On the Outage Theorem and its Converse for Discrete Memoryless Channels

Tarik Muharemović and Behnaam Aazhang

Abstract

We prove both the outage theorem and its converse for block fading discrete memoryless channel. If P_{out} is the outage probability for the rate R, then any transmission rate below R, together with any average error probability greater than P_{out} is achievable. Conversely, for any code with rate slightly greater than R, the error probability is greater than any number smaller than P_{out}, for large codeword lengths. Two cases are considered depending on whether the transmitter has channel state information (CSI) or not.

I. Introduction

Block or "quasi-static" fading channels remain in a given fading state during a single codeword transmission, or possibly longer. For a fixed input distribution P_X, the mutual information between input and output random variables becomes a random quantity, since it is a function of the fading state. Outage event is defined as the event where the mutual information falls below attempted transmission rate R. This outage event is entirely specified by the channel realization. Sometimes we refer to this event as "channel beeing in outage." The Outage Probability at the attempted transmission rate R is simply the probability of the outage event, minimized over all input distributions. Essentially, it is a quantity designed to monitor the failure of transmission. The receiver *always* has CSI since channel remains constant throughout the codeword transmission.

Simple case is the one where the transmitter also has CSI. It can adapt its codebook to the channel realization in order to maximize instantenuous mutual information. Outage event then becomes the one where actual channel capacity falls below R. The achievability part of the capacity theorem guarantees successfull transmission if the channel is not in outage since in this case the channel capacity is greater than R. Hence, the source should be able to transmit at a rate R, with errors occurring almost only in the case that the channel is in outage. This is the direct or the "achievability" part of the outage theorem and is essentially a counterpart to the "achievability" part of the capacity theorem.

At this point it is not clear whether the source can actually have a successful transmission even if the channel is in outage. In order to claim this we essentially need the strong converse to the capacity theorem which holds for large block lengths. It states that a failed transmission will occur if the actual channel capacity falls below R. Obviously, the exsistence of a strong converse is imperative for this discussion.

If the transmitter doesn't have the CSI, then the relationship between outage events and errors is not quite as trivial. The transmitter cannot adapt its codebook to the instanteous channel realization, and the problem becomes choosing the codebook which works for "most" of the channels. The "compound" channel capacity theorem becomes a crucial part of establishing the fact that the average error probability can be made smaller or equal than the Outage Probability [3].

However, this still doesn't establish the Outage Probability as a tight *lower* bound to the average error probability. This is the main contribution of this document, but only for DMC since out proof relies on the method of types.

It was shown in [2] that in the multiple antenna block fading channel the SNR exponent in the error probability expression has to be greater or equal than the SNR exponent in the outage probability expression. However, from here we can't hastily conclude that the error probability itself has to be greater that the Outage Probability. Evaluation of outage probabilities, especially in multiple input multiple output channels, is done in [5], [4] and others. Coding theorem and its converse has been proven in [5] for coded diversity channels, but the optimizing input distribution was the same regardless of the channel realization. This is much similar to the case where the transmitter is equipped with CSI. The derivation below relies on the existing channel coding theorems and is therefore brief as compared to [5]. Direct part of the outage theorem, which is achievabilty of error probabilities greater than P_{out} when the transmitter does not have CSI is outlined in [3], but the converse is not provided for.

Any DMC is entirely specified by the input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$, and the probability transition matrix $w : \mathcal{X} \to \mathcal{Y}$. We adopt a somewhat non-standard notation for the mutual information between input and the output random variables, namely $I(P_X; w)$ rather than $I(X, Y)$, since we want to emphasize dependence on the input distribution P_X as well as the probability transition matrix w. We are interested in the case when the probability transition matrix w is itself random, and we therefore denote it as W. Such a channel is sometimes referred to as "quasi-static" fading channel, since W does not change during a codeword transmission. The average error probability of a particular code $\mathcal{C}$ for a given W is denoted as $P_e(\mathcal{C}, W)$ and is a random variable itself. This error probability is further averaged over all realizations of W, namely:

$$P_e(\mathcal{C}) = E_W[P_e(\mathcal{C}, W)]. \tag{1}$$

We emphasize that W is essentially a probability transition matrix, which is itself random. For the purposes of this discussion W need not be discrete valued. The mutual information becomes a random quantity itself denoted as $I(P_X; W)$. The *Outage Probability* P_{out} given the transmission rate R is defined as:

$$P_{out} = \inf P_W[I(P_X; W) < R], \tag{2}$$

where infimum is taken over all distributions P_X if there is no CSI at the transmitter, and over all conditional distributions $P_{X|W}$ if there is. We will prove both the direct part and the converse to the coding theorems for this channel using the method of types. Briefly, type of a codeword $\underline{x} \in \mathcal{X}^n$ is its empirical distribution. We heavily rely on channel coding theorems as stated in [1], pages 102-105. To clarify the notation, all sets denoted as $[\cdot]$, are with respect to channel realization w, $|\cdot|$ denotes alphabet size of a set, and Υ is the indicator function. $\Upsilon_{\mathcal{A}}(w) = 1$ if $w \in \mathcal{A}$, and 0 otherwize.

II. CSI at the transmitter

This case is mathematically simpler, and therefore we choose to discuss it first. Since the transmitter has access to CSI, optimization is done over input distributions conditioned on the channel observation, $P_{X|W}$. Obviously, optimum transmit strategy is to choose a distribution which achieves the capacity for that particular channel realization. The outage probability is given as

$$P_{out} = \inf_{P_{X|W}} P_W[I(P_X; W) < R] = P_W[\sup_{P_{X|W}} I(P_X; W) < R] = P_W[C(W) < R], \tag{3}$$

where $C(W)$ is the random capacity of the resulting channel W. We first discuss the direct part of the outage theorem, and then we follow with the converse.

A. Direct Part of the Outage Theorem

We first require direct part of the capacity theorem for ordinary DMC channels, as stated in [1]. We have explicitely added the channel dependence, w. This theorem deals with achieving rates below $C(w)$.

Theorem 1: [1] $\forall \epsilon, \tau > 0$, to every DMC w there exists a code $\mathcal{C}(w)$ of length n, and the average error probability less than ϵ so that the rate of the code is at least $C(w) - \tau$, provided that $n \geq n_0(|\mathcal{X}|, |\mathcal{Y}|, \epsilon, \tau)$.

There are numerous proofs for this theorem, and one of them is given in [1]. The direct part of the outage theorem for block fading DMC is a consequence of this.

Theorem 2: Let W be a block fading DMC, and let P_{out} be the outage probability for W at the rate R, as defined above. Then $\forall\ \tau, \lambda > 0$ there is a $\mathcal{C} = \mathcal{C}(w)$, a family of codes of rate $R - \tau$, designed for the transmitter with CSI w, so that whenever code length satisfies $n \geq n_0(|\mathcal{X}|, |\mathcal{Y}|, \lambda, \tau)$ we have:

$$P_e(\mathcal{C}) = E_W[P_e(\mathcal{C}(W), W)] \leq P_{out} + \lambda. \tag{4}$$

Proof: We break up channel realizations into outage, $\{w : C(w) < R\} = [C(w) < R]$, and its complement $[C(w) \geq R]$. Suppose that the channel is not in outage. For, $n \geq n_0$ the transmitter has a codebook of rate $R - \tau$, and average error probability less than ϵ. On the other hand, if the channel is in outage, error probability is upper bounded by 1. Let $\epsilon = \lambda$, in the previous theorem, and we have:

$$P_e(\mathcal{C}) = E_W[P_e\left(\mathcal{C}(W), W\right)]$$

$$= E_W[P_e(\mathcal{C}(W), W)\ \Upsilon_{[C(w)<R]}(W)] + E_W[P_e(\mathcal{C}(W), W)\ \Upsilon_{[C(w)\geq R]}(W)]$$

$$\leq E_W[\Upsilon_{[C(w)<R]}(W)] + E_W[\epsilon\ \Upsilon_{[C(w)\geq R]}(W)]$$

$$\leq P_{out} + \epsilon = P_{out} + \lambda \tag{5}$$

for $n \geq n_0$. Hence the transmitter can transmit at a rate $R - \tau$, and achieve error probabilities less than $P_{out} + \lambda$. Loosely speaking, average error probabillity can be made smaller than the Outage Probability. ■

B. Converse to the Outage Theorem

But just how small can this average error probability be? Converse part of the outage theorem will state that they cannot go much below P_{out}. To prove this, we first need the following strong converse to the channel coding theorem as stated and proved in [1]. It states that if one attempts a transmission rate even slightly above the capacity, then the error probability for a long code is very close to one.

Theorem 3: [1] $\forall(\epsilon,\tau)\in(0,1)$, if $\mathcal{C}(w)$ is a code of length n for a DMC channel w with average codeword error probability less than ϵ, then for $n \geq n_0(|\mathcal{X}|,|\mathcal{Y}|,\epsilon,\tau)$ we have

$$\frac{1}{n}\log|\mathcal{C}(w)| < \sup_P I(P,w)+\tau = C(w)+\tau. \tag{6}$$

It follows that for $n \geq n_0(|\mathcal{X}|,|\mathcal{Y}|,\epsilon,\tau)$, if the rate of the code is larger than $C(w)+\tau$, then the average error probability is greater than ϵ. Note that ϵ can be made arbitrary close to 1, and also that n_0 is uniform in channel (probability transition matrix) w, as long as the alphabet size doesn't change. The impact of this result in a block fading DMC is that the average error probability cannot be much smaller than the Outage Probability.

Theorem 4: Let W be a block fading DMC, and let P_{out} be the outage probability for W at the rate R, as defined above. Then $\forall\ \tau,\lambda>0$ if $\mathcal{C}=\mathcal{C}(w)$ is any family of codes of rate $R+\tau$ designed for the transmitter with CSI w, whenever code length satisfies $n \geq n_0(|\mathcal{X}|,|\mathcal{Y}|,\lambda,\tau)$, the average error probability is bounded as:

$$P_e(\mathcal{C}) = E_W[P_e(\mathcal{C}(W),W)] \geq P_{out}-\lambda. \tag{7}$$

Proof: Let τ,λ be given. Choose $\epsilon\in(0,1)$ so that $\epsilon P_{out} \geq P_{out}-\lambda$, and let $n \geq n_0(|\mathcal{X}|,|\mathcal{Y}|,\epsilon,\tau)$. Every $\mathcal{C}(w)$ is of the same rate, namely $R+\tau$. The above Theorem (3) states that whenever $C(w)+\tau$ is less than this rate, it follows that $P_e(\mathcal{C}(w),w)>\epsilon$. Therefore $C(w)+\tau<R+\tau$ implies that $P_e(\mathcal{C}(w),w)>\epsilon$.

$$\begin{aligned} P_e(\mathcal{C}) = E_W[P_e(\mathcal{C}(W),W)] &\geq E_W[P_e(\mathcal{C}(W),W)\ \Upsilon_{[C(w)<R]}(W)] \\ &\geq \epsilon E_W[\Upsilon_{[C(w)<R]}(W)] = \epsilon P_{out} \geq P_{out}-\lambda. \end{aligned} \tag{8}$$

for $n \geq n_0(|\mathcal{X}|,|\mathcal{Y}|,\epsilon,\tau)$. Since λ is arbitrary small, it follows that average error probability cannot be smaller than P_{out}, asymptotically in the length of the code. ■

III. No CSI at the transmitter

Many communications systems are designed without the transmitter having access to the instantenious channel realization. In this case the transmitter cannot "adapt" its codebook, and therefore the Outage Probability is defined as:

$$P_{out} = \inf_{P_X} P_W[I(P_X;W)<R]. \tag{9}$$

Once the optimizing distribution is chosen, or a distribution which brings the optimizing function "close" to P_{out}, then a transmission codebook will be chosen from this distribution. Obviously, error probabilites of P_{out} will be achievable. We formalize this point below. However, it is not immediately clear whether error probabilities can be lower than P_{out}, because the transmitter codebook may not necessarily be typical to a single distribution. We will use the fact that in any given codebook of finite length, there are exponentially many codewords but only polynomially many types.

A. Direct Part of the Outage Theorem

Subsequent discussion relies heavily on the "compound" channel coding theorem. A compound channel with input set $\mathcal{X}$, and output set $\mathcal{Y}$ is nothing more than a family $w \in \mathcal{W}$ of channels $w : \mathcal{X} \to \mathcal{Y}$. Note that the input alphabet is the same for every $w \in \mathcal{W}$, as well as the output alphabet. The capacity of a compound channel is:

$$C = \sup_{P_X} \inf_{w \in \mathcal{W}} I(P_X; w). \tag{10}$$

Our goal is to prove the equivalent of Theorem 2, for the case of no CSI at the transmitter. This proof has been outlined in [3]. For that purpose let $\lambda, \tau > 0$. Since $P_{out} = \inf_{P_X} P_W[I(P_X; W) < R]$, there exists some $P_X^\star$ so that

$$P_W[I(P_X^\star; W) < R] < P_{out} + \lambda/2. \tag{11}$$

Define $\mathcal{W}$ to be the set of all channel realizations "favorable" to $P_X^\star$, meaning that $\mathcal{W} = \{w : I(P_X^\star; w) \geq R\} = [I(P_X^\star; w) \geq R]$. We have that:

$$C = \sup_{P_X} \inf_{w \in \mathcal{W}} I(P_X; w) \geq \inf_{w \in \mathcal{W}} I(P_X^\star; w) \geq R. \tag{12}$$

Since the compound channel capacity for $w \in \mathcal{W}$ is not less than R, for $n \geq n_0(|\mathcal{X}|, |\mathcal{Y}|, \lambda/2, \tau)$ there is a code $\mathcal{C}$ of rate $R - \tau$, with the average error probability less than $\lambda/2$, for all $w \in \mathcal{W}$. Now we upper bound the average error probability as:

$$P_e(\mathcal{C}) = E_W[P_e(\mathcal{C}, W)]$$

$$\stackrel{(a)}{=} E_W[P_e(\mathcal{C}, W)\ \Upsilon_{[I(P^\star,w)<R]}(W)] + E_W[P_e(\mathcal{C}, W)\ \Upsilon_{[I(P^\star,w)\geq R]}(W)]$$

$$\stackrel{(b)}{\leq} E_W[\Upsilon_{[I(P^\star,w)<R]}(W)] + E_W[P_e(\mathcal{C}, W)\ \Upsilon_{\mathcal{W}}(W)]$$

$$\stackrel{(c)}{\leq} (P_{out} + \lambda/2) + (\lambda/2)E_W[\Upsilon_{\mathcal{W}}(W)] \leq P_{out} + \lambda. \quad (13)$$

Here, (a) is trivial, the first term in (b) is because any error probability is less than one and the second term in (b) is just the definition of $\mathcal{W}$. The first term in (c) is by equation (11), and the second term in (c) is because if $w \in \mathcal{W}$, then $P_e(\mathcal{C}, w) \leq \lambda/2$. We arrive at the following theorem.

Theorem 5: Let W be a block fading DMC, and let P_{out} be the outage probability for W at the rate R, as defined above. Then $\forall\ \tau, \lambda > 0$ there is code $\mathcal{C}$ of rate $R - \tau$ so that

$$P_e(\mathcal{C}) = E_W[P_e(\mathcal{C}, W)] \leq P_{out} + \lambda. \quad (14)$$

Hence the transmitter can transmit at a rate $R-\tau$, with average error probability arbitrary close to outage.

B. Converse to the Outage Theorem

Here we essentially want to make a claim similar to the one in Theorem 4, but in this case the transmitter doesn't have the CSI. We cannot immediately use the converse to the capacity theorem, as we did in the proof of the Theorem 4, since in this case the expression for outage does not involve capacity, but only the mutual information. The following theorem which was proven in [1] will be applicable, though not immediately.

Theorem 6: [1] For every $(\epsilon, \tau) \in (0, 1)$, if $\mathcal{C}$ is a code of length n for the DMC w, whose average error probability is less than ϵ, *and all codewords belong to type* P, then:

$$\frac{1}{n} \log(|\mathcal{C}|) < I(P, w) + \tau, \quad (15)$$

whenever $n \geq n_0(|\mathcal{X}|, |\mathcal{Y}|, \epsilon, \tau)$.

Type of a codeword is its empirical distribution [1]. It counts the relative frequency of occurence of a given symbol in a codeword. If all the codewords in a codebook have the same type P, we say that the codebook itself has a type P. However, we want to show that for *any* codebook the Outage Probability is an asymptotical lower bound for the average error probability. All the codewords in a given codebook need not have the same type, and that is what prevents us from using the Theorem 6 directly.

Another note on Theorem 6 is that n_0 is independent of w and P. Contrapositive statement to the theorem says that if the code rate is greater than $I(P,w)+\tau$, then its average error probability is greater than ϵ, for $n \geq n_0(|\mathcal{X}|, |\mathcal{Y}|, \epsilon, \tau)$. Next few lemmas are relevant for any codebook, and together with the Theorem 6 they will enable us to prove the claim equivalent to Theorem 4, but now for the case of no CSI at the transmitter.

Lemma 1: Let $\mathcal{C}$ be a codebook of rate $R+\tau$ with $|\mathcal{C}| = 2^{n(R+\tau)}$ codewords, $\mathcal{C}_P$ be the set of codewords of type P, and $\tau > 0$. We have:

$$1 \geq \sum_{P:|\mathcal{C}_P|>2^{n(R+\tau/2)}} \frac{|\mathcal{C}_P|}{|\mathcal{C}|} \geq 1 - 2^{-n\tau/2}(n+1)^{\mathcal{X}}. \tag{16}$$

The middle term therefore converges to 1 as $n \to \infty$, because the exponential decay is stronger than the polynomial increase.

Proof: The left inequality is obvious, since $\mathcal{C}_P$ form a disjoint partition of $\mathcal{C}$, as we let P vary. To prove the right inequality, we observe that for a code of length n there are not more than $(n+1)^{\mathcal{X}}$ different types, since any symbol from $\mathcal{X}$ can occur in at most n places, or not occur at all. The number of terms in the sum $\sum_P |\mathcal{C}_P|$ is therefore at most $(n+1)^{\mathcal{X}}$, therefore:

$$\sum_{P:|\mathcal{C}_P|>2^{n(R+\tau/2)}} |\mathcal{C}_P| + 2^{n(R+\tau/2)}(n+1)^{\mathcal{X}} \geq \sum_P |\mathcal{C}_P| = |\mathcal{C}| = 2^{n(R+\tau)}. \tag{17}$$

The result follows by dividing through with $|\mathcal{C}| = 2^{n(R+\tau)}$. ■

Lemma 2 (Disjoint Partition Lower Bound): If $P_e(\mathcal{C})$ is the average error probability of the code $\mathcal{C}$, then:

$$P_e(\mathcal{C}) \geq \sum_P \frac{|\mathcal{C}_P|}{|\mathcal{C}|} P_e(\mathcal{C}_P). \tag{18}$$

Proof: Let $\underline{x} \in \mathcal{C}_P$ be a codeword, and let $P_e(\mathcal{C}|\underline{x})$ be the error probability given that $\underline{x}$ is transmitted. Since $\mathcal{C}_P \subset \mathcal{C}$, we have that $P_e(\mathcal{C}|\underline{x}) \geq P_e(\mathcal{C}_P|\underline{x})$. Summing up over all $\underline{x} \in \mathcal{C}$ we have:

$$P_e(\mathcal{C}) = \frac{1}{|\mathcal{C}|} \sum_{\underline{x}\in\mathcal{C}} P_e(\mathcal{C}|\underline{x}) = \frac{1}{|\mathcal{C}|} \sum_P \sum_{\underline{x}\in\mathcal{C}_P} P_e(\mathcal{C}|\underline{x})$$

$$\geq \frac{1}{|\mathcal{C}|}\sum_P \sum_{\underline{x}\in\mathcal{C}_P} P_e(\mathcal{C}_P|\underline{x}) = \frac{1}{|\mathcal{C}|}\sum_P |\mathcal{C}_P| P_e(\mathcal{C}_P). \tag{19}$$

The result is now immediate. ■

We are now ready to state and prove the following result.

Theorem 7: Let W be a block fading DMC, and let P_{out} be the outage probability for W at the rate R, as defined above. Then $\forall\ \tau, \lambda > 0$ if $\mathcal{C}$ is a code of rate $R+\tau$, then whenever code length satisfies $n \geq n_1(|\mathcal{X}|, |\mathcal{Y}|, \lambda, \tau, R)$, we have:

$$P_e(\mathcal{C}) = E_W[P_e(\mathcal{C}, W)] \geq P_{out} - \lambda. \tag{20}$$

Proof: Let τ, λ be given and choose $\epsilon \in (0,1)$ so that $\epsilon^2 P_{out} \geq P_{out} - \lambda$. Choose $n'_0(\epsilon, \tau, R)$ so that the convergence from Lemma 1 is between 1 and ϵ. Choose $n_0(|\mathcal{X}|, |\mathcal{Y}|, \epsilon, \tau/2)$ so that the statements from Theorem 6 hold. Let $n_1 = \max\{n_0, n'_0\}$. When $n \geq n_1$, we have that:

$$\begin{aligned}
P_e(\mathcal{C}) &\overset{(a)}{\geq} \sum_P \frac{|\mathcal{C}_P|}{|\mathcal{C}|} P_e(\mathcal{C}_P) \overset{(b)}{\geq} \sum_{P:|\mathcal{C}_P|>2^{n(R+\tau/2)}} \frac{|\mathcal{C}_P|}{|\mathcal{C}|} P_e(\mathcal{C}_P) \\
&\overset{(c)}{\geq} \sum_{P:|\mathcal{C}_P|>2^{n(R+\tau/2)}} \frac{|\mathcal{C}_P|}{|\mathcal{C}|} E_W[P_e(\mathcal{C}_P, W)\ \Upsilon_{[I(P,w)+\tau/2<R+\tau/2]}(W)] \\
&\overset{(d)}{\geq} \sum_{P:|\mathcal{C}_P|>2^{n(R+\tau/2)}} \frac{|\mathcal{C}_P|}{|\mathcal{C}|} E_W[\epsilon\ \Upsilon_{[I(P,w)<R]}(W)] \\
&\overset{(e)}{\geq} \left(\sum_{P:|\mathcal{C}_P|>2^{n(R+\tau/2)}} \frac{|\mathcal{C}_P|}{|\mathcal{C}|}\right) \epsilon\ \inf_P E_W[\Upsilon_{[I(P,w)<R]}(W)] \\
&\overset{(f)}{\geq} \epsilon^2 P_{out} \overset{(g)}{\geq} P_{out} - \lambda.
\end{aligned} \tag{21}$$

Here, (a) is by *Lemma 1.* (b), and (c) are by trivial lower bounding. Theorem 6 is relevant for (d), since all codewords in a given sub-codebook $|\mathcal{C}_P|$ are of the same type P, and of rate $R+\tau/2$. Since $n \geq n_0(|\mathcal{X}|, |\mathcal{Y}|, \epsilon, \tau/2)$, for any w for which $I(P,w)+\tau/2 < R+\tau/2$, we have that $P_e(\mathcal{C}_P, w) \geq \epsilon$. (e) is because we have replaced all the terms in the sum from (d) with something (possibly) lower. *Lemma 2* is relevant for (f), since $n > n'_0$. implies that the bracketed term is greater than ϵ. (g) is simply by our choice of ϵ. ■

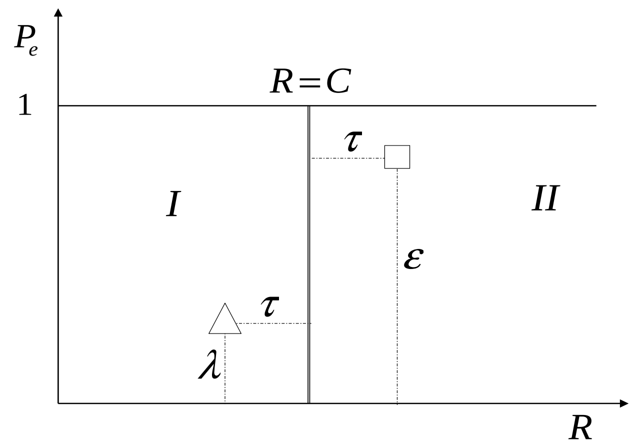

Fig. 1. Codeword error probability vs. transmission rate. Graphic representation of the Capacity theorem, together with the Converse.

IV. Discussion and Conclusions

Here we provide a graphical interpretation of the above results. All statements that follow are valid for large code lengths. In general, one would like to communicate at a maximum possible rate with the smallest possible average error probability. Fundamental theorems of information theory put bounds and tradeoffs on these two quantities.

Fig. 1 interprets the capacity theorem and its converse for any DMC channel. Theorem 1 essentially says that the region I from the Fig. 1 is achievable. However, in order to establish that the line $R = C$ is a fundamental bound, it is necessary to establish the Theorem 3. It says that the region II in Fig. 1 is not achievable. Only after Theorem 3 is established, we have completed a tradeoff between the transmission rate and the average error probability. No claims are made about those exact points on the line $R = C$.

Fig. 2 interprets the Outage theorem for a block fading DMC channel. Theorems 2 and 5 essentially say the the region I from the Fig. 2 is achievable. However, in order for the curve $P_e = P_{out}(R)$ to represent a fundamental tradeoff, a claim about the points on the right of this curve has to be made. This is where Theorems 4 and 7 become relevant. They essentially state the the region II from in the Fig. 1 is not achievable. No claims are made about those exact points on the line $P_e = P_{out}(R)$.

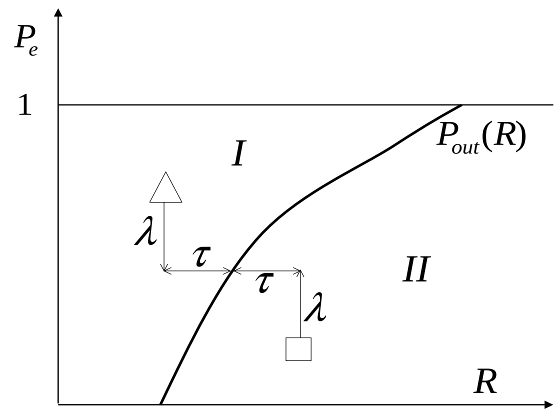

Fig. 2. Codeword error probability vs. transmission rate. Graphic representation of the Outage theorem, together with the Converse.

References

[1] I. Csiszár and J. Körner, *Information Theory.* New York: Academic, 1981.

[2] L. Zheng and D. Tse "Divesity and Multiplexing: A Fundamental Tradeoff in Multiple Antenna Systems" *to appear in IEEE Trans. Inform. Theory,*

[3] E. Biglieri, J. Proakis and S. Shamai (Shitz), "Fading Channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory,* vol. 44, pp. 2619-2692, Oct. 1998.

[4] E. Biglieri, G. Caire and G. Taricco, "Limiting Performance of Block-Fading Channels with Multiple Antennas," *IEEE Trans. Inform. Theory,* vol. 47, pp. 1273-1289, May. 2001.

[5] E. Malkamäki and H. Leib, "Coded diversity on Block-Fading channels," *IEEE Trans. Inform. Theory,* vol. 45, pp. 771-781, March 1999.

[6] L. H. Ozarow and S. Shamai and A. D. Wyner, "Information Theoretic Considerations for Cellular Mobile Radio," *IEEE Trans. on Vehic. Tech.,* vol. 43, pp. 359-378, May 1994.

[7] R. Knopp and P. A. Humblet, "On Coding for Block Fading Channels," *IEEE Trans. Inform. Theory,* vol. 46, pp.189-205, Jan. 2000.

[8] R. G. Gallager *Information Theory and Reliable Communication,* John Willey and Sons, 1965.