Detection of Unknown Constant Magnitude Signals in Time-Varying Channels

Daniel Romero and Roberto López-Valcarce
Dept. Signal Theory & Communications
University of Vigo, Spain
email: {dromero, valcarce}@gts.tsc.uvigo.es

Submitted to: SPECIAL SESSION ON COGNITIVE RADIO, CIP 2012

***Abstract*—Spectrum sensing constitutes a key ingredient in many cognitive radio paradigms in order to detect and protect primary transmissions. Most sensing schemes in the literature assume a time-invariant channel. However, when operating in low Signal-to-Noise Ratio (SNR) conditions, observation times are necessarily long and may become larger than the coherence time of the channel. In this paper the problem of detecting an unknown constant-magnitude waveform in frequency-flat time-varying channels with noise background of unknown variance is considered. The channel is modeled using a basis expansion model (BEM) with random coefficients.**

Adopting a generalized likelihood ratio (GLR) approach in order to deal with nuisance parameters, a non-convex optimization problem results. We discuss different possibilities to circumvent this problem, including several low complexity approximations to the GLR test as well as an efficient fixed-point iterative method to obtain the true GLR statistic. The approximations exhibit a performance ceiling in terms of probability of detection as the SNR increases, whereas the true GLR test does not. Thus, the proposed fixed-point iteration constitutes the preferred choice in applications requiring a high probability of detection.

I. Introduction

Signal activity detection constitutes a key functional block in many signal processing systems like those in sonar [1], radar [2] and spectrum sensing for cognitive radio [3] among others. The traditional approach assumes that the channel is time-invariant so that some performance loss is incurred when implemented for actual channels. This loss will depend on how far the channel is from being time-invariant. In certain applications, it may be necessary to adopt a time-varying channel model form the very beginning.

Examples of these include, for instance, narrowband communications, where the symbol period can be in the order of the coherence time of the channel. This effect is particularly important in the case of acoustic communications such as those in underwater environments, typically affected by large Doppler spreads [1]. For spectrum sensing, detectors shall typically fulfil stringent requirements in terms of high probability of detection at low SNR so that long observation windows are needed [3]. Consequently, the channel is less likely to be accurately approximated as time-invariant.

Herein we assume a frequency-flat channel following a basis expansion model (BEM) [4], [5] with random coefficients. Although formally any orthonormal basis can be employed, we focus on the Fourier basis for concreteness. Moreover, this allows us to model the channel as a low-pass random process, as it frequently occurs in practice [6], by choosing the basis functions associated with the lowest frequencies.

We assume that the signal to be detected has constant magnitude (CM). This introduces additional structure in the problem as well as mathematical tractability. Many practical signals have this property since it relaxes linearity requirements on power amplifiers. Examples include frequency shift keying (FSK) modulation, continuous-phase modulation (CPM) or Gaussian minimum shift keying (GMSK) modulation as used in GSM. An important application in Cognitive Radio is the detection of wireless microphones: these typically employ analog frequency modulation (FM), resulting in a CM waveform.

The detector is derived according to the generalized likelihood ratio (GLR) rule [7], where the unknown parameters are substituted by their maximum likelihood (ML) estimates. The starting point is the detector for known CM signals from [8]. It turns out that the GLR for unknown CM signals is the result of maximizing the test statistic from [8] with respect to the transmitted signal, subject to the CM property. This work addresses some alternatives to this (non-convex) problem together with the computation of the exact solution in order to provide the user with a performance/complexity tradeoff.

Notation: superscripts $\cdot^*$, $\cdot^T$ and $\cdot^H$ denote the conjugate, transpose, and conjugate transpose respectively. All vectors are assumed to be column vectors unless otherwise stated. $\boldsymbol{I}_N$ is the identity matrix of dimension N. The notation $(\boldsymbol{v})_k$ and $(\boldsymbol{A})_{k,l}$ refers to the k-th component of vector $\boldsymbol{v}$, and the (k,l) entry of matrix $\boldsymbol{A}$, respectively. $\mathrm{diag}\{\boldsymbol{v}\}$ is a diagonal matrix with the components of vector $\boldsymbol{v}$ on its diagonal. On the other hand, $\mathrm{diag}\{\boldsymbol{A}\}$ is a vector with the diagonal elements of matrix $\boldsymbol{A}$. The ℓ^n-norm of vector $\boldsymbol{v}$ is represented as $||\boldsymbol{v}||_n$. Finally, $\boldsymbol{0}_N$ and $\boldsymbol{1}_N$ denote $N \times 1$ vectors with all zeros and all ones respectively.

The rest of the paper is structured as follows. The system

This work was supported by the the European Regional Development Fund (ERDF) and the Spanish Government under projects DYNACS (TEC2010-21245-C02-02/TCM), COMONSENS (CONSOLIDER-INGENIO 2010 CSD2008-00010), and FPU grant AP2010-0149, and by the Galician Regional Government under projects "Consolidation of Research Units" 2009/62 and 2010/85.

model is introduced in Sec. II, and the derivation of the optimization problem corresponding to the GLR test is given in Sec. III. Alternative low-complexity detectors are presented in Sec. IV, whereas an efficient numerical method to compute the GLR statistic is given in Sec. V. Simulation results are shown in Sec. VI, and Sec. VII provides some final remarks.

II. System Model

An unknown, length-N signal $\boldsymbol{x}^* \in \mathbb{C}^N$ is to be detected after propagating through a frequency-flat, time-varying channel $\boldsymbol{h} \in \mathbb{C}^N$ in the presence of additive white Gaussian noise. Depending on the hypothesis active, the observation $\boldsymbol{y} \in \mathbb{C}^N$ is given by

$$\mathcal{H}_0: \quad \boldsymbol{y} = \sigma \boldsymbol{n}, \qquad \mathcal{H}_1: \quad \boldsymbol{y} = \alpha \boldsymbol{X}\boldsymbol{h} + \sigma \boldsymbol{n}, \tag{1}$$

where α is a constant accounting for the transmitted power and path loss, σ^2 is the noise power and $\boldsymbol{X} \doteq \operatorname{diag}\{\boldsymbol{x}^*\}$. The signal $\boldsymbol{x}^*$ is assumed to have unit constant magnitude. Under this model, α, σ and $\boldsymbol{X}$ are unknown deterministic parameters. The noise vector $\boldsymbol{n}$ is zero-mean circularly symmetric complex Gaussian, with $\mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^H\} = \boldsymbol{I}_N$.

The time variations of the channel are modeled using a BEM with K orthonormal basis functions:

$$\boldsymbol{h} = \sum_{k=1}^{K} c_k \boldsymbol{f}_k = \boldsymbol{F}\boldsymbol{c} \tag{2}$$

where the c_k are the random coefficients, collected in the vector $\boldsymbol{c} \doteq [c_1 \ \dots \ c_K]^T$, and $\boldsymbol{f}_K \in \mathbb{C}^N$ are the basis functions, arranged as the columns of the matrix $\boldsymbol{F} \doteq [\boldsymbol{f}_1 \ \dots \ \boldsymbol{f}_K]$. The vector $\boldsymbol{c}$ is regarded as zero-mean circularly symmetric complex Gaussian with $\mathrm{E}\{\boldsymbol{c}\boldsymbol{c}^H\} = \frac{N}{K}\boldsymbol{I}_K$. The factor N/K ensures that $\mathrm{E}\{||\boldsymbol{h}||_2^2\} = N$.

Although any orthonormal basis can be considered, we will focus on that consisting of the K elements with lowest frequency in the Fourier basis [4], [9]. Thus, if the Fourier basis is given by the columns of $\boldsymbol{W} \doteq [\boldsymbol{w}_0 \ \dots \ \boldsymbol{w}_{N-1}]$ where

$$\boldsymbol{w}_l \doteq \frac{1}{\sqrt{N}} \left[1 \ e^{j\frac{2\pi}{N}l1} \ \dots \ e^{j\frac{2\pi}{N}l(N-1)}\right]^T, \tag{3}$$

then, with K odd and $s \doteq \frac{K-1}{2}$, $\boldsymbol{F}$ is given by

$$\boldsymbol{F} = [\boldsymbol{w}_0 \ \boldsymbol{w}_1 \ \dots \ \boldsymbol{w}_s \ \boldsymbol{w}_{N-s} \ \dots \ \boldsymbol{w}_{N-1}]. \tag{4}$$

For future reference, observe that

$$\boldsymbol{F}\boldsymbol{F}^H = \sum_{k=1}^{K} \boldsymbol{f}_k \boldsymbol{f}_k^H = \boldsymbol{w}_0\boldsymbol{w}_0^H + 2\sum_{k=1}^{s} \mathrm{Re}\{\boldsymbol{w}_k\boldsymbol{w}_k^H\} \tag{5}$$

is a matrix of real coefficients; this follows from the fact that $\boldsymbol{w}_{N-l} = \boldsymbol{w}_l^*$, $1 \le l \le s$. On the other hand, $\boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{W}\boldsymbol{D}\boldsymbol{W}^H$, where $\boldsymbol{D} \doteq \operatorname{diag}\{[\boldsymbol{1}_{s+1}^T \ \boldsymbol{0}_{N-K}^T \ \boldsymbol{1}_s^T]^T\}$. The autocorrelation matrix of the channel is therefore

$$\boldsymbol{R}_{\boldsymbol{h}} \doteq \mathrm{E}\{\boldsymbol{h}\boldsymbol{h}^H\} = \frac{N}{K}\boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{W}\left[\frac{N}{K}\boldsymbol{D}\right]\boldsymbol{W}^H, \tag{6}$$

and it follows that the power spectral density of the channel tap is (asymptotically) given by the diagonal of the matrix $(N/K)\boldsymbol{D}$. Therefore, the channel is a low-pass random process of bandwidth $2\pi s/N$, which is proportional to the maximum Doppler shift. This process has unit power, i.e., $\mathrm{E}\{|(\boldsymbol{h})_k|^2\} = 1$, since $\boldsymbol{R}_{\boldsymbol{h}}$ has ones on its main diagonal.

Finally, let us define the SNR as $\rho \doteq \frac{\alpha^2}{\sigma^2}$ and the bandwidth occupancy ratio as $b \doteq K/N$.

III. GLR Test Derivation

The GLR test [7] is a generalization of the Neyman-Pearson test where nuisance parameters are replaced by their ML estimates under the corresponding hypothesis. In case that $\boldsymbol{x}$ is known, the GLR test is given by:

$$L_{\boldsymbol{x}}(\boldsymbol{y}) \doteq \frac{\max\limits_{\alpha^2\ge 0,\sigma^2\ge 0} p(\boldsymbol{y};\boldsymbol{x},\alpha^2,\sigma^2|\mathcal{H}_1)}{\max\limits_{\sigma^2\ge 0} p(\boldsymbol{y};\sigma^2|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma_x, \tag{7}$$

where $p(\cdot|\mathcal{H}_i)$ denotes the probability density function of $\boldsymbol{y}$ under hypothesis $\mathcal{H}_i$ (note that $\boldsymbol{y}$ is conditionally Gaussian), and γ_x is the threshold. Therefore, when $\boldsymbol{x}$ is not known, the GLR test can be obtained as

$$L(\boldsymbol{y}) = \max_{\boldsymbol{x}\in\mathcal{M}^N} L_{\boldsymbol{x}}(\boldsymbol{y}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma, \tag{8}$$

where $\mathcal{M}^N \doteq \{\boldsymbol{x} \in \mathbb{C}^N : |(\boldsymbol{x})_i| = 1\}$ is the set of unit-magnitude signals of length N. The log-GLR statistic for the test in (7) is derived in [8] and can be written as

$$\log L_{\boldsymbol{x}}(\boldsymbol{y}) = \begin{cases} N \cdot \log \frac{b\hat{p}+(1-b)\hat{q}}{\hat{p}^b\hat{q}^{1-b}}, & \text{if } \hat{p} > \hat{q}, \\ 0, & \text{otherwise,} \end{cases} \tag{9}$$

where

$$\hat{p} \doteq \frac{1}{K}||\boldsymbol{F}^H\boldsymbol{X}^H\boldsymbol{y}||_2^2 \tag{10}$$

and

$$\hat{q} \doteq \frac{1}{N-K}||\boldsymbol{G}^H\boldsymbol{X}^H\boldsymbol{y}||_2^2. \tag{11}$$

The $N \times (N-K)$ matrix $\boldsymbol{G}$ comprises the $N-K$ columns of the Fourier basis $\boldsymbol{W}$ not included in $\boldsymbol{F}$. We can think of $\hat{p}$ as an estimator of the observed average power within the Doppler bandwidth after removing the influence of the information signal. Recall that $\boldsymbol{X}^H\boldsymbol{X} = \boldsymbol{I}_N$ since $\boldsymbol{x}$ is CM. On the other hand, $\hat{q}$ is an estimate of the out-of-band average power. Note that the observed energy is given by $||\boldsymbol{y}||_2^2 = K\hat{p} + (N-K)\hat{q}$.

It is readily checked that (9) depends on the data only through the ratio $\hat{p}/\hat{q}$ and, moreover, it is a monotonically nondecreasing function of $\hat{p}/\hat{q}$ in the range $\hat{p}/\hat{q} \ge 1$. Hence, maximizing $\log L_{\boldsymbol{x}}(\boldsymbol{y})$ w.r.t. $\boldsymbol{x}$ amounts to maximizing $\hat{p}/\hat{q}$. Since this ratio can be expressed as

$$\frac{\hat{p}}{\hat{q}} = (N-K)\frac{\hat{p}}{||\boldsymbol{y}||_2^2 - K\hat{p}}, \tag{12}$$

which is a non-decreasing function of $\hat{p}$ for fixed $||\boldsymbol{y}||_2^2$, the problem boils down to maximizing $\hat{p}$. Defining $\boldsymbol{Y} \doteq \operatorname{diag}\{\boldsymbol{y}^*\}$ enables us to rewrite (10) as

$$K\hat{p} = ||\boldsymbol{F}^H\boldsymbol{Y}^H\boldsymbol{x}||^2 = \boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}, \tag{13}$$

where $\boldsymbol{A} \doteq \boldsymbol{Y}\boldsymbol{F}\boldsymbol{F}^H\boldsymbol{Y}^H$ is an $N \times N$ positive semidefinite matrix with $\mathrm{rank}\,(\boldsymbol{A}) = K$. Consequently, the problem reduces to that of maximizing $\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}$ subject to the CM constraint $\boldsymbol{x} \in \mathcal{M}^N$:

$$\hat{\boldsymbol{x}} = \arg \max_{\boldsymbol{x} \in \mathcal{M}^N} \boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}. \tag{14}$$

Problem (14) can be solved in closed form when $K = 1$. In that case, $\boldsymbol{A} = \boldsymbol{a}\boldsymbol{a}^H$ is a rank-1 matrix, where $\boldsymbol{a} = \boldsymbol{Y}\boldsymbol{f}_1$, so that $\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x} = |\boldsymbol{a}^H\boldsymbol{x}|^2 \leq ||\boldsymbol{a}||_1^2$, with equality when $\angle(\boldsymbol{x})_i = \angle(\boldsymbol{a})_i + C, \forall i$, where C is a constant. Specifically, if we use the Fourier basis as in (4), then $\sqrt{N}\boldsymbol{f}_1 = \boldsymbol{1}_N$ and the detector becomes that for CM signals in time-invariant channels [10].

To the best of our knowledge, no closed-form expression exists for $\hat{\boldsymbol{x}}$ in (14) when $K > 1$ and, consequently, one has to either resort to numerical methods, or find suitable approximations. The latter are appealing since they are likely to present a reduced computational complexity when compared to numerical procedures. For this reason, in Sec. IV we explore different possibilities along these lines. The computation of the true GLR statistic will developed in turn in Sec. V.

IV. Low-Complexity Detectors

We now investigate five approaches based on different approximations and bounds of the cost function in (14). Remarkably, all these approaches yield the correct solution to the problem if $K = 1$, although they are not necessarily equivalent if $K > 1$.

1) Largest eigenvalue of $\boldsymbol{A}$: Since $||\boldsymbol{x}||_2^2 = N$ for all $\boldsymbol{x} \in \mathcal{M}^N$, one can relax the constraint in (14) to obtain the following problem:

$$\text{maximize} \quad \boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x} \tag{15}$$

$$\text{subject to} \quad ||\boldsymbol{x}||_2^2 = N. \tag{16}$$

The maximum is given by $\lambda_1(\boldsymbol{A}) \cdot N$, where $\lambda_1(\cdot)$ denotes the largest eigenvalue. This constitutes an upper bound on $\hat{\boldsymbol{x}}^H\boldsymbol{A}\hat{\boldsymbol{x}}$. Hence, one can think of approximating $\hat{p}$ as $\hat{p} \approx \hat{p}_{\text{L1}} \doteq \frac{N}{K}\lambda_1(\boldsymbol{A})$.

2) ℓ^1-Norm of principal eigenvector: Since the original problem is easily solvable when $K = 1$, consider a rank-1 approximation of $\boldsymbol{A}$ in terms of its principal unit-norm eigenvector $\boldsymbol{v}_1$, i.e., $\boldsymbol{A}$ as $\boldsymbol{A} \approx \lambda_1\boldsymbol{v}_1\boldsymbol{v}_1^H$. This leads to

$$\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x} \approx \lambda_1(\boldsymbol{A}) \cdot |\boldsymbol{v}_1^H\boldsymbol{x}|^2 \leq \lambda_1 \cdot ||\boldsymbol{v}_1||_1^2 \tag{17}$$

with equality attained for $\angle(\boldsymbol{x})_i = \angle(\boldsymbol{v}_1)_i + C, \forall i$. As a result, one can approximate $\hat{p}$ as $\hat{p} \approx \hat{p}_{\text{L1n1}} \doteq \frac{1}{K}\lambda_1(\boldsymbol{A})||\boldsymbol{v}_1||_1^2$.

3) Phase relaxation: If we write the vector $\boldsymbol{x}$ as $\boldsymbol{x} = [e^{j\theta_1}\ e^{j\theta_2}\ \ldots\ e^{j\theta_N}]^T$, then it is clear that

$$\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x} = \sum_{k=1}^{N}\sum_{l=1}^{N} a_{kl}e^{j(\theta_l - \theta_k)} \tag{18}$$

where $a_{kl} \doteq (\boldsymbol{A})_{k,l}$. Since $\boldsymbol{A}$ is Hermitian, (18) can be rewritten in terms of the elements on and below the diagonal:

$$\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x} = \sum_{k=1}^{N} a_{kk} + 2\,\mathrm{Re}\left\{\sum_{k=2}^{N}\sum_{l=1}^{k-1} a_{kl}e^{-j\theta_{kl}}\right\}, \tag{19}$$

where we have defined $\theta_{kl} \doteq \theta_k - \theta_l$. Although only N degrees of freedom are available to choose the values of θ_{kl}, we can relax this requirement and regard all of the $\theta_{kl},\ l < k$ as free parameters, thus having $N(N-1)/2$ degrees of freedom. In doing so, the θ_{kl} maximizing (19) are given by $\theta_{kl} = \angle a_{kl}$, resulting in the following bound:

$$\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x} \leq \sum_{k=1}^{N}\sum_{l=1}^{N} |a_{kl}|. \tag{20}$$

Therefore, it is reasonable to approximate $\hat{p}$ as:

$$\hat{p} \approx \hat{p}_{\text{PR}} \doteq \frac{1}{K}\sum_{k=1}^{N}\sum_{l=1}^{N} |a_{kl}| \tag{21}$$

4) Phases of principal eigenvector: If we were to approximate $\hat{\boldsymbol{x}}$ rather than $\hat{p}$, a similar argument to that provided in Sec. IV-1 could be applied. Consequently, one can take $(\boldsymbol{x}_{\text{PPE}})_i = (\boldsymbol{v}_1)_i / |(\boldsymbol{v}_1)_i|$, where $\boldsymbol{v}_1$ is the principal eigenvector of $\boldsymbol{A}$. Substituting the resulting vector in (13) another approximation for $\hat{p}$ arises:

$$\hat{p} \approx \hat{p}_{\text{PPE}} \doteq \frac{1}{K}\boldsymbol{x}_{\text{PPE}}^H\boldsymbol{A}\boldsymbol{x}_{\text{PPE}} \tag{22}$$

5) Semidefinite Relaxation: As suggested in [11], upon noting that $\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x} = \mathrm{Tr}\left\{\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^H\right\} = \mathrm{Tr}\left\{\boldsymbol{A}\tilde{\boldsymbol{X}}\right\}$ where $\tilde{\boldsymbol{X}} \doteq \boldsymbol{x}\boldsymbol{x}^H$, one can relax the (non-convex) rank-1 constraint on $\tilde{\boldsymbol{X}}$. This results in the following semidefinite program (SDP):

$$\text{maximize} \quad \mathrm{Tr}\left\{\boldsymbol{A}\tilde{\boldsymbol{X}}\right\} \tag{23}$$

$$\text{subject to} \quad \mathrm{diag}\left\{\tilde{\boldsymbol{X}}\right\} = \boldsymbol{1}_N \tag{24}$$

$$\tilde{\boldsymbol{X}} \in S_+ \tag{25}$$

with S_+ the cone of positive semidefinite matrices. A solution can be efficiently found using any convex solver, and then we can retrieve $\boldsymbol{x} \in \mathcal{M}^N$ from the optimal $\tilde{\boldsymbol{X}}$, e.g., by taking the phases of its principal eigenvector, as suggested in [11].

Disappointingly, the solution of the SDP above is not unique (although all of them result in the same value of the cost $\mathrm{Tr}\left\{\boldsymbol{A}\tilde{\boldsymbol{X}}\right\}$), so that the particular one attained by any specific algorithm will be highly dependent on the initialization as well as the algorithm parameters. Thus, although the SDP cost associated to two different solutions $\tilde{\boldsymbol{X}}_1$ and $\tilde{\boldsymbol{X}}_2$ is the same, after mapping these to vectors $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in \mathcal{M}^N$, the cost of the original problem can be highly different, i.e. $\boldsymbol{x}_1^H\boldsymbol{A}\boldsymbol{x}_1$ is not necessarily close to $\boldsymbol{x}_2^H\boldsymbol{A}\boldsymbol{x}_2$. This is illustrated in Fig. 1, where the cost of the original non-convex problem is represented *vs* the cost of the SDP. Every cluster of points corresponds to a particular realization of the matrix $\boldsymbol{A}$ whereas each point is the result of solving the SDP with a different initialization. The fact that in many instances the points in each cluster are spread vertically is an indication of the phenomenon described above. This has an effect on the performance of the resulting detectors, as we have observed via simulations: the probability of detection strongly depends on the initializations

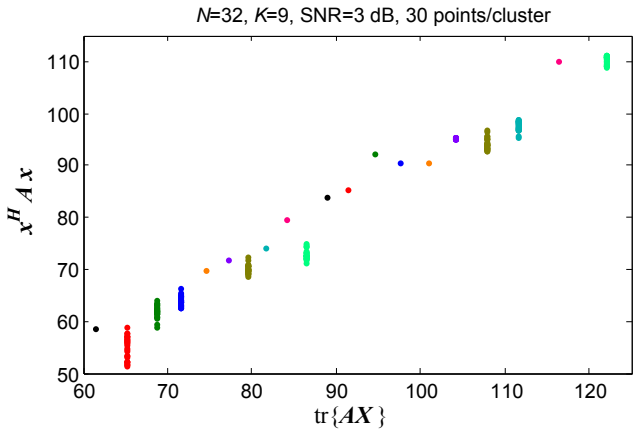

Fig. 1. Cost of the original problem in terms of the optimal cost of the SDP (23).

and parameters of the convex solver, resulting in a quite erratic and unpredictable behavior. In view of this, this approach will not be pursued further.

V. Computation of the true GLR

Although the matrix $\boldsymbol{A}$ is positive semidefinite, and thus $\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}$ is a convex function of $\boldsymbol{x}$, the feasible set of (14) is clearly non-convex. Therefore, conventional algorithms for solving this kind of problems cannot be used. On the other hand, since computational efficiency is required for the algorithm to be implemented in real time, one can think of employing a simple method for local search such as a gradient ascent or a fixed-point iterative algorithm.

A nice property of gradient ascent methods is the fact that they are guaranteed to converge to a (local) maximum, provided that the stepsize sequence is chosen adequately. Unfortunately, for the problem at hand, stepsize sequence design is highly dependent on the (unknown) SNR. In order to sidestep the problem of stepsize tuning, we abandon this approach in favor of fixed-point iterations.

These schemes aim at solving a set of equations by iteratively updating each variable so that one of the equations is satisfied. In this case the set of equations is built upon setting the gradient of the cost to zero. In order to compute the gradient, consider the unconstrained problem that results from substituting $\boldsymbol{x} = [e^{j\theta_1}\ e^{j\theta_2}\ \ldots\ e^{j\theta_N}]^T$ in $\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}$, and then taking the derivatives with respect to $\boldsymbol{\theta} \doteq [\theta_1\ \theta_2\ \ldots\ \theta_N]^T$. Recalling that $\boldsymbol{X} \doteq \operatorname{diag}\{\boldsymbol{x}^*\}$, this results in

$$\nabla_{\boldsymbol{\theta}} = 2\operatorname{Im}\{\boldsymbol{X}\boldsymbol{A}\boldsymbol{x}\}\,. \tag{26}$$

The iteration we propose is listed as Algorithm 1, where the vector $\boldsymbol{a}_i^H$ denotes the i-th row in $\boldsymbol{A}$. Several remarks are in order regarding Algorithm 1:

- One can readily check that at any fixed point of this iteration, the gradient (26) must vanish.
- Observe that, unlike the gradient descent algorithm, no parameter tuning is required.

Algorithm 1 Fixed-Point iteration

1: Set initial vector $\boldsymbol{x}$
2: **repeat**
3: **for** $i = 1$ to N **do**
4: Set $\theta_i = \angle(\boldsymbol{a}_i^H\boldsymbol{x})$
5: Set $(\boldsymbol{x})_i = \exp\{j\theta_i\}$
6: **end for**
7: **until** stopping criterion is satisfied
8: Set $\hat{p} = \frac{1}{K}\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}$

- Variables θ_i and $(\boldsymbol{x})_i$ are overwritten at each iteration of the outer loop.
- There are two nested loops instead of just one. This is for stability reasons: each component in $\boldsymbol{x}$ is updated one at a time rather than all together. In other words, to compute $(\boldsymbol{x})_i$ during the k-th iteration of the outer loop, the values of $(\boldsymbol{x})_1, (\boldsymbol{x})_2 \ldots (\boldsymbol{x})_{i-1}$ corresponding to the outer k-th iteration and the values of $(\boldsymbol{x})_i, (\boldsymbol{x})_{i+1} \ldots (\boldsymbol{x})_N$ corresponding to the $k-1$-th iteration are used in the right hand side of the expression in line 4. This prevents oscillations without introducing additional complexity.
- No stability problems have been observed in all cases tested with this scheme, which invariably converges to a maximum of the original cost.
- In the simulations, the stopping criterion considered is whether the norm of the gradient exceeds some certain tolerance ϵ.

As mentioned before, the optimization problem is not convex, and local maxima may exist. Thus, the choice of the initial estimate is critical in order to find the global maximum. To this end, one can consider the approaches presented in Sec. IV. Since we need an approximation for the vector $\hat{\boldsymbol{x}}$ rather than a bound for the cost, we can disregard the approaches from Secs. IV-1, IV-2 and IV-3. Regarding the approximation in IV-4, it can be shown that if the matrix $\boldsymbol{F}\boldsymbol{F}^H$ is real-valued, then the gradient vanishes when evaluated at the CM vectors obtained by retaining the phases of the components of any eigenvector of $\boldsymbol{A}$ associated to a non-null eigenvalue. This happens, in particular, when $\boldsymbol{F}$ has the form of (4). Unfortunately, these CM vectors are not, in general, maxima of the cost function but saddle points, as can be checked numerically after evaluating the Hessian matrix at those points. Therefore, these vectors are not suitable initializations for the fixed-point algorithm since no correction will be applied on the initial iterate. On the other hand, the approximation in Sec. IV-5 is not reliable as initializer either for the reasons discussed in that section.

Having dismissed all approximations in Sec. IV, we must consider other alternatives. Following the same reasoning as in Sec. IV-3, an initial estimate $\boldsymbol{\theta}$ could be found by projecting the solution of the phase-relaxed problem onto the feasible set. This can be accomplished by arranging an overdetermined system of equations with the relationships $\theta_{kl} = \theta_k - \theta_l, k > l$

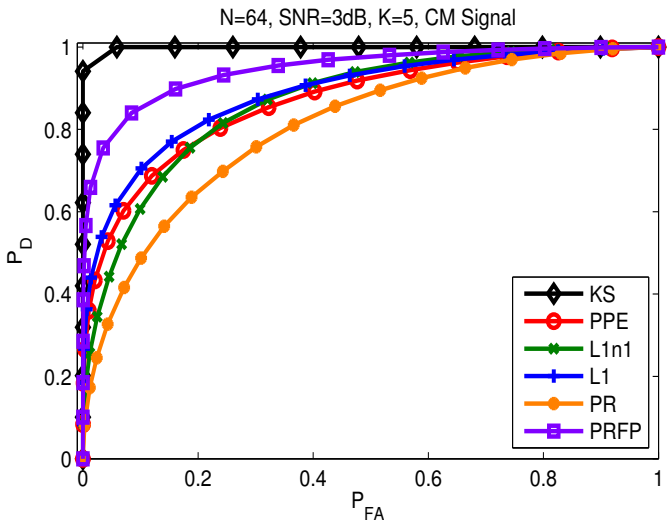

Fig. 2. ROC curves of the different detectors. $N = 64$ samples, SNR = 3 dB, $K = 5$ (Doppler spread $= \pi \times 0.0625$ rad/sample).

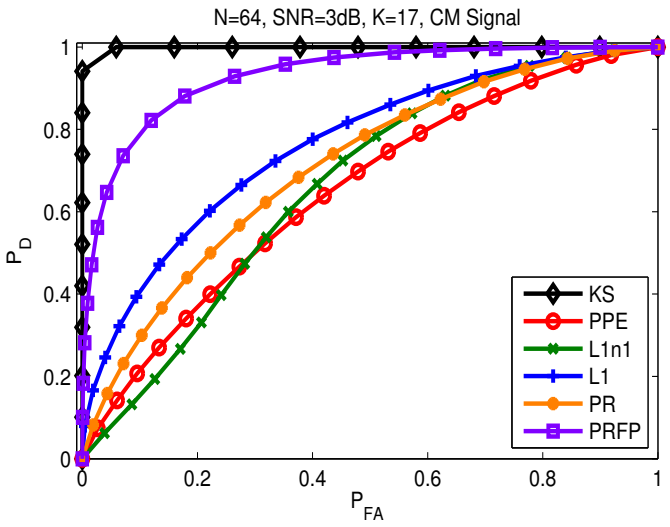

Fig. 3. ROC curves of the different detectors. $N = 64$ samples, SNR = 3 dB, $K = 17$ (Doppler spread $= \pi \times 0.250$ rad/sample).

and solving it using least squares (LS):

$$\hat{\boldsymbol{\theta}} = \arg\min_{\theta} ||\boldsymbol{M}\boldsymbol{\theta} - \boldsymbol{\theta}_{\rm PR}||^2, \tag{27}$$

where $\boldsymbol{\theta}_{\rm PR}$ is an $N(N-1)/2$-element vector with the $\theta_{kl},\ k > l$, with $\theta_{kl} = \angle a_{kl}$, and $\boldsymbol{M}$ is an $N(N-1)/2 \times N$ matrix where the i-th row is a vector of the form

$$\boldsymbol{m}_i^T = [\ 0 \quad \cdots \quad 0 \quad 1 \quad 0 \quad \cdots \quad 0 \quad -1 \quad 0 \quad \cdots \quad 0\ ] \tag{28}$$

with the non-null coefficients placed in such a way that the aforementioned conditions are imposed. Note that the sum of the columns of $\boldsymbol{M}$ is the zero vector; this reflects the fact that the relationships above are invariant to any constant added to the $\theta_k\ \forall k$. To sidestep this rank deficiency issue, we can fix $\theta_1 = 0$ so that we can drop the first column in $\boldsymbol{M}$ and the first row in $\boldsymbol{\theta}$. In that case $\boldsymbol{M}$ has full rank and the solution of (27) is unique.

VI. Simulation Results

We illustrate now the performance of the true GLR test as well as the detectors based on the approximations from Sec. IV, for several values of Doppler spread and SNR. The signal is generated as $(\boldsymbol{x})_i = e^{ju_i}$ where the phases $\{u_i\}$ are i.i.d. and uniformly distributed in $(0, 2\pi)$. In the implementation of the true GLR detector, the stopping criterion for the fixed-point iteration with *phase relaxation*-based initialization (PRFP) uses a tolerance value $\epsilon = 10^{-3}$.

We start by considering a BEM channel as described in Sec. II where the basis functions are the columns of the matrix $\boldsymbol{F}$ in (4). Clearly, the maximum Doppler frequency (Doppler spread) of such a channel is given by $\omega_d = \frac{2\pi s}{N}$.

With SNR = 3 dB and using $N = 64$ samples, Figs. 2 and 3 show the Receiver Operating Characteristic (ROC) curves [7], for $K = 5$ and $K = 17$ respectively, of the true GLR detector ("PRFP") and the detectors derived in Sec. IV-1 ("L1"), Sec. IV-2 ("L1n1"), Sec. IV-3 ("PR") and Sec. IV-4 ("PPE"). In addition, the curve labeled as "KS" corresponds to the GLR detector with knowledge of the transmitted signal [8], which is given by (9). This constitutes a performance upper bound for all of the other detectors. We observe that the true GLR test (PRFP detector) exhibits the best performance among those without knowledge of the signal, as could be expected. The non-GLR detectors of this group perform quite similarly and suffer from a higher performance loss relative to the true GLR test as K (and hence the Doppler spread) is increased.

In order to illustrate the influence of the SNR on the detectors, Figs. 4 and 5 show the probability of detection $P_{\rm D}$ for fixed false alarm rate $P_{\rm FA} = 0.1$ *vs* SNR, with the same value of $N = 64$, and $K = 5$ and 17 respectively. Note that the performance loss of the true GLR detector due to the lack of knowledge about the transmitted signal is about 10 dB. Another interesting effect is that there exists a performance ceiling for the non-GLR detectors, that is, for given N and K there exists an upper bound on the $P_{\rm D}$ that cannot be exceeded even as the SNR becomes arbitrarily large.

Finally, Fig. 6 shows the performance of the true GLR detector and illustrates two different effects. On the one hand, when the autocorrelation of the channel tap does not correspond to a flat PSD within the Doppler bandwidth, as assumed in the system model of Sec. II, some deviation in the detector behavior is expected. To test this, we use the popular dense scatterer model (also known as Jakes model [6], [12]). Under this model the vector $\boldsymbol{h}$ is zero-mean Gaussian distributed, with a Toeplitz covariance matrix $\mathrm{E}\left\{\boldsymbol{h}\boldsymbol{h}^H\right\}$ with (k, l) element given by $J_0(\omega_d(k-l))$. It is seen that the performance of the detector is quite robust to PSD mismatches.

On the other hand, the GLR detector assumes knowledge about the Doppler spread; hence, it makes sense to investigate the effect of model mismatches on this parameter. When ω_d

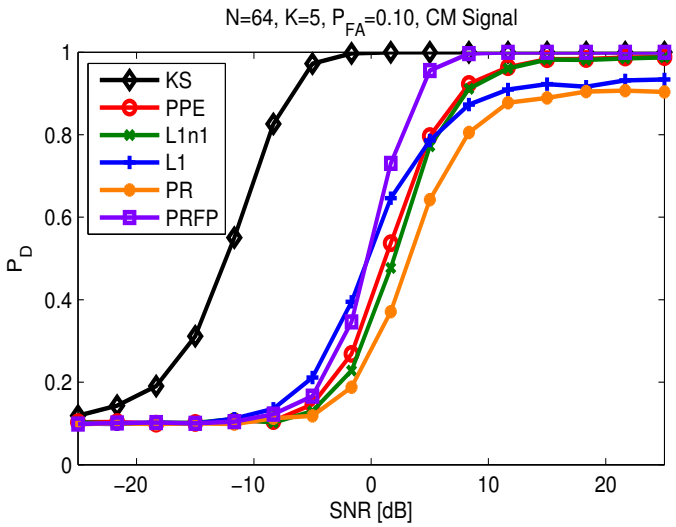

Fig. 4. Probability of detection *vs* SNR for $P_{\rm FA} = 0.1$. $N = 64$ samples, $K = 5$ (Doppler spread $= \pi \times 0.0625$ rad/sample).

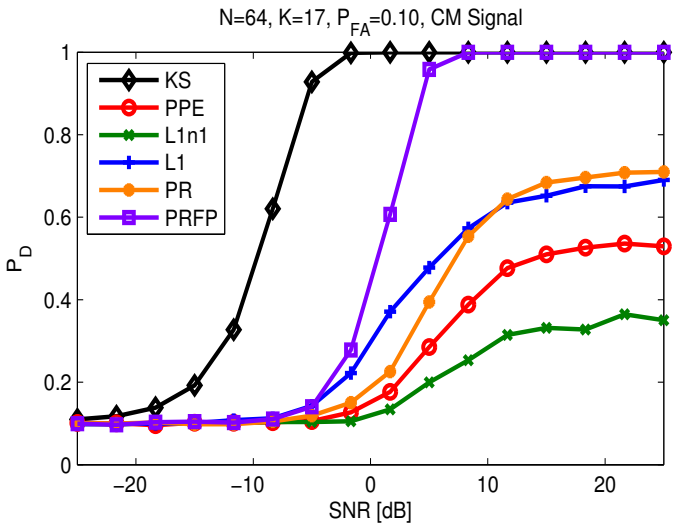

Fig. 5. Probability of detection *vs* SNR for $P_{\rm FA} = 0.1$. $N = 64$ samples, $K = 17$ (Doppler spread $= \pi \times 0.250$ rad/sample).

is known, a reasonable choice for K satisfies $\omega_d = \frac{K-1}{2}\,\frac{2\pi}{N}$. However, the actual value of the channel Doppler spread may not be available to the detector. Thus, Fig. 6 shows the probability of detection *vs* the actual Doppler spread, for several values of K assumed in the detection. A significant performance loss is observed as soon as the true Doppler spread exceeds the assumed one, whereas the penalty for overestimating this parameter is much milder.

VII. Conclusions

We have extended the detector for known CM signals in time-varying BEM channels to the more challenging scenario where the signal is unknown. Several low-complexity alternatives to the GLR test have been proposed, together with an iterative algorithm to compute the true GLR statistic (a fixed-point iteration initialized with the solution of a LS problem).

Through simulations it has been observed that the probability of detection of the low-complexity detectors remains bounded away from 1 even as the SNR goes to infinity, an effect which is more pronounced as the Doppler spread of the channel increases. This significant performance loss may preclude the use of these schemes in practice. Fortunately, the true GLR detector does not suffer from this drawback and constitutes an appealing alternative, performing well even under channel correlation not corresponding to the adopted model as long as the true channel Doppler spread is not underestimated.

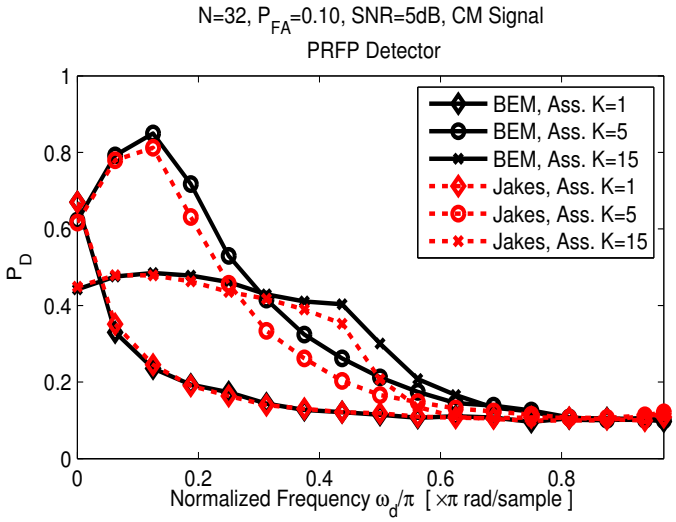

Fig. 6. Probability of detection of the true GLR test *vs.* the Doppler spread of the channel for different values of K used by the detector.

References

[1] J.-P. Marage and Y. Mori, *Sonar and underwater acoustics*, Wiley-ISTE, 2010.

[2] M. I. Skolnik, *Introduction to radar systems*, McGraw-Hill, 3rd edition, 2002.

[3] D. Cabric, "Addressing feasibility of cognitive radios," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 85–93, 2008.

[4] Michail K. Tsatsanis and Georgios B. Giannakis, "Modelling and equalization of rapidly fading channels," *Int. J. Adaptive Control Signal Process.*, vol. 10, no. 2-3, pp. 159–176, 1996.

[5] G. B. Giannakis and C. Tepedelenlioğlu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, no. 10, pp. 1969–1986, Oct. 1998.

[6] F. Hlawatsch and G. Matz, Eds., *Wireless communications over rapidly time-varying channels*, Academic Press, Oxford, UK, 2011.

[7] S.M. Kay, *Fundamentals of Statistical Signal Processing, Vol. II: Detection Theory*, Prentice-Hall, 1998.

[8] D. Romero, R. López-Valcarce, and G. Leus, "Generalized matched filter detector for fast fading channels," in *Proc. ICASSP*, 2012, Accepted.

[9] G. Leus, "On the estimation of rapidly time varying channels," in *Proc. EUSIPCO*, 2004, pp. 2227–2230.

[10] M. Derakhtian, AA Tadaion, S. Gazor, and MM Nayebi, "Invariant activity detection of a constant magnitude signal with unknown parameters in white Gaussian noise," *IET Commun.*, vol. 3, no. 8, pp. 1420–1431, 2009.

[11] F. Gholam, J. Via, and I. Santamaria, "Beamforming design for simplified analog antenna combining architectures," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2373 –2378, jun 2011.

[12] A. Goldsmith, *Wireless communications*, Cambridge Univ Pr, 2005.