Journal of Communications and Information Networks
Vol. 1, No. 1, Jun. 2016
DOI: 10.11959/j.issn.2096-1081.2016.001

Review paper

# The impact of uncertain channel models on wireless communication

BIGLIERI Ezio

Universitat Pompeu Fabra, Barcelona 08018, Spain

**Abstract:** It is mainly deals with models of wireless communication channels, and in particular with the difficulty of finding an exact model which is both mathematically easy to deal with and physically accurate. It is argued that a certain amount of uncertainty necessarily following the use of inaccurate channel models should be accepted, and suitable tools used to evaluate its effects on analysis and design. With our approach, lower and upper bounds on required performance parameters are derived under no assumption of exact knowledge of the underlying probability distributions, and model uncertainty effects are propagated throughout calculations. Sepcial attention is given to the derivation of upper and lower bounds on system performance when this is determined by random variables whose dependence is not exactly known.

**Key words:** wireless channels, channel modeling, epistemic uncertainty, fading effects

**Citation**: BIGLIERI E. The impact of uncertain channel models on wireless communication[J]. Journal of communications and information networks, 2016, 1(1): 1-13.

## 1 Introduction

A good deal of research activity in wireless communications has been devoted to finding accurate statistical models of the channel. A major problem here is caused by the fact that no single model (e.g., Rayleigh or Rice distributions modeling fading effects) can be accurate enough for a wide variety of channels. Considerable efforts have been spent in the search for general classes of probability measures, viz., pdfs (probability density functions) or cdfs (cumulative distribution functions) that are physically justified and flexible enough to fit a large mass of experimental results. Nonetheless, wireless system analysis and design should in some way account for the uncertainty intrinsic in the use of inaccurate channel models. This uncertainty adds to that caused by the randomness of system behavior, but differs from it in a substantial way, so that we should treat differently the uncertainty due to randomness and the "epistemic" uncertainty due to ignorance. This distinction is not new, and has generated several techniques to deal with this second type of uncertainty: among them, we recall random-set theory[1,2], fuzzy-set theory[3], Dempster-Shafer theory[4], and probability boxes[5] (relations among these techniques, and some equivalence results, are discussed for example in Refs.[2,6,7]). In this paper we focus on mathematical models for wireless channels. While most previous works in this area were primarily focused on the un-

Manuscript received Feb. 17, 2016; accepted May 03, 2016
This work is supported by Project TEC2015-66228-P.

certainties due to randomness under the optimistic assumption that the channel models were perfectly known, here we study the impact of epistemic uncertainty on performance analysis. Our key point here is that one should avoid performance analyses relying on unwarranted distribution assumptions, and at the same time evaluate the cost of the inaccuracy of a physical model. Citing verbatim from Ref.[8], "it is better to have a correct analysis that honestly distinguishes between variability and incertitude than an analysis that depends on unjustified assumptions and wishful thinking. If the price of a correct assessment is broad uncertainty as a recognition or admission of limitations in our scientific knowledge, then we must pay that price."

This paper is organized as follows: Section 2 serves as motivation, and describes two simple yet general problems to which the theory expounded here can be applied, viz., the computation of error probabilities and outage probabilities. Section 3 shows how bounds on cdfs can be used or combined in computations, and it describes moment bounds, i.e., methods to derive upper and lower bounds to cdfs and to averages when only a few moments of the underlying RVs (Random Variables) are known. Section 4 shows how moment bounds can be applied to cope with certain sources of inaccuracy in cognitive radio using linear, quadratic, or linear-quadratic spectrum sensing. Section 5 describes a method to derive a "worst" distribution when the exact one is unknown, but lies "not too far" from the one being used. Section 6 deals with upper and lower performance bounds obtained when the statistical dependence among the random variables used in the model is totally unknown or only partially known.

## 2 Two applications

Although the theories described in this paper have various other applications, here we focus, for motivation's sake, on two specific problems, simple enough to be often amenable to closed-form solutions.

The first one is the calculation of a performance parameter of interest, say $H$, that can be expressed as the expectation

$$H = \mathcal{E}_Z[G(Z)] \ , \tag{1}$$

where $G(\cdot)$ is a known function, and $Z$ is a RV whose pdf is only inaccurately known.

The formulation above encompasses for example the calculation of the error probability $h$ of uncoded binary antipodal transmission over a channel affected by additive white Gaussian noise, SNR (Signal-to-Noise Ratio) equal to $snr$ and a fading whose envelope is modeled by the RV $R$, under the assumption of perfect channel state information at the receiver. In this case, Eq.(1) applies with

$$G(Z) = Q(\sqrt{2snr\,Z}) \ , \tag{2}$$

where $Q$ is the Gaussian tail function and $Z$=$R^2$.

The second problem of interest refers to a nonergodic channel affected by fading with random envelope $R$ and additive white Gaussian noise. The parameter $H$ of interest is now the information outage probability $h_{\text{out}}$, i.e., the probability that the transmission rate $\rho$, measured in bits per dimension pair, exceeds the instantaneous mutual information of the channel at signal-to-noise ratio $snr$. This is given by[9]

$$h_{\text{out}} = \mathcal{P}[\text{lb}\,(1+R^2 snr) < \rho] \ , \tag{3}$$

which can be rewritten in the form

$$\begin{aligned} h_{\text{out}} &= \mathcal{P}[R^2 < (2^{\rho}-1)/snr] \\ &= F_{R^2}((2^{\rho}-1)/snr)) \ , \end{aligned} \tag{4}$$

where the cdf $F_{R^2}$ is inaccurately known. Notice that the calculation of $h_{\text{out}}$ can be reduced to a special case of Eq.(1), with $Z$=$R^2$ and the function $G$ being the indicator of the interval

$$\mathcal{I} \triangleq (0, \sqrt{(2^{\rho}-1)/snr}) \ . \tag{5}$$

With a nonergodic channel, $h_{\text{out}}$ is the information-theoretical rate limit which cannot be exceeded by the word error probability of any coding scheme, and hence can be utilized for estimating the error probability of coded systems with powerful codes. The comparison of the two parameters $h$ and $h_{out}$ yields an indication of how coding can be beneficial for transmission over a given channel.

## 3 Bounds on distribution functions

We consider first the situation where sharp upper and lower bounds to the cdf of a RV, denoted $\overline{F}$ and $\underline{F}$, respectively, can be obtained on the basis of the incomplete knowledge available about the RV itself. A key point here is that the width of the gap between the bounds yields a quantitative indication of the effects of the model uncertainty on the distribution of the RV and on the performance parameters derived from it. Thus, a wide gap would reflect, rather than a weakness of the theory, a large amount of epistemic uncertainty.

As an example of how these bounds can be used, consider again the calculation of Eq.(1) in terms of the cdf $F_Z(z)$. If $Z$ is a continuous RV with range $\mathbf{R}^+$ and $G(\infty)=0$ and $g(z)\triangleq G'(z)\leqslant 0$, we can rewrite Eq.(1). Integration by parts yields

$$H=-\int_0^{\infty}F_Z(z)g(z)\mathrm{d}z \quad , \tag{6}$$

where this time we assume that the cdf $F_Z(z)$ is inaccurately known. If upper and lower bounds to $F_Z(z)$ are known, then we obtain upper and lower bounds to $H$ as follows. With obvious notations:

$$\overline{h}=-\int_0^{\infty}\overline{F}_Z(z)g(z)\mathrm{d}z \quad ,$$

$$\underline{h}=-\int_0^{\infty}\underline{F}_Z(z)g(z)\mathrm{d}z \quad , \tag{7}$$

With $G(z)$ as in Eq.(2), Eq.(7) hold with

$$g(z)=-\sqrt{\frac{snr}{4\pi z}}\mathrm{e}^{-snr\, z} \quad . \tag{8}$$

### 3.1 Using parameter intervals

A simple case of cdf bounds occurs when a model cdf can be assessed with reasonable accuracy, except for an uncertainty about the exact values of its parameters, which are known only within an interval (see, e.g., Ref.[10] for the presentation of an exceedingly general class of parametrizable distributions to be used as fading models). In this case bounds can be generated by determining the envelope of all cdfs whose parameters lie in that interval. As an example, Fig.1 shows the p-box generated by Nakagami cdfs[11] whose parameters are $m\in(0.6, 2.0)$, $\Omega\in(1.2, 1.8)$.

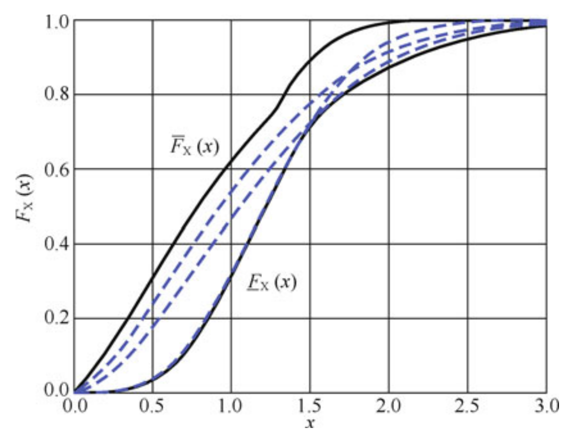


Figure 1 Upper and lower bounds $\overline{F}_X$, $\underline{F}_X$ generated by Nakagami-$m$ cdfs with parameters $m\in(0.6, 2.0)$, $\Omega\in(1.2, 1.8)$. The dashed curves correspond to three Nakagami-$m$ cdfs with randomly selected values of $m$ and $\Omega$ within their intervals.

Once bounds are derived, they can be aggregated in several ways, for example using binary operations in the set $\{+, -, \times, \div\}$. Consider first the case of two independent RVs whose upper and lower bounds to cdfs are given, and assume again that the RVs involved in the calculations take their values on $\mathbf{R}^+$. We denote by $\oplus$, $\ominus$, $\otimes$, and $\oslash$, respectively, the "generalized convolutions" that combine $F_X$ and $F_Y$ to generate $F_{X\circ Y}$, so that if for example $Z=X\times Y$ with $X$ and $Y$ independent we have $F_{X\times Y}(z)=(F_X\otimes F_Y)(z)=\int_0^z F_X(z/t)\mathrm{d}F_Y(t)$. In these conditions it can be easily proved[7] that the upper and lower bounds for the cdfs, denoted $\overline{F}$ and $\underline{F}$, respectively, are given by

$$\begin{aligned}
&\underline{F}_{X+Y}=\underline{F}_X\oplus\underline{F}_Y,\quad \overline{F}_{X+Y}=\overline{F}_X\oplus\overline{F}_Y,\\
&\underline{F}_{X-Y}=\underline{F}_X\ominus\overline{F}_Y,\quad \overline{F}_{X-Y}=\overline{F}_X\ominus\underline{F}_Y,\\
&\underline{F}_{X\times Y}=\underline{F}_X\otimes\underline{F}_Y,\quad \overline{F}_{X\times Y}=\overline{F}_X\otimes\overline{F}_Y,\\
&\underline{F}_{X\div Y}=\underline{F}_X\oslash\overline{F}_Y,\quad \overline{F}_{X\div Y}=\overline{F}_X\oslash\underline{F}_Y.
\end{aligned}\tag{9}$$

### 3.2 Moment bounds

Here we assume that the inaccurately modeled RV $Z$ is known through a (possibly small) number of its moments, i.e., of expected values of known functions of $Z$. Formally, we look for upper and lower bounds to $H$ in Eq.(1) under the constraint that the values of some moments of $Z$ are exactly known, and the range $\mathcal{Z}$ of $Z$ is known. It can be expected that the more moments are known, the tighter the interval within which the exact value of $H$ is confined.

In this case, geometric moment-bound theory (see, e.g., Refs.[12-15] and references therein) allows one to obtain sharp upper and lower bounds to the values of $H$. To keep our treatment simple, we examine here only the two-dimensional case of moment bounds. Thus, let $k_1(z)$ and $k_2(z)$ be two continuous functions defined over a finite $\mathcal{Z}$. The moment space of $Z$, denoted $\mathcal{M}$, is defined as the (closed, bounded, and convex) set of the pairs

$$\left(\int_{\mathcal{Z}}k_1(z)\mathrm{d}H(z),\int_{\mathcal{Z}}k_2(z)\mathrm{d}H(z)\right),\tag{10}$$

as $H(\cdot)$ runs over all cdfs defined over $\mathcal{Z}$. The main result we need is the following[13,15]: $\mathcal{M}$ is the convex hull of the curve $\mathcal{C}\triangleq\{(k_1(z),k_2(z))\,|\,z\in\mathcal{Z}\}$ in the two-dimensional Euclidean space $\mathbf{R}^2$. For example, by choosing $k_1(z)=z^2$ and $k_2(z)=G(z)$, the expected value of $G(z)$ can be identified with the second coordinate of $\mathcal{M}$. If the first coordinate is chosen as the known value of $\mathcal{E}Z^2$, then upper and lower bounds to $\mathcal{E}G(X)$ are obtained by direct evaluation of the upper and lower envelopes of $\mathcal{M}$ (see next section for an application). Moment bound theory yields bounds that are sharp, i.e., such that there exists a pair of RVs $Z_1$, $Z_2$ that have moments $\mathcal{E}[k_i(Z_j)]$, $i,j=1,2$, and whose cdfs yield exactly these upper and lower bounds. (It is also possible to derive bounds for a RV whose moment values are not known only known exactly, but only within intervals[16].)

If the parameter of interest is to be evaluated through the cdf $F_Z(z)$, as in Eq.(6), then one can obtain upper and lower bounds to it (see Ref.[17] for the relevant theory and bibliographic references). As an example, assume that the mean $\mu_1$ and the variance $\sigma_X^2\triangleq\mathcal{E}[X-\mu_1]^2$ are known for a RV $X$ taking values in the finite interval $[a,b]$ (we must have $\sigma_X^2<(b-\mu_1)(\mu_1-a)$ for consistency). Then we have, for $x\leqslant\overline{x}_1$,

$$0\leqslant F_X(x)\leqslant\frac{\sigma_X^2}{(\mu_1-x)^2+\sigma_X^2},\tag{11}$$

While for $\overline{x}_1\leqslant x\leqslant\underline{x}_2$,

$$\frac{(x-\mu_1)(b-\mu_1)+\sigma_X^2}{(x-a)(b-a)}\leqslant F_X(x)\tag{12}$$

and

$$F_X(x)\leqslant 1-\frac{(\mu_1-x)(\mu_1-a)+\sigma_X^2}{(b-x)(x-a)}.\tag{13}$$

Finally, for $\underline{x}_2\leqslant x\leqslant b$,

$$\frac{(x-\mu_1)^2}{(x-\mu_1)^2+\sigma_X^2}\leqslant F_X(x)\leqslant 1,\tag{14}$$

where

$$\begin{aligned}
&\underline{x}_1=a,\qquad \overline{x}_1=\mu_1-\frac{\sigma_X^2}{b-\mu_1},\\
&\underline{x}_2=\mu_1+\frac{\sigma_X^2}{\mu_1-a},\quad \overline{x}_2=b,
\end{aligned}\tag{15}$$

Fig.2 shows upper and lower moment bounds to the cdf of a RV defined in [0, 1] with expected value $m_1$=0.2 and variance $\sigma_X^2$=0.125.

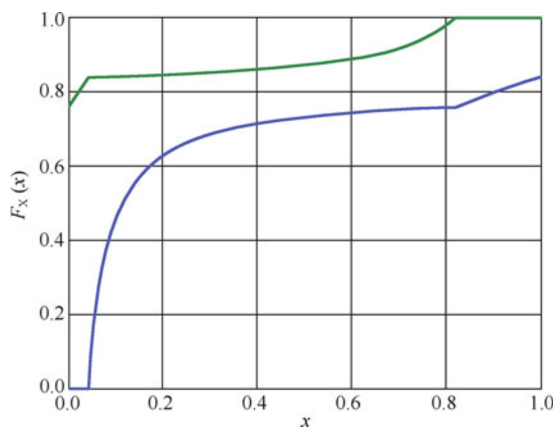


Figure 2 Upper and lower moment bounds to the cdf of a RV $X$, defined in [0,1], that is known only through its expected value $\mu_1$=0.2 and variance $\sigma_X^2$=0.125

## 4 Application to spectrum sensing

Moment bound theory can be applied to spectrum sensing in cognitive radio[18,19]. The situation here is one in which the presence of a primary signal in an observed signal should be detected, when the observation is affected by an interference $i(t)$ that cannot be modeled exactly. We assume in particular that only the amplitude range and maximum power of $i(t)$ are known. Moment bound theory allows us to characterize all probability distributions of the interference satisfying the known constraints, and derive those yielding maximum and minimum values for the probability of false alarm $h_{\text{FA}}$ and of missed detection $h_{\text{MD}}$.

### 4.1 Coherent sensor

Assume first that the primary signal is deterministic and known (corresponding to a known preamble). In order to assess the presence of the primary-user signal, the secondary-user sensor compares against a threshold $\theta$ the maximum-likelihood statistics

$$Y \triangleq \frac{1}{N}\sum_{n=0}^{N-1}\Re\{y_n x_n^*\} = \frac{1}{N}\sum_{n=0}^{N-1}\Re\{\varepsilon\,|x_n|^2 + i_n x_n^* + z_n x_n^*\}\,, \tag{16}$$

where $N$ denotes the number of samples observed, $y_n$ denote the samples of the observed signal, $z_n$ the samples of complex Gaussian additive noise, $i_n$ the samples of the interfering signal, and $\varepsilon$ takes value 1 if the primary signal $x_n$ is present, and 0 otherwise. The false-alarm and missed-detection probabilities have the form

$$P_{\text{FA}}(I) = \mathcal{P}(Y > \theta \mid \varepsilon = 0) = \mathcal{E}_I\left[Q\left(\sqrt{N}\,\frac{\theta - 1}{\sqrt{P\sigma_z^2}}\right)\right], \tag{17}$$

$$P_{\text{MD}}(I) = \mathcal{P}(Y < \theta \mid \varepsilon = 1) = \mathcal{E}_I\left[Q\left(\sqrt{N}\,\frac{P - \theta + I}{\sqrt{P\sigma_z^2}}\right)\right], \tag{18}$$

where $\mathcal{E}_I$ denotes expectation wrt $I$, $I$ is the interference term

$$I \triangleq \frac{1}{N}\sum_{n=0}^{N-1}\Re\{i_n x_n^*\} \tag{19}$$

and $P$ is the average signal power

$$P \triangleq \frac{1}{N}\sum_{n=0}^{N-1}|x_n|^2\,. \tag{20}$$

We finally assume for simplicity that $N$=1, $P$=1, $I\in[-1, 1]$, and that $I$ has an even probability density function (this implies in particular that its mean value is zero). Fig.3 shows the evolution of the convex hull of $P_{\text{FA}}$ with the noise power (here $\theta$=0.5). The convex hull contains all possible values of $P_{\text{FA}}$ for a given value of $\sigma_I^2 = \mathcal{E}[I^2]$.

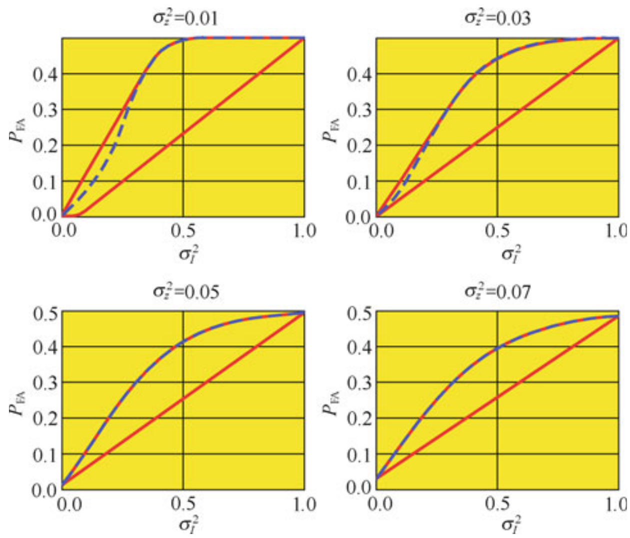


Figure 3 Coherent sensing and its probability of false alarm $P_{FA}$. Evolution of $P_{FA}$ uncertainty with noise power for $\theta$=0.5

Similar results hold for $P_{MD}$ (see Ref.[19] for further details).

### 4.2 Energy sensor

This is the simplest spectrum sensor. The presence of a primary signal is detected by comparing the measured energy against a suitable threshold, which must depend on the noise floor. The decision metric $Y$ is built again from a sequence of $N$ received signal samples

$$Y=\frac{1}{N}\sum_{n=0}^{N-1}|y_n|^2 . \tag{21}$$

We assume, as commonly done, that the noise samples are independent, circularly symmetric zero-mean complex Gaussian random variables with variance $\mathcal{E}[|z_n|^2]=\sigma_z^2$, i.e., $z_n\sim\mathcal{N}(0,\sigma_z^2)$. The simplest decision strategy consists of comparing $Y$ against a suitably optimized threshold $\theta$. Since the decision threshold depends on the observed-signal model, it becomes vulnerable to modeling inaccuracies.

We assume a deterministic primary signal with average power $P$. In this situation, the probability of a false alarm conditioned on the interference is

$$P_{FA}(Z,I)=\mathcal{E}_{Z,I}\mathcal{P}\left(\frac{1}{N}\sum_{n=0}^{N-1}|i_n+z_n|^2>\theta\right)$$
$$=\mathcal{E}_{Z,I}\mathcal{P}\left(\frac{2}{\sigma_z^2}\sum_{n=0}^{N-1}|i_n+z_n|^2>\frac{2N}{\sigma_z^2}\theta\right) .$$

The random variable appearing before the inequality sign in last equation has a noncentral chi-square distribution with $2N$ degrees of freedom and noncentrality parameter

$$\lambda_0\triangleq\frac{2}{\sigma_z^2}\sum_{n=0}^{N-1}|i_n|^2 . \tag{22}$$

Therefore,

$$P_{FA}(I)=\mathcal{E}_I Q_N(\sqrt{\lambda_0},\sqrt{2N\theta/\sigma_z^2}) , \tag{23}$$

where $Q_N$ denotes the generalized Marcum Q-function[11]. The conditional probability of missed detection can be derived in a similar way, to yield

$$P_{MD}(I)=1-\mathcal{E}_I Q_N(\sqrt{\lambda_1},\sqrt{2N\theta/\sigma_z^2}) , \tag{24}$$

where now

$$\lambda_1\triangleq\frac{2}{\sigma_z^2}\sum_{n=0}^{N-1}|x_n+i_n|^2 . \tag{25}$$

Sharp upper and lower bounds to $P_{FA}$ and $P_{MD}$ can be obtained under the assumption that $\mathcal{E}[i_n]$=0, which allows us to obtain the moments

$$\mathcal{E}[\lambda_0]=2N\frac{\sigma_I^2}{\sigma_z^2} \text{ and } \mathcal{E}[\lambda_1]=2N\frac{P+\sigma_I^2}{\sigma_z^2} . \tag{26}$$

Thus, the moment space relevant to our problem is obtained as the convex hull of the curve $(x,Q(\sqrt{x},\sqrt{2N\theta/\sigma_z^2}))$ (for $P_{FA}$) and $(x,1-Q(\sqrt{x},\sqrt{2N\theta/\sigma_z^2}))$ (for $P_{MD}$). See Refs.[18,19] for further details).

### 4.3 Linear–quadratic sensor

Coherent sensing uses a matched-filter detector in the presence of full knowledge of the signal that may be transmitted by the primary user, while energy sensing assume that that knowledge is missing. An intermediate situation occurs when the primary signal is imperfectly known. In this case, as advocated in Ref.[20], a linear-quadratic detector may be used. The signal observed by the spectrum sensor during a sensing interval of duration $N$ has the vector form

$$\boldsymbol{y}=\varepsilon\boldsymbol{x}+\boldsymbol{z} \quad , \tag{27}$$

where $\boldsymbol{x}$ is the primary-user signal, $\boldsymbol{z}$ the noise, and $\varepsilon$ takes on value 1 if a primary signal is included in the observation, and 0 otherwise. The vectors in Eq.(27) have $N$ real components. By indicating with the notation $\boldsymbol{g}\sim\mathcal{N}(\boldsymbol{m}, \boldsymbol{R})$ the fact that the random vector $\boldsymbol{g}$ has a Gaussian probability density function with mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{R}$, a standard assumption for Eq.(27) is $\boldsymbol{z}\sim\mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_z)$ (where the covariance $\boldsymbol{R}_z$ is assumed to have full rank).

Decision is made again by comparing the statistic $Y$, a suitable function of the observed signal, against a threshold $\theta$. Assuming that $\boldsymbol{x}$ is incompletely known, we use a LQ (Linear-Quadratic) detector which has the matched-filter and the energy one as special cases. Its performance approaches that of the linear detector when the uncertainty on the primary signal is small, and that of the quadratic detector in the opposite case. A simple model for the uncertainty assumes that $\boldsymbol{x}$ is the sum of a perfectly known signal $\boldsymbol{s}$ and a disturbance $\boldsymbol{i}$ whose probability distribution is only known within an "uncertainty set," which includes distributions whose first moments are known, so that moment bound theory can be used[20].

To choose the detector parameters under the assumed uncertainty of the model (which does not allow the "natural" choice of using for $Y$ the likelihood ratio) a generalized signal-to-noise ratio is used, called deflection. Thus, moment bound theory is used to scrutinize the implications of the model mismatch and to evaluate the robustness of the LQ statistics to signal-model variations. To do this, while accepting that $\boldsymbol{x} = \boldsymbol{s} + \boldsymbol{i}$, with s a known signal, we assume that a limited knowledge of the distribution of $\boldsymbol{i}$ is available, for example in the form of its range and variance (we also assume that it has mean zero), and study how the detector performs as that distribution varies in the uncertainty set dened by those constraints. Under these conditions, after observing that the probability of false alarm does not depend on $\boldsymbol{i}$, we may write

$$P_D=\mathcal{E}_I[P_D(\boldsymbol{i})] \quad , \tag{28}$$

where $I$ denotes the actual distribution of $\boldsymbol{i}$, $\mathcal{E}_I$ expectation with respect to $I$, and $P_D(\boldsymbol{i})$ the detection probability conditioned on $\boldsymbol{i}$. The extent of variation of $P_D$ as $I$ runs in the uncertainty set tells us how robust the detector is (see Ref.[20] for further details).

## 5 Finding the worst distribution within a set

Assume again that Eq.(1) must be computed, while the exact pdf of $Z$ is unknown and approximated by the pdf $f_0$. It is assumed that $f_0$ is "reasonably close" to the exact density, where the measure of closeness is chosen to be the K-L (Kullback-Leibler) divergence[21]. The solution of an optimization problem allows one to determine the worst distribution having a given K-L divergence from the nominal distribution, and assess the system performance when the former is used in lieu of the latter.

The mathematical problem of evaluating a performance metric vs. the K-L divergence between the nominal and the worst distribution is described and solved in Ref.[22] as the convex optimization problem

$$(\pi) \qquad \max_f \int G(z)f(z)\mathrm{d}z \quad ,$$
$$\text{s.t.} \quad \int \log\frac{f(z)}{\tilde{f}(z)}f(z)\mathrm{d}z \leqslant \delta \quad ,$$
$$\int f(z)\mathrm{d}z = 1 \quad .$$

(Condition $f(z) \geqslant 0$ should be added unless automatically satisfied by the solution of ($\pi$))

The optimizing $f(z)$ is

$$f^\star(x)=\frac{\mathrm{e}^{\nu^\star G(x)}f_0(x)}{\xi(\nu^\star)} \quad , \tag{29}$$

where

$$\xi(x) \triangleq \int \mathrm{e}^{\nu G(x)} f_0(x)\,\mathrm{d}x \ , \tag{30}$$

$\nu^\star$ is the solution of

$$\nu \frac{\xi'(\nu)}{\xi(\nu)} - \log \xi(\nu) = \delta \tag{31}$$

and

$$\xi'(\nu) \triangleq \frac{\mathrm{d}\xi(\nu)}{\mathrm{d}\nu} = \int G(x)\mathrm{e}^{\nu G(x)} f_0(x)\mathrm{d}x \ . \tag{32}$$

The resulting maximum value of $\mathcal{E}_Z[G(Z)]$, denoted $h_{\max}$, is given by

$$h_{\max} = \frac{\xi'(\nu^\star)}{\xi(\nu^\star)} \ , \tag{33}$$

that is, $h_{\max}$ is the slope of the logarithmic derivative of $\xi(\nu)$ at $\nu = \nu^\star$.

## 5.1 Applications

The general expression for the error probability of uncoded binary antipodal modulation with equally likely signals having a common signal-to-noise ratio equal to *snr*, under the assumption of ergodic Rayleigh fading with amplitude $R$, additive white Gaussian noise, and perfect channel state information at the receiver, is

$$h_0 = \frac{1}{2}\left[1 - \sqrt{\frac{snr}{1+snr}}\right] , \tag{34}$$

Fig.4 depicts the behavior of $h_0$ and $h_{\max}$ vs. SNR for two values of $\delta$. It is seen that for large values of SNR the curve slope (the "diversity") becomes logarithmic, showing that the performance loss is mainly due to the model uncertainty, while for small SNR it is approximately dictated only by the SNR.

### 5.1.1 Outage probability

As for outage probability, in the special case of Rayleigh-distributed fading we obtain

$$h_{\text{out},0} = 1 - \exp[-(2^\rho - 1)/snr] \ .$$

An indication of the robustness of the coding/modulation choice, one may evaluate the ratio between the performance metric value corresponding to the worst probability distribution and the one corresponding to the nominal distribution. The higher this ratio for a given value of $\delta$, the lower the robustness of the design. Tab.1 shows the ratios $\eta_{\text{uncod.}} \triangleq h_{\max}/h_0$ and $\eta_{\text{cod.}} \triangleq h_{\text{out, max}}/h_{\text{out, 0}}$ for different signal-to-noise ratios and values of the K-L divergence $\delta$. The first ratio indicates the robustness of uncoded transmission, while the second indicates that of a system using a near-optimal code. It is seen that for small $\delta$ addition of error-control coding does not add much to robustness, which on the contrary is increased when $\delta$ is large. Robustness is lower for large values of signal-to-noise ratio, indicating that in that regime the performance is dictated by model uncertainty rather than by noise (see Ref.[22] for further details about this approach).

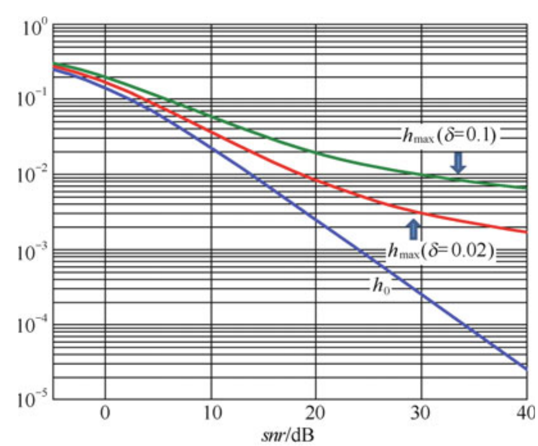


Figure 4 Error probability of uncoded binary antipodal modulation with equally likely signals having a common signal-to-noise ratio equal to *snr*. $h_0$ is the error probability under the assumption of ergodic Rayleigh fadingwith amplitude $R$. $h_{\max}$ is the worst error probability with model uncertainty $\delta$

Table 1 Values of $\eta_{\text{uncod.}} \triangleq h_{\max}/h_0$ and $\eta_{\text{cod.}} \triangleq h_{\text{out, max}}/h_{\text{out, 0}}$ for different signal-to-noise ratios and values of the K-L divergence $\delta$

| $snr$/dB | $\delta = 0.1$ | | $\delta = 5.0$ | |
|---|---|---|---|---|
| | $\eta_{\text{uncod.}}$ | $\eta_{\text{cod.}}$ | $\eta_{\text{uncod.}}$ | $\eta_{\text{cod.}}$ |
| 0 | 1.38 | 1.32 | 3.26 | 1.58 |
| 10 | 2.52 | 2.62 | 17.0 | 10.5 |
| 20 | 7.8 | 8.0 | 132 | 100 |

## 6 Dependence bounds

Here we examine a situation in which the model uncertainty arises from a RV $Z$ that is a combination of other RVs that are not independent, with their dependence being unknown or only partially known. It may happen that independence is assumed for ease of treatment only, in which case it is important to assess the possible performance penalty caused by a wrong independence assumption. In this case, bounds to cdfs are obtained by bounding the copulas connecting marginal cdfs to their joint cdf. To study how this can be done, we first summarize a few facts about copulas, then we describe the resulting dependence bounds resulting from copula theory. This approach allows one to determine the width of the performance range caused by possibly unwarranted independence assumptions.

A copula[23,24] is a function that links the marginal cdfs of $d$ random variables $X_i$, $i=1,\cdots,d$, to their joint cdf. Using copulas, the joint cdf is identified through two separate entities, one describing the marginal cdfs and the other describing the dependence structure. For simplicity's sake, we consider two-dimensional copulas first.

Specifically, if $F_{XY}$ denotes a two-dimensional cdf with marginals $F_X$, $F_Y$, then a function $K$, called a copula, exists such that

$$F_{XY}(x, y)=K(F_X(x), F_Y(y)) \quad . \tag{35}$$

Hence, the copula $K$ contains all the information about the dependence of $X$ and $Y$.

The dual $K^{\partial}$ of copula $K$ is defined as

$$K^{\partial}(a, b) \triangleq a+b-K(a, b) \quad . \tag{36}$$

Given its definition, much can be obtained in the study of copulas by examining different dependences occurring among RVs uniform in [0, 1], which we denote writing $U \sim \mathcal{U}(0, 1)$.

**Comonotonicity:** The copula

$$K(u_1, u_2)=\min(u_1, u_2) \tag{37}$$

is always attained if $U_2=T(U_1)$, where $T$ is an a.s. monotonic increasing transformation. RVs of this type are called comonotonic.

**Independence:** Independence occurs with

$$K(u_1, u_2) = u_1 u_2 \quad . \tag{38}$$

**Countermonotonicity:** The copula

$$K(u_1, u_2) = \max(u_1+u_2-1,0) \tag{39}$$

refers to RV with perfect negative dependence: $U_2=T(U_1)$, with $T$ strictly decreasing. RVs of this type are called countermonotonic.

The following general inequalities can be proved[23]. For every copula $K$ and all $\{x, y\}\in[0, 1]\times[0, 1]$:

$$W(x, y) \leqslant K(x, y) \leqslant M(x, y) \quad , \tag{40}$$

where both $W$ and $M$ are copulas, defined as

$$W(x, y) \triangleq \max(x+y-1,0) \quad , \tag{41}$$

$$M(x, y) \triangleq \min(x, y) \quad . \tag{42}$$

If $T_X$ and $T_Y$ are strictly increasing a.s. on ran $X$ and $Y$, then[25]

$$K_{T_X(X),\, T_Y(Y)}(x, y) = K_{XY}(x, y) \quad . \tag{43}$$

This equality shows how copulas capture the properties of a joint cdf which are invariant under a.s. strictly increasing transformations. If $T_X$ and $T_Y$ are strictly decreasing a.s. on ran $X$ and ran $Y$, respectively, then[25]

$$K_{T_X(X),\, Y}(x, y) = y - K_{XY}(1-x, y) \quad , \tag{44}$$

$$K_{X,\, T_Y(Y)}(x, y) = x - K_{XY}(x, 1-y) \quad , \tag{45}$$

$$K_{T_X(X),\, T_Y(Y)}(x, y) = x+y-1+K_{XY}(1-x, 1-y) \quad . \tag{46}$$

Combining Eq.(35) with Eq.(40)—Eq.(42) we obtain the Fréchet–Hoeffding bounds on a joint cdf in terms of its marginals[7,23,26]:

$$\begin{aligned} F_{XY}(x,y) &\geqslant \max[F_X(x)+F_Y(y)-1,0] \ , \\ F_{XY}(x,y) &\leqslant \min[F_X(x),F_Y(y)] \ . \end{aligned} \tag{47}$$

The upper bound is achieved when $Y$ is a.s. an increasing function of $X$, while the lower bound is achieved when $Y$ is a.s. a decreasing function of $X$.

The two-dimensional bounds Eq.(40) can be generalized to $d$-dimensional RVs:

$$\max\left(\sum_{i=1}^{d} u_i + 1 - d, 0\right) \leqslant K(u) \leqslant \min(u_1,\cdots,u_d) \ , \tag{48}$$

while the $d$-dimensional Fréchet–Hoeffding bounds Eq.(47) are

$$\begin{aligned} &\max\left(\sum_{i=1}^{d} F_{X_i}(x_i) + 1 - d, 0\right) \\ &\leqslant F_{X_1,\cdots,X_d}(X_1,\cdots,X_d) \\ &\leqslant \min(F_{X_1},\cdots,F_{X_d}) \ . \end{aligned} \tag{49}$$

The upper Fréchet-Hoeffding bound in Eq.(49) is a copula, while the lower bound is not a copula for $d$>2. Thus, while the upper bound in Eq.(49) is achieved by comonotonic RVs, the lower bound is achieved in general only for $d$=2.

## 6.1 Operations on RVs

We examine first a RV $Z$ obtained as a composition of two RVs $X$ and $Y$. The following key result holds[26]: Let $X$, $Y$ denote two RVs defined on the extended real line $\mathbf{R}^* \triangleq \mathbf{R} \cup \{-\infty,\infty\}$, and $\mathcal{L}$ the set of binary operations mapping $\mathbf{R}^*\times\mathbf{R}^*$ to $\mathbf{R}^*$ to which are nondecreasing in each place and continuous except possibly at (0, ∞) and (∞, 0). If $Z \triangleq X \circ Y$, where $\circ \in \mathcal{L}$ and $\underline{K}_{XY}$ is any lower bound on copula $K_{XY}$, then two functions $\mathrm{ldb}_{\underline{K}_{XY}}$ (the “lower dependence bound”) and $\mathrm{udb}_{\underline{K}_{XY}}$ (the “upper dependence bound”) exist such that, $\forall_{z} \in \mathbf{R}^*$,

$$\begin{aligned} F_Z(z) &\geqslant \mathrm{ldb}_{\underline{K}_{XY}}(F_X,F_Y,\circ)(z) \ , \\ F_Z(z) &\leqslant \mathrm{udb}_{\underline{K}_{XY}}(F_X,F_Y,\circ)(z) \ , \end{aligned} \tag{50}$$

where[7]

$$\begin{aligned} \mathrm{ldb}_{\underline{K}_{XY}}(F_X,F_Y,\circ)(z) &\triangleq \sup_{x\circ y=z} \underline{K}_{XY}(F_X(x),F_Y(y)) \ , \\ \mathrm{udb}_{\underline{K}_{XY}}(F_X,F_Y,\circ)(z) &\triangleq \inf_{x\circ y=z} \underline{K}^{\partial}_{XY}(F_X(x),F_Y(y)) \ . \end{aligned} \tag{51}$$

For example, the special case of a sum of RVs yields the following result, valid on the real line $\mathbf{R}$[7]:

$$\begin{aligned} F_{X+Y}(z) &\geqslant \sup_{x+y=z} \max[F_X(x)+F_Y(y)-1,0] \ , \\ F_{X+Y}(z) &\leqslant \inf_{x+y=z} \min[F_X(x)+F_Y(y),1] \ . \end{aligned} \tag{52}$$

This bound is sharp, i.e., cannot be further improved[7].

The results above can be generalized to operations involving more than two RVs by exploiting their associativity[27]. For example, explicit equations for the sum of RVs are[28]

$$\begin{aligned} F(z) &\geqslant \sup_{x_1+\cdots+x_d=z} \max\left(\sum_{i=1}^{d} F_i(x_i) - (d-1), 0\right) \ , \\ F(z) &\leqslant \inf_{x_1+\cdots+x_d=z} \min\left(\sum_{i=1}^{d} F_i(x_i), 1\right) \ . \end{aligned} \tag{53}$$

## 6.2 Order statistics

We examine first the case of two RVs $X_1$, $X_2$ with joint cdf $F_{X_1X_2}(x_1, x_2)$ and their order statistics $X_{\min}$ and $X_{\max}$[29]. Since we have, with obvious notations,

$$F_{\max}(x) = F_{X_1X_2}(x,x) \ , \tag{54}$$

using the Fréchet-Hoeffding bound Eq.(47) with $x$=$y$ we obtain

$$\max(F_{X_1}+F_{X_2}-1,0) \leqslant F_{\max} \leqslant \min(F_{X_1},F_{X_2}) \ . \tag{55}$$

Result Eq.(55) can be generalized to the extremes of $d$ RVs, and even to more general order statistics[29].

## 6.3 Example 1

Consider a block fading channel with $d$ blocks[9]. Using independent Gaussian symbols on the $d$ blocks, the outage probability is given by

$$p_{\text{out}}(\rho) = \mathcal{P}\left(\frac{1}{d}\sum_{i=1}^{d} C(R_i) \leqslant \rho\right) \ , \tag{56}$$

where $\rho$ is the average transmission rate, and $C(R_i) \triangleq \mathrm{lb}(1+R_i^2 snr)$ is the instantaneous mutual information of the block with fading amplitude $R_i$ and signal-to-noise ratio $snr$. Eq.(56) can be given the form

$$p_{\mathrm{out}}(\rho)=F_Z(\rho) \ , \qquad (57)$$

where

$$Z \triangleq \sum_{i=1}^{d} X_i \qquad (58)$$

and

$$X_i \triangleq \frac{1}{d} C(R_i) \ . \qquad (59)$$

Under the assumption that the RVs $R_i$ have a common Rayleigh distribution, we obtain, for the cdf of $X_i$,

$$F(x_i) = 1 - \exp\left(-\frac{2^{dx_i}-1}{snr}\right). \qquad (60)$$

To examine the effect on $p_{\mathrm{out}}$ of the lack of independence of the fading across blocks, we determine the upper and lower dependence bounds of $F_Z$, viz.,

$$\overline{F}_Z(z) = \inf_{x_1+\cdots+x_d=z} \min\left[\sum_{i=1}^{d} F(x_i), 1\right] \ ,$$

$$\underline{F}_Z(z) = \sup_{x_1+\cdots+x_d=z} \max\left[\sum_{i=1}^{d} F(x_i) - d + 1, 0\right] \ , \qquad (61)$$

After some algebra, we obtain

$$\overline{F}_Z(z) = F(z) \ ,$$

$$\underline{F}_Z(z) = \max[dF(z/d) - d + 1, 0] \ . \qquad (62)$$

Numerical results are shown in Fig.5.

### 6.4 Example 2

Consider a wireless transmission system with diversity $d$, Rayleigh fading, and selection combining[9]. Usual analyses assume independence of the diversity branches, which may be violated in practical systems. The dependence bounds of Subsection 6.2 can be applied here, with $F_{X_i}(x) = 1 - \mathrm{e}^{-x}, i = 1, \cdots, d$, the cdf of the square fading amplitude. We obtain, for $x \geqslant 0$,

$$\max(0, 1 - d\mathrm{e}^{-x}) \leqslant F_{\max}(x) \leqslant 1 - \mathrm{e}^{-x} \ , \qquad (63)$$

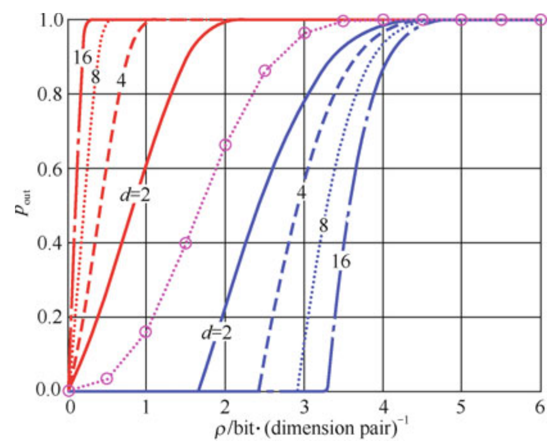


Figure 5 Dependence upper and lower bounds for the outage probability of a block fading channel with Rayleigh fading, signal-to-noise ratio $snr$=5 dB, and $d$ blocks, $d$=2, 4, 8, and 16. The middle curve shows $p_{\mathrm{out}}$ for $d$=2 and independent fading[9]. The value at which the lower bound equals zero is $\rho_0=\mathrm{lb}(1+snr \log d)$

with the upper bound corresponding to the same fading amplitude in each branch, and hence no diversity. The corresponding error probability for binary antipodal transmission are obtained from Eq.(6) and Eq.(8), which yield, after some algebra,

$$P(e) = \frac{1}{2} \sum_{k=0}^{d} \binom{d}{k} (-1)^k \sqrt{\frac{snr}{k+snr}} \ , \qquad (64)$$

while the upper bound is

$$\overline{P}(e) = \frac{1}{2}\left(\sqrt{\frac{snr}{1+snr}}\right) \ , \qquad (65)$$

and the lower bound

$$\underline{P}(e) = \frac{1}{2}\left[1 - d\sqrt{\frac{snr}{1+snr}} - \mathrm{erf}(\sqrt{snr \log d}) + d\sqrt{\frac{snr}{1+snr}} \mathrm{erf}(\sqrt{(1+snr)\log d})\right] \ . \qquad (66)$$

What is especially relevant in this example is that the upper dependence bound carries essentially no new information, as it only tells us that the worst dependence among diversity branches corresponds to the sheer absence of diversity. Thus, to obtain useful dependence bounds one should add some information about what is known about the dependence. This aspect is covered in Ref.[30], where two ways of gathering and using partial dependence information to tighten the dependence bounds are discussed and compared for pairs $X$, $Y$ of RVs: One consists of using a measure of dependence like Kendall's *tau* or Spearmans' *rho*[23,31,32], while the other one is based on assuming positive quadrant dependence, i.e., assuming that the probability that $X$ and $Y$ be simultaneously small (large) is at least as great as it would be were they independent[23]. Yet another approach was taken in Ref.[33], where a parametric family of copulas (the Clayton copulas) was chosen to model the dependence of the RVs of interest.

## 7 Conclusions

We have examined how interval-type bounds on pdfs or cdfs can make one able to handle problems arising when a wireless communication system performance must be assessed in the presence of model uncertainties. Several techniques are described, and a few applications discussed.

## References

[1] BIGLIERI E, GROSSI E, LOPS M. Random-set theory and wireless communications[J]. Foundations and trends in communications and information theory, 2012, 7(4): 317-462.

[2] MAHLER R P S. Statistical multisource-multitarget information fusion[M]. Norwood: Artech House, 2007.

[3] ZADEH L A. Fuzzy sets[J]. Information and control, 1965, 8(65), 338-353.

[4] SHAFER G. A mathematical theory of evidence[M]. Princeton: Princeton University Press, 1976.

[5] BIGLIERI E. Dealing with uncertain models in wireless communications[C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Shanghai, China, c2016: SP-COM-P7.4.

[6] REGAN H M, FERSON S, BERLEANT D. Equivalence of methods for uncertainty propagation of real-valued random variables[J]. International journal of approximate reasoning, 2004, 36(1): 1-30.

[7] WILLIAMSON R C. Probabilistic arithmetic[D]. Brisbane: University of Queensland, 1989.

[8] FERSON S, KREINOVICH V, GINZBURG L, et al. Constructing probability boxes and dempster-shafer structures[R]. Albuquerque: Sandia National Laboratory Report SAND 2002-4015, 2003.

[9] BIGLIERI E. Coding for wireless channels[M]. New York: Springer, 2005.

[10] BIGLIERI E, YAO K, YANG C-A. Fading models from spherically invariant processes[J]. IEEE transactions on wireless communications, 2015, 14(10): 5526-5538.

[11] SIMON M K, ALOUINI M-S. Digital communication over fading channels[M]. 2nd ed. Hoboken: John Wiley & Sons, 2005.

[12] DRESHER M. Moment spaces and inequalities[J]. Duke mathematical journal, 1953, 20(2): 261-271.

[13] KEMPERMAN J H B. Moments in mathematics: geometry of the moment problem[C]//Proceedings of Symposia in Applied Mathematics, 37, American Mathematical Society, San Antonio, USA, c1987: 16-53.

[14] KREĬN M G, NUDEL'MAN A A. The Markov moment problem and extremal problems[M]. Providence: American Mathematical Society, 1977.

[15] YAO K, TOBIN R M. Moment space upper and lower error bounds for digital systems with intersymbol interference[J]. IEEE trans. inform. theory, 1976, IT-22(1): 65-74.

[16] BIGLIERI E. Probability of error for digital systems with inaccurately known interference[J]. IEEE trans. inform. theory, 1984, 30(2): 443-446.

[17] BIGLIERI E, TARICCO G. Bounds on the distribution of a random variable whose first moments are known[J]. Archiv für Elektronik und Übertragungstechnik, 1987, 41(6): 330-336.

[18] BIGLIERI E. Spectrum sensing under uncertain channel modeling[J]. Journal of communications and networks, 2012, 14(3): 225-229.

[19] BIGLIERI E, GOLDSMITH A J, GREENSTEIN L J, et al. Principles of cognitive radio[M]. Cambridge: Cambridge University Press, 2013.

[20] BIGLIERI E, LOPS M, Linear-quadratic detectors for spectrum sensing[J]. Journal of communications and networks, 2014, 16(5): 485-492.

[21] COVER T M, THOMAS J A. Elements of information theory[M]. New York: J. Wiley & Sons, 1991.

[22] BIGLIERI E, ALRAJEH N. The robustness of coding and modulation for body-area networks[J]. Journal of communications and networks, 2014, 16(3): 264-269.

[23] NELSEN R B. An introduction to copulas[M]. 2nd ed. New York: Springer, 2006.

[24] SCHMIDT T. Copulas—from theory to applications in finance: coping with copulas[M]. London: Risk Books, 2007.

[25] ALSINA C, FRANK M J, SCHWEIZER B. Associative functions: triangular norms and copulas[M]. Singapore: World Scientific Publ. Co., 2006.

[26] WILLIAMSON R C, DOWNS T. Probabilistic arithmetic. I.

Numerical methods for calculating convolutions and dependency bounds[J]. International journal of approximate reasoning, 1990, 4(2): 89-158.
[27] FRANK M J, NELSEN R B, SCHWEIZER B. Best-possible bounds for the distribution of a sum—A problem of Kolmogorov[J]. Probability theory and related fields, 1987, 74(2): 199-211.
[28] DENUIT M, GENEST C, MARCEAU É. Stochastic bounds on sums of dependent risks[J]. Insurance: mathematics and economics, 1999, 25, 85-104.
[29] RYCHLIK T. Distributions with fixed marginals and related topics: order statistics of variables with given marginal distributions[B]. IMS Lecture Notes, Monograph Series, 1996, 28: 297-306.
[30] BIGLIERI E, LAI I-W. The impact of independence assumptions on wireless communication analysis[C]//Proceedings of IEEE International Symposium on Information Theory, Barcelona, spain, c2016.
[31] NELSEN R B, QUESADA-MOLINA J J, RODRÍGUEZ-LALLENA J A, et al. Bounds on bivariate distribution functions with given margins and measures of association[J]. Communications in statistics—theory and methods, 2001, 30(6): 1155-1162.
[32] NELSEN R B, ÚBEDA-FLORES M. A comparison of bounds on sets of joint distribution functions derived from various mesures of association[J]. Communications in statistics—theory and methods, 2004, 33(10): 2299-2305.
[33] RITCEY J A, Copula models for wireless fading and their impact on wireless diversity combining[C]//Proceedings of the Forty-First Asilomar Conference on Signals, Systems and Computers, Pacific Grove, USA, c2007.

## About the author

**BIGLIERI Ezio** was born in Aosta, Italy. He received the Dr.Ing. degree in electronic engineering from Politecnico di Torino, Italy. He is now an Adjunct Professor with Universitat Pompeu Fabra, Barcelona, Spain, and the EE Department of UCLA. His research interests include coding and modulation for wireless channels. (Email: e.biglieri@ieee.org)