One-Shot Capacity of Discrete Channels

Rui A. Costa
Instituto de Telecomunicações
Faculdade de Ciências da
Universidade do Porto, Portugal
Email: rfcosta@fe.up.pt

Michael Langberg
Computer Science Division
Open University of Israel
Email: mikel@openu.ac.il

João Barros
Instituto de Telecomunicações
Faculdade de Engenharia da
Universidade do Porto, Portugal
Email: jbarros@fe.up.pt

***Abstract*—Shannon defined channel capacity as the highest rate at which there exists a sequence of codes of block length n such that the error probability goes to zero as n goes to infinity. In this definition, it is implicit that the block length, which can be viewed as the number of available channel uses, is unlimited. This is not the case when the transmission power must be concentrated on a single transmission, most notably in military scenarios with adversarial conditions or delay-tolerant networks with random short encounters. A natural question arises: how much information can we transmit in a single use of the channel? We give a precise characterization of the one-shot capacity of discrete channels, defined as the maximum number of bits that can be transmitted in a single use of a channel with an error probability that does not exceed a prescribed value. This capacity definition is shown to be useful and significantly different from the zero-error problem statement.**

I. Introduction

Shannon's notion of channel capacity [1] is asymptotic in the sense that the number of channels uses (or, equivalently, the block length of the code) can be arbitrarily large. The rate of a code is defined as the ratio between the number of input symbols and the number of channel uses required to transmit them. A rate is said to be achievable if there exists a sequence of codes (of that rate) with block length n, whose error probability goes to zero asymptotically as n goes to infinity. The behavior of the channel capacity with a limited number of channel uses is less well understood. A typical approach is to consider the rate at which the error probability decays to zero, which motivates the study of *error exponents* [2],[3],[4]. Beyond the aforementioned definitions, it is also reasonable to ask what rates can be achieved when the error probability must be precisely zero. In [5], Shannon assumes once again that the channel is available as many times as necessary and defines the zero-error capacity as the supremum of the independence numbers of the extensions of the confusion graph [6], [7].

A different question is how much information can we convey in a single use of the channel or, in other words, what is the *one-shot capacity* of the channel. The question arises for example when the transmission power must be concentrated on a single transmission, most notably in military scenarios with adversarial conditions or delay-tolerant networks with random short encounters. As we have seen, classical definitions of capacity do not encompass these scenarios.

The one-shot capacity problem can also be viewed as a special instance of the single-letter coding problem since, in both problems, the encoder must assign to every source output symbol one channel input symbol (and not a sequence of them). However, to the best of our knowledge, studies on single-letter coding use optimality criteria based on an unlimited number of channel uses. For instance, [8] characterizes optimal single-letter source-channel codes, with respect to the rate-distortion and capacity-cost functions, which are of asymptotic nature. For comparison, in our study of the one-shot capacity, only a single channel use is considered.

Using a combinatorial approach, the zero-error one-shot capacity of a given channel was considered in [6], [7]. More specifically, [6], [7] construct a certain undirected graph G corresponding to the channel at hand; and characterize the zero-error one-shot capacity by the size of the maximum independent set in G.

In this work we generalize the results and combinatorial framework of [6], [7] to capture communication that allows an error probability below ϵ; namely, we study the ϵ*-error one-shot* capacity. We note that preliminary results on the ϵ-error one-shot capacity appear in [9], which uses smooth min-entropy and a probabilistic approach to develop bounds for the one-shot capacity (also called, single-serving channel capacity). Our work differs from [9] in that we characterize the exact value of the ϵ-error one-shot capacity by means of classical combinatorics.

Our main contributions are as follows:

- *Problem Formulation:* We provide a rigorous mathematical framework for analyzing the ϵ-capacity of discrete channels subject to a one-shot constraint. We consider two different metrics of performance: maximum error probability and average error probability.
- *Operational Interpretation:* We illustrate the practical relevance of the one-shot capacity by means of examples where the zero-error one-shot capacity and the ϵ-error one-shot capacity present significantly distinct behaviors.
- *Combinatorial description of the One-Shot Capacity of Discrete Channels:* We cast the capacity in terms of the properties of a special graph G derived from the channel. For maximum error, we describe the one-shot capacity through the independence number of G, whereas for average error we consider the maximum size of *sparse* sets in G.
- *Complexity Analysis:* We show that the problem of computing the one-shot capacity is NP-Hard.

The remainder of the paper is organized as follows. In Section II, we give a formal definition of the problem at hand, namely the concepts of ϵ-maximum and ϵ-average one-shot capacity. In Section III we present a non-trivial example of a class of channels for which the one-shot capacity is relevant. Our main result for maximum error one-shot capacity is stated and proved in Section IV. In Section V, we prove that computing the ϵ-maximum one-shot capacity is NP-Hard. Finally, in Section VI we discuss the case of ϵ-average one-shot capacity and Section VII concludes the paper.

II. Problem Statement

We start our problem statement with the usual definition of a discrete channel.

Definition 1: A discrete *channel* is composed of an input alphabet $\mathcal{X}$, an output alphabet $\mathcal{Y}$ and the transition probabilities $\mathcal{P}(Y = y|X = x)$.

We will refer to such a channel as "the channel described by $P_{Y|X}$". Next, we present the definition of a one-shot communication scheme over a discrete channel.

Definition 2: A *one-shot communication scheme* over a $P_{Y|X}$ channel is composed of a codebook $\underline{\mathcal{X}} \subseteq \mathcal{X}$, and a decoding function $\gamma : \mathcal{Y} \to \underline{\mathcal{X}}$.

We will refer to such a communication scheme as the "$(\underline{\mathcal{X}}, \gamma)$ pair". It is natural to view the set $\underline{\mathcal{X}}$ as the set of messages to be transmitted over the channel. Our figure of merit is the probability of error in the decoding process. We consider two different metrics: maximum and average error probability.

Definition 3: The *maximum error probability* associated with a pair $(\underline{\mathcal{X}}, \gamma)$ is defined as

$$\epsilon_{\underline{\mathcal{X}},\gamma} = \max_{x\in\underline{\mathcal{X}}} \mathcal{P}(\gamma(Y) \neq x|X = x).$$

Definition 4: The *average error probability* associated with a pair $(\underline{\mathcal{X}}, \gamma)$ is defined as

$$\bar{\epsilon}_{\underline{\mathcal{X}},\gamma} = \frac{1}{|\underline{\mathcal{X}}|} \sum_{x\in\underline{\mathcal{X}}} \mathcal{P}(\gamma(Y) \neq x|X = x).$$

We are now ready to define the one-shot capacity of a discrete channel. From an intuitive point of view, we are intrigued by the maximum number of distinct messages (the size of the codebook $\underline{\mathcal{X}}$) that can be transmitted in a single use of the channel, while ensuring that the error probability (maximum or average) does not exceed a prescribed value ϵ. We must first define an admissible $(\underline{\mathcal{X}}, \gamma)$ pair.

Definition 5: Consider a channel described by $P_{Y|X}$ and let $\epsilon \in [0, 1]$. The pair $(\underline{\mathcal{X}}, \gamma)$ is *maximum-ϵ-admissible* if $\epsilon_{\underline{\mathcal{X}},\gamma} \leq \epsilon$. The set of all ϵ-admissible pairs is denoted by $\mathcal{A}_\epsilon$.

Definition 6: Consider a channel described by $P_{Y|X}$ and let $\epsilon \in [0, 1]$. The pair (P_X, γ) is *average-ϵ-admissible* if $\bar{\epsilon}_{\underline{\mathcal{X}},\gamma} \leq \epsilon$. The set of all average-ϵ-admissible pairs is denoted by $\bar{\mathcal{A}}_\epsilon$.

The notion of single-serving capacity is outlined in [9] as "*the maximum number of bits that can be transmitted in a single use of $P_{Y|X}$, such that every symbol can be decoded by an error of at most ϵ*". We formalize this notion by defining the ϵ-maximum one-shot capacity as follows:

Definition 7: Consider a channel described by $P_{Y|X}$. For $\epsilon \in [0, 1]$, the ϵ*-maximum one-shot channel capacity,* denoted by C_ϵ, is defined as

$$C_\epsilon = \max_{(\underline{\mathcal{X}},\gamma)\in\mathcal{A}_\epsilon} \log(|\underline{\mathcal{X}}|).^1$$

Similarly, we can define the ϵ-average one-shot capacity as follows:

Definition 8: Consider a channel described by $P_{Y|X}$. For $\epsilon \in [0, 1]$, the ϵ*-average one-shot channel capacity,* denoted by C_ϵ, is defined as

$$\bar{C}_\epsilon = \max_{(\underline{\mathcal{X}},\gamma)\in\bar{\mathcal{A}}_\epsilon} \log(|\underline{\mathcal{X}}|)$$

Our goal is to provide a precise characterization of the one-shot capacity.

III. Practical Relevance of the One-Shot Capacity

So far, we have formally defined the concept of the ϵ-error one-shot capacity. One question that naturally arises is the following: does the ϵ-error one-shot capacity significantly differ from the zero-error one-shot capacity, for small values of ϵ? In other words, are there classes of channels for which allowing a small error probability enables the transmission of a significantly larger number of bits than in the zero-error case? In this section, we present a class of channels for which the answer to the previous questions is yes, which asserts for the practical relevance of the ϵ-error one-shot capacity notion. Our examples use the maximum error criterion (and thus imply the gap for average error also).

Example 1: Consider a channel with input alphabet $\mathcal{X} = \{0, 1, \ldots, n-1\}$, output alphabet $\mathcal{Y} = \{0, 1, \ldots, n-1\}$, and let $0 < e_1 < e_2 < \ldots < e_{n-1} \leq 1$. Let $\mathcal{P}(Y = 0|X = 0) = 1$. For each $i \in \mathcal{X} \setminus \{0\}$, let the transition probability distribution be

$$P(Y = y|X = i) = \begin{cases} 1 - e_i & \text{if } y = i \\ e_i & \text{if } y = 0 \\ 0 & \text{otherwise} \end{cases}$$

In Fig. 1(a), we present an example of a channel in this class. Notice that, given that all symbols are "confusable" (i.e. $\mathcal{P}(Y = 0|X = i) > 0,\ \forall i \in \mathcal{X}$), the zero-error one-shot capacity of this channel is zero, $\forall n$. However, by allowing a small error probability, we are able to transmit a significant number of bits.

Lemma 1: The ϵ-maximum one-shot capacity of the channel described in *Example 1* satisfies

$$C_\epsilon = \log(i + 1) \text{ for } i \text{ s.t. } e_i \leq \epsilon < e_{i+1}, \tag{1}$$

where $e_0 = 0$ and $e_n = 1$.

Proof: We start by proving that the ϵ-maximum one-shot capacity, C_ϵ, is lower bounded by (1). Let $e_i \leq \epsilon < e_{i+1}$, for some $i \in \{1, \ldots, n-1\}$ (the case $i = 0$ is trivial, since by definition $C_\epsilon \geq 0$). Consider the codebook $\underline{\mathcal{X}} = \{1, \ldots, i+1\}$ and the following decoding function:

$$\gamma(y) = \begin{cases} y & \text{if } y \in \{1, \ldots, i+1\} \\ i+1 & \text{otherwise} \end{cases}$$

[1] Throughout this work, all the logarithms are considered in base 2.

For $x \in \{1, \dots, i\}$, we have that $\gamma^{-1}(x) = \{x\}$ (where $\gamma^{-1}(x) = \{y \in \mathcal{Y} : \gamma(y) = x\}$) and, thus, $\mathcal{P}(\gamma(Y) \neq x | X = x) = e_x \leq \epsilon$, because we have that $e_i \leq \epsilon < e_{i+1}$ and $e_x \leq e_i$ (because $x \leq i$). With respect to $x = i+1$, we have that $\gamma^{-1}(i+1) = \{0, i+1, \dots, n-1\}$. Moreover, $\mathcal{P}(Y = i+1 | X = i+1) = 1 - e_{i+1}$ and $\mathcal{P}(Y = 0 | X = i+1) = e_{i+1}$. Therefore,

$$\mathcal{P}(\gamma(Y) \neq i+1 | X = i+1) = 1 - \sum_{y \in \gamma^{-1}(i+1)} \mathcal{P}(Y = y | X = i+1) = 0.$$

Hence, we have constructed a pair $(\underline{\mathcal{X}}, \gamma)$ for which $|\underline{\mathcal{X}}| = i+1$ and $\epsilon_\gamma = \max_{x \in \underline{\mathcal{X}}} \mathcal{P}(\gamma(Y) \neq x | X = x) \leq \epsilon$.

Now, we show that C_ϵ is upper bounded by (1). Let $e_i \leq \epsilon < e_{i+1}$, for some $i \in \{0, \dots, n-1\}$, and let $(\underline{\mathcal{X}}, \gamma)$ be a pair for which $C_\epsilon = \log |\underline{\mathcal{X}}|$. Notice that for $x \in \{0\} \cup \{i+1, \dots, n-1\}$, we have $\mathcal{P}(Y = 0 | X = x) \geq \epsilon$. Therefore, if $x \in \underline{\mathcal{X}} \cap (\{0\} \cup \{i+1, \dots, n-1\})$, we must have $0 \in \gamma^{-1}(x)$. Thus, since γ is a function, we have that $|\underline{\mathcal{X}} \cap (\{0\} \cup \{i+1, \dots, n-1\})| \leq 1$, which implies that $|\underline{\mathcal{X}}| \leq i+1$, thus concluding our proof. ■

The previous example shows that, by allowing for a small probability of error in the decoding process, we are able to transmit a significantly higher number of bits in one use of the channel, in comparison with the case where no errors are allowed. In the case illustrated in Fig. 1(a), we have that the ϵ-maximum one-shot capacity verifies

$$C_\epsilon = \begin{cases} 0 & \text{if } \epsilon < 0.01 \\ 1 & \text{if } 0.01 \leq \epsilon < 0.02 \\ \log(3) & \text{if } \epsilon \geq 0.02 \end{cases}$$

IV. The Case of Maximum Error Probability

In this section, we present a combinatorial description of the one-shot capacity under maximum error ϵ. We start by defining the graph that will help us obtain the desired description. For that, we first need to use the following definition which associates with each input symbol x a set of output symbols denoted by $D_\epsilon(x)$.

Definition 9: For each $x \in \mathcal{X}$, let

$$D_\epsilon(x) = \left\{ D \subset \mathcal{Y} : \sum_{y \in D} \mathcal{P}(Y = y | X = x) \geq 1 - \epsilon \right\}.$$

We can view $D_\epsilon(x)$ as the set of all possible inverse images of x through a decoding function γ (i.e. all possible $\gamma^{-1}(x)$), with γ verifying $\epsilon_{\underline{\mathcal{X}},\gamma} \leq \epsilon$. We are now ready to present the definition of the maximum-one-shot graph of the channel described by $P_{Y|X}$.

Definition 10: The *maximum-one-shot graph* of the channel described by $P_{Y|X}$ is the graph G_ϵ (with node set V and edge set E_ϵ) constructed as follows:

- the nodes are the elements of the form (x, D) with $x \in \mathcal{X}$ and $D \in D_\epsilon(x)$;
- two nodes (x, D) and (x', D') are connected if and only if $x = x'$ or $D \cap D' \neq \emptyset$.

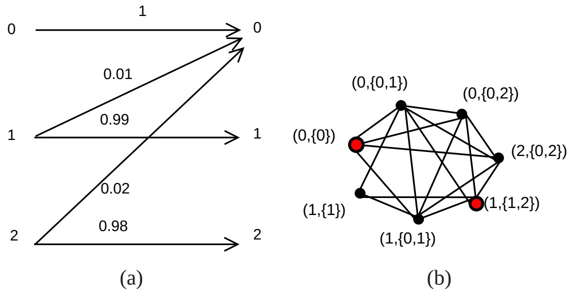

Fig. 1. In (a), we present an instance of the class of channels defined in *Example 1*, with $n = 3$, $e_1 = 0.01$ and $e_2 = 0.02$. In (b), we present the maximum-one-shot graph of the channel described in (a), for $\epsilon = 0.01$. The nodes in red form a maximum independent set, which by *Theorem 1* implies that the ϵ-maximum one-shot capacity of the channel in (a) is $\log(2) = 1$. For the sake of clarity, we excluded the nodes $(0, \{0,1,2\})$, $(1, \{0,1,2\})$ and $(2, \{0,1,2\})$, since these nodes are connected to every other node in the graph and, thus, are not part of any maximum independent set.

In Fig. 1(b), we present the maximum-one-shot graph of the channel in Fig. 1(a). Due to the definition of $D_\epsilon(x)$, in the maximum-one-shot graph, nodes represent all the possible $\gamma^{-1}(x)$ (such that γ is ϵ-admissible). To obtain a proper decoding function from the maximum-one-shot graph, we need to find an *independent set*, since a connection between two nodes represents the incompatibility of two inverse images.

Definition 11: Consider a graph $G = (V, E)$. An *independent set* I_G in G is a set of nodes $v \in V$ in which no two nodes are connected by an edge.

Definition 12: Consider a graph $G = (V, E)$. A *maximum independent set* I_G is a largest independent set and its cardinality is called the *independence number* of the graph G, and it is denoted by $\alpha(G)$.

Using these definitions, we are now able to state our main result, which relates the one-shot capacity with the independence number of the previously defined graph.

Theorem 1: Consider a channel described by $P_{Y|X}$ and the corresponding maximum-one-shot graph $G_\epsilon = (V, E_\epsilon)$, with $\epsilon \in [0, 1)$. The ϵ-maximum one-shot capacity satisfies

$$C_\epsilon = \log(\alpha(G_\epsilon)).$$

We prove this theorem by establishing first that one can transmit a codebook of size at least $\alpha(G_\epsilon)$ with a single use of the channel. We then show that this is the best one can do.

Lemma 2: $C_\epsilon \geq \log(\alpha(G_\epsilon))$.

Proof: Let $G_\epsilon = (V, E_\epsilon)$ be the maximum-one-shot graph of the channel and let I_{G_ϵ} be a maximum independent set in G_ϵ. Let $\mathcal{X}^*$ be the set of symbols in $\mathcal{X}$ that are represented in I_{G_ϵ}, i.e.

$$\mathcal{X}^* = \{x \in \mathcal{X} : \exists D = (y_1, \dots, y_k) \text{ such that } (x, D) \in I_{G_\epsilon}\}.$$

For each $x \in \mathcal{X}^*$, let $d(x)$ be the set of output symbols that are represented in the same node as x in I_{G_ϵ}, i.e.

$$d(x) = \{y_1, \dots, y_k : (x, D) \in I_{G_\epsilon} \text{ with } D = (y_1, \dots, y_k)\}.$$

Notice that, since I_{G_ϵ} is an independent set in G_ϵ and all pairs of nodes of the form $(x, y_1, \dots, y_k)$ and $(x, y'_1, \dots, y'_{k'})$ are connected in G_ϵ, we have that $d(x)$ is unique and properly defined. Let $\mathcal{Y}^* = \{y \in \mathcal{Y} : \exists x \in \mathcal{X}^* \text{ such that } y \in d(x)\}$.

Now, consider the decoder $\gamma(.)$ constructed as follows:

- for $y \in \mathcal{Y}^*$, we set $\gamma(y) = x'$, where $x' \in \mathcal{X}^*$ is such that $y \in d(x')$;

- for $y \notin \mathcal{Y}^*$, we set $\gamma(y) = x^*$, where x^* is some symbol in $\mathcal{X}^*$.

We have that I_{G_ϵ} is an independent set in G_ϵ. Thus, for every $y \in \mathcal{Y}^*$, there is only one $x' \in \mathcal{X}^*$ such that $y \in d(x')$. Therefore, the function $\gamma(.)$ is well-defined, i.e. $\forall y \in \mathcal{Y}, \exists! x \in \mathcal{X}^* : \gamma(y) = x$. We also have that $\forall x \in \mathcal{X}^*, \exists y \in \mathcal{Y} : \gamma(y) = x$, which is equivalent to $\gamma(\mathcal{Y}) = \mathcal{X}^*$. Let $\underline{\mathcal{X}} = \mathcal{X}^*$. We have that $\alpha(G_\epsilon) = |I_{G_\epsilon}| = |\mathcal{X}^*|$ and, therefore, $|\underline{\mathcal{X}}| = \alpha(G_\epsilon)$.

Now, we need to analyze the error probability of the pair $(\underline{\mathcal{X}}, \gamma)$ previously constructed. Let $x \in \underline{\mathcal{X}}$ and let $\gamma^{-1}(x) = \{y \in \mathcal{Y} : \gamma(y) = x\} = \{y_1, \ldots, y_k\}$. We have that

$$\mathcal{P}(\gamma(Y) \neq x | X = x) = 1 - \sum_{i=1}^{k} \mathcal{P}(Y = y_i | X = x).$$

Notice that, by the construction of $\gamma(\cdot)$, we have that, for $D = (y_1, \ldots, y_k)$, (x, D) is a node of I_{G_ϵ} and, therefore, a node in G_ϵ. Thus, by the definition of the maximum-one-shot graph G_ϵ, we have that $(y_1, \ldots, y_k) \in D_\epsilon(x)$, which is equivalent to $\sum_{i=1}^{k} \mathcal{P}(Y = y_i | X = x) \geq 1 - \epsilon$. Therefore, we have that $\mathcal{P}(\gamma(Y) \neq x | X = x) \leq \epsilon$, and this inequality is not dependent on the choice of $x \in \underline{\mathcal{X}}$. Therefore, we have that $\forall x \in \underline{\mathcal{X}}$, $\mathcal{P}(\gamma(Y) \neq x | X = x) \leq \epsilon$, which is equivalent to $\epsilon_{\underline{\mathcal{X}}, \gamma} \leq \epsilon$. Thus, we have constructed a pair $(\underline{\mathcal{X}}, \gamma)$ such that $\epsilon_{\underline{\mathcal{X}}, \gamma} \leq \epsilon$ and $|\underline{\mathcal{X}}| = \alpha(G_\epsilon)$, which implies that $C_\epsilon \geq \log(\alpha(G_\epsilon))$. ■

We proved that one can transmit $\alpha(G_\epsilon)$ symbols with a single use of the channel. Now, we prove that it is not possible to transmit more than that.

Lemma 3: $C_\epsilon \leq \log(\alpha(G_\epsilon))$.

Proof: Let $(\underline{\mathcal{X}}, \gamma)$ be a pair such that $|\underline{\mathcal{X}}| = 2^{C_\epsilon}$ and $\epsilon_{\underline{\mathcal{X}}, \gamma} \leq \epsilon$. Let $\gamma^{-1}(x) = \{y \in \mathcal{Y} : \gamma(y) = x\}$. Since $\epsilon_{\underline{\mathcal{X}}, \gamma} \leq \epsilon$, we have that $\forall x \in \underline{\mathcal{X}}$, $\sum_{i=1}^{k} \mathcal{P}(Y = y_i | X = x) \geq 1 - \epsilon$, where $\{y_1, \ldots, y_k\} = \gamma^{-1}(x)$. Thus, $(y_1, \ldots, y_k) \in D_\epsilon(x)$ and, therefore, for $D = (y_1, \ldots, y_k)$, (x, D) is a node in the maximum-one-shot graph $G_\epsilon = (V, E)$.

Now, notice that, since $\gamma(.)$ is a function with $\mathcal{Y}$ as domain, we have that $\forall x_1 \neq x_2 \in \underline{\mathcal{X}}$, $\gamma^{-1}(x_1) \cap \gamma^{-1}(x_2) = \emptyset$. Therefore, the set $I = \{(x, D) : x \in \underline{\mathcal{X}}$ and $D = \gamma^{-1}(x)\}$ is an independent set in G_ϵ, which implies that $|I| \leq \alpha(G_\epsilon)$. Since $|I| = |\underline{\mathcal{X}}|$, we have that $|\underline{\mathcal{X}}| \leq \alpha(G_\epsilon)$ and, therefore, $2^{C_\epsilon} \leq \alpha(G_\epsilon)$. ■

V. Complexity of the Computation of the One-Shot Capacity

Up to now we have shown that the ϵ-error one-shot capacity can be characterized by the independence number of the graph G_ϵ. Computing the independence number is known to be an NP-Hard problem. However, it may be the case that the graphs G_ϵ we obtain in our reduction are of a *simple* nature allowing us to find their independence number efficiently. In what follows we show that this is not the case.

Theorem 2: The computation of the ϵ-maximum one-shot capacity is NP-Hard for $\epsilon < 1/3$.

Proof: We will prove the NP-Hardness of the ϵ-maximum one-shot capacity problem by reducing the independent set problem in 3-regular graphs (a known NP-Hard problem [10]) to an instance of the ϵ-maximum one-shot capacity problem, for $\epsilon < 1/3$. The reduction technique is similar to the one used in [6].

Consider a 3-regular graph $G = (V, E)$. We will construct a communication channel driven from this graph as follows: the input alphabet is $\mathcal{X} = V$, the output alphabet is $\mathcal{Y} = E$ and the transition probability distribution is given by

$$\mathcal{P}(Y = y | X = x) = \begin{cases} 1/3 & \text{if } x \text{ is an endpoint of } y \\ 0 & \text{otherwise} \end{cases}$$

Notice that $\mathcal{P}_{Y|X}$ is well-defined, since G is a 3-regular graph (each node has degree 3) and, therefore, $\forall x \in \mathcal{X}$, $\sum_{y \in \mathcal{Y}} \mathcal{P}(Y = y | X = x) = 1$. For each $x \in \mathcal{X}$, let $d(x) = \{y \in \mathcal{Y} : \mathcal{P}(Y = y | X = x) > 0\}$. Notice that $|d(x)| = 3$ and $\forall y \in d(x), \mathcal{P}(Y = y | X = x) = 1/3$.

Now, we shall focus on the ϵ-maximum one-shot capacity of the previously constructed channel. Let $\epsilon < 1/3$. Let us now construct the maximum-one-shot graph $G_\epsilon = (V_\epsilon, E_\epsilon)$. The node set is composed of elements of the form $(x, y_1, \ldots, y_k)$ such that $\sum_{i=1}^{k} \mathcal{P}(Y = y_i | X = x) \geq 1 - \epsilon \, (> 2/3)$. Two nodes $(x, y_1, \ldots, y_k)$ and $(x', y'_1, \ldots, y'_{k'})$ are connected in G_ϵ if and only if $x = x'$ or $\exists i, j$ such that $y_i = y'_j$. As $\epsilon < 1/3$, notice that for any node (x, D) in G_ϵ it holds that $d(x) \subseteq D$.

We now show that $\alpha(G_\epsilon)$=$\alpha(G)$. This suffices to prove our assertion since computing the independence number of G is NP-Hard and the ϵ-maximum one-shot capacity is equal to the (logarithm of the) independence number of G_ϵ. Namely, we prove that a maximum independent set in G_ϵ corresponds to a maximum independent set in the original 3-regular graph G, and vice-versa.

Let I_ϵ be an independent set in G_ϵ. Consider the set $I = \{x : \exists D \text{ s.t. } (x, D) \in I_\epsilon\}$. It holds that $|I| = |I_\epsilon|$. Moreover, for any two nodes (x, D) and (x', D') in I_ϵ it holds that $d(x) \subseteq D$, $d(x') \subseteq D'$ and $D \cap D' = \phi$. This implies that $d(x) \cap d(x') = \phi$, which in turn implies that x and x' are not connected by an edge. We conclude that I is an independent set in G.

For the other direction, Let I be an independent set in G. Consider the set $I_\epsilon = \{(x, d(x)) : x \in I\}$. It holds that $|I| = |I_\epsilon|$. Moreover, for any two nodes x and x' in I it holds that $d(x) \cap d(x') = \phi$. This implies that $(x, d(x))$ and $(x', d(x'))$ are not connected by an edge in G_ϵ. We conclude that I_ϵ is an independent set in G_ϵ. ■

VI. The Case of Average Error Probability

In this section, we devote our attention to the ϵ-average one-shot capacity (*Definition 8* in Section II).

Definition 13: For each $x \in \mathcal{X}$, we define

$$D(x) = \left\{ D \subset \mathcal{Y} : \sum_{y \in D} \mathcal{P}(Y = y | X = x) > 0 \right\}.$$

We are now ready to present the definition of the *average*-one-shot graph of the channel described by $P_{Y|X}$.

Definition 14: The *average-one-shot graph* of the channel described by $P_{Y|X}$ is the weighted graph G_ϵ (with node set V and edge set E_ϵ), constructed as follows:

- the nodes are the elements of the form (x, D) with $x \in \mathcal{X}$ and $D \in D(x)$;
- two nodes (x, D) and (x', D') are connected by an infinite weight edge if and only if $x = x'$ or $D \cap D' \neq \emptyset$.
- all other pairs of nodes (x, D) and (x, D') are connected by an edge of weight $\mathcal{P}(Y \notin D|X = x) + \mathcal{P}(Y \notin D'|X = x')$.

The previous definition provides us a tool to describe a relationship between the ϵ-average one-shot capacity and sparse sets in the average-one-shot graph.

Definition 15: Consider a weighted graph $G = (V, E)$. An *ϵ-sparse set* I_ϵ in G is a set of nodes $v \in V$ for which the weight of edges in the subgraph induced on I_ϵ is at most $\epsilon|I_\epsilon|(|I_\epsilon| - 1)$. For example, a 0-sparse set is an independent set.

Definition 16: Consider a graph $G = (V, E)$. A *maximum ϵ-sparse set* I_ϵ is a largest set in G that is ϵ-sparse, its cardinality is called the *ϵ-sparse number* of the graph G, and is denoted by $\alpha_\epsilon(G)$.

We are now ready to present our main result related to the ϵ-average one-shot capacity.

Theorem 3: Let $\epsilon \in [0, 1]$. Consider a channel described by $P_{Y|X}$ and let $G_\epsilon = (V, E)$ be the average-one-shot graph. The ϵ-average one-shot capacity is given by $\bar{C}_\epsilon = \log(\alpha_\epsilon(G_\epsilon))$.

Proof: Let $G_\epsilon = (V, E)$ be the average-one-shot graph of the channel and let I_ϵ be a maximum ϵ-sparse set in G_ϵ. Let $\mathcal{X}^*$ be the set of symbols in $\mathcal{X}$ that are represented in I_ϵ, i.e. $\mathcal{X}^* = \{x \in \mathcal{X} : \exists D \text{ such that } (x, D) \in I_\epsilon\}$. Notice that I_ϵ cannot contain two vertices (x, D) and (x, D') (as they share an edge of infinite weight), or two vertices (x, D) and (x', D') with $D \cap D' \neq \phi$ (for the same reason).

For each $x \in \mathcal{X}^*$, let $(x, D_x) \in I_\epsilon$. Let $\mathcal{Y}^* = \cup_{x \in \mathcal{X}^*} D_x$. Now, consider the decoder $\gamma(.)$ constructed as follows:

- for $y \in D_x$, we set $\gamma(y) = x$.
- for $y \notin \mathcal{Y}^*$, we set $\gamma(y) = x^*$, where x^* is some symbol in $\mathcal{X}^*$.

We have that I_ϵ is an ϵ-sparse set in G.

$$\begin{aligned}
\epsilon &\geq \frac{1}{|I_\epsilon|(|I_\epsilon| - 1)} \sum_{x \neq x' \in \mathcal{X}^*} \mathcal{P}(Y \notin D_x|X = x) + \\
&\qquad\qquad + \mathcal{P}(Y \notin D_{x'}|X = x') \\
&= \frac{|I_\epsilon| - 1}{|I_\epsilon|(|I_\epsilon| - 1)} \sum_{x \in \mathcal{X}^*} \mathcal{P}(Y \notin D_x|X = x) \\
&= \frac{1}{|I_\epsilon|} \sum_{x \in \mathcal{X}^*} \mathcal{P}(\gamma(Y) \neq x|X = x)
\end{aligned}$$

This implies that $\bar{C}_\epsilon \geq \log|I_\epsilon| = \log(\alpha_\epsilon(G))$. For the other direction, let $(\underline{\mathcal{X}}, \gamma)$ be a pair in $\bar{\mathcal{A}}_\epsilon$ such that $\bar{C}_\epsilon = \log|\underline{\mathcal{X}}|$. Let $I_\epsilon = \{(x, \gamma^{-1}(x)) \mid x \in \underline{\mathcal{X}}\}$. Clearly, $|I_\epsilon| = |\underline{\mathcal{X}}|$. We now show (very similar to the analysis above) that I_ϵ is ϵ-sparse. Namely, first notice that I_ϵ does not contain any infinite weight edges (as γ is a decoding function). Moreover

$$\begin{aligned}
\epsilon &\geq \frac{1}{|\underline{\mathcal{X}}|} \sum_{x \in \underline{\mathcal{X}}} \mathcal{P}(\gamma(Y) \neq x|X = x) \\
&= \frac{1}{|I_\epsilon|} \sum_{x \in I_\epsilon} \mathcal{P}(\gamma(Y) \neq x|X = x) \\
&= \frac{|I_\epsilon| - 1}{|I_\epsilon|(|I_\epsilon - 1|)} \sum_{x \in I_\epsilon} \mathcal{P}(Y \notin \gamma^{-1}(x)|X = x) \\
&= \frac{1}{|I_\epsilon|(|I_\epsilon| - 1)} \sum_{x \neq x' \in I_\epsilon} \mathcal{P}(Y \notin D_x|X = x) + \\
&\qquad\qquad + \mathcal{P}(Y \notin D_{x'}|X = x')
\end{aligned}$$

We conclude that $\log \alpha_\epsilon(G) \geq \log|\underline{\mathcal{X}}| = \bar{C}_\epsilon$, which concludes our proof. ■

VII. Conclusions

Intrigued by the capacity of discrete channels that can be used only once, we elaborated on the ϵ-one-shot capacity, defined as the maximum number of bits that can be transmitted with one channel use while assuring that the decoding error probability is not greater than ϵ. Based on this definition, we introduced the concept of the ϵ-one-shot graph associated with a discrete channel and provided an exact characterization of the ϵ-one-shot capacity through combinatorial properties of the ϵ-one-shot graph. Using this formulation, we prove that computing the ϵ-one-shot capacity (for $\epsilon < 1/3$) is NP-Hard.

The practical relevance of the concept we present in this paper was discussed through a non-trivial example of a class of discrete channels for which the zero-error capacity is null, but allowing for small error probability enables the transmission of a significant number of bits in a single use of the channel.

Acknowledgment

This work was partially funded by the Fundação para a Ciência e Tecnologia (FCT, Portuguese Foundation for Science and Technology) under grants SFRH-BD-27273-2006 and PTDC/EIA/71362/2006 (WITS project), and by ISF grant 480/08.

References

[1] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[2] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, “Lower bounds to error probability for coding on discrete memoryless channels,” *Information and Control*, vol. 10, no. 1, pp. 65–103, 1967.

[3] R. Gallager, “A simple derivation of the coding theorem and some applications,” *IEEE Transactions on Information Theory*, vol. 11, no. 1, pp. 3–18, January 1965.

[4] I. Csiszar and J. G. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, Inc., Orlando, FL, USA, 1982.

[5] C. E. Shannon, “The zero error capacity of a noisy channel,” *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8 – 19, September 1956.

[6] L. Lovász, “On the shannon capacity of a graph,” *IEEE Transactions on Information Theory*, vol. IT-25, pp. 1–7, January 1979.

[7] J. Korner and A. Orlitsky, “Zero-error information theory,” *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2207 – 2229, October 1998.

[8] M. Gastpar, B. Rimoldi, and M. Vetterli, “To code, or not to code: Lossy source-channel communication revisited,” *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[9] R. Renner, S. Wolf, and J. Wullschleger, “The single-serving channel capacity,” in *IEEE International Symposium on Information Theory*, Seattle, WA, July 2006, pp. 1424–1427.

[10] M. R. Garey, D. S. Johnson, and L. Stockmeyer, “Some simplified np-complete graph problems,” *Theory of Computer Science*, vol. 1, pp. 237–267, 1976.