Outage Probability Under Channel Distribution Uncertainty

Ioanna Ioannou, Charalambos D. Charalambous and Sergey Loyka

***Abstract*—Outage probability of a class of block-fading (MIMO) channels is considered under channel distribution uncertainty, when the channel or its distribution are not known but the latter is known to belong to a class of distributions where each member is within a certain distance from a nominal distribution. Relative entropy is used as a measure of distance between distributions. Compound outage probability defined as min (over the input distribution) -max (over the channel distribution class) outage probability is introduced and investigated, which generalizes the standard outage probability to the case of partial channel distribution information. Compound outage probability characterization via one-dimensional convex optimization, its properties and approximations are given. It is shown to have a two-regime behavior: when the nominal outage probability decreases, the compound outage first decreases linearly down to a certain threshold and then only logarithmically (i.e. very slowly), so that no significant further decrease is possible. The input distribution optimized for the nominal channel distribution is shown to be also optimal for the whole class of distributions. The effect of swapping the distributions in relative entropy is investigated and an error floor effect is established. The obtained results hold for a generic channel model (arbitrary nominal fading and noise distributions).**

I. Introduction

Wireless channel capacity depends significantly on the channel state information available at the transmitter and the receiver as well as the fading statistics experienced by the channel [1]. Since channel state information (CSI) is obtained via channel measurements, its accuracy may be limited due to variability and difficult propagation conditions (e.g. low SNR). The transmitter CSI is further limited due to limitations of the feedback channel (if any). This situation can be modeled via a compound channel model, where the true channel is not known but it is known to belong to a certain (limited) class of channels and the corresponding compound channel capacity theorems have been established [1][2]. The compound MIMO channel capacity under the spectral norm constraints have been studied recently in [3]. While compound channel capacity theorems treat all channels in the class equally and build a code that performs well on any such channel, the corresponding capacity is typically limited by the worst channel in the class and may be low, even though most channels in class are good and the worst channel is realized with low probability, i.e. it is a conservative performance indicator. To avoid this problem, a concept of composite channel has been introduced [2][5], where each channel in a class has associated probability measure, so that bad low-probability channels do not penalize significantly the performance metric. The corresponding channel capacity theorems can be proved via the concept of information density [4][5] or using the compound channel approach [1].

I. Ioannou and C.D. Charalambous are with the ECE Department, University of Cyprus, 75 Kallipoleos Avenue, P.O. Box 20537, Nicosia, 1678, Cyprus, e-mail: aioannak@yahoo.gr, chadcha@ucy.ac.cy

S. Loyka is with the School of Information Technology and Engineering, University of Ottawa, Ontario, Canada, K1N 6N5, e-mail: sergey.loyka@ieee.org

In this paper, we consider a situation where even the channel distribution information is not available at the transmitter; rather, the transmitter knows that the distribution belongs to a certain class centered around a nominal distribution (known to the transmitter). This models a practical scenario where the channel distribution information is obtained from multiple but limited measurements, so that the true distribution is known only with limited accuracy. This also models a dynamic scenario where the channel distribution information obtained from past measurements may be outdated. The uncertainty in this information may also be related to the limitation of the feedback channel used to supply the information to the transmitter. We assume a quasi-static (block-fading) scenario so that the CSI at the receiver is irrelevant [1]. This channel model is quite generic: we do not assume any particular nominal channel distribution and even the channel noise can be arbitrary so that the results are general too. Relative entropy between two distributions is used as a measure of distance, so that the distribution uncertainty class includes all distributions within certain relative entropy distance of the nominal one. Similar approach was adopted in [6] to study the ergodic capacity under channel distribution uncertainty and in [7] to investigate an optimal control of stochastic uncertain systems.

Since the channel is block-fading, the outage probability is the main performance metric, which we term ”compound outage probability” to emphasize that it applies to a class of fading distributions (i.e. ”compound distribution”) rather than any particular one. This parallels the concept of compound channel, where a code is designed to operate on any member in the class. In our case, a code is designed to operate for any channel distribution in the class, so that the compound outage probability involves maximization over all feasible channel distributions and minimization over the transmitted signal distribution (subject to the power constraint), and the corresponding compound outage capacity can be derived from it.

The channel model is introduced in Section II. Compound outage probability is defined and investigated in Section III, which includes its closed-form characterization in Theorem 1 (as one-dimensional convex optimization problem) and the worst channel distribution (which is a piece-wise constant scaling of the nominal distribution). Remarkably, the compound outage probability depends only on the nominal one and the relative entropy distance, all other details being irrelevant. Properties of the compound outage probability are given in Propositions 1-3, and its two-regime asymptotic behavior is identified in Section III-B. Specifically, as the nominal outage probability decreases (say by increasing the SNR), the compound outage probability first decreases linearly, but after a certain threshold - only logarithmically, i.e. very slowly, so that significant decrease is not possible anymore. Optimizing the input signal distribution in this regime does not bring in significant improvement either so that any reasonable distribution (e.g. isotropic signalling in MIMO channels) will do as well. Compact, closed-from approximations are obtained for the compound outage probability in these two regimes using the tools of asymptotic analysis. Theorem 2 shows that the input signal distribution optimal for the nominal channel distribution is also optimal for the whole class, so that numerous known optimal distributions can be "recycled". Since relative entropy is not symmetric, Section IV investigates the impact of this asymmetry on the compound outage probability. Swapping the distributions (nominal and true) is shown to result in the error floor effect: the compound outage probability is bounded away from zero, does not matter how low the nominal outage (or how high the SNR) is. The error floor depends on the relative entropy distance: it decreases with it and when it is small, they are equal, so that the relative entropy distance is an adequate measure of fading uncertainty in the non-ergodic regime.

II. Channel Model and Outage Probability

Let us consider a generic discrete-time baseband multiple-input multiple-output channel where $\mathbf{x}$ and $\mathbf{y}$ are the input (transmitted) and output (received) vectors (or sequences), and $\mathbf{H}$ denotes channel state. In the general case, the channel is described by the conditional probability distribution $W(\mathbf{y}|\mathbf{x},\mathbf{H})$ of $\mathbf{y}$ given $\mathbf{x}$ and $\mathbf{H}$, and the mutual information (per channel use) supported by the channel for a given distribution of $\mathbf{x}$ and channel state $\mathbf{H}$ is $I(\mathbf{x};\mathbf{y}|\mathbf{H})$. We assume that the channel is block-fading (non-ergodic), i.e. a particular channel realization $\mathbf{H}$ is selected in the beginning and stays fixed for the whole duration of codeword transmission; next codeword will see a different channel realization [1]. Channel fading distribution is described by its probability density function $f(\mathbf{H})$. Most of our results will hold in this generic scenario, which includes as special cases frequency-selective (inter-symbol interference) or frequency-flat (no ISI) Gaussian MIMO channels.

We will not assume any particular fading and noise distribution so that our results are general and apply to *any* such distribution. The transmitted signal, receiver noise and the channel are assumed to be independent of each other. We also assume that the transmitter does not know the channel but only has a partial knowledge of its distribution. A popular Gaussian MIMO channel is a special case in this model.

Main performance metrics in the block-fading regime are outage probability and outage capacity [1] [2]. Outage probability is the probability that the channel is not able to support the target rate R. When the transmitter knows the channel distribution (but not the channel itself), the outage probability is

$$P_{out}(R) = \min_{\rho(\mathbf{x})} \Pr\{I(\mathbf{x};\mathbf{y}|\mathbf{H}) < R\} \tag{1}$$

where $\rho(\mathbf{x})$ is the distribution of $\mathbf{x}$ subject to the total power constraint $\mathbb{E}[\mathbf{x}^\dagger\mathbf{x}] \leq P_T$, $\Pr\{I(\mathbf{x};\mathbf{y}|\mathbf{H}) < R\}$ is the outage probability for a given $\rho(\mathbf{x})$ and the minimization is over all possible distributions of the input $\mathbf{x}$ satisfying the power constraint. Using the outage probability, outage capacity can also be found as the maximum possible rate subject to the outage probability constraint. Finally, one may also consider the outage probability and capacity for a given (fixed) $\rho(\mathbf{x})$. Operational meaning of the outage capacity/probability follows from the compound channel capacity theorems [1][2]; see also [4][5] for a modern treatment using the concept of information density.

III. Compound Outage Probability

Consider the scenario where the transmitter has only partial channel distribution information. Namely, it knows that the channel probability density function (PDF) $f(\mathbf{H})$ is within a certain distance of the nominal one $f_0(\mathbf{H})$. We use the relative entropy as a measure of the distance between two distributions, so that all feasible distributions f satisfy the following inequality:

$$D(f||f_0) = \int f \ln \frac{f}{f_0} d\mathbf{H} \leq d \tag{2}$$

where $D(f||f_0)$ is the relative entropy or Kullback-Leibler distance between the distributions, and d is the maximum possible distance in the uncertainty set to which f belongs; both d and $f_0(\mathbf{H})$ are known to the transmitter. Throughout the paper we assume that $d < \infty$. In this scenario, the definition in (1) does not apply (since the true channel distribution is not known) but can be generalized to

$$P_{out}^* = \min_{\rho(\mathbf{x})} \max_{D(f||f_0)\leq d} \Pr\{I(\mathbf{x};\mathbf{y}|\mathbf{H}) < R\} \tag{3}$$

where the max is over all feasible (satisfying (2)) channel distributions f. Its operational meaning also follows from the compound channel capacity theorems [1][2] or from [4][5], since the optimal input distribution does not depend on the true channel distribution, but only on the nominal one and the relative entropy distance d, both known to the transmitter. This problem setup models a practical situation where the channel distribution information is obtained from measurements or physical modeling, which are never perfect. It also accounts for the fact that the estimated channel distribution

[1]With a slight modification in notations, this block-fading model can also be extended to the case where each codeword sees a finite number of channel realizations and our results will hold in that case as well.

[2]It can be further shown that the outage probability is the best achievable average codeword error probability [4][5].

may change with time in dynamic scenarios. We term P_{out}^* in (3) "compound outage probability" since it is a performance measure of a class of channel distributions rather than a single distribution. This approach parallels the work on compound channel capacity [2][3] where the channel is not known to the transmitter but it is known to belong to a certain class.

To characterize the compound outage probability P_{out}^*, we adopt a two-step approach: first, we characterize the outage probability for a given input distribution, i.e. no minimization in (3), which also represent a practical situation where this distribution is set a priory; then it is minimized over all feasible input distributions.

A. Step 1: Compound Outage for a Given Input Distribution

When the input distribution $\rho(\mathbf{x})$ is fixed a priory, the compound outage probability is

$$P_{out} = \max_{D(f||f_0)\leq d} \Pr\{I(\mathbf{x};\mathbf{y}|\mathbf{H}) < R\} \tag{4}$$

Its characterization is strikingly simple in the generic scenario, i.e. for any noise and nominal fading distribution.

Theorem 1. *For a given input distribution $\rho(\mathbf{x})$ and arbitrary nominal fading distribution f_0, the outage probability in* (4) *can be expressed as*

$$P_{out} = \min_{s\geq 0}[s\ln(1+(e^{1/s}-1)\varepsilon)+sd] \tag{5}$$

where

$$\varepsilon = \int\limits_{I(\mathbf{x};\mathbf{y}|\mathbf{H})<R} f_0 d\mathbf{H} \tag{6}$$

is the nominal outage probability (i.e. the outage probability under the nominal channel distribution). The worst channel distribution f^* (the maximizer in (4)*) is given by*

$$f^* = \frac{(e^{1/s^*}-1)\ell(\mathbf{H})+1}{(e^{1/s^*}-1)\varepsilon+1} f_0 \tag{7}$$

where s^* is the minimizing s in (5)*, and $\ell(\mathbf{H})$ is the indicator of the outage set: $\ell(\mathbf{H})=1$ if $I(\mathbf{x};\mathbf{y}|\mathbf{H})<R$ and 0 otherwise.*

Proof: since the problem is convex with zero duality gap, follows from the KKT conditions, see [11] for details. ■

Note that Theorem 1 effectively reduces the infinite-dimensional optimization problem in (4) (the optimization there is over the set of all admissible distributions f) to one-dimensional convex optimization in (5), which can be efficiently solved using numerical algorithms. It is remarkable that the nominal distribution enters the compound outage probability in (5) only via the nominal outage probability ε, all other its details being irrelevant, i.e. two different nominal distributions with the same nominal outage probability will produce the same compound outage probability.

Note that the maximizing density f^* in (7) mimics the nominal one f_0 in a piece-wise constant manner:

$$\frac{f^*}{f_0} = \begin{cases} \frac{e^{1/s^*}}{(e^{1/s^*}-1)\varepsilon+1}, & \text{if } \mathbf{H}\in\mathcal{O} \\ \frac{1}{(e^{1/s^*}-1)\varepsilon+1} & \text{if } \mathbf{H}\notin\mathcal{O} \end{cases} \tag{8}$$

where $\mathcal{O}=\{\mathbf{H}:\ell(\mathbf{H})=1\}$ is the outage set, so that the right-hand side of (8) is independent of $\mathbf{H}$ in each set and f^* is a scaled up version of f_0 in the outage set and scaled down otherwise.

A number of properties of the compound outage probability are given below (proofs are omitted due to the page limit and are available in a full version of this paper [11]).

Proposition 1. *For a given input distribution $\rho(\mathbf{x})$, the compound outage probability $P_{out}(d)$ as a function of distance d has the following properties:*

1) $P_{out}(d)$ is concave in $d\geq 0$.

2) $P_{out}(d=0)=\varepsilon$, i.e. the compound outage probability equals the nominal one when $d=0$.

3) $P_{out}(d)$ is a non-decreasing function of d, that is

$$P_{out}(d_1)\leq P_{out}(d_2), \quad 0\leq d_1<d_2<\infty. \tag{9}$$

and the equality holds if and only if $P_{out}(d_1)=P_{out}(d_2)=1,0$ i.e. $P_{out}(d)$ is a strictly increasing function of the distance d unless $P_{out}=1,0$. □

Proposition 2. *The compound outage probability P_{out} has the following properties:*

1) $P_{out}=1$ if and only if $\varepsilon=1$.

2) $P_{out}=0$ if and only if $\varepsilon=0$.

3) $P_{out}\geq\varepsilon$, and the equality holds if and only if $d=0$ or $\varepsilon=0,1$. □

While in general the compound and nominal outage probabilities can differ significantly, the former takes on a limiting value (either 0 or 1) if and only if the latter does so.

Proposition 3. *The compound outage probability in* (5) *is a strictly-increasing, concave function of the nominal outage ε, i.e.*

$$P_{out}(\varepsilon_1)<P_{out}(\varepsilon_2), \quad 0\leq\varepsilon_1<\varepsilon_2\leq 1. \tag{10}$$

with the boundary conditions $P_{out}(\varepsilon=0)=0$, $P_{out}(\varepsilon=1)=1$. □

B. Asymptotic Regimes

We now consider the compound outage in (5) in two limiting regimes: 1) The uncertainty-dominated regime $\varepsilon\to 0$ and fixed d, and 2) The nominal-outage dominated regime $d\to 0$ and fixed ε. The following approximations can be obtained using the standard tools of asymptotic analysis [9].

Proposition 4. *The outage probability P_{out} in* (5) *in the low nominal outage regime, $\varepsilon\to 0$ and fixed $d>0$, is as follows:*

$$P_{out} = \frac{d}{\ln\frac{d}{\varepsilon}-\ln\ln\frac{d}{\varepsilon}}(1+o(1)). \tag{11}$$

Note from (11) that the main contribution to P_{out} is coming from d (i.e. the uncertainty) rather than ε (i.e. the nominal outage) since $\ln(d/\varepsilon)$ is a slowly-varying function of ε, so that variations from the nominal channel distribution dominate the outage events. Also note that the relative entropy distance d is directly related to the compound outage probability, which indicates that it is this distance that should be used as a measure of accuracy in estimating the channel distribution from measurements or physical modeling since it is directly related to the system performance.

Let us now consider the nominal outage-dominated regime.

Proposition 5. *In the low channel distribution uncertainty regime, $d \to 0$ and fixed ε, the compound outage probability is*

$$P_{out} = \varepsilon + \sqrt{2d(1-\varepsilon)\varepsilon} + o(\sqrt{d}) \tag{12}$$

Comparing Proposition 4 and Proposition 5, one concludes that indeed there are two regimes in the behavior of $P_{out}(\varepsilon)$, as illustrated in Fig. 1:

1) The uncertainty dominated regime (nominal outage is neglegible), when $\varepsilon \ll d < 1$ so that

$$P_{out} \approx \frac{d}{\ln\frac{d}{\varepsilon} - \ln\ln\frac{d}{\varepsilon}} \sim \frac{d}{\ln(1/\varepsilon)} \tag{13}$$

where $\sim$ means "scales as", so that P_{out} depends linearly on d but only logarithmically (i.e. very slowly) on ε.

2) The nominal outage-dominated regime (uncertainty is negligible), when $d \ll \varepsilon < 1$ and

$$P_{out} \approx \varepsilon + \sqrt{2d(1-\varepsilon)\varepsilon} \sim \varepsilon \tag{14}$$

i.e. d contributes very little to the outage probability.

These two regimes immediately suggest some design guidelines related to the outage probability. In the uncertainty dominated regime, the main way to reduce outage probability is via decreasing the uncertainty of the channel distribution, e.g. via improved channel measurements or modeling; reducing the nominal outage probability is not efficient here, so that minimizing it via the optimal input distribution is not worth the effort - any reasonable distribution (e.g. isotropic signalling in MIMO channels) will do as well. This approach, however, will bring little improvement in the nominal outage-dominated regime, where the only way to reduce the outage probability is via improving systems performance under the nominal fading, e.g. by increasing the SNR or optimizing the input distribution. These conclusions hold for any nominal channel distribution (e.g. not limited to i.i.d. Rayleigh) and for any noise (not only Gaussian).

This two-regime behavior can also be linked to the way channel distribution is obtained from measurements: a finite number of fading channel realizations are measured and the empirical channel distribution is derived based on it. However, the relative accuracy of this empirical distribution is always lower at the distribution tails, where fewer measurement points are available. On the other hand, in the uncertainty-dominated regime, the nominal outage probability is low when the average SNR is high so that a nominal outage event takes place when the channel is very weak, i.e. at the distribution tail, so that the inaccuracy in the channel distribution estimation plays a dominant role there. Ultimately, low compound outage probability can only be achieved by insuring sufficiently high accuracy of the estimated distribution tail (small d), i.e. when a sufficient number of independent measurements fall into that region.

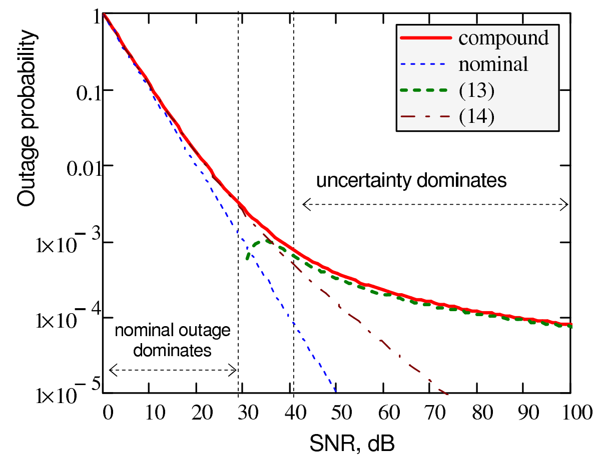

Fig. 1. Two-regime behavior of the compound outage probability. Its approximations in (13) and (14) and nominal outage $\varepsilon = 1/\mathrm{SNR}$ (set for convenience as in Rayleigh fading at high SNR) are also shown; $d = 10^{-3}$. Decreasing the compound outage probability beyond about $10^{-3}(\approx d)$ requires exponentially high SNR and is not practical (it takes 60 dB extra to go from 10^{-3} to 10^{-4}, while normally, i.e. without uncertainty, it would take only 10 dB).

C. Step 2: Minimizing over the Input Distribution

Using Theorem 1, we are now in a position to characterize the compound outage probability in (3).

Theorem 2. *Consider a class of fading channels in* (2)*. Its compound outage probability in* (3) *can be found from*

$$P_{out}^* \;=\; \min_{s\geq 0}[s\ln(1+(e^{1/s}-1)\varepsilon^*)+sd] \tag{15}$$

where $\varepsilon^* = \min_{\rho(\mathbf{x})}\varepsilon$ is the optimized nominal outage probability, so that the outage-minimizing input distribution for the class of channel distributions in (2) *and for the nominal distribution f_0 are the same,*

$$\arg\min_{\rho(\mathbf{x})} P_{out} = \arg\min_{\rho(\mathbf{x})} \varepsilon \tag{16}$$

Proof: use Theorem 1 and observe that $\ln(\cdot)$ is a monotonic function, $\min_{\rho(\mathbf{x})}$ and $\min_{s\geq 0}$ can be swaped, so that the minimization of the compound outage over $\rho(\mathbf{x})$ is equivalent to the minimization of the nominal outage. ■

A significance of Theorem 2 is that the optimal input distribution is the same as for the nominal channel, so that a significant number of known results apply directly to the compound fading channel as well, i.e. no new search of optimal input distributions is required.

When the compound outage P_{out}^* in (15) is considered as a function of the distance d, $P_{out}^*(d)$, its properties mimic those in Proposition 1 with the substitution $\varepsilon \to \varepsilon^*$. Also, the results and conclusions of Section III-B hold under this substitution. In particular, optimizing the input distribution is worth the effort only in the nominal outage dominated regime.

IV. The Impact of Asymmetry

Since relative entropy is not symmetric, i.e. $D(f||f_0) \neq D(f_0||f)$, we consider in this section the constraint $D(f_0||f) \leq d$ to see the impact of the order on the obtained results. The main results are summarized below (see [11] for details).

Proposition 6. *For a given input distribution, the compound outage under the distribution class $D(f_0||f) \leq d$,*

$$P_{out} = \max_{D(f_0||f) \leq d} \Pr\{I(\mathbf{x};\mathbf{y}|\mathbf{H}) < R\} \tag{17}$$

is bounded as follows:

$$P_{out} \geq 1 - e^{-d} + e^{-d}\varepsilon \geq 1 - e^{-d} \tag{18}$$

The bounds in (18) *are tight when $\varepsilon \to 0$ and fixed d,*

$$P_{out} = 1 - e^{-d} + o(1), \tag{19}$$

and when $d \to 0$ and fixed ε,

$$P_{out} = \varepsilon + \sqrt{2\varepsilon(1-\varepsilon)d} + o\left(\sqrt{d}\right) \to \varepsilon \tag{20}$$

When $d \ll 1$,

$$P_{out} \geq d + \varepsilon \tag{21}$$

An important conclusion is immediate from (18): $P_{out}(\varepsilon = 0) \geq 1 - e^{-d}$, i.e. there is an error floor effect in the behavior of $P_{out}(\varepsilon)$: even though $\varepsilon \to 0$ (e.g. by $\mathrm{SNR} \to \infty$), $P_{out} \nrightarrow 0$. From (21), $P_{out} \geq d$, i.e. cannot be made smaller than the relative entropy distance d, does not matter how large the SNR (or how small the nominal outage) is. This is in contrast to (11), where $P_{out} \to 0$ when $\varepsilon \to 0$, even though logarithmically slowly (i.e. no error floor). The absence of error floor in the latter case should not however be overestimated, since the convergence $P_{out} \to 0$ is logarithmically slow in ε, i.e. requires exponentially large SNR, so that for all practical purposes, P_{out} also saturates around d, as was indicated in Section III-B. Note also that (21) places d and ε on equal footing, re-enforcing our earlier conclusion that d is an adequate measure of fading uncertainty in the non-ergodic regime.

Comparing (20) to (12), we conclude that the compound outage probability is the same for the $D(f\|f_0) \leq d$ and $D(f_0\|f) \leq d$ uncertainty sets in the low uncertainty regime while the same cannot be said about the low nominal outage regime (compare (19) to (11)).

V. Conclusion

Compound outage probability of a class of fading (MIMO) channels with partial channel distribution information has been introduced and studied. This concept generalize well-known and widely used concepts of outage probability and capacity with completely known channel distribution to the case where only its partial knowledge is available. Relative entropy distance is used as a measure of uncertainty, which is shown to be related directly to the compound outage probability so that it is an adequate measure of the fading uncertainty in the non-ergodic regime. A number of properties, bounds and approximations of the compound outage probability are given. The input distribution optimized for the nominal outage probability is shown to be also optimal for the compound one. These results hold for an arbitrary nominal channel distribution (e.g. Rayleigh, Rician, Nakagami, Log-Normal, correlated and/or non-identically distributed, etc.) and also for arbitrary noise (not only Gaussian).

References

[1] E. Biglieri, J. Proakis, and S. Shamai, “Fading channels: Information-theoretic and communications aspects,” *IEEE Trans. Inform. Theory*, vol. 44, No. 6, pp. 2619-2692, Oct. 1998.

[2] A. Lapidoth and P. Narayan, “Reliable Communication Under Channel Uncertainty,” *IEEE Trans. Inform. Theory*, vol. 44, No. 6, Octomber 1998.

[3] S. Loyka, C.D. Charalambous, On the Capacity of a Class of MIMO Channels Subject to Normed Uncertainty, 2008 IEEE Int. Symp. Info. Theory, Toronto, Canada, July 2008.

[4] S. Verdu, T.S. Han, ”A General Formula for Channel Capacity”, *IEEE Trans. Info. Theory*, vol. 40, no. 4, pp. 1147-1157, July 1994.

[5] M. Effros, A. Goldsmith, Y. Liang, ”Generalizing Capacity: New Definitions and Capacity Theorems for Composite Channels,” *IEEE Trans. Info. Theory*, vol. 56, no. 7, pp. 3069-3087, July 2010.

[6] C. D. Charalambous, D. Stojan, and C. Constantinou, ”Capacity of the Class of MIMO Channels With Incomplete CDI: Properties of Mutual Information for a Class of Channels,” *IEEE Trans. Info. Theory*, vol. 55, pp. 3725-3734, Aug. 2009.

[7] C. D. Charalambous, and F. Rezaei, ”Stochastic Uncertain System Subject to Relative Entropy Constraints: Induced Norms and Monotonicity Properties of Minimax Games”, *IEEE Trans. Automatic Control*, vol.52, no.4, pp. 647-663, Apr. 2007.

[8] S. Boyd, L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[9] M.A. Efgrafov, *Asymptotic Expansions and Entire Functions*, Moscow: GITTL, 1957.

[10] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[11] I. Ioannou, C.D. Charalambous, S. Loyka, ”Compound Outage Probability and Capacity of a Class of Fading MIMO Channels with Channel Distribution Uncertainty,” *IEEE Trans. Info. Theory*, submitted, 2011. (available at ArXiv.org)