Linear–Quadratic Detectors for Spectrum Sensing

Ezio Biglieri and Marco Lops

Abstract: **Spectrum sensing for cognitive-radio applications may use a matched-filter detector (in the presence of full knowledge of the signal that may be transmitted by the primary user) or an energy detector (when that knowledge is missing). An intermediate situation occurs when the primary signal is imperfectly known, in which case we advocate the use of a linear–quadratic detector. We show how this detector can be designed by maximizing its deflection, and, using moment-bound theory, we examine its robustness to the variations of the actual probability distribution of the inaccurately known primary signal.**

Index Terms: **Cognitive radio, linear-quadratic (LQ) detectors, spectrum sensing.**

I. INTRODUCTION AND MOTIVATION OF THE WORK

Spectrum sensing, one of the major functions of interweaved cognitive radio [4], detects and classifies *spectrum holes*, i.e., regions of the spectrum space that can be opportunistically used by secondary users. The signal observed by the spectrum sensor has the vector form

$$y = \epsilon x + n \qquad (1)$$

where x is the primary-user signal, n the noise, and ϵ takes on value 1 if a primary signal is included in the observation, and 0 otherwise. The vectors in (1) have N real components, corresponding to discrete signal samples (in our context, N is also called *sensing time*). By indicating with the notation $g \sim \mathcal{N}(m, R)$ the fact that the random vector g has a Gaussian probability density function with mean m and covariance matrix R, a standard assumption for (1) is $n \sim \mathcal{N}(0, R_n)$ (where the covariance R_n is assumed to have full rank).

A spectrum sensor decides between the two hypotheses $\mathcal{H}_0 : \epsilon = 0$ and $\mathcal{H}_1 : \epsilon = 1$. Decision is made by comparing the statistic Y, a suitable function of the observed signal, against a threshold θ. The two probabilities of interest here are the false-alarm and detection probabilities, whose definitions are

$$P_{\text{FA}} \triangleq \mathbb{P}\{Y > \theta \mid \mathcal{H}_0\}, \qquad (2)$$

$$P_{\text{D}} \triangleq \mathbb{P}\{Y > \theta \mid \mathcal{H}_1\}. \qquad (3)$$

The choice of the statistic Y depends on how much information about x and n is available to the detector, and on the tolerable complexity of the calculations entailed in the decision process. Two common choices of Y refer to situations in which x

has a structure which is perfectly known (coherent, or matched-filter linear detector) or totally unknown (energy quadratic detector) [4]. The situation we examine in this paper is intermediate between those two: Here, assuming that x is incompletely known, we use a detector which has the matched-filter and the energy one as special cases: the *linear–quadratic* (LQ) detector. 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. To illustrate our findings, we use a simple model for the uncertainty, assuming that x is the sum of a perfectly known signal s and a disturbance i whose probability distribution is only known within an "uncertainty set," which includes distributions whose first moments are known, reflecting a common approach to partial statistical modeling through moments. (This model can be viewed as a variation on the theme of the one proposed in [23], in which availability of side information on the minimum primary-signal strength is also assumed. See also [21].) A way of describing the philosophy underlying our approach is by observing the difference between *decision making under risk*, which occurs when a perfect statistical model of the observation is available and *decision making under ignorance*, which occurs when there is uncertainty on the model to be used.

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) we maximize a generalized signal-to-noise ratio (SNR), called *deflection*. Next, we examine how this detector performs under several scenarios, and discuss its robustness to the variation of the actual probability distribution of the unknown signal. The results described in this paper are related to the "robust decision design" problem (e.g., [17] and references therein), which we approach by assuming a specific structure for the receiver and examining its robustness to probability distributions differing from the Gaussian one. In fact, when the distribution of signal and/or noise is unknown, likelihood ratio cannot be used for optimum detection. If this is the case, it seems a reasonable solution to assume a fixed form for the detection statistic and optimize its performance. The rationale behind this choice will be discussed in the following.

Robustness is evaluated here by deriving sharp upper and lower bounds to the performance of the detector as the probability distribution of i ranges through the uncertainty set. Among other information, robustness analysis provides the designer with a tool yielding the conditions under which unsatisfactory performance is due to model uncertainty rather than to noise effects (other examples of this situation are examined, in different contexts, in [10] and [20]).

II. LINEAR AND LINEAR–QUADRATIC SENSING

A general expansion of the functional which maps the observations $y_1, \cdots, y_N$ to the statistic Y can be obtained in the form

of a Volterra expansion [3] and [16]:

$$Y = S(y_1, y_2, \cdots, y_N)$$
$$= w^{(0)} + \sum_i w_i^{(1)} y_i + \sum_i \sum_j w_{ij}^{(2)} y_i y_j$$
$$+ \sum_i \sum_j \sum_k w_{ijk}^{(3)} y_i y_j y_k + \cdots. \qquad (4)$$

One may choose the "Volterra coefficients" $w^{(0)}$, $w_i^{(1)}$, etc., which maximize a suitable cost function, as the *deflection* to be described below. The use of a Volterra expansion typically allows one to express the optimum statistic in terms of the moments of the random variables $y_1, \cdots, y_N$. In particular, one may choose the order of the Volterra expansion (4) so that only the known moments of $y_1, \cdots, y_N$ are involved. Now, it seldom occurs that many moments are available with sufficient accuracy. Thus, a reasonable approach to this situation of uncertainty would be to *assume* a given structure for the moments (typically, a Gaussian structure in which all moments can be computed from first- and second-order moments), and evaluate the performance when the observed samples are actually not Gaussian. This is the approach we take in the following, where for simplicity we focus our attention on a second-order expansion.

Coherent sensing is the most natural spectrum sensing technique under the assumption that the primary signal x is perfectly known to the detector. The corresponding decision statistic Y is built as a linear function of the observed vector y:

$$Y_{\mathrm{L}} = w^T R_n^{-1} y. \qquad (5)$$

Assume instead that no prior knowledge of x is available. In this case hypothesis testing becomes a composite problem, and a computable decision statistic can be obtained through the generalized likelihood-ratio test (GLRT), viz.,

$$Y_{\mathrm{GLRT}} = \max_{x \in \mathbb{R}^N} \frac{f_{y|\mathbb{H}_1, x}(y \mid \mathbb{H}_1, x)}{f_{y|\mathbb{H}_0}(y \mid \mathbb{H}_0)}$$
$$= \max_{x \in \mathbb{R}^N} \frac{f_n(y - x)}{f_n(y)}$$

with f denoting probability density functions. This leads to the quadratic test statistic

$$Y_{\mathrm{Q}} = y^T R_n^{-1} y. \qquad (6)$$

In particular, when the primary signal structure is totally unknown and n is white, one may use as Y the measure of the energy contents of the observed signal, which yields

$$Y_{\mathrm{Q}} = \|y\|^2. \qquad (7)$$

The more general statistic we advocate here, where only partial knowledge of x is assumed, encompasses (5) and (7) as special cases, and consists of a LQ function Y_{LQ} of y. This may be thought of as obtained by truncating to the second-order term the Volterra-series expansion of a generic nonlinear decision functional of y.[1] It has the form

$$Y_{\mathrm{LQ}} = c + w^T y + y^T W y \qquad (8)$$

[1] Gardner [8] discusses the concept of *structurally constrained receivers*, which are based on a combination of a simple ad hoc procedure with an optimization procedure, and yield the best performance for some classes of problems.

where c is a constant (whose value is irrelevant to the detector performance), and the superscript ^{T} denotes transposition. The "optimum" values w_o and W_o of the two relevant parameters in (8), the vector w and the matrix W, may be chosen as those maximizing a generalized SNR known as *deflection*[2]

$$D \triangleq \frac{[\mathbb{E}(y \mid \mathcal{H}_1) - \mathbb{E}(y \mid \mathcal{H}_0)]^2}{\mathbb{V}(y \mid \mathcal{H}_0)} \qquad (9)$$

where $\mathbb{E}$ denotes expectation, and $\mathbb{V}$ denotes variance.

The solution to this problem is illustrated in [15] and, for the complex case, in [5]. It requires knowledge of the fourth-order statistics of the random variables involved, which we do not assume to be available for x (more generally, one could use a truncated version of a Volterra series including terms beyond second-order, which would need exact knowledge of higher-order moments for the optimization of its parameters [16].) We search instead for a solution assuming only knowledge of second-order statistics. This can be obtained by assuming a Gaussian distribution for x, or, more generally, a "Gaussian-like" distribution, characterized by zero third-order moments and a relation between second-order and fourth-order moments typical of Gaussian distributions (see [15] for details. Reference [8] shows an example of a Gaussian-like, but non-Gaussian, probability density function. Observe also that having null thrid-order moments uncouples the equations yielding the optimum linear and quadratic parts, so that the optimum LQ detector is obtained by using independently calculated optimal linear and quadratic systems [15]). The resulting solution has a closed form. Specifically, assuming that

$$x \sim \mathcal{N}(s, R_i) \qquad (10)$$

which corresponds to having $x = s + i$, with s a known deterministic signal and i a Gaussian disturbance, we have [15] and [19]

$$w_o = R_n^{-1} s, \qquad (11)$$
$$W_o = R_n^{-1} R_i R_n^{-1} \qquad (12)$$

and hence

$$Y_{\mathrm{LQ}} = s^T R_n^{-1} y + y^T R_n^{-1} R_i R_n^{-1} y. \qquad (13)$$

Denoting by 0 and O the null vector and the null matrix, respectively, we can see from (11)–(13) that
(a) w_o corresponds to the whitened matched filter.
(b) If $i = 0$ (corresponding to a deterministic primary signal), then $R_i = O$, and hence the optimum statistic Y is linear.
(c) If $s = 0$ (corresponding to a zero-mean Gaussian primary signal), then the optimum statistic is quadratic, which yields the energy-detector statistic Y_{Q} when n and i are white.

[2] Also called "detection index" [6] or "output SNR" [11]. The most sensible optimization criterion would be to optimize the receiver operating characteristics, which appears to be a formidable task. The rationale behind the choice of the deflection as a cost function for the optimization of the LQ detector offering both tractability and pratical utility is discussed in [14] (see also [1]). Other possible second-order cost functions related to deflection are categorized in [9] and [19]. Baker [2] derives relations between the optimum deflection criterion test statistics and the log-likelihood ratio. In [1], the deflection is optimized for a purely quadratic statistic, viz., $Y_{\mathrm{Q}} = y^T W y$, with results consistent with those presented here.

(d) The "full" LQ statistic is optimum only if $s \neq 0$ and $R_i \neq O$.

Defining the vector $\mathbf{u} \triangleq R_n^{-1} y$, completing the square in (13), and removing an irrelevant additive constant, we may write the LQ statistic in the new form[3]

$$Y_{\text{LQ}} = \left\| \frac{1}{2} R_i^{-1/2} s + R_i^{1/2} R_n^{-1} y \right\|^2. \tag{14}$$

In the special case of white n and i, viz., $R_n = \sigma_n^2 i$ and $R_i = \sigma_i^2 i$ with $\sigma_i^2 > 0$, the LQ statistic assumes the exceedingly simple form

$$Y_{\text{LQ}} = \|\gamma^2 s + y\|^2 \tag{15}$$

where

$$\gamma^2 \triangleq \sigma_n^2 / 2\sigma_i^2. \tag{16}$$

This parameter quantifies in a way the amount of uncertainty on the distribution of i by comparing the variance of the noise against that of i: Large values of γ^2 indicate small uncertainty, and hence suggest the use of a coherent detector, while a small γ^2 would naturally lead to the energy detector, as immediately reflected by the structure of (13). We may also observe the expression of the resulting maximum deflection, which yields

$$D_{\max} = \frac{\|x\|^2}{\sigma_n^2} + \frac{\sigma_i^2}{\sigma_n^2} \frac{\|x\|^2}{\sigma_n^2} \tag{17}$$

$$= \left(1 + \frac{1}{2\gamma^2}\right) \frac{\|x\|^2}{\sigma_n^2} \tag{18}$$

and shows the two separate contributions of the linear and quadratic part of the detector.

Since the derivation of (11) and (12) was made under the assumption of Gaussian-like distribution for i, which might not be valid in practice, the LQ-detector must be scrutinized to examine its behavior with distributions differing from the one assumed. Thus, after examining its "optimum" behavior, we shall proceed to derive upper and lower bounds to its performance when the distribution of i ranges in an uncertainty set, as defined by the partial knowledge of the distribution itself.

III. PERFORMANCE OF LQ DETECTOR WITH GAUSSIAN-LIKE DISTRIBUTION

From now on, and purely for simplicity's sake, we pursue our analysis referring only to the case corresponding to (15). The performance of the decision statistic is now evaluated by computing

$$P_{\text{FA-LQ}} = \mathbb{P} \left[\|\gamma^2 s + n\|^2 > \theta \right] \tag{19}$$

$$= Q_{N/2} \left(\sqrt{\lambda_0}, \sqrt{\theta/\sigma_n^2} \right), \tag{20}$$

[3]With a notational abuse that should not lead to ambiguities, we denote by the same symbol two equivalent statistics, i.e., statistics leading to the same detector performance—for example, obtained by adding or multiplying by a constant term.

$$P_{\text{D-LQ}} = \mathbb{P} \left[\|(1 + \gamma^2)s + i + n\|^2 > \theta \right] \tag{21}$$

$$= Q_{N/2} \left(\sqrt{\lambda_1}, \sqrt{\theta/(\sigma_i^2 + \sigma_n^2)} \right) \tag{22}$$

where $Q_.(\cdot, \cdot)$ denotes the generalized Marcum Q-function

$$Q_m(a, b) \triangleq \frac{1}{a^{m-1}} \int_b^\infty x^m \exp\left(-\frac{x^2 + a^2}{2}\right) I_{m-1}(ax)\,dx \tag{23}$$

($I_\nu(\cdot)$ is the modified Bessel function of the first kind and order ν) and

$$\lambda_0 \triangleq \frac{\gamma^4}{\sigma_n^2} \|s\|^2, \tag{24a}$$

$$\lambda_1 \triangleq \frac{(1 + \gamma^2)^2}{\sigma_i^2 + \sigma_n^2} \|s\|^2. \tag{24b}$$

We compare the performance of the LQ detector to that of the linear[4] detector, which has

$$P_{\text{FA-L}} = \mathbb{P} \left[s^T n > \theta \right] = Q\left(\frac{\theta}{\sqrt{\|s\|^2 \sigma_n^2}} \right), \tag{25}$$

$$P_{\text{D-L}} = \mathbb{P} \left[s^T(s + i + n) > \theta \right] = Q\left(\frac{\theta - \|s\|^2}{\sqrt{\|s\|^2(\sigma_i^2 + \sigma_n^2)}} \right) \tag{26}$$

where $Q(\cdot)$ denotes the Gaussian tail function.

Figs. 1 and 2 illustrate the improvement over the linear detector obtained by using a LQ statistic with a Gaussian i. The calculations leading to these figures assumed a primary signal $s = 1$, where 1 denotes the all-1 N-vector. It is seen that the improvement obtained depends on the value of γ^2: A small γ^2, corresponding to a relatively high energy in the partially known component i, makes the statistic (15) close to Y_Q, which justifies the improvement on the linear detector. Conversely, a large value of γ^2, corresponding to a small amount of uncertainty in modeling y, makes the improvement introduced by the quadratic term marginal, as expected.

A. Sample Complexity

We now examine, following [20], how the sensing time N depends on the SNR and on γ^2 for a given performance level. Assuming that P_{FA} and P_{D} can be given the form

$$P_{\text{FA}} = Q\left(\sqrt{N} \frac{\theta - A}{\sqrt{B}} \right), \tag{27}$$

$$P_{\text{D}} = Q\left(\sqrt{N} \frac{\theta - C}{\sqrt{D}} \right) \tag{28}$$

and eliminating θ from (27) and (28), we obtain

$$N = (C - A)^{-2} \left[\sqrt{B} Q^{-1}(P_{\text{FA}}) - \sqrt{D} Q^{-1}(P_{\text{D}}) \right]^2. \tag{29}$$

With the LQ detector, assume a Gaussian i and $s = s1$, so that $\text{SNR} \triangleq s^2/\sigma_n^2$. We observe that

$$\mathbb{E} \|\gamma^2 s + n\|^2 = N\sigma_n^2(\gamma^4 \text{SNR} + 1), \tag{30}$$

$$\mathbb{V} \|\gamma^2 s + n\|^2 = N\sigma_n^4(4\gamma^4 \text{SNR} + 2) \tag{31}$$

[4]A comparison with the quadratic detector would be unfair, since its use does not assume any information on the primary signal structure.
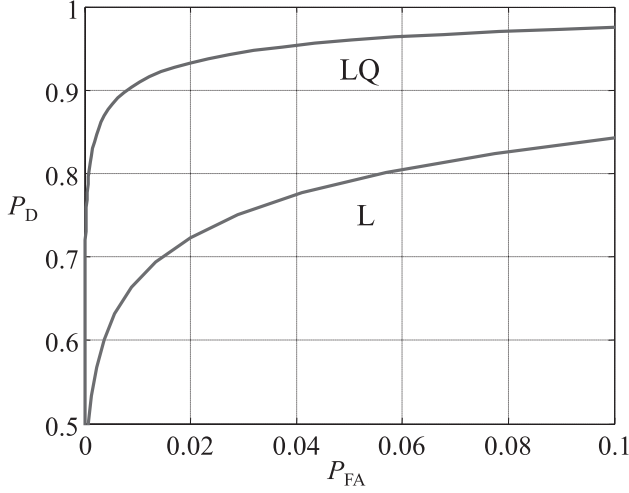
Fig. 1. Receiver operating characteristics of linear (L) and LQ statistic with Gaussian i, $N = 10$, $s = 1$, $\sigma_n^2 = 1$, and $\sigma_i^2 = 2.5$ (and hence $\gamma^2 = 0.2$).

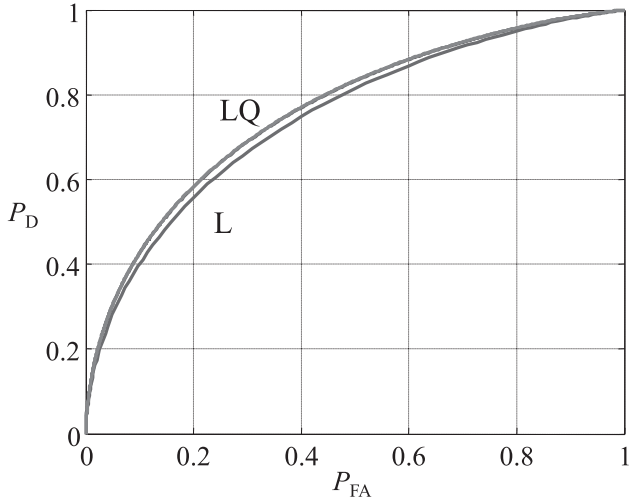Fig. 2. Receiver operating characteristics of linear (L) and LQ statistic with with Gaussian i, $N = 10$, $s = 1$, $\sigma_n^2 = 10$, and $\sigma_i^2 = 2.5$ (and hence $\gamma^2 = 2$).

and

$$\mathbb{E}\left\|(1+\gamma)^2 s + i + n\right\|^2$$
$$= N\sigma_n^2[(1+\gamma^2)^2\text{SNR} + 1 + 1/2\gamma^2], \quad (32)$$

$$\mathbb{V}\left\|(1+\gamma)^2 s + i + n\right\|^2$$
$$= N\sigma_n^4[4(1+\gamma^2)^2(1+1/2\gamma^2)\text{SNR} + 2(1+1/4\gamma^4)]. \quad (33)$$

Assuming that N is large enough to justify a Gaussian approximation for $\|\gamma^2 s + n\|^2$ and $\|(1+\gamma)^2 s + i + n\|^2$, we may approximate $P_{\text{FA-LQ}}$ and $P_{\text{D-LQ}}$ in the form (27), (28), where

$$A = \gamma^4\text{SNR} + 1, \tag{34a}$$
$$B = 4\gamma^4\text{SNR} + 2, \tag{34b}$$
$$C = (1+\gamma^2)^2\text{SNR} + 1 + 1/2\gamma^2, \tag{34c}$$
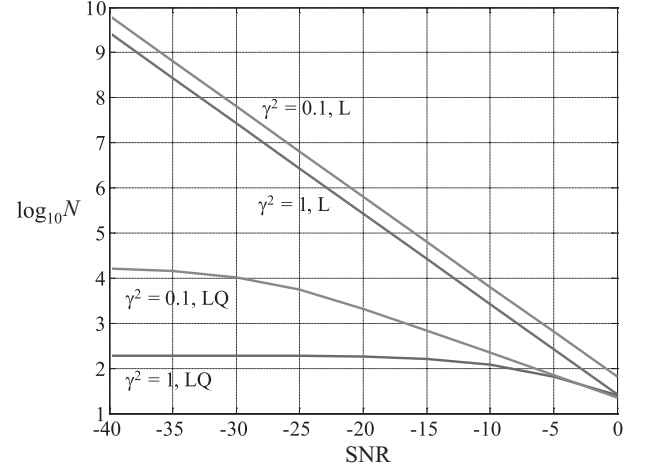$$D = 4(1+\gamma^2)^2(1+1/2\gamma^2)\text{SNR} + 2(1+1/4\gamma^4) \tag{34d}$$

Fig. 3. Sample size N vs. the $\text{SNR}(\triangleq s^2/\sigma_n^2)$ for the linear and LQ detector with $s = s1$, two different values of γ^2, and $P_{\text{FA}} = 1 - P_{\text{D}} = 0.01$.

and hence

$$N \approx$$
$$\left[(1+2\gamma^2)\text{SNR} + 1/2\gamma^2\right]^{-2} \left[\sqrt{4\gamma^2\text{SNR} + 2}\, Q^{-1}(P_{\text{FA-LQ}})\right.$$
$$\left. - \sqrt{4(1+\gamma^2)(1+1/2\gamma^2)\text{SNR} + 2(1+1/4\gamma^4)}\, Q^{-1}(P_{\text{D-LQ}})\right]^2. \tag{35}$$

It is observed that the first factor in the RHS of the above equation is roughly constant for small SNR, while varies as SNR^{-2} for large values of SNR, while the second factor increases as SNR. Thus, for very low SNR the value of N remains about constant with SNR, while it decreases as SNR^{-1} for larger SNR (notice also that for relatively large values of SNR the accuracy of (35) is questionable, as the Gaussian approximation leading to it may not be valid).

The linear detector has

$$A = 0, \tag{36a}$$
$$B = s^2\sigma_n^2, \tag{36b}$$
$$C = s^2, \tag{36c}$$
$$D = s^2(\sigma_i^2 + \sigma_n^2) \tag{36d}$$

and hence

$$N = \text{SNR}^{-2}\left[Q^{-1}(P_{\text{FA-L}}) - \sqrt{1+1/2\gamma^2}\, Q^{-1}(P_{\text{D-L}})\right]^2, \tag{37}$$

so that, for a given target pair $P_{\text{FA-L}}$, $P_{\text{D-L}}$, N varies as SNR^{-2}.

Fig. 3 shows the behavior of N as a function of SNR for the linear and LQ detector.

B. Robustness of LQ Detector: A First Stab

The analysis carried on *supra* was assuming that, in addition to i being a Gaussian vector, the variance σ_i^2 of its components were known, so that the value of γ^2 parametrizing the detector was computed using the exact values of σ_n^2 and σ_i^2. A more realistic assumption is that the value of σ_i^2 is only approximately
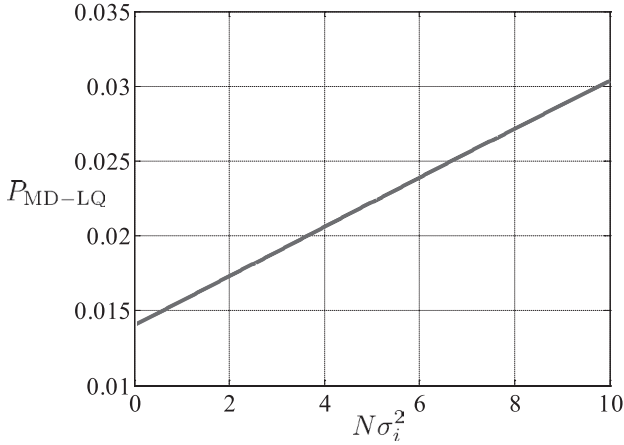
Fig. 4. Probability of misdetection $P_{\mathrm{MD-LQ}}$ for a LQ detector with $\boldsymbol{s} = \boldsymbol{1}$, $N = 20$, $\sigma_n^2 = 1$, and $\hat{\sigma}_i^2 = 1/N$ for different actual values of σ_i^2. The value of the threshold θ is chosen so that $P_{\mathrm{MD-LQ}} = P_{\mathrm{FA-LQ}} = 0.016$.

Fig. 5. Receiver operating characteristics of linear (L) and LQ statistic with Gaussian (LQ(G)) and uniformly distributed $\boldsymbol{i}$ [LQ(U)—obtained by computer simulation]. Here $N = 10$, $\boldsymbol{s} = \boldsymbol{1}$, $\sigma_n^2 = 1$, and $\sigma_i^2 = 2.5$ (and hence $\gamma^2 = 0.2$).

known through its estimate $\hat{\sigma}_i^2$, so that the detector using an estimated value of γ^2 is likely to be mismatched. We now examine this situation by evaluating the effect on $P_{\mathrm{D-LQ}}$ of a mismatched σ_i^2 (the value of $P_{\mathrm{FA-LQ}}$ will not be affected by the mismatch. For convenience, here we use the probability of misdetection $P_{\mathrm{MD-LQ}} \triangleq 1 - P_{\mathrm{D-LQ}}$). For a fair analysis, we consider that adding $\boldsymbol{i}$ to $\boldsymbol{s}$ increases its power, so that the probability of misdetection would be improved by a larger uncertainty term σ_i^2 unless the observed signal power is kept constant. This is done by replacing for $\|\boldsymbol{s}\|^2$ in (24a) and (24b), under the assumption $\boldsymbol{s} = s\boldsymbol{i}$, the term $N\|s^2 - \sigma_i^2\|^2$, where σ_i^2 is the actual value of the variance of the components of $\boldsymbol{i}$, which may differ from the value $\hat{\sigma}_i^2$ estimated and used to determine γ^2. Fig. 4 illustrates the effect of such mismatch. In this figure, $s = 1$, $N = 20$, $\sigma_n^2 = 1$, and $\hat{\sigma}_i^2 = 1/N$, so that $\gamma^2 = 10$.

IV. INACCURATE MODEL OF I: DETECTOR ROBUSTNESS

The structure of Y_{LQ} was chosen in Section II under a Gaussian-like assumption on $\boldsymbol{x}$ and a given value of γ^2. Now, if the assumption on the statistics of $\boldsymbol{x}$ is not valid, the LQ detector is mismatched, and hence its performance may be degraded. One may examine this performance with a number of different models for $\boldsymbol{x}$: for example, Fig. 5 shows the receiver operating characteristics with uniformly distributed $\boldsymbol{i}$. We observe that in this situation the performance of the LQ detector is not degraded by the mismatch (actually, it is even improved, reflecting the fact that maximization of the deflection does not necessarily imply optimization of P_{D} and P_{FA}).

A more accurate scrutiny of the implications of the model mismatch leads to the evaluation of the robustness of the LQ statistics to signal-model variations. To do this, while still accepting that $\boldsymbol{x} = \boldsymbol{s} + \boldsymbol{i}$, with $\boldsymbol{s}$ 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 defined by those constraints. Under these conditions, after observing that the proba-
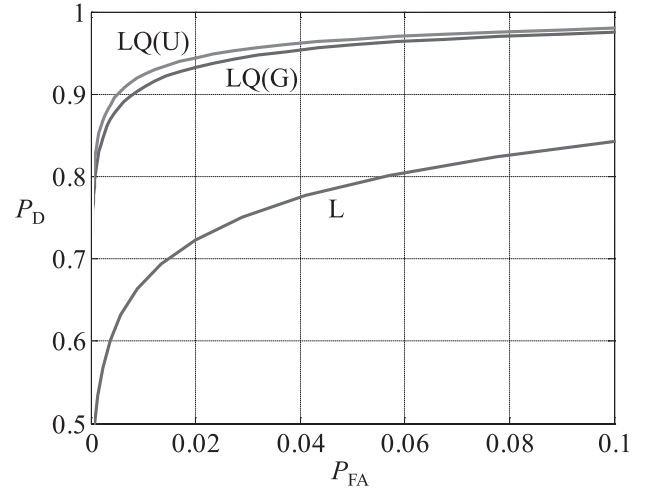
bility of false alarm does not depend on $\boldsymbol{i}$, we may write

$$P_{\mathrm{D}} = \mathbb{E}_{\mathrm{I}} P_{\mathrm{D}}(\boldsymbol{i}) \tag{38}$$

where I denotes the actual distribution of $\boldsymbol{i}$, $\mathbb{E}_{\mathrm{I}}$ expectation with respect to I, and $P_{\mathrm{D}}(\boldsymbol{i})$ 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.

A. Moment Bounds

Our study of the LQ detector robustness consists of finding sharp upper and lower bounds to (38) as I runs through all the possible distributions of $\boldsymbol{i}$ satisfying the set of constraints imposed by the physical aspects of the problem. Some of these constraints take the form of generalized moments of $\boldsymbol{i}$ (i.e., expected values of known functions of $\boldsymbol{i}$), while others may involve what is known about the structure of the distribution of I. Formally, the problem to be solved is

$$\sup_{\mathcal{I}} \ \mathbb{E}_{\mathrm{I}} P_{\mathrm{D}}(\boldsymbol{i}), \qquad \text{s.t.} \quad \mathbb{E}_{\mathrm{I}} \, \mathbf{k}(\boldsymbol{i}) = \boldsymbol{\mu}$$

along with its equivalent version with inf in lieu of sup. Here, $\mathcal{I}$ is the subset of all possible probability distributions satisfying the given constraints.

A simple, yet important special situation occurs when $\mathcal{I}$ is the unconstrained set of all possible distributions with finite support. In this case, if few generalized moments of $\boldsymbol{i}$ are exactly known, or even known within a certain interval, geometric moment-bound theory (e.g., [7], [12], [13], [25], and references therein) allows one to obtain sharp upper and lower bounds to the values of P_D. Here we assume the knowledge of range and variance of $\boldsymbol{i}$. The geometric moment-bound theory relevant here is summarized by the following fact. Let Z denote a random variable with range in the finite interval $\mathcal{Z}$ and unknown cumulative distribution function (CDF). Let $k_1(z)$ and $k_2(z)$ be two continuous functions defined over $\mathcal{Z}$. The *moment space* of Z, denoted $\mathcal{M}$,

is defined as the (closed, bounded, and convex) set of the pairs

$$\left(\int_{\mathcal{Z}} k_1(z)\, dG(z), \int_{\mathcal{Z}} k_2(z)\, dG(z) \right) \tag{39}$$

as $G(\cdot)$ runs over all CDFs defined over $\mathcal{Z}$. The main result we need is the following [12], [25]: $\mathcal{M}$ is the convex hull of the curve $\mathcal{C} \triangleq \{(k_1(z), k_2(z)) \mid z \in \mathcal{Z}\}$ in $\mathbb{R}^2$. Explicitly, by choosing $k_1(z) = z^2$ and $k_2(z) = f(z)$, the expected value of $f(Z)$ can be identified with the second coordinate of $\mathcal{M}$. If the first coordinate is chosen as the known value of $\mathbb{E}\, Z^2$, then upper and lower bounds to $\mathbb{E}\, f(X)$ are obtained by direct evaluation of the upper and lower envelopes of $\mathcal{M}$.

B. Robustness of Linear Detector

Consider the linear detector first. We have

$$P_{\mathrm{D-L}}(\boldsymbol{i}) = \mathbb{P}[\boldsymbol{s}^T\boldsymbol{n} > \theta - \|\boldsymbol{s}\|^2 - \boldsymbol{s}^T\boldsymbol{i}] \tag{40}$$

$$= Q\left(\frac{\theta - \|\boldsymbol{s}\|^2 - \mu(\boldsymbol{i})}{\sqrt{\|\boldsymbol{s}\|^2 \sigma_n^2}} \right) \tag{41}$$

where $\mu(\boldsymbol{i}) \triangleq \boldsymbol{s}^T\boldsymbol{i}$ is a random variable with range $(\mu_{\min}, \mu_{\max})$, mean zero, and variance $\|\boldsymbol{s}\|^2 \sigma_i^2$. Assuming again $\boldsymbol{s} = \boldsymbol{1}$ and the components of $\boldsymbol{i}$ confined in $(-a, a)$ (so that $\sigma_i^2 \in (0, a^2)$), and a symmetric probability density function for $\mu(\boldsymbol{i})$, the curve whose convex hull yields sharp upper and lower bounds to $P_{\mathrm{D-L}}$ is

$$\mathcal{C}_{\mathrm{D-L}} \triangleq \left(z^2, f(z) \mid z^2 \in (0, Na^2) \right) \tag{42}$$

where

$$f(z) \triangleq \frac{1}{2}\left[Q\left(\frac{\theta - N - z}{\sqrt{N\sigma_n^2}} \right) + Q\left(\frac{\theta - N + z}{\sqrt{N\sigma_n^2}} \right) \right]. \tag{43}$$

C. Robustness of LQ Detector

Consider next the LQ detector. From (21) we obtain

$$P_{\mathrm{D-LQ}}(\boldsymbol{i}) = \mathbb{P}[\|\mathbf{u}\|^2 > \theta/\sigma_n^2 \mid \boldsymbol{i}] \tag{44}$$

where $\mathbf{u} \triangleq \sigma_n^2 \|(1+\gamma^2)\boldsymbol{s} + \boldsymbol{i} + \boldsymbol{n}\|^2/\sigma_n^2$ has, conditionally on $\boldsymbol{i}$, a noncentral χ^2 distribution with noncentrality parameter

$$\lambda(\boldsymbol{i}) = \frac{1}{\sigma_n^2} \|(1+\gamma^2)\boldsymbol{s} + \boldsymbol{i}\|^2. \tag{45}$$

Consequently,

$$P_{\mathrm{D-LQ}}(\boldsymbol{i}) = Q_{N/2}(\sqrt{\lambda(\boldsymbol{i})}, \sqrt{\theta/\sigma_n^2}). \tag{46}$$

The random variable $\lambda(\boldsymbol{i})$ has, under the the assumption that $\boldsymbol{i}$ has mean zero, a known mean value

$$\mathbb{E}[\lambda(\boldsymbol{i})] = \frac{1}{\sigma_n^2}[(1+\gamma^2)^2\|\boldsymbol{s}\|^2 + N\sigma_i^2]. \tag{47}$$

The range of $\lambda(\boldsymbol{i})$ (which we assume to be finite) is denoted $(\lambda_{\min}, \lambda_{\max})$. For example, if $\boldsymbol{s} = \boldsymbol{1}$ and the components of $\boldsymbol{i}$ take values in $(-a, +a)$, with $a \geq 1 + \gamma^2$, we have from (45)

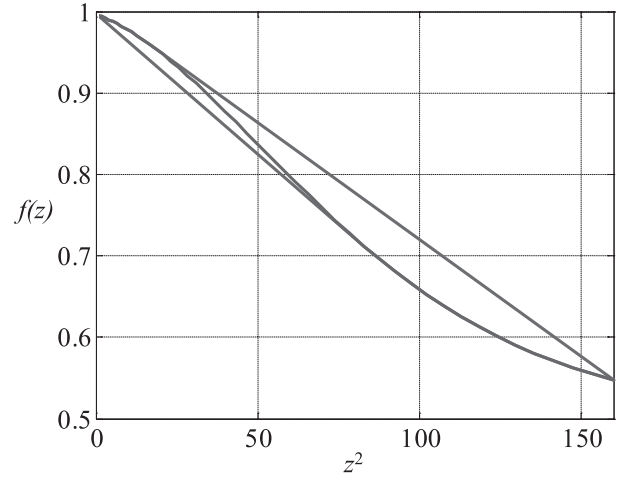$$\lambda_{\min} = 0, \qquad \lambda_{\max} = \frac{N}{\sigma_n^2}(1 + \gamma^2 + a)^2.$$

Fig. 6. Moment space of $P_{\mathrm{D-L}}$, $N = 10$, $\boldsymbol{s} = \boldsymbol{1}$, $\sigma_n^2 = 1$, $\sigma_i^2 = 4$, and $a = 4$. The curve $P_{\mathrm{D-L}}$ and the boundaries of its convex hull are shown.

Thus, the upper and lower bounds for $P_{\mathrm{D-LQ}}$ at $\sigma_i^2 \in (0, a^2)$ are given by the values of the upper and lower extremes of the convex hull of the curve

$$\mathcal{C}_{\mathrm{D-LQ}} \triangleq \left(z, Q_{N/2}\left(\sqrt{z}, \sqrt{\theta/\sigma_n^2} \right) \mid z \in (\lambda_{\min}, \lambda_{\max}) \right) \tag{48}$$

corresponding to the abscissa $\mathbb{E}[\lambda(\boldsymbol{i})] = N[(1+\gamma^2)^2 + \sigma_i^2]/\sigma_n^2$.

D. Comparisons

Fig. 6 shows the moment space of the detection probability with linear statistic. Increasing the sample size N, besides increasing $P_{\mathrm{D-L}}$, yields a narrower moment space. This observation may be used to determine the sample size, whose choice influences the performance of the spectrum sensor as well as its robustness. Robustness analysis provides a tool to decide when unacceptable detector performance is caused by low SNR, or rather by an insufficiently accurate model of system parameters.

Figs. 7 and 8 show two moment spaces of the detection probability with LQ statistic. As for the linear case, increasing the sample size N, besides increasing the detection probability, yields a narrower moment space. Comparing the moment spaces of linear and LQ detector, one may observe that, as discussed above, increasing the power of the interference increases the probability of detection of the LQ detector, while decreases the one of the linear detector.

E. A Generalization

We observe here that, while the bounds achieved are *sharp*, i.e., they correspond to distributions that are achievable under the constraints assigned and hence cannot be further tightened, the choice of a wide set $\mathcal{I}$ might provide loose, and hence pessimistic, bounds. For example, in the cases examined above the extremal distributions turn out to be discrete, which may not be a realistic model. Thus, one may want to rule out distributions being ill-fitted to the specific problem and hence making the moment bounds unreasonably loose. For computational purposes, it is convenient to restrict the underlying distribution to belong to a *convex* $\mathcal{I}$ [18], [22]. This is because a number
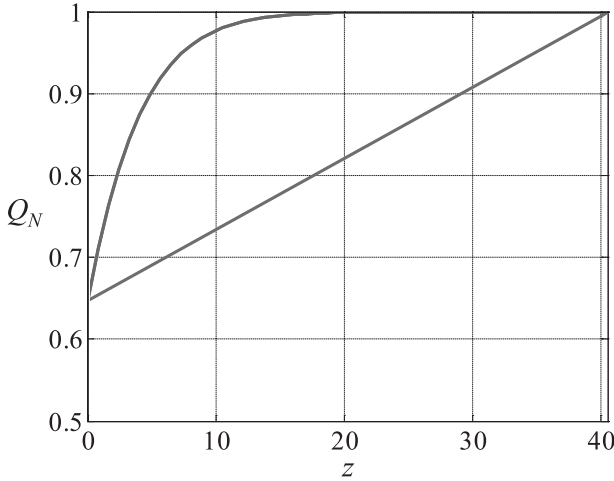
Fig. 7. Moment space of $P_{\mathrm{D-LQ}}$, $N = 8$, $\boldsymbol{s} = \mathbf{1}$, $\sigma_n^2 = 1$, $\sigma_i^2 = 4$, and $a = 1.125$. The curve $P_{\mathrm{D-LQ}}$ and the boundaries of its convex hull are shown.
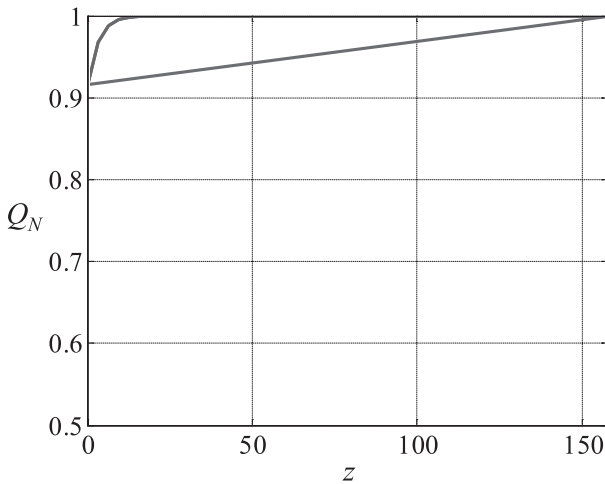

Fig. 8. Same as in Fig. 7, but with $N = 12$.

of constrained moment bounds characterized by a convex constraint set $\mathcal{I}$ can be solved as semidefinite programs. Since the intersection of two convex sets is also convex, these constraints can also be combined together. Unimodal, symmetric, and monotone convex distributions may be considered. In this framework, the authors of [24] present a semidefinite programming formulation of the problem of deriving bounds to the probability $\mathbb{P}[\boldsymbol{x}^T \mathbf{A}\boldsymbol{x} + 2\mathbf{b}^T\boldsymbol{x} + c < 0]$, where $\boldsymbol{x}$ is an n-dimensional real random vector with known first and second moments $\mathbb{E}\,\boldsymbol{x}$ and $\mathbb{E}\,\boldsymbol{x}\boldsymbol{x}^T$.

V. CONCLUSIONS

We have examined the robustness of a LQ detector for spectrum sensing when the primary signal can only be imperfectly modeled, and hence the detector may be mismatched. The detector was designed by maximizing a generalized SNR. The robustness of the detector to variations of the signal distribution was studied using geometric moment-bound theory.

REFERENCES

[1] C. R. Baker, "Optimum quadratic detection of a random vector in Gaussian noise," *IEEE Trans. Commun. Technol.*, vol. COM-14, no. 6, pp. 802–805, Dec. 1966.

[2] C. R. Baker, "On the deflection of a quadratic–linear test statistic," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 1, pp. 16–21, Jan. 1969.

[3] S. Benedetto and E. Biglieri, *Principles of Digital Transmission with Wireless Applications*. New York, NY: Kluwer, 1999.

[4] E. Biglieri *et al.*, *Principles of Cognitive Radio*. Cambridge, UK: Cambridge University Press, 2012.

[5] P. Chevalier and B. Picinbono, "Complex linear–quadratic systems for detection and array processing," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2631–2634, Oct. 1996.

[6] H. Cox, "Optimum arrays and the Schwartz inequality," *J. Acoust. Soc. Am.*, vol. 45, no. 1, pp. 228–232, 1969.

[7] M. Dresher, "Moment spaces and inequalities," *Duke Math. J.*, vol. 20, pp. 261–271, June 1953.

[8] W. A. Gardner, "Structurally constrained receivers for signal detection and estimation," *IEEE Trans. Commun.*, vol. COM-24, no. 6, pp. 578–592, June 1976.

[9] W. A. Gardner, "A unifying view of second-order measures of quality for signal classification," *IEEE Trans. Commun.*, vol. COM-28, no. 6, pp. 807–816, June 1980.

[10] I. Ioannou, C. D. Charalambous, and S. Loyka, "Outage probability under channel distribution uncertainty," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6825–6838, Nov. 2012.

[11] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*. New York: Springer-Verlag, 1988.

[12] J. H. B. Kemperman, "Geometry of the moment problem," in *Proc. Symposia in Applied Mathematics*, vol. 37, American Mathematical Society, 1987, pp. 16–53.

[13] M. G. Kreĭn and A. A. Nudel'man, *The Markov Moment Problem and Extremal Problems*. American Mathematical Society, 1977.

[14] B. Picinbono, "On deflection as a performance criterion in detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 3, pp. 1072–1081, July 1995.

[15] B. Picinbono and P. Duvaut, "Optimal linear-quadratic systems for detection and estimation," *IEEE Trans. Inf. Theory*, vol. 34, no. 2, pp. 304–311, Mar. 1988.

[16] B. Picinbono and P. Duvaut, "Geometrical properties of optimal Volterra filters for signal detection," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1061–1068, Sept. 1990.

[17] H. V. Poor, "Robust decision design using a distance criterion," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 5, pp. 575–587, Sept. 1980.

[18] I. Popescu, "A semidefinite programming approach to optimal moment bounds for convex classes of distributions," *Mathematics of Operations Research*, vol. 30, no. 3, pp. 632–657, Aug. 2005.

[19] J. D. Taft and N. K. Bose, "Quadratic–linear filters for signal detection," *IEEE Trans. Signal Process.*, vol. 39, no. 11, Nov. 1991.

[20] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 4–17, Feb. 2008.

[21] G. Taricco and E. Biglieri, "Space–time decoding with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1874–1888, July 2005.

[22] J. E. Smith, "Generalized Chebychev inequalities: Theory and applications in decision analysis," *Operations Research*, vol. 43, no. 5, pp. 807–825, Sept.–Oct. 1995.

[23] J. Unnikrishnan and V. V. Veeravalli, "Cooperative sensing for primary detection in cognitive radio," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 18–27, Feb. 2008.

[24] L. Vandenberghe, S. Boyd, and K. Comanor, "Generalized Chebyshev bounds via semidefinite programming," *SIAM Review*, vol. 49, no. 1, pp. 52–64, 2007.

[25] K. Yao and R. M. Tobin, "Moment space upper and lower error bounds for digital systems with intersymbol interference," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 1, pp. 65–74, Jan. 1976.

Ezio Biglieri was born in Aosta, Italy. He is now an Adjunct Professor with the Electrical Engineering Department of University of California at Los Angeles (UCLA), and with Universitat Pompeu Fabra, Barcelona, Spain. He was elected three times to the Board of Governors of the IEEE Information Theory Society, and in 1999 he was the President of the Society. He served as Editor in Chief for the IEEE Transactions on Information Theory, the IEEE Communications Letters, the European Transactions on Telecommunications, and the Journal of Communications and Networks. He is serving on the Scientific Board of the French company Sequans Communications, and, till 2012, he was a Member of the Scientific Council of the Groupe des Ecoles des Télécommunications (GET), France. Since 2011 he has been a member of the Scientific Advisory Board of CHIST-ERA (European Coordinated Research on Long-term Challenges in Information and Communication Sciences & Technologies ERA-Net). He is a Life Fellow of the IEEE. In 1992, he received the IEE Benefactors Premium from the Institution of Electrical Engineers (U.K.) for a paper on trellis-coded modulation. In 2000, he received, jointly with John Proakis and Shlomo Shamai, the IEEE Donald G. Fink Prize Paper Award and the IEEE Third-Millennium Medal. In 2001, he received a Best Paper Award from WPMC01, Aalborg, Denmark, and the IEEE Communications Society Edwin Howard Armstrong Achievement Award. He received twice (in 2004 and 2012) the Journal of Communications and Networks Best Paper Award. In 2012 he received from the IEEE Information Theory Society the Aaron D. Wyner Distinguished Service Award, and from EURASIP the Athanasios Papoulis Award for outstanding contributions to education in communications and information theory.

Marco Lops was born in Naples, Italy, March, 16, 1961. He received the "Laurea" and Ph.D. degrees from "Federico II" University, Naples. From 1989 to 1991, he was an Assistant Professor, and from 1991 to 2000 an Associate Professor, with "Federico II" University. Since March 2000, he has been a Professor at the University of Cassino, Italy, and, during 2009-2011, he was also with ENSEEIHT, Toulouse, France. In Fall 2008, he was a Visiting Professor with the University of Minnesota, Minneapolis, and in Spring 2009, with Columbia University, New York. His research interests are in detection and estimation.