatrix} \mathbf{P}_{2n+1|0}^{-1} & \mathbf{0} \\ -\mathbf{P}_{2n+1|0}^{-1}\mathbf{P}_{2n+1|1}\mathbf{P}_{2n+1|0}^{-1} & 2\mathbf{P}_{2n+1|0}^{-1} \end{bmatrix}, \tag{55}$$

which follow from (1) and (21). Computing the first 2^{2n+1} entries of $\boldsymbol{\omega}_{2n+2|0}$ (i.e., the first half), using Definition II.1(b), (54) and (55), we obtain

$$\begin{bmatrix} \boldsymbol{\omega}_{2n+2|0}^{(1)} \\ \boldsymbol{\omega}_{2n+2|0}^{(2)} \end{bmatrix} = \mathbf{P}_{2n+1|0}^{-1} \left(\mathbf{P}_{2n+1|0} \circ \log_2 \mathbf{P}_{2n+1|0} \right) \mathbf{1}_{2n+1}. \tag{56}$$

By definition, the right hand side of (56) is $\boldsymbol{\omega}_{2n+1|0}$. Hence, under consideration of (44), we have

$$\boldsymbol{\omega}_{2n+1|0} = \begin{bmatrix} \boldsymbol{\omega}_{2n|0} \\ \boldsymbol{\omega}_{2n|0} - 2 \cdot \mathbf{1}_{2n} \end{bmatrix}. \tag{57}$$

It remains to express $\boldsymbol{\omega}_{2n|0}$ in (57) in terms of $\boldsymbol{\omega}_{2n-1|0}$. By the same argument as just used, the first half of the vector $\boldsymbol{\omega}_{2n|0}$ equals $\boldsymbol{\omega}_{2n-1|0}$. Since $\boldsymbol{\omega}_{2n|0}$ is a palindrome, the second half of $\boldsymbol{\omega}_{2n|0}$ equals $\tilde{\mathbf{I}}_{2n-1} \cdot \boldsymbol{\omega}_{2n-1|0}$. Hence,

$$\boldsymbol{\omega}_{2n|0} = \begin{bmatrix} \boldsymbol{\omega}_{2n-1|0} \\ \tilde{\mathbf{I}}_{2n-1} \cdot \boldsymbol{\omega}_{2n-1|0} \end{bmatrix}. \tag{58}$$

By replacing $\boldsymbol{\omega}_{2n|0}$ in (57) with (58), we get (45). The initial value $\boldsymbol{\omega}_1 = \begin{bmatrix} 0 & -2 \end{bmatrix}^T$ follows from (57) and noting that $\boldsymbol{\omega}_{0|0} = 0$. ■

Remark II.8. *The recursions derived in Lemma II.5 and II.6 are with respect to initial state $s_0 = 0$. They can be transformed to recursions with respect to initial state $s_0 = 1$ using (27) and (29).*

D. Proof of the Main Result

In this section, we evaluate (11) based on Lemma II.6. In particular, we find exact solutions to the optimization problem (3)-(5) for every $n \in \mathbb{N}$.

Theorem II.9. *Consider the convex optimization problem (3) to (5). The absolute maximum for input blocks of even length $2n$ is*

$$C_{2n}^{\uparrow} = \frac{1}{2} \log_2 \left(\frac{5}{2} \right) \ \textit{b/u}, \tag{59}$$

where $n \in \mathbb{N}$. For input blocks of odd length $2n-1$, the absolute maximum is

$$C_{2n-1}^{\uparrow} = \frac{1}{2n-1} \left[\log_2 \left(\frac{5}{4} \right) + (n-1) \cdot \log_2 \left(\frac{5}{2} \right) \right] \ \textit{b/u}, \tag{60}$$

where $n \in \mathbb{N}$.

Proof: Without loss of generality, we assume that the initial state is $s_0 = 0$. Recall (11), which for input blocks of length $2n-k$ reads

$$C^{\uparrow}_{2n-k} = \frac{1}{2n-k} \log_2 \left[\mathbf{1}^T_{2n-k} \exp_2 \left(\boldsymbol{\omega}_{2n-k|0} \right) \right] \text{ b/u}, \tag{61}$$

where $n \in \mathbb{N}$ and $k = 0, 1$. For $n = 1$, a straightforward computation shows using (44) and (45) in (61), that $C^{\uparrow}_1 = \log_2 \left(\frac{5}{4} \right)$ b/u and $C^{\uparrow}_2 = \frac{1}{2} \log_2 \left(\frac{5}{2} \right)$ b/u. Now assume that (59) and (60) hold for some n. In particular, suppose that

$$\mathbf{1}^T_{2n} \exp_2 \left(\boldsymbol{\omega}_{2n|0} \right) = \left(\frac{5}{2} \right)^n \tag{62}$$

and

$$\mathbf{1}^T_{2n-1} \exp_2 \left(\boldsymbol{\omega}_{2n-1|0} \right) = \frac{5}{4} \left(\frac{5}{2} \right)^{n-1}. \tag{63}$$

We now show that (59) and (60) hold if n is replaced by $n+1$. Using the recursions derived in Lemma II.6, we have

$$\begin{aligned} \mathbf{1}^T_{2n+2} \exp_2 \left(\boldsymbol{\omega}_{2n+2|0} \right) &= \mathbf{1}^T_{2n} \left[2 \exp_2 \left(\boldsymbol{\omega}_{2n|0} \right) + 2 \exp_2 \left(\boldsymbol{\omega}_{2n|0} - 2 \cdot \mathbf{1}_{2n} \right) \right] \\ &= \left(2 + 2 \cdot 2^{-2} \right) \mathbf{1}^T_{2n} \exp_2 \left(\boldsymbol{\omega}_{2n|0} \right) \end{aligned} \tag{64}$$

and

$$\begin{aligned} \mathbf{1}^T_{2n+1} \exp_2 \left(\boldsymbol{\omega}_{2n+1|0} \right) &= \mathbf{1}^T_{2n-1} \Big[\exp_2 \left(\boldsymbol{\omega}_{2n-1|0} \right) + \exp_2 \left(\tilde{\mathbf{I}}_{2n-1} \boldsymbol{\omega}_{2n-1|0} \right) \\ &\quad + \exp_2 \left(\boldsymbol{\omega}_{2n-1|0} - 2 \cdot \mathbf{1}_{2n-1} \right) + \exp_2 \left(\tilde{\mathbf{I}}_{2n-1} \boldsymbol{\omega}_{2n-1|0} - 2 \cdot \mathbf{1}_{2n-1} \right) \Big] \\ &= \mathbf{1}^T_{2n-1} \left[2 \exp_2 \left(\boldsymbol{\omega}_{2n-1|0} \right) + 2 \exp_2 \left(\boldsymbol{\omega}_{2n-1|0} - 2 \cdot \mathbf{1}_{2n-1} \right) \right] \qquad (65) \\ &= \left(2 + 2 \cdot 2^{-2} \right) \mathbf{1}^T_{2n-1} \exp_2 \left(\boldsymbol{\omega}_{2n-1|0} \right). \qquad (66) \end{aligned}$$

Equation (65) holds since $\mathbf{1}^T_{2n-1} \exp_2 \left(\tilde{\mathbf{I}}_{2n-1} \boldsymbol{\omega}_{2n-1|0} \right) = \mathbf{1}^T_{2n-1} \exp_2 \left(\boldsymbol{\omega}_{2n-1|0} \right)$ due to the fact that a multiplication with $\tilde{\mathbf{I}}_{2n-1}$ just permutes the entries of $\boldsymbol{\omega}_{2n-1|0}$ (see Lemma II.3). Finally, using (64), (66) and the induction hypotheses (62), (63) in (61), we obtain

$$\begin{aligned} C^{\uparrow}_{2n+2} &= \frac{1}{2n+2} \log_2 \left[\mathbf{1}^T_{2n+2} \exp_2 \left(\boldsymbol{\omega}_{2n+2|0} \right) \right] \\ &= \frac{1}{2n+2} \log_2 \left[\left(2 + 2 \cdot 2^{-2} \right) \mathbf{1}^T_{2n} \exp_2 \left(\boldsymbol{\omega}_{2n|0} \right) \right] \end{aligned}$$

$$= \frac{1}{2}\log_2\left(\frac{5}{2}\right) \text{ b/u}$$

and

$$\begin{aligned} C^{\uparrow}_{2n+1} &= \frac{1}{2n+1}\log_2\left[\mathbf{1}^T_{2n+1}\exp_2\left(\boldsymbol{\omega}_{2n+1|0}\right)\right] \\ &= \frac{1}{2n+1}\log_2\left[\left(2+2\cdot 2^{-2}\right)\mathbf{1}^T_{2n-1}\exp_2\left(\boldsymbol{\omega}_{2n-1|0}\right)\right] \\ &= \frac{1}{2n+1}\left[\log_2\left(\frac{5}{4}\right)+n\cdot\log_2\left(\frac{5}{2}\right)\right] \text{ b/u}. \end{aligned}$$

■

Remark II.10. *Observe that* $\lim_{n\to\infty} C^{\uparrow}_{2n+1} = \frac{1}{2}\log_2\left(\frac{5}{2}\right)$ *b/u where convergence is from below. Hence, we have*

$$\max_{n\in\mathbb{N}} C^{\uparrow}_n = \frac{1}{2}\log_2\left(\frac{5}{2}\right) \text{ b/u}.$$

Unfortunately, the distributions corresponding to (59) and (60) involve negative "probabilities" — otherwise the capacity of the trapdoor channel would have been established. We elaborate this issue in the following remark.

Remark II.11. *Note that the non-negativity of condition (9) does not hold for all* $k = 1, \ldots, 2^n$. *This can be verified as follows. For a trapdoor channel* $\mathbf{P}_{n|0}$, *condition (9) reads in matrix vector notation as*

$$\left[d_k\right]_{1\le k\le 2^n} = \left(\mathbf{P}^{-1}_{n|0}\right)^T \exp_2\left(\boldsymbol{\omega}_n\right). \tag{67}$$

We now compute the second last row of $\left(\mathbf{P}^{-1}_{n|0}\right)^T$ *by the following recursive scheme. Applying the matrix inversion lemma in the form of (21) to* $\mathbf{P}_{n|0}$, *which is written as in (1), and subsequently taking the transpose, then replacing the right bottom block of this matrix, which is* $2\left(\mathbf{P}^{-1}_{n-1|0}\right)^T$, *with the just obtained matrix times two (where* $n-1$ *is replaced by* $n-2$*), then applying the same procedure to the right bottom block of* $2\left(\mathbf{P}^{-1}_{n-1|0}\right)^T$ *and so on until the right bottom block equals* $2^{n-1}\left(\mathbf{P}^{-1}_{1|0}\right)^T$ *shows that the second last row of* $\left(\mathbf{P}^{-1}_{n|0}\right)^T$ *equals* $\begin{bmatrix}0 & \cdots & 0 & 2^{n-1} & -2^{n-1}\end{bmatrix}$. *Further, by Lemma II.6, the second to last entry and the last entry of* $\boldsymbol{\omega}_n$ *equals* -2 *and* 0 *if* $n\in\mathbb{N}$ *is even, and* -4 *and* -2 *if* $n\in\mathbb{N}$ *is odd. Inserting into (67) yields*

$$d_{2^n-1} = \begin{cases} -3\cdot 2^{n-3} & \text{if } n \text{ is even} \\ -3\cdot 2^{n-5} & \text{if } n \text{ is odd.} \end{cases}$$

Hence, condition (9) does not hold for all $k = 1, \ldots, 2^n$.

III. The Trapdoor Channel as a Fractal

Unlike in Section II, we do not focus on the problem of deriving and attaining the capacity of a given problem in this section. We rather reinterpret a given information theoretic model, namely the trapdoor channel, in various ways by intentionally not using information theoretic tools. The approach is motivated by the fact that the capacity of the trapdoor channel, an open problem since 1961, seems to be difficult to solve using standard tools from information theory. The considerable effort, e.g. taken in Section II, to solve the optimization problem (3) to (5) resulted only in an upper bound on the capacity of the trapdoor channel. On the other hand, the trapdoor channel exhibits lots of structure (see Lemma II.5 and Lemma II.6), which might give the capacity if exploited properly. In the following, we present two novel views on the trapdoor channel. Based on the underlying stochastic matrices (1) and (2), the trapdoor channel is described geometrically as a fractal or algorithmically as a recursive procedure. By deriving the underlying iterated function system (IFS), we show that the trapdoor channel with input blocks of length n can be regarded as the n^{th} element of a sequence of shapes approximating a fractal.

A. Prerequisites

We review the idea of *iterated function systems* and *fractals*. For a comprehensive introduction to the subject, see e.g. [8]. A fractal is a geometric pattern which exhibits self-similarity at every scale. A systematic way for generating a fractal starts with a complete metric space (M, d). The space to which the fractal belongs is, however, not M but the space of non-empty compact subsets of M, denoted as $\mathcal{H}(M)$. A suitable choice for a metric for $\mathcal{H}(M)$ is the Hausdorff distance $h_d(A, B) \stackrel{def}{=} \max\{d(A, B), d(B, A)\}$ where $d(A, B) \stackrel{def}{=} \max_{x \in A} \min_{y \in B} d(x, y)$ and $d(B, A) \stackrel{def}{=} \max_{x \in B} \min_{y \in A} d(x, y)$, $A, B \in \mathcal{H}(M)$. It is then guaranteed that $(\mathcal{H}(M), h_d)$ is a complete metric space and that every *contraction*[1] $\varphi : M \to M$ on (M, d) becomes a contraction mapping $\varphi : \mathcal{H}(M) \to \mathcal{H}(M)$ on $(\mathcal{H}(M), h_d)$ defined by $\varphi(A) = \{\varphi(x) : x \in A\}$ for all $A \in \mathcal{H}(M)$.

[1]Let (M, d) be a metric space. Recall that a mapping $\varphi : M \to M$ is a contraction if there exists a contractivity factor s, $0 < s < 1$, such that $d(\varphi(x), \varphi(y)) \leq s \cdot d(x, y)$ for all $x, y \in M$.

The following definition and theorem provides a method for generating fractals.

Definition III.1. *[8, Chapter 3.7] A hyperbolic iterated function system (IFS) consists of a complete metric space (M, d) together with a finite set of contraction mappings $\varphi_n : M \to M$, with contractivity factors s_n for $n = 1, 2, \ldots, N$. The notation for the IFS is $\{M; \varphi_n\, n = 1, 2, \ldots, N\}$ and its contractivity factor is $s = \max\{s_n : n = 1, 2, \ldots, N\}$.*

The fixed point of a hyperbolic IFS, also called the *attractor* or *self-similar set* of the IFS, is a (deterministic) fractal and results from iterating the IFS with respect to any $A \in \mathcal{H}(M)$. This is the content of the following theorem.

Theorem III.2. *[8, Chapter 3.7] Let $\{M; \varphi_n\, n = 1, 2, \ldots, N\}$ be an IFS with contractivity factor s. Then the transformation $\Phi : \mathcal{H}(M) \to \mathcal{H}(M)$ defined by*

$$\Phi(A) = \bigcup_{n=1}^{N} \varphi_n(A) \tag{68}$$

for all $A \in \mathcal{H}(M)$, is a contraction mapping on the complete metric space $(\mathcal{H}(M), h_d)$ with contractivity factor s. Its unique fixed point, $A^\star \in \mathcal{H}(M)$, obeys

$$A^\star = \Phi(A^\star) = \bigcup_{n=1}^{N} \varphi_n(A^\star),$$

and is given by $A^\star = \lim_{k\to\infty} \Phi^{\circ k}(A)$ for any $A \in \mathcal{H}(M)$.

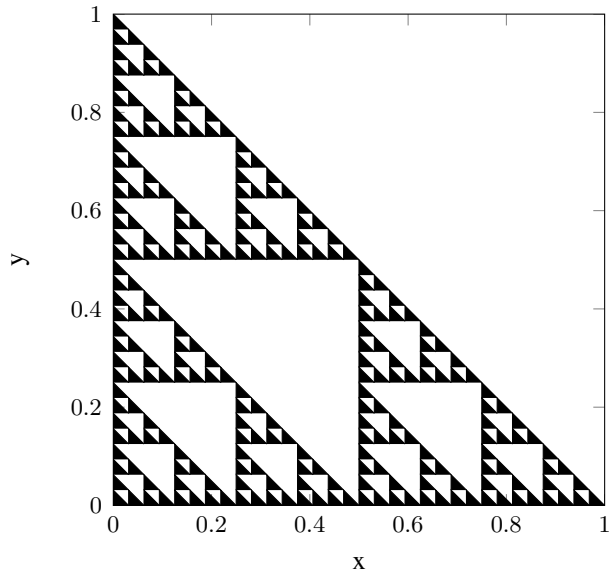
Many well-known fractals, e.g., the *Koch snowflake*, the *Cantor set*, the *Mandelbrot set*, etc., can be generated using Definition III.1 and Theorem III.2. A segment of the Mandelbrot set is shown on the cover of the information theory book by Cover and Thomas [9]. Another representative, the *Sierpinski triangle*, is introduced in the following example. We will later see that this fractal is related to the trapdoor channel.

Example III.3. (Sierpinski triangle) *Consider the IFS*

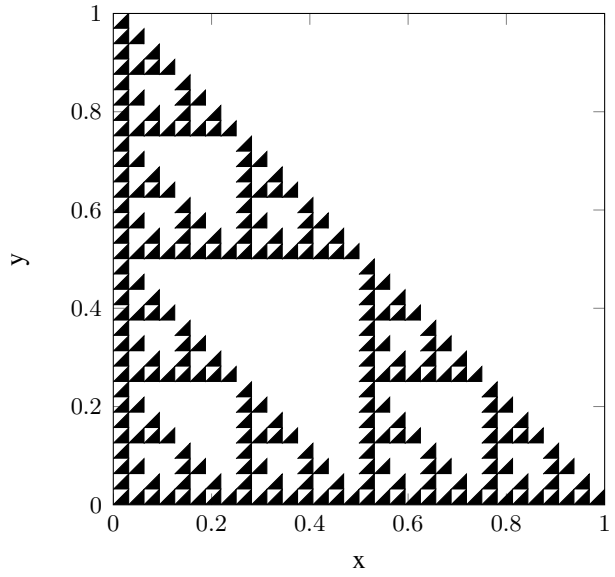
$$\left\{[0,1]^2; \varphi_1(x,y) = \left(\frac{x+1}{2}, \frac{y}{2}\right), \varphi_2(x,y) = \left(\frac{x}{2}, \frac{y+1}{2}\right), \varphi_3(x,y) = \left(\frac{x}{2}, \frac{y}{2}\right)\right\}. \tag{69}$$

The affine transformations φ_n, $n = 1, 2, 3$, scale any $A \in \mathcal{H}([0,1]^2)$ by a factor of 0.5. Additionally, φ_1 and φ_2 introduce translations by 0.5 into the x- and y-direction, respectively. The Sierpinski triangle is approximated arbitrarily close by iterating $\Phi(A)$ for any $A \in \mathcal{H}([0,1]^2)$.

Fig. 2 shows the result after performing four iterations of (69). The initial shape A in Fig. 2a is a triangle with corner points $(0,0),(1,0),(0,1)$ and in Fig. 2b a triangle with corner points $(0,0),(1,1),(1,0)$. As one performs more and more iterations, both sets converge to the same attractor $A^\star$.



(a) The initial shape is a triangle with corner points $(0,0),(1,0),(0,1)$.



(b) The initial shape is a triangle with corner points $(0,0),(1,1),(1,0)$.

Fig. 2: Sierpinski triangle after four iterations.

B. The Underlying Iterated Function System

In this section, we derive a hyperbolic IFS for the trapdoor channel. Instead of working with $\mathbf{P}_{n|s_0}$, we take a geometric approach, i.e., $\mathbf{P}_{n|s_0}$ will be mapped to the unit cube $[0,1]^3 \subset \mathbb{R}^3$.

Definition III.4. *Let $\mathcal{M}$ denote the set $\left\{\mathbf{P}_{n|s_0} : n \in \mathbb{N}_0, s_0 = 0, 1\right\}$ of trapdoor channel matrices. The function $\rho : \mathcal{M} \to [0,1]^3$ represents each $\mathbf{P}_{n|s_0}$ as a shape in $[0,1]^3$ according to*

$$\mathbf{P}_{n|s_0} \mapsto \left(x, y, \left[\mathbf{P}_{n|s_0}\right]_{i,j}\right), \quad \textit{for all } 1 \leq i, j \leq 2^n \tag{70}$$

where $(i-1)\cdot 2^{-n} < x < i \cdot 2^{-n}$ and $1 - j \cdot 2^{-n} < y < 1 - (j-1)\cdot 2^{-n}$.

Each entry $\left[\mathbf{P}_{n|s_0}\right]_{i,j}$ of $\mathbf{P}_{n|s_0}$ is identified with a square of side length 2^{-n}, which lies at a distance of $\left[\mathbf{P}_{n|s_0}\right]_{i,j}$ from the xy-plane. The alignment of the square corresponding to $\left[\mathbf{P}_{n|s_0}\right]_{i,j}$ with respect to the other squares in $\rho(\mathbf{P}_{n|s_0})$ is in accordance with the alignment of $\left[\mathbf{P}_{n|s_0}\right]_{i,j}$ with respect to the other entries of $\mathbf{P}_{n|s_0}$. Fig. 3 depicts the representations $\rho(\mathbf{P}_{1|0})$ and $\rho(\mathbf{P}_{1|1})$. Each of the four regions within $\rho\left(\mathbf{P}_{n|0}\right)$ and $\rho\left(\mathbf{P}_{n|1}\right)$ corresponds to one of the conditional probabilities $0, 0.5$ and 1. The following proposition expresses $\rho\left(\mathbf{P}_{n+1|0}\right)$ and $\rho\left(\mathbf{P}_{n+1|1}\right)$ recursively in terms of $\rho\left(\mathbf{P}_{n|0}\right)$ and $\rho\left(\mathbf{P}_{n|1}\right)$.

Lemma III.5. *The representations $\rho\left(\mathbf{P}_{n+1|0}\right)$ and $\rho\left(\mathbf{P}_{n+1|1}\right)$ of $\mathbf{P}_{n+1|0}$ and $\mathbf{P}_{n+1|1}$ satisfy the recursion laws*

$$\rho\left(\mathbf{P}_{n+1|0}\right) = \frac{1}{2} \cdot \left\{\rho\left(\mathbf{P}_{n|0}\right) + \mathbf{e}_x,\ \rho\left(2 \cdot \mathbf{P}_{n|0}\right) + \mathbf{e}_y,\ \rho\left(\mathbf{P}_{n|1}\right)\right\} \tag{71}$$

$$\rho\left(\mathbf{P}_{n+1|1}\right) = \frac{1}{2} \cdot \left\{\rho\left(2 \cdot \mathbf{P}_{n|1}\right) + \mathbf{e}_x,\ \rho\left(\mathbf{P}_{n|1}\right) + \mathbf{e}_y,\ \rho\left(\mathbf{P}_{n|0}\right) + \mathbf{e}_x + \mathbf{e}_y\right\} \tag{72}$$

for all $n \in \mathbb{N}_0$.

Proof: Recursions (71) and (72) are a consequence of the structure of (1) and (2). We just outline the derivation of (71). The first term $\frac{1}{2} \cdot \left\{\rho\left(\mathbf{P}_{n|0}\right) + \mathbf{e}_x\right\}$ on the right hand side of (71) represents the lower right corner of (1). Observe that $\left[\mathbf{P}_{n+1|0}\right]_{i,j}$ is equal to $\frac{1}{2}\left[\mathbf{P}_{n|0}\right]_{i-2^n, j-2^n}$ for all $2^n < i, j, \leq 2^{n+1}$. Hence, scaling the three dimensions of $\rho\left(\mathbf{P}_{n|0}\right)$ by a factor of $\frac{1}{2}$ and shifting the result by $\frac{1}{2}$ along the x-direction yields a representation of the lower right corner of (1) according to Definition III.4. Similarly, the second term $\frac{1}{2} \cdot \left\{\rho\left(2 \cdot \mathbf{P}_{n|0}\right) + \mathbf{e}_y\right\}$ of (71) represents the upper left corner of (1). Note that each entry $\left[\mathbf{P}_{n+1|0}\right]_{i,j}$ is equal to $\left[\mathbf{P}_{n|0}\right]_{i,j}$ for all $1 \leq i, j, \leq 2^n$. Hence, scaling the x- and y-coordinates of $\rho\left(\mathbf{P}_{n|0}\right)$ by a factor of $\frac{1}{2}$ and

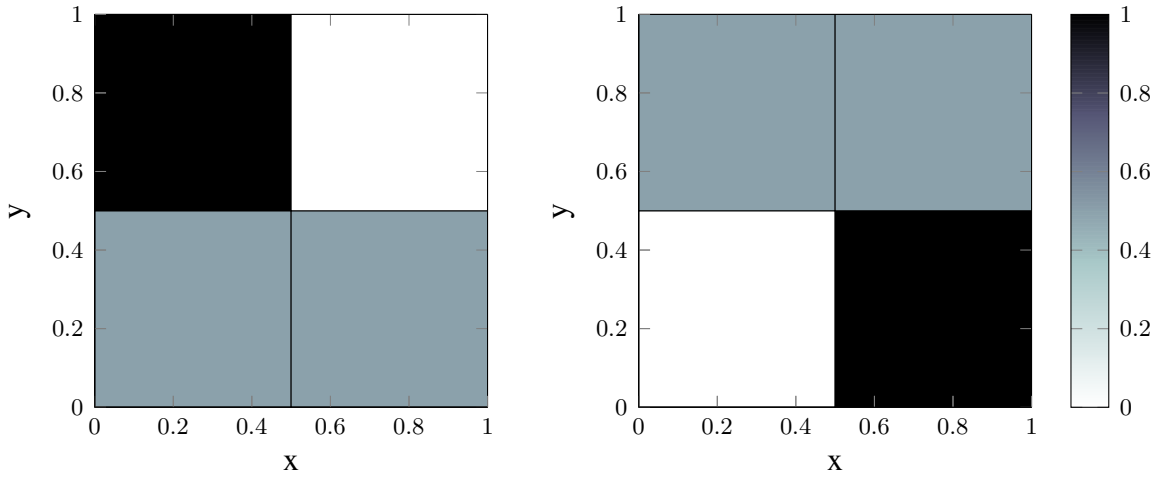


Fig. 3: Color map of $\rho(\mathbf{P}_{1|0})$ and $\rho(\mathbf{P}_{1|1})$.

shifting the resulting figure by $\frac{1}{2}$ along the y-direction yields a representation of the upper left corner $\mathbf{P}_{n|0}$ of (1) according to Definition III.4. The last term $\frac{1}{2} \cdot \rho\left(\mathbf{P}_{n|1}\right)$ of (71) represents the lower left corner of (1). Clearly, each entry $\left[\mathbf{P}_{n+1|0}\right]_{i,j}$ is equal to $\frac{1}{2}\left[\mathbf{P}_{n|1}\right]_{i-2^n,j}$ for all $2^n < i \leq 2^{n+1}$ and $1 \leq j \leq 2^n$. Hence, scaling all coordinates of $\rho\left(\mathbf{P}_{n|1}\right)$ by a factor of $\frac{1}{2}$ yields a representation of the lower left corner of (1) according to Definition III.4. ■

Recursions (71) and (72) are used below to obtain an IFS for the trapdoor channel. Recall from Theorem III.2 that an IFS is initialized with a single shape. Hence, it would be desirable that the right hand side of (71) depends only on $\mathbf{P}_{n|0}$ and the right hand side of (72) only on $\mathbf{P}_{n|1}$. The following proposition introduces an affine transformation which turns $\rho\left(\mathbf{P}_{n|0}\right)$ into $\rho\left(\mathbf{P}_{n|1}\right)$ and vice versa.

Lemma III.6. *Let* $\tau : [0,1]^3 \to [0,1]^3$ *be defined as* $\tau(x,y,z) = (-x+1, -y+1, z)$. *Then*

$$\rho\left(\mathbf{P}_{n|1}\right) = \tau \circ \rho\left(\mathbf{P}_{n|0}\right) \tag{73}$$

$$\rho\left(\mathbf{P}_{n|0}\right) = \tau \circ \rho\left(\mathbf{P}_{n|1}\right), \tag{74}$$

for all $n \in \mathbb{N}_0$.

Proof: Equation (74) follows from (73) by noting that $\tau \circ \tau = (x,y,z)$. It remains to prove (73), which is done by induction. Observe that the affine transformation τ corresponds to a counter-clockwise rotation through 180 degrees about the z-axis and a translation by 1 along

the x- and y-direction. Using this property, (73) for $n = 1$ is readily verified from Fig. 3. Now assume that the assertion holds for some $n > 1$. A direct computation of $\tau \circ \rho\left(\mathbf{P}_{n+1|0}\right)$, using the right hand side of (71) and the induction hypotheses (73) and (74), shows that $\tau \circ \rho\left(\mathbf{P}_{n+1|0}\right)$ is equal to $\rho\left(\mathbf{P}_{n+1|1}\right)$. This is demonstrated for the first function in (71). Observe that

$$\begin{aligned}\tau \circ \frac{1}{2}\left\{\rho\left(\mathbf{P}_{n|0}\right) + \mathbf{e}_x\right\} &= \frac{1}{2}\left\{\left(-x+1, -y+1, \left[\mathbf{P}_{n|s_0}\right]_{i,j}\right) + \mathbf{e}_y\right\} \\ &= \frac{1}{2}\left\{\tau \circ \rho\left(\mathbf{P}_{n|0}\right) + \mathbf{e}_y\right\} \\ &= \frac{1}{2}\left\{\rho\left(\mathbf{P}_{n|1}\right) + \mathbf{e}_y\right\},\end{aligned}$$

where the last step follows from the induction hypothesis. ■

We can now state the final recursion laws. A combination of Lemma III.5 and Lemma III.6, i.e., replacing $\rho\left(\mathbf{P}_{n|1}\right)$ in (71) with (73), $\rho\left(\mathbf{P}_{n|0}\right)$ in (72) with (74) and using Definition III.4, yields the following theorem.

Theorem III.7. *The representations $\rho\left(\mathbf{P}_{n+1|0}\right)$ and $\rho\left(\mathbf{P}_{n+1|1}\right)$ of $\mathbf{P}_{n+1|0}$ and $\mathbf{P}_{n+1|1}$ satisfy the following recursions:*

$$\rho\left(\mathbf{P}_{n+1|0}\right) = \left\{\begin{aligned}\phi_1(x,y,z) &= \left(\frac{x+1}{2}, \frac{y}{2}, \frac{\left[\mathbf{P}_{n|0}\right]_{i,j}}{2}\right), \\ \phi_2(x,y,z) &= \left(\frac{x}{2}, \frac{y+1}{2}, \left[\mathbf{P}_{n|0}\right]_{i,j}\right), \\ \phi_3(x,y,z) &= \left(-\frac{x-1}{2}, -\frac{y-1}{2}, \frac{\left[\mathbf{P}_{n|0}\right]_{i,j}}{2}\right)\end{aligned}\right\}, \tag{75}$$

$$\rho\left(\mathbf{P}_{n+1|1}\right) = \left\{\begin{aligned}\psi_1(x,y,z) &= \left(\frac{x+1}{2}, \frac{y}{2}, \left[\mathbf{P}_{n|1}\right]_{i,j}\right), \\ \psi_2(x,y,z) &= \left(\frac{x}{2}, \frac{y+1}{2}, \frac{\left[\mathbf{P}_{n|1}\right]_{i,j}}{2}\right), \\ \psi_3(x,y,z) &= \left(-\frac{x}{2}+1, -\frac{y}{2}+1, \frac{\left[\mathbf{P}_{n|1}\right]_{i,j}}{2}\right)\end{aligned}\right\}, \tag{76}$$

where $(i-1)\cdot 2^{-n} < x < i \cdot 2^{-n}$ and $1 - j \cdot 2^{-n} < y < 1 - (j-1)\cdot 2^{-n}$ and $1 \le i, j \le 2^n$.

Remark III.8. *The restrictions of ϕ_i and ψ_i, $1 \le i \le 3$, to the x- and y-dimensions are contraction mappings resulting in two hyperbolic IFS with a unique attractor each. An approximation of the attractor for $s_0 = 0$ is shown in the plots on the right side of Fig. 4. There is also a*

relation to the Sierpinski triangle. Observe that ϕ_i *and* ψ_i*,* $1 \leq i \leq 2$*, when restricted to the* xy*-plane, are equal to* φ_1*,* φ_2 *in (69).*

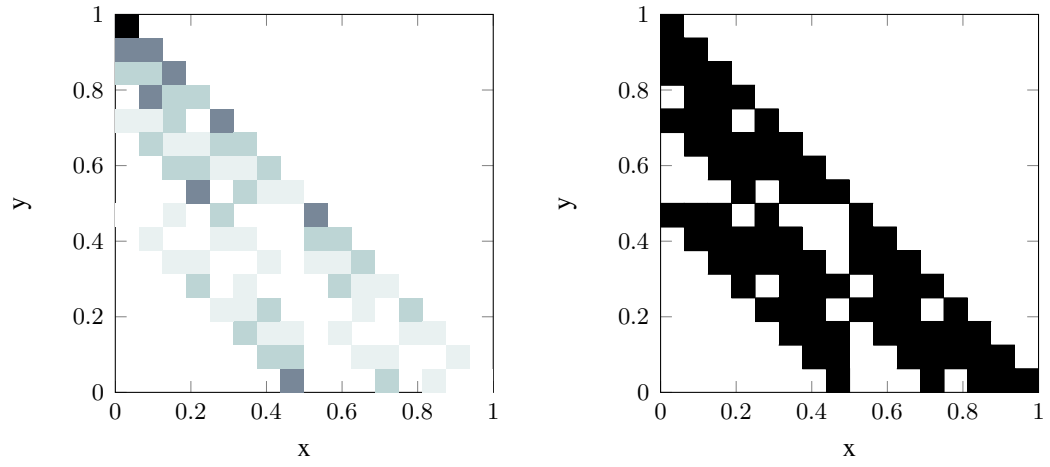
Remark III.9. *Recall that* $\mathbf{P}_{0|0} = 1$ *and* $\mathbf{P}_{0|1} = 1$*. Then* $\lim_{n\to\infty} \rho\left(\mathbf{P}_{n|s_0}\right)$ *for* $s_0 = 0$ *can be approximated arbitrarily close by iterating (according to Theorem III.2)*

$$\left\{[0,1]^3; \phi_1 = \left(\frac{x+1}{2}, \frac{y}{2}, \frac{z}{2}\right), \phi_2 = \left(\frac{x}{2}, \frac{y+1}{2}, z\right), \phi_3 = \left(-\frac{x-1}{2}, -\frac{y-1}{2}, \frac{z}{2}\right)\right\} \tag{77}$$

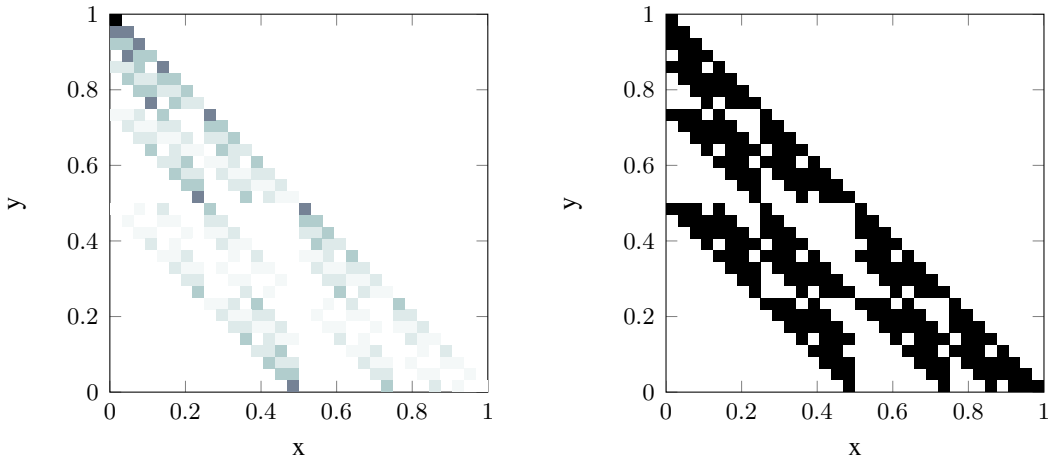
and for $s_0 = 1$

$$\left\{[0,1]^3; \psi_1 = \left(\frac{x+1}{2}, \frac{y}{2}, z\right), \psi_2 = \left(\frac{x}{2}, \frac{y+1}{2}, \frac{z}{2}\right), \psi_3 = \left(-\frac{x}{2}+1, -\frac{y}{2}+1, \frac{z}{2}\right)\right\}, \tag{78}$$

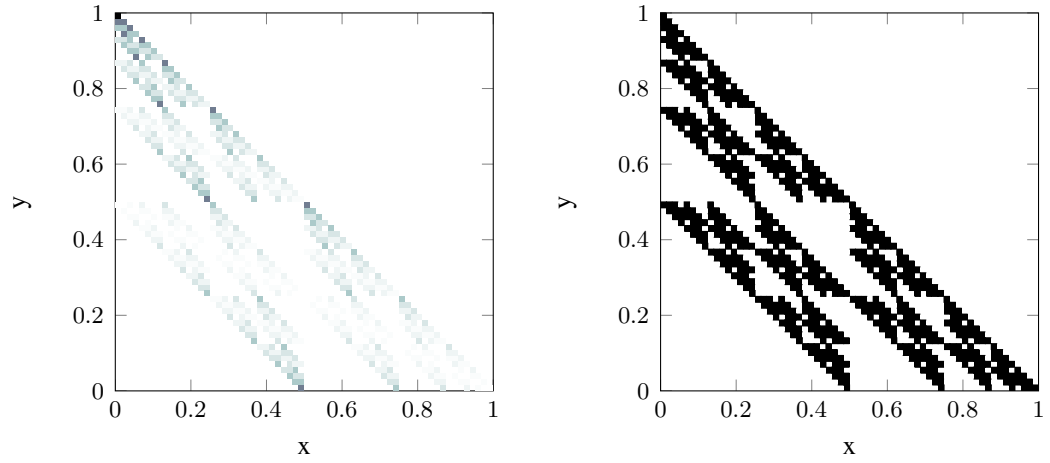
where the initial shape can be any $A \in \mathcal{H}([0,1]^3)$ *such that the restriction of* A *to the* z*-dimension equals one. Fig. 4 depicts three, four, and five iterations of (77) with an initial shape* $\{(x,y,z) \in [0,1]^3 : z = 1\}$.



(a) Three iterations.



(b) Four iterations.



(c) Five iterations.

Fig. 4: Three, four, and five iterations of (77) and its projections onto the xy-plane. The initial shape is $\{(x, y, z) \in [0, 1]^3 : z = 1\}$. The color scale is the same as in Fig. 3.

IV. Algorithmic View of the Trapdoor Channel

A. Remarks on the Permutation Nature

The trapdoor channel has been called a permuting channel [4], where the output is a permutation of the input [5]. We point out that in general not all possible permutations of the input are feasible and that not every output is a permutation of the input. The reason that not all permutations are feasible is that the channel actions are causal, i.e., an input symbol at time n cannot become a channel output at a time instance smaller than n. Consider, for instance, the string 101 which, when applied to a trapdoor channel with initial state 0, cannot give rise to an output 110. Next, not every output is a permutation of the input because at a certain time instance the initial state might become an output symbol and, therefore, the resulting output sequence might not be compatible with a permutation of the input. For illustration purposes, consider again the string 101 and initial state 0. Two feasible outputs are 010 and 001, which are not permutations of 110.

B. The Algorithm

We now present an algorithm that fully characterizes the trapdoor channel and resembles the recursion of generating all permutations of a given string.The following recursive procedure GENERATEOUTPUTS computes the set of feasible output sequences and their likelihoods given an input sequence and an initial state. It gives a complete characterization of the trapdoor channel. Generating outputs and their corresponding likelihoods for a particular input sequence might be instrumental for designing codes. In the following, we adopt the standard convention that the first element of a string corresponds to index 0.

procedure GENERATEOUTPUTS($in, out, state, prob$)
  if $in = \emptyset$ **then**
    $set \leftarrow \{out, prob\}$
  else if $in[0] = state$ **then**
    $out \leftarrow out + in[0]$
    $set \leftarrow$ GENERATEOUTPUTS($in.substr(1), out, state, prob$)
  else
    $out \leftarrow out + in[0]$

$set \leftarrow$ GENERATEOUTPUTS($in.substr(1), out, state, 0.5 \cdot prob$)
$out[out.length() - 1] \leftarrow state$ ▷ $in[0]$ is removed from the end of out
$set \leftarrow$ GENERATEOUTPUTS($in.substr(1), out, in[0], 0.5 \cdot prob$)
end if
return set
end procedure

The four variables $in, out, state, prob$ have the following meaning: in denotes the part of the input string that has not been processed yet, out indicates the part of one particular output string that has been generated so far, $state$ refers to the current channel state, $prob$ denotes the likelihood of out. The procedure is initialized with the complete input string and the initial state of the channel; out is initially empty while $prob$ equals 1. The first if statement checks the simple case of the recursion, namely whether the input string has been processed completely. If yes, the corresponding output out and its likelihood $prob$ is stored and returned in set. Otherwise, we have to distinguish whether the next input symbol $in[0]$ is equal to the current state or not. If yes, the next output takes the value of $in[0]$ with probability 1 (or of $state$, but both are equal), i.e., $out \leftarrow out + in[0]$, and the procedure GENERATEOUTPUTS is applied recursively to the unprocessed part $in.substr(1)$ of the input string, i.e., the substring of in with indices greater than 0. Clearly, $state$ and $prob$ do not change and are passed directly to the recursive call. In the other case, namely when $in[0]$ is not equal to the current state, the next output symbol will have a probability of 0.5 to be either $in[0]$ or $state$. If $in[0]$ becomes the next channel output, the following state remains the same. Then the remaining input string $in.substr(1)$ is processed by the recursive call GENERATEOUTPUTS($in.substr(1), out, state, 0.5 \cdot prob$). However, if $state$ becomes the channel output, then the following state will be $in[0]$ and the remaining input string is processed by GENERATEOUTPUTS($in.substr(1), out, in[0], 0.5 \cdot prob$).

Observe that a recursive implementation of the algorithm is needed to process inputs of any length, which is not the case if only iterative control structures are used. We emphasize that each of the three recursive calls of the algorithm resembles a recursion for generating all permutations of a string (see, e.g., [10, ch. 8.3]). This gives an algorithmic justification why some output sequences are permutations of the input sequence.

V. Discussion

We focused on the convex optimization problem (3) to (5) where the feasible set is larger than the probability simplex. An absolute maximum of the n-letter mutual information was established for any $n \in \mathbb{N}$ by using the method of Lagrange multipliers via [2, Theorem 3.3.3]. The same absolute maximum $\frac{1}{2}\log_2\left(\frac{5}{2}\right) \approx 0.6610$ b/u results for all even n and the sequence of absolute maxima corresponding to odd block lengths converges from below to $\frac{1}{2}\log_2\left(\frac{5}{2}\right)$ b/u as the block length increases. Unfortunately, all absolute maxima are attained outside the probability simplex. Hence, instead of establishing the capacity of the trapdoor channel, we have shown only that $\frac{1}{2}\log_2\left(\frac{5}{2}\right)$ b/u is an upper bound on the capacity. To the best of our knowledge, this upper bound is the tightest known bound. Notably, this upper bound is strictly smaller than the feedback capacity [5]. Moreover, the result gives an indirect justification that the capacity of the trapdoor channel is attained on the boundary of the probability simplex. Subsequently, two different views on the trapdoor channel were presented. We first derived the IFS of the trapdoor channel, which was motivated by the observation that standard approaches from information theory have failed so far to derive its capacity. Subsequently, we described the trapdoor channel by means of a recursive procedure. The procedure, which generates all feasible output sequences and their likelihoods given a particular input sequence, might be helpful to construct codes for the trapdoor channel.

Acknowledgment

The author is supported by the German Ministry of Education and Research in the framework of the Alexander von Humboldt-Professorship and would like to thank Prof. Haim Permuter who suggested to use [2, Theorem 3.3.3]. Moreover, the author wishes to thank Prof. Gerhard Kramer and Prof. Tsachy Weissman for helpful discussions.

References

[1] D. Blackwell, *Information Theory*, E. F. Beckenbach, Ed. McGraw-Hill Book Co., New York, 1961, vol. Modern Mathematics for the Engineer.

[2] R. Ash, *Information Theory*. Interscience Publishers, 1965.

[3] K. Kobayashi and H. Morita, "An input/output recursion for the trapdoor channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Lausanne, Switzerland, Jun. 30–Jul. 5 2002, p. 423.

[4] R. Ahlswede and A. H. Kaspi, "Optimal coding strategies for certain permuting channels." *IEEE Trans. Inf. Theory*, vol. 33, no. 3, pp. 310–314, 1987.

[5] H. H. Permuter, P. Cuff, B. VanRoy, and T. Weissman, "Capacity of the trapdoor channel with feedback," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3150–3165, Jul. 2008.

[6] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, Inc., 1968.

[7] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. The John Hopkins University Press, 1996.

[8] M. Barnsley, *Fractals Everywhere*. Academic Press, Inc., 1988.

[9] T. M. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[10] E. Roberts, *Programming Abstractions in C++*. Prentice Hall, 2014.