Optimal and Sub-Optimal Spectrum Sensing of OFDM Signals in Known and Unknown Noise Variance

Erik Axell and Erik G. Larsson

Erik Axell and Erik G. Larsson, Optimal and Sub-Optimal Spectrum Sensing of OFDM Signals in Known and Unknown Noise Variance, 2011, IEEE Journal on Selected Areas in Communications, (29), 2, 290-304.

http://dx.doi.org/10.1109/JSAC.2011.110203

Optimal and Sub-Optimal Spectrum Sensing of OFDM Signals in Known and Unknown Noise Variance

Erik Axell and Erik G. Larsson

Abstract—**We consider spectrum sensing of OFDM signals in an AWGN channel. For the case of completely known noise and signal powers, we set up a vector-matrix model for an OFDM signal with a cyclic prefix and derive the optimal Neyman-Pearson detector from first principles. The optimal detector exploits the inherent correlation of the OFDM signal incurred by the repetition of data in the cyclic prefix, using knowledge of the length of the cyclic prefix and the length of the OFDM symbol. We compare the optimal detector to the energy detector numerically. We show that the energy detector is near-optimal (within 1 dB SNR) when the noise variance is known. Thus, when the noise power is known, no substantial gain can be achieved by using any other detector than the energy detector.**

For the case of completely unknown noise and signal powers, we derive a generalized likelihood ratio test (GLRT) based on empirical second-order statistics of the received data. The proposed GLRT detector exploits the non-stationary correlation structure of the OFDM signal and does not require any knowledge of the noise power or the signal power. The GLRT detector is compared to state-of-the-art OFDM signal detectors, and shown to improve the detection performance with 5 dB SNR in relevant cases.

Index Terms—**spectrum sensing,signal detection, OFDM, cyclic prefix, subspace detection, second-order statistics**

I. Introduction

A. Background

THE INTRODUCTION of cognitive radios in a primary user network will inevitably have an impact on the primary system, for example in terms of increased interference. Cognitive radios must be able to detect very weak primary user signals, to be able to keep the interference power at an acceptable level [1], [2]. Therefore, one of the most essential parts of cognitive radio is spectrum sensing.

One of the most basic sensing schemes is the energy detector [3]. This detector is optimal if both the signal and the noise are white Gaussian, and the noise variance is known. However, all man-made signals have some structure. This structure is intentionally introduced by the channel coding, the modulation and by the insertion of pilot sequences. Many modulation schemes give rise to a structure in the form of cyclostationarity (cf. [4], [5]), that may be used for signal detection [6]. A cyclostationary signal has a cyclic autocorrelation function that is nonzero at some nonzero cyclic frequency. The cyclostationarity property is also inherent at the same cyclic frequency in the cyclic spectral density, or the cyclic spectrum, of the signal. The detectors proposed in [6] use the cyclic autocorrelation and the cyclic spectrum in the time- and frequency domain respectively, to detect if the received signal is cyclostationary for a given cyclic frequency.

Many of the current and future technologies for wireless communication, such as WiFi, WiMAX, LTE and DVB-T, use OFDM signaling (cf. [7], [8]). Therefore it is reasonable to assume that cognitive radios must be able to detect OFDM signals. The structure of OFDM signals with a cyclic prefix (CP) gives a well known and useful cyclostationarity property [9]. Detectors that utilize this property have been derived, for example in [10], [11], [12] using the autocorrelation property, and in [13] using multiple cyclic frequencies. The detector proposed in [13] is an extension of the one in [6], to multiple cyclic frequencies. None of these detectors are derived based on statistical models for the received data that capture the non-stationarity of an OFDM signal, and they are not optimal in the Neyman-Pearson sense. We will show that it is possible to obtain much better detection performance.

In practice, the detector will have imperfect or no knowledge of parameters such as the noise power, the signal power and the synchronization timing of the transmitted signal. It is well known that the performance of the energy detector quickly deteriorates if the noise power is imperfectly known (cf. [14]). Any parameter uncertainties lead to fundamental limits on the detection performance, if not treated carefully [15]. More specifically, it was shown in [15] that any uncertainties in the model assumptions will have as consequence that robust detection is impossible at SNRs below a certain SNR wall. However, the problem of SNR walls can be mitigated by taking the imperfections into account. For example, it was also shown in [15] that noise calibration can improve the detector robustness.

In this work we consider the detection of an OFDM signal with a CP of known length. The proposed detectors can either be used stand-alone, or they can constitute building blocks in a larger spectrum sensing architecture. For example, a cognitive radio may look for different kinds of primary user signals in

Manuscript received 30 November 2009; revised 24 May 2010. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 216076. This work was also supported in part by the Swedish Research Council (VR) and the Swedish Foundation for Strategic Research (SSF). E. Larsson is a Royal Swedish Academy of Sciences (KVA) Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation. Parts of this work were presented at the 2nd International Workshop on Cognitive Information Processing, 2010.

The authors are with Linköping University, Sweden (e-mail: axell@isy.liu.se).

Digital Object Identifier 10.1109/JSAC.2011.110203.

many different frequency bands, where OFDM with CP is one example of a signal to be detected. The detectors used in each frequency band, and for each type of primary user signal may be different. Multiple detectors can then be run simultaneously, and a data fusion unit can be used to collect sensing decisions from all different detectors in order to make joint decisions on what spectrum that is free and occupied, respectively. Note also that many standards based on OFDM allow for multiple possible CP lengths (cf. [8]). To cope with this several versions of the proposed detectors, one for each CP length, can be run in parallel and the (soft) decisions be fused together. Moreover, the proposed detector can distinguish between an OFDM signal with a specific CP length from any other kind of signal. This can be an advantage in the presence of competing secondary users, where the proposed detector can discriminate between the primary and the secondary user signals.

B. Contributions

This paper contains two main contributions. First, we derive the optimal Neyman-Pearson detector for OFDM signals from first principles. In particular we give a closed-form expression for its test statistic, for the case when the noise power and the signal power are perfectly known. The optimal detector that we present exploits knowledge of the lengths of the OFDM symbol and its CP. This detector can be directly implemented in practice, provided that sufficiently accurate estimates of the noise and signal powers are available. The case when the noise and signal powers are unknown, is also dealt with (see the paragraph below). The optimal detector is also useful in that it provides an upper bound on the performance of other, suboptimal detectors. For example, we show numerically that when the signal and noise powers are known, the energy detector is near-optimal (within 1 dB SNR) for OFDM signals.

Second, we derive a computationally efficient detector based on a generalized likelihood ratio test (GLRT), operating on empirical second-order statistics of the received signal. A GLRT is a standard test, that takes as test statistic a likelihood ratio where the unknown parameters are replaced with their maximum-likelihood estimates (cf. [18, p. 92]). The so-obtained detector does not need any knowledge of the noise power or the signal power. We compare this detector to state-of-the-art methods [10], [11], [12]. The most relevant comparison is that with the detector of [10], which also works without knowing neither the signal variance nor the noise variance. We show that our proposed method can outperform the detector of [10] with 5 dB SNR in relevant cases. We also make comparisons when the noise variance supplied to the detectors is erroneous. We show that even for small errors of the noise power, the proposed detector is superior to all compared detectors which assumes perfect knowledge of the noise variance.

II. Model

We consider a discrete-time (sampled) complex baseband model. Assume that $\mathbf{x}$ is a received vector of length N that consists of an OFDM signal plus noise, i.e.

$$\mathbf{x} = \mathbf{s} + \mathbf{n},$$

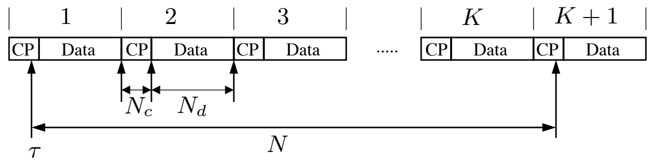

Fig. 1. Model for the N samples of the received OFDM signal.

where $\mathbf{s}$ is a sequence of K consecutively transmitted OFDM symbols, and $\mathbf{n}$ is a noise vector. The noise $\mathbf{n}$ is assumed to be i.i.d. zero-mean circularly symmetric complex Gaussian with variance σ_n^2, that is, $\mathbf{n} \sim \mathcal{C}N(\mathbf{0}, \sigma_n^2\mathbf{I})$. Each OFDM symbol consists of a data sequence of length N_d, and a cyclic prefix (CP) of length N_c ($\leq N_d$). Like in most related literature (cf. [10], [13]) we consider an AWGN channel, in order to study the most important fundamental aspects of OFDM signal detection.

In practice one cannot know exactly when to start the detection. That is, the receiver is not synchronized to the transmitted signal that is to be detected. Let τ be the synchronization mismatch, in other words the time when the first sample is observed. That is, $\tau = 0$ corresponds to perfect synchronization. We assume that the transmitted signal consists of an infinite sequence of OFDM symbols, so that detection can equivalently start within any symbol. Then, it is only useful to consider synchronization mismatches within one OFDM symbol, that is in the interval $0 \leq \tau < N_c + N_d$. In a perfectly synchronized case ($\tau = 0$) we would observe a number (K) of *complete* OFDM symbols, in order to fully exploit the structure of the signal. Without loss of generality, we assume that the total number of samples in the vector $\mathbf{x}$ is $N = K(N_c + N_d)$. This implies that $\mathbf{x}$ will in general (when $\tau > 0$) contain samples from $K+1$ OFDM symbols, as shown in Figure 1.

The model and methods that we present are valid for any value of K. Generally, the detection performance will improve with increasing K. However, choosing a very large K in practice may cause problems. For example if the A/D converter at the receiver is not synchronized (which it is generally *not*) to the D/A converter at the transmitter, there will be serious sampling errors due to the clock drift. There might also be problems with Doppler effects and carrier frequency offsets, if the number of samples is too large. Thus, our model is mostly useful for quite moderate values of K. In addition, for large values of K, our proposed detector can be run on a smaller number of OFDM symbols and then the soft decisions can be combined.

III. Optimal Neyman-Pearson Detector

In what follows, we provide a derivation of the optimal Neyman-Pearson detector from first principles. The key observation for deducing the optimal detector is that the OFDM signal lies in a certain subspace, owing to the structure introduced by the repetition of data in the CP. If the synchronization mismatch τ were known, this subspace would be perfectly known. The theory of detection of a signal in a known subspace has been extensively analyzed, both in white and colored noise [16]. In realistic scenarios, τ will be unknown.

Since the received signal depends on the synchronization mismatch τ, the signal subspace will be only partially known in general.

We start by formulating a vector-matrix model for the received signal. Let $\mathbf{q}_i$ be the N_d-vector of data associated with the ith OFDM symbol. This data vector is the output of the IFFT operation, used to create the OFDM data. An OFDM symbol $\mathbf{s}_i$ is then obtained by repeating the last N_c elements of $\mathbf{q}_i$ at the beginning of the symbol. This operation can be written in matrix form as

$$\mathbf{s}_i = \mathbf{U}\mathbf{q}_i,$$

where

$$\mathbf{U} = \begin{bmatrix} \mathbf{0}_{N_c \times N_d - N_c} & \mathbf{I}_{N_c} \\ \multicolumn{2}{c}{\mathbf{I}_{N_d}} \end{bmatrix} \in \mathbb{R}^{(N_c+N_d)\times N_d}.$$

Here $\mathbf{0}_{n\times m}$ denotes the $n \times m$ all-zero matrix, and $\mathbf{I}_n$ denotes the $n \times n$ identity matrix.

In a perfectly synchronized scenario ($\tau = 0$) we only need to consider samples from OFDM symbols $1, \dots, K$. In all other cases ($\tau = 1, \dots, N_c + N_d - 1$), the received signal $\mathbf{x}$ will contain samples from symbols $2, \dots, K$ and from parts of symbols 1 and $K+1$. Thus, we let the generated data $\mathbf{q}$ consist of $K+1$ data blocks, although $\mathbf{x}$ only consists of $K(N_c+N_d)$ samples. That is, we let $\mathbf{q} \triangleq [\mathbf{q}_1^T \mathbf{q}_2^T \dots \mathbf{q}_{K+1}^T]^T$ be a vector of length $(K+1)N_d$, consisting of the data that correspond to $K + 1$ OFDM symbols. Furthermore, let $\mathbf{T}$ be the following block-diagonal matrix, where the "diagonal" consists of $K+1$ instances of the matrix $\mathbf{U}$:

$$\mathbf{T} \triangleq \begin{bmatrix} \mathbf{U} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{U} & \mathbf{0} & \cdots \\ \vdots & & \ddots & \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{U} \end{bmatrix} \in \mathbb{R}^{(K+1)(N_c+N_d)\times(K+1)N_d}.$$

Then, a vector consisting of $K+1$ OFDM symbols is created by the multiplication $\mathbf{Tq}$. The received signal $\mathbf{s}$ contains samples $\tau, \dots, K(N_c+N_d)+\tau-1$ of the transmitted signal vector $\mathbf{Tq}$. That is, $\mathbf{s}$ is equal to $\mathbf{Tq}$ but the first τ samples and the last $N_c + N_d - \tau$ samples are excluded. This implies that the received signal $\mathbf{s}$ can be written

$$\mathbf{s} = \mathbf{T}_\tau \mathbf{q},$$

where $\mathbf{T}_\tau$ is the $K(N_c+N_d) \times (K+1)N_d$ matrix obtained by deleting the first τ rows and the last $N_c+N_d-\tau$ rows of $\mathbf{T}$.

Figure 2 shows an example of the created signal $\mathbf{Tq}$ and the received data samples for $N_c = 2$, $N_d = 4$ and $K = 2$. Note that the matrix $\mathbf{T}_\tau$ will have a few all-zero columns corresponding to the data samples that are not received. The rank of $\mathbf{T}_\tau$ is equal to the number of unique and independent data samples that are observed. In the perfectly synchronized case ($\tau = 0$), KN_d independent data samples are observed. In the unsynchronized case ($\tau \neq 0$), there will be fewer correlated samples, and thus more independent data. Consider the example shown in Figure 2, when $N_c = 2$, $N_d = 4$ and $K = 2$. Here, the number of unique samples and the rank of

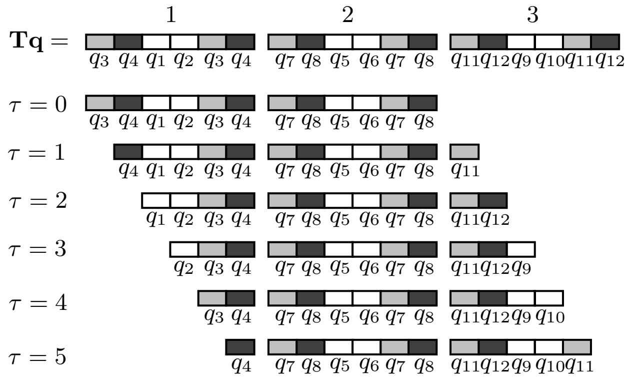

Fig. 2. Example of received data samples for different τ, for $N_c = 2$, $N_d = 4$ and $K = 2$.

$\mathbf{T}_\tau$ is

$$\text{rank}(\mathbf{T}_\tau) = \begin{cases} 8 & \tau = 0, \\ 9 & \tau = 1, \\ 10 & \tau = 2, \\ 10 & \tau = 3, \\ 10 & \tau = 4, \\ 9 & \tau = 5. \end{cases}$$

In general, it can be shown that the number of unique data samples, and thus the rank of $\mathbf{T}_\tau$, is $KN_d + \mu(\tau)$, where

$$\mu(\tau) \triangleq \begin{cases} \tau & 0 \le \tau \le N_c, \\ N_c & N_c \le \tau \le N_d, \\ N_c + N_d - \tau & N_d \le \tau \le N_d + N_c - 1. \end{cases} \tag{1}$$

The function $\mu(\tau)$ is the number of repeated samples that are lost due to imperfect synchronization.

Assuming a sufficiently large IFFT, the data vector $\mathbf{q}$ can be assumed to be Gaussian by the central limit theorem. That is, $\mathbf{q} \sim \mathcal{C}N(\mathbf{0}, \sigma_s^2\mathbf{I})$, where σ_s^2 is the variance of the complex signal samples. Conditioned on τ, the distribution of the signal $\mathbf{s}$ is zero-mean Gaussian with covariance matrix

$$E\left[\mathbf{s}\mathbf{s}^H|\tau\right] = E\left[\mathbf{T}_\tau\mathbf{q}(\mathbf{T}_\tau\mathbf{q})^H\right] = \sigma_s^2\mathbf{T}_\tau\mathbf{T}_\tau^T.$$

That is, $\mathbf{s}|\tau \sim \mathcal{C}N(\mathbf{0}, \sigma_s^2\mathbf{T}_\tau\mathbf{T}_\tau^T)$.

We wish to detect whether there is a signal present or not. That is, we want to discriminate between the following two hypotheses:

$$\begin{aligned} H_0 &: \mathbf{x} = \mathbf{n}, \\ H_1 &: \mathbf{x} = \mathbf{s} + \mathbf{n}. \end{aligned} \tag{2}$$

We start by considering detection when σ_n^2 and σ_s^2 are perfectly known.

A. Known σ_n^2 and σ_s^2

In this subsection, we derive the optimal Neyman-Pearson detector, for the unsynchronized case when τ is unknown. Under H_0, the received vector contains only noise. That is,

$$p(\mathbf{x}|H_0) = \frac{1}{\pi^N\sigma_n^{2N}} \exp\left(-\frac{\|\mathbf{x}\|^2}{\sigma_n^2}\right),$$

where, as before N is the length of the received vector (total number of received samples).

Under H_1, the received vector contains an OFDM signal plus noise, and the first sample is received at time τ. Since τ is unknown, we model it as a random variable, and obtain the marginal distribution:

$$p(\mathbf{x}|H_1) = \sum_{\tau=0}^{N_c+N_d-1} P(\tau|H_1)p(\mathbf{x}|H_1,\tau).$$

We assume that τ is completely unknown, and model this by taking τ uniformly distributed over the interval $[0, N_c + N_d - 1]$, so that

$$P(\tau|H_1) = \frac{1}{N_c + N_d}.$$

From the derivation in Section II, we know that $\mathbf{s}|\tau \sim \mathcal{CN}(\mathbf{0}, \sigma_s^2 \mathbf{T}_\tau \mathbf{T}_\tau^T)$, and thus $\mathbf{x}|H_1, \tau \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_\tau)$, where

$$\mathbf{Q}_\tau \triangleq \sigma_n^2 \mathbf{I} + \sigma_s^2 \mathbf{T}_\tau \mathbf{T}_\tau^T.$$

That is,

$$p(\mathbf{x}|H_1,\tau) = \frac{1}{\pi^N \det(\mathbf{Q}_\tau)} \exp(-\mathbf{x}^H \mathbf{Q}_\tau^{-1}\mathbf{x}).$$

The optimal Neyman-Pearson test is

$$\begin{aligned}\Lambda_{\text{optimal}} &\triangleq \log\left(\frac{p(\mathbf{x}|H_1)}{p(\mathbf{x}|H_0)}\right)\\ &= \log\left(\frac{\sum_{\tau=0}^{N_c+N_d-1} \frac{1}{(N_c+N_d)\pi^N \det(\mathbf{Q}_\tau)} \exp(-\mathbf{x}^H \mathbf{Q}_\tau^{-1}\mathbf{x})}{\frac{1}{\pi^N \sigma_n^{2N}} \exp(-\frac{\|\mathbf{x}\|^2}{\sigma_n^2})}\right)\\ &= \log\left(\sum_{\tau=0}^{N_c+N_d-1} \frac{1}{\det(\mathbf{Q}_\tau)} \exp\left(-\mathbf{x}^H \left(\mathbf{Q}_\tau^{-1} - \frac{1}{\sigma_n^2}\mathbf{I}\right)\mathbf{x}\right)\right)\\ &\quad + \log\left(\frac{\sigma_n^{2N}}{N_c+N_d}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\text{optimal}}.\end{aligned} \tag{3}$$

where η_{optimal} is a detection threshold. There appears to be no closed-form expression for the distribution of this test statistic, so the threshold has to be computed empirically. The computation of the threshold for a given probability of false alarm also requires knowledge of σ_n^2 and σ_s^2, as well as the CP length N_c and the data length N_d.

For collaborative detection [17], the detectors should deliver soft decisions that quantify how reliable the decision is. If the signals received by a number of collaborative sensors are independent, and the soft information is expressed in terms of LLRs, then the optimal fusion rule is to add the LLRs together. If no reliable soft decisions are available, then one is referred to using suboptimal schemes that combine hard decisions via AND- or OR-type voting rules [17]. For the optimal detector derived here, the soft decision is the LLR value Λ_{optimal}.

To compute the LLR (3), we need to compute $\det(\mathbf{Q}_\tau)$ and $\mathbf{x}^H \left(\mathbf{Q}_\tau^{-1} - \frac{1}{\sigma_n^2}\mathbf{I}\right)\mathbf{x}$. A direct computation of these quantities can be very burdensome if N is large. However, the computations can be significantly simplified by exploiting the sparse structure of $\mathbf{Q}_\tau$, as shown in Appendices A and B. The simplified computations of $\mathbf{x}^H \left(\mathbf{Q}_\tau^{-1} - \frac{1}{\sigma_n^2}\mathbf{I}\right)\mathbf{x}$ and $\det(\mathbf{Q}_\tau)$ are shown in (29)-(30) and (31) respectively.

B. Special cases

So far, we have considered the general unsynchronized case, when σ_n^2 and σ_s^2 are known, but τ is unknown. In this section we derive the optimal detector for some special cases, that will be used as benchmarks for the detection performance.

1) Known τ: First consider the case when the synchronization mismatch, τ, is known. Note that τ being known is not equivalent to $\tau = 0$, because of the end effects. For example, as already mentioned, the rank of the matrix $\mathbf{T}_\tau$ depends on τ. The rank of $\mathbf{T}_\tau$ achieves its smallest value for $\tau = 0$. Thus, the dimension of the signal subspace is smallest for $\tau = 0$, and then the signal should be easier to detect. For short OFDM symbols (small N_c and N_d) and/or small number of symbols K, these end effects can be quite large.

If τ is known (but not necessarily $\tau = 0$), then $\mathbf{x}|H_1 \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_\tau)$ under hypothesis H_1. In this case, the LLR can be written

$$\begin{aligned}\Lambda_{\text{synch}} &= \log\left(\frac{p(\mathbf{x}|H_1)}{p(\mathbf{x}|H_0)}\right)\\ &= \log\left(\frac{\frac{1}{\pi^N \det(\mathbf{Q}_\tau)} \exp(-\mathbf{x}^H \mathbf{Q}_\tau^{-1}\mathbf{x})}{\frac{1}{\pi^N \sigma_n^{2N}} \exp(-\frac{\|\mathbf{x}\|^2}{\sigma_n^2})}\right)\\ &= \log\left(\frac{\sigma_n^{2N}}{\det(\mathbf{Q}_\tau)}\right) - \mathbf{x}^H \left(\mathbf{Q}_\tau^{-1} - \frac{1}{\sigma_n^2}\mathbf{I}\right)\mathbf{x}.\end{aligned} \tag{4}$$

In practice, the detector will not be able to perfectly synchronize to the received signal. However, the performance of the synchronized detector can be used as an upper bound on the performance of the optimal detector in the unsynchronized case.

2) No cyclic prefix, $N_c = 0$ (energy detection): Consider the case when $N_c = 0$ (no CP), so that there is no structure in the signal that can be used. Then $\mathbf{T}_\tau \mathbf{T}_\tau^T = \mathbf{I}$, and $\mathbf{x}|H_1 \sim \mathcal{CN}(\mathbf{0}, (\sigma_n^2 + \sigma_s^2)\mathbf{I})$. In this case, the LLR is

$$\log\left(\frac{p(\mathbf{x}|H_1)}{p(\mathbf{x}|H_0)}\right) = \log\left(\frac{\frac{1}{\pi^N (\sigma_n^2+\sigma_s^2)^N} \exp(-\frac{\|\mathbf{x}\|^2}{\sigma_n^2+\sigma_s^2})}{\frac{1}{\pi^N \sigma_n^{2N}} \exp(-\frac{\|\mathbf{x}\|^2}{\sigma_n^2})}\right).$$

By removing all constants that are independent of the received vector $\mathbf{x}$, we obtain the test statistic

$$\Lambda_e = \|\mathbf{x}\|^2 = \sum_{i=0}^{N-1} |x_i|^2. \tag{5}$$

Hence, the optimal detector is in this case the energy detector, also known as the radiometer [3]. It is optimal if there is no knowledge about the signal, but the noise variance σ_n^2 is known. The energy detector is simple and widely used. Therefore it will be used as a benchmark to the optimal detector derived in Section III-A, that utilizes the knowledge of the lengths of the CP and the data. The performance of the energy detector is well known, cf. [18]. The probability of false alarm P_{FA} is given by

$$P_{\text{FA}} = \Pr(\Lambda_e > \eta_e | H_0) = 1 - F_{\chi^2_{2N}}\left(\frac{2\eta_e}{\sigma_n^2}\right). \tag{6}$$

Thus, given a false alarm probability, we can derive the threshold η_e from

$$\eta_e = F^{-1}_{\chi^2_{2N}}\left(1 - P_{\text{FA}}\right)\frac{\sigma_n^2}{2}. \tag{7}$$

The probability of detection P_d is then given by

$$\begin{aligned} P_{\text{D}} &= \Pr\left(\Lambda_e > \eta_e | H_1\right) = 1 - F_{\chi^2_{2N}}\left(\frac{2\eta_e}{\sigma_n^2 + \sigma_s^2}\right) \\ &= 1 - F_{\chi^2_{2N}}\left(\frac{F^{-1}_{\chi^2_{2N}}\left(1 - P_{\text{FA}}\right)}{1 + \frac{\sigma_s^2}{\sigma_n^2}}\right). \end{aligned} \tag{8}$$

It is clear from (7) that σ_n^2 is the only parameter that needs to be known at the detector to set the decision threshold (P_{FA} is a design parameter). Thus, the energy detector does not require σ_s^2 to be known.

3) Longest possible CP and perfect synchronization, $N_c = N_d$, $\tau = 0$: If the CP has the same length as the data ($N_c = N_d$) and the receiver is perfectly synchronized ($\tau = 0$), then each signal sample is repeated and both versions experience independent noise. That is, $\mathbf{U} = [\mathbf{I}_{N_d}\ \mathbf{I}_{N_d}]^T$ and $\mathbf{s} = [\mathbf{q}_1^T \mathbf{q}_1^T \mathbf{q}_2^T \mathbf{q}_2^T \ldots \mathbf{q}_K^T \mathbf{q}_K^T]^T$. This is approximately true also for $\tau \neq 0$ if the number of samples is large ($K \gg 1$). This scenario is not realistic in practice, but it provides a feeling for how much the CP structure of the signal can ultimately improve the detection performance relative to that of the energy detector.

Consider two received samples, corresponding to the two identical versions of a signal sample. Both samples contain i.i.d. noise added to the same signal value. Then the average of the two samples is a sufficient statistic for detection. Thus, when all data samples are repeated it is optimal to take the pairwise average of the samples corresponding to identical signal values. Let y_i be the average of the two received samples corresponding to identical signal values, and let $n_i^{(1)}$ and $n_i^{(2)}$ be the noise experienced by the first and the second version of the signal sample respectively. Then, the detection problem becomes

$$\begin{aligned} H_0 &: \ y_i = \frac{1}{2}\left(n_i^{(1)} + n_i^{(2)}\right),\ i = 0, \ldots, \frac{N}{2} - 1 \\ H_1 &: \ y_i = q_i + \frac{1}{2}\left(n_i^{(1)} + n_i^{(2)}\right),\ i = 0, \ldots, \frac{N}{2} - 1. \end{aligned}$$

We get $N/2$ independent samples y_i, where $y_i|H_0 \sim \mathcal{CN}(0, \sigma_n^2/2)$ and $y_i|H_1 \sim \mathcal{CN}(0, \sigma_s^2 + \sigma_n^2/2)$. Thus, the optimal detector is the same as for $N_c = 0$ in Section III-B2, i.e. the energy detector, but with half as many samples and half the noise variance (twice the SNR). Since this is also an energy detector, σ_n^2 needs to be known, but not σ_s^2. Replacing N with $N/2$ and σ_n^2 with $\sigma_n^2/2$ in (8), we get the relation between the probability of false alarm and the probability of detection as

$$P_{\text{D}} = 1 - F_{\chi^2_N}\left(\frac{F^{-1}_{\chi^2_N}\left(1 - P_{\text{FA}}\right)}{1 + 2\frac{\sigma_s^2}{\sigma_n^2}}\right). \tag{9}$$

C. Unknown σ_n^2 and σ_s^2

When σ_n^2 and σ_s^2 are unknown, the optimal strategy is to eliminate them from the problem by marginalization. That is, compute

$$p(\mathbf{x}|H_1) = \int_{\sigma_n^2, \sigma_s^2} \sum_{\tau=0}^{N_c+N_d-1} p(\mathbf{x}|H_1, \tau, \sigma_n^2, \sigma_s^2) \times P(\tau|H_1) p(\sigma_n^2, \sigma_s^2|H_1) d\sigma_n^2 d\sigma_s^2 \tag{10}$$

and

$$p(\mathbf{x}|H_0) = \int_{\sigma_n^2} p(\mathbf{x}|H_0, \sigma_n^2) p(\sigma_n^2|H_0) d\sigma_n^2. \tag{11}$$

We need to choose proper a priori distributions for σ_n^2 and σ_s^2 to get $p(\sigma_n^2, \sigma_s^2|H_1, \tau)$ and $p(\sigma_n^2|H_0)$, and then compute the integrals (10)-(11). It is not clear how these a priori distributions should be chosen. One possibility is to choose a non-informative distribution, for example the gamma distribution as we used in [19] to express lack of knowledge of the noise power. For most sensible distributions, the integrals are very hard to compute analytically. Therefore, for the case of unknown σ_n^2, σ_s^2, we proceed by instead using generalized likelihood-ratio tests.

IV. Detection based on second-order statistics

In this section, we propose a detector that exploits the structure of the OFDM signal by using empirical second-order statistics of the received data. The approach is inspired by the works of [10], [11], [12], which also use second-order statistics although in a highly suboptimal manner, see Section IV-D for a discussion. The case of most interest is when σ_n^2 and σ_s^2 are unknown, and we start our treatment with this assumption.

A. GLRT-approach for unknown σ_n^2 and σ_s^2

The repetition of data in the CP gives the OFDM signal a nonstationary correlation structure. We will propose a detector based on the generalized likelihood-ratio test (GLRT), that exploits this structure. Without loss of generality we assume throughout this section that the number of received samples is $N = K(N_c + N_d) + N_d$. Note that this is slight redefinition of N compared to Section III, to simplify notation. Define the sample value products

$$r_i \triangleq x_i^* x_{i+N_d},\ \ i = 0, \ldots, K(N_c + N_d) - 1. \tag{12}$$

The expected value of r_i of an OFDM signal is non-zero, for the data that is repeated in the CP of each OFDM symbol. This property will be used for detection. The received vector $\mathbf{x}$ consists of K consecutive OFDM symbols. Moreover, we know that if $s_i = s_{i+N_d}$ $(= q_{i+\tau} = q_{i+N_d+\tau})$, $0 \le i < N_c + N_d$, then $s_{i+k(N_c+N_d)} = s_{i+N_d+k(N_c+N_d)}$, $k = 1, \ldots, K-1$. Analogously, if $s_i = q_{i+\tau}$ and $s_{i+N_d} = q_{i+N_d+\tau}$ are independent ($q_{i+\tau} \neq q_{i+N_d+\tau}$), then $s_{i+k(N_c+N_d)}$ and $s_{i+N_d+k(N_c+N_d)}$ are also independent. Thus, we define

$$R_i \triangleq \frac{1}{K}\sum_{k=0}^{K-1} r_{i+k(N_c+N_d)},\ \ i = 0, \ldots, N_c + N_d - 1. \tag{13}$$

Under H_0, all the averaged sample value products R_i are identically distributed. Under H_1, there will be N_c consecutive values of R_i (starting with R_τ) that have a different distribution than the other N_d values. Figure 3 illustrates this for a

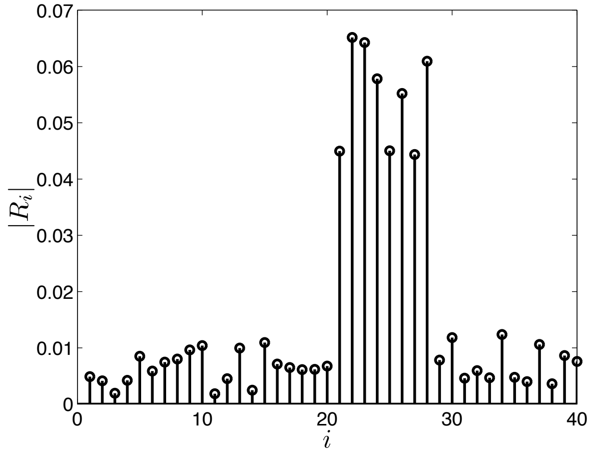

Fig. 3. Example of the correlation structure of a noise-free OFDM signal. $N_c = 8$, $N_d = 32$, $K = 50$, $\tau = 20$.

noise-free OFDM signal with $N_c = 8$, $N_d = 32$, $K = 50$ and $\tau = 20$. Since R_i is complex-valued, the figure shows $|R_i|$. It is clear that the 8 samples corresponding to the CP are significantly larger than the other. The aim of our proposed method is to detect whether R_i are i.i.d. or whether their statistics depend on i as explained above and as illustrated in Figure 3. Essentially, our proposed method implements a form of change detection. We propose a detector based on a GLRT that deals with the difficulty of not knowing τ, σ_s, σ_n. Let $\mathbf{R} \triangleq [R_0, \ldots, R_{N_c+N_d-1}]^T$. The GLRT is then

$$\Lambda_{\text{GLRT}} \triangleq \log\left(\frac{\max\limits_{\tau,\sigma_n^2,\sigma_s^2} p\left(\mathbf{R}|H_1, \tau, \sigma_n^2, \sigma_s^2\right)}{\max\limits_{\sigma_n^2} p\left(\mathbf{R}|H_0, \sigma_n^2\right)}\right) = \max_{\tau} \log\left(\frac{p\left(\mathbf{R}|H_1, \tau, \widehat{\sigma_n^2}, \widehat{\sigma_s^2}\right)}{p\left(\mathbf{R}|H_0, \widehat{\sigma_n^2}\right)}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\text{GLRT}}, \tag{14}$$

where $\widehat{\theta}$ denotes the maximum-likelihood (ML) estimate of the parameter θ.

To simplify the derivation of the joint distribution and the ML estimates, we assume that the variables R_i are approximately independent. Then, the likelihood function can be approximated as

$$p(\mathbf{R}|H_k, \tau, \sigma_n^2, \sigma_s^2) \approx \prod_{i=0}^{N_c+N_d-1} p(R_i|H_k, \tau, \sigma_n^2, \sigma_s^2)$$

and we only need to derive the marginal distributions of R_i. Since R_i is a complex-valued random variable, we must consider its real and imaginary parts separately. Let $\overline{a}$ and $\widetilde{a}$ denote the real and imaginary parts of a respectively. Then, $R_i = \overline{R}_i + j\widetilde{R}_i$, where

$$\overline{R}_i = \frac{1}{K}\sum_{k=0}^{K-1} \overline{r}_{i+k(N_c+N_d)}, \quad i = 0, \ldots, N_c + N_d - 1,$$

$$\widetilde{R}_i = \frac{1}{K}\sum_{k=0}^{K-1} \widetilde{r}_{i+k(N_c+N_d)}, \quad i = 0, \ldots, N_c + N_d - 1.$$

TABLE I
FIRST AND SECOND ORDER MOMENTS OF R_i.

Moment	$\mathbf{H_0}$	$\mathbf{H_1}$: $i \notin S_\tau$	$\mathbf{H_1}$: $i \in S_\tau$
$E\left[\overline{R}_i\middle|\cdot\right]$	0	0	σ_s^2
Var $\left[\overline{R}_i\middle|\cdot\right]$	$\frac{\sigma_n^4}{2K}$	$\frac{\left(\sigma_s^2+\sigma_n^2\right)^2}{2K}$	$\frac{\sigma_s^4+\frac{\sigma_n^4}{2}+\sigma_s^2\sigma_n^2}{K}$
$E\left[\widetilde{R}_i\middle|\cdot\right]$	0	0	0
Var $\left[\widetilde{R}_i\middle|\cdot\right]$	$\frac{\sigma_n^4}{2K}$	$\frac{\left(\sigma_s^2+\sigma_n^2\right)^2}{2K}$	$\frac{\frac{\sigma_n^4}{2}+\sigma_s^2\sigma_n^2}{K}$
Cov $\left[\overline{R}_i, \widetilde{R}_i\middle|\cdot\right]$	0	0	0

The terms $r_{i+k(N_c+N_d)}$ and $r_{i+l(N_c+N_d)}$ of the sum (13) are i.i.d. for $k \neq l$ by construction. Hence, R_i is a sum of i.i.d. random variables. Let $\mathbf{R}_i \triangleq \left[\overline{R}_i \; \widetilde{R}_i\right]^T$. Then, for large K, by the central limit theorem (cf. [20, pp. 108–109]), $\mathbf{R}_i$ has the two-dimensional Gaussian distribution

$$\mathbf{R}_i \sim \mathcal{N}\left(\begin{bmatrix} E\left[\overline{R}_i\right] \\ E\left[\widetilde{R}_i\right] \end{bmatrix}, \begin{bmatrix} \text{Var}\left[\overline{R}_i\right] & \text{Cov}\left[\overline{R}_i, \widetilde{R}_i\right] \\ \text{Cov}\left[\overline{R}_i, \widetilde{R}_i\right] & \text{Var}\left[\widetilde{R}_i\right] \end{bmatrix}\right). \tag{15}$$

A similar approximation was also used in [11]. This approximation approaches the true distribution with increasing K. Moreover, while the distribution of the product of two Gaussian variables has a relatively heavy tail, it is still exponentially decreasing (cf. [21, pp. 49–53]). Thus, the convergence to the Gaussian distribution should be quite fast, and the approximation should be acceptable even for rather small K. Moreover, this approximation allows us to derive the log-likelihood ratio in closed form, as shown in the following.

The structure of the OFDM signal incurs that the equality $s_i = s_{i+N_d}$ holds for N_c consecutive variables R_i, and that s_i and s_{i+N_d} are independent for all the other N_d variables. Let S_τ denote the set of consecutive indices for which $s_i = s_{i+N_d}$, given the synchronization mismatch τ. The expectations, variances, and covariances of $\overline{R}_i$ and $\widetilde{R}_i$ respectively are derived in Appendices C–E and summarized in Table I. Inserting the statistics of R_i from Table I in (15) yields

$$\begin{cases} \mathbf{R}_i | \{H_0\} \sim \mathcal{N}\left(\mathbf{0}, \frac{\sigma_n^4}{2K}\mathbf{I}\right), \\ \mathbf{R}_i | \{H_1, i \notin S_\tau\} \sim \mathcal{N}\left(\mathbf{0}, \frac{\left(\sigma_s^2+\sigma_n^2\right)^2}{2K}\mathbf{I}\right), \\ \mathbf{R}_i | \{H_1, i \in S_\tau\} \\ \sim \mathcal{N}\left(\begin{bmatrix} \sigma_s^2 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{\sigma_s^4+\frac{\sigma_n^4}{2}+\sigma_s^2\sigma_n^2}{K} & 0 \\ 0 & \frac{\frac{\sigma_n^4}{2}+\sigma_s^2\sigma_n^2}{K} \end{bmatrix}\right). \end{cases} \tag{16}$$

Detection is most crucial at low SNR ($\sigma_n^2 \gg \sigma_s^2$). We use this low-SNR approximation in the remainder of this section to simplify the computations of the ML estimates of the unknown parameters. A similar approximation was used in [10]. Define $\sigma_1^2 \triangleq \frac{\sigma_n^4}{2K}$. Then, at low SNR, the variances of $\overline{R}_i$ and $\widetilde{R}_i$ are approximately equal to σ_1^2 in all cases, and (16) simplifies to

$$\begin{cases} \mathbf{R}_i | \{H_0\} \sim \mathcal{N}\left(\mathbf{0}, \sigma_1^2 \mathbf{I}\right), i = 0, \ldots, N_c + N_d - 1, \\ \mathbf{R}_i | \{H_1, i \notin S_\tau\} \sim \mathcal{N}\left(\mathbf{0}, \sigma_1^2 \mathbf{I}\right), \\ \mathbf{R}_i | \{H_1, i \in S_\tau\} \sim \mathcal{N}\left(\begin{bmatrix} \sigma_s^2 \\ 0 \end{bmatrix}, \sigma_1^2 \mathbf{I}\right). \end{cases} \tag{17}$$

Under the assumptions made, the likelihood functions can approximately be written as

$$\begin{aligned} p(\mathbf{R}|H_0, \sigma_1^2) &\approx \prod_{i=0}^{N_c+N_d-1} p(\mathbf{R}_i|H_0, \sigma_1^2) \\ &\approx \prod_{i=0}^{N_c+N_d-1} \frac{1}{2\pi\sigma_1^2} \exp\left(-\frac{|R_i|^2}{2\sigma_1^2}\right) \\ &= \left(2\pi\sigma_1^2\right)^{-(N_c+N_d)} \exp\left(-\frac{1}{2\sigma_1^2} \sum_{i=0}^{N_c+N_d-1} |R_i|^2\right), \end{aligned} \tag{18}$$

and

$$\begin{aligned} p(\mathbf{R}|H_1, \tau, \sigma_1^2, \sigma_s^2) &\approx \prod_{i=0}^{N_c+N_d-1} p(\mathbf{R}_i|H_1, \tau, \sigma_1^2, \sigma_s^2) \\ &\approx \prod_{i \in S_\tau} \frac{\exp\left(-\frac{|R_i - \sigma_s^2|^2}{2\sigma_1^2}\right)}{2\pi\sigma_1^2} \prod_{j \notin S_\tau} \frac{\exp\left(-\frac{|R_j|^2}{2\sigma_1^2}\right)}{2\pi\sigma_1^2} \\ &= \left(2\pi\sigma_1^2\right)^{-(N_c+N_d)} \exp\left(-\frac{\sum_{i \in S_\tau} \left|R_i - \sigma_s^2\right|^2 + \sum_{j \notin S_\tau} |R_j|^2}{2\sigma_1^2}\right). \end{aligned} \tag{19}$$

It can be shown that the ML estimates are

$$\begin{aligned} &\widehat{\sigma_s^2} | \{H_1, \tau\} = \frac{1}{N_c} \sum_{i \in S_\tau} \overline{R}_i, \\ &\widehat{\sigma_1^2} | \{H_1, \tau\} \\ &= \frac{1}{2(N_c + N_d)} \left(\sum_{i \in S_\tau} \left| R_i - \frac{1}{N_c} \sum_{k \in S_\tau} \overline{R}_k \right|^2 + \sum_{j \notin S_\tau} |R_j|^2\right), \\ &\widehat{\sigma_1^2} | \{H_0\} = \frac{1}{2(N_c + N_d)} \left(\sum_{i=0}^{N_c+N_d-1} |R_i|^2\right). \end{aligned} \tag{20}$$

If we insert the ML estimates (20) and the likelihood functions (18)-(19) in (14), and remove all constants that are independent of R_i, we obtain the test

$$\max_\tau \frac{\sum_{i=0}^{N_c+N_d-1} |R_i|^2}{\sum_{k \in S_\tau} \left| R_k - \frac{1}{N_c} \sum_{i \in S_\tau} \overline{R}_i \right|^2 + \sum_{j \notin S_\tau} |R_j|^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\text{GLRT}}. \tag{21}$$

This test is computationally efficient. We only need to compute the empirical averages R_i from (12) and (13), then compute the likelihood ratio (21) for each τ, $0 \le \tau < N_c + N_d$, and take the maximum. Again, there appears to be no closed form expression for the distribution of the test statistic. Hence, the decision threshold has to be computed empirically. It should be noted that this is a constant false alarm rate (CFAR) test, meaning that the threshold can be computed for a fixed probability of false alarm independent of the SNR.

Any knowledge of the parameters σ_n^2, σ_s^2 or τ can easily be incorporated in the proposed detector by inserting the corresponding true parameter value into (14). See Subsections IV-B and IV-C for a brief discussion. If the synchronization mismatch τ is known, then the maximization in (21) can be omitted.

Note that although the expected value of the correlation is real-valued, the test statistic (21) depends on both the real and the imaginary parts of R_i. This is so because of the unknown noise power. Both $\overline{R}_i$ and $\widetilde{R}_i$ add information to the ML-estimate $(\widehat{\sigma_1^2})$ of the noise power.

B. Special case: Known σ_n^2 and unknown σ_s^2

If σ_n^2 is known, it can be directly inserted into (14) instead of the ML estimate. Knowledge of σ_n^2 does not change the ML estimate of σ_s^2 given by (20). After some algebra, we get

$$\begin{aligned} &\max_\tau \log\left(\frac{p\left(\mathbf{R}|H_1, \sigma_n^2, \widehat{\sigma_s^2}\right)}{p\left(\mathbf{R}|H_0, \sigma_n^2\right)}\right) \\ &\propto \max_\tau \left(\sum_{i \in S_\tau} \overline{R}_i\right)^2. \end{aligned} \tag{22}$$

The detector (22) may be compared with the energy detector, since both only need to know σ_n^2 in order to set the decision threshold.

Note that when the noise power is known (or rather $\sigma_1^2 = \frac{\sigma_n^4}{2K}$ is known), and the low-SNR approximation (17) is used, the test statistic depends only on the real parts $\overline{R}_i$. This is so because the imaginary parts $\widetilde{R}_i$ have the same distribution under both hypotheses, and $\overline{R}_i$ and $\widetilde{R}_i$ are uncorrelated.

C. Special case: Known σ_n^2 and σ_s^2

If both σ_n^2 and σ_s^2 are known, they can be directly inserted into (14), instead of the ML estimates (20). In this case we do not need to use the low-SNR approximation, since both σ_n^2 and σ_s^2 are known. Using the distributions (16) and some algebra, the LLR is given by

$$\begin{aligned} &\max_\tau \log\left(\frac{p\left(\mathbf{R}|H_1, \sigma_n^2, \sigma_s^2\right)}{p\left(\mathbf{R}|H_0, \sigma_n^2\right)}\right) \\ &\propto \max_\tau \left(\frac{1}{\sigma_1^2} \sum_{k=0}^{N_c+N_d-1} |R_k|^2 - \frac{1}{\gamma_1^2} \sum_{i \notin S_\tau} |R_i|^2 \right. \\ &\qquad \left. - \sum_{j \in S_\tau} \left(\frac{\left(\overline{R}_j - \sigma_s^2\right)^2}{\overline{\gamma_1}^2} + \frac{\widetilde{R}_j^2}{\widetilde{\gamma_1}^2}\right)\right), \end{aligned} \tag{23}$$

where

$$\gamma_1^2 \triangleq \mathrm{Var}\left[\overline{R}_i|H_1, i \notin S_\tau\right] = \mathrm{Var}\left[\widetilde{R}_i|H_1, i \notin S_\tau\right] = \frac{1}{2K}\left(\sigma_s^2+\sigma_n^2\right)^2,$$

$$\overline{\gamma_1}^2 \triangleq \mathrm{Var}\left[\overline{R}_i|H_1, i \in S_\tau\right] = \frac{1}{K}\left(\sigma_s^4+\frac{\sigma_n^4}{2}+\sigma_s^2\sigma_n^2\right),$$

$$\widetilde{\gamma_1}^2 \triangleq \mathrm{Var}\left[\widetilde{R}_i|H_1, i \in S_\tau\right] = \frac{1}{K}\left(\frac{\sigma_n^4}{2}+\sigma_s^2\sigma_n^2\right).$$

Note that the proposed GLRT detector with complete knowledge of the parameters is *not* equivalent to the optimal genie detector (3). Therefore, the detector in (23) is suboptimal. However, it is interesting to use for comparison purposes, since a comparison between (23) and (3) provides a feeling for how much performance that is lost by basing the detection on the second-order statistics R_i instead of on the received raw data $\mathbf{x}$.

In this case, even though σ_n^2 is known, the test statistic depends on both $\overline{R}_i$ and $\widetilde{R}_i$. This is so because the low-SNR approximation is not used, but instead the true moments as shown in Table I are used. Then, even though $E[\widetilde{R}_i|\cdot] = 0$ in all cases (under H_0 and under H_1 both for $i \in S_\tau$ and $i \notin S_\tau$), the variance $\mathrm{Var}[\widetilde{R}_i|\cdot]$ is different in the different cases. Thus, the imaginary part adds information, since the distribution is not exactly the same in all cases.

D. Benchmarks

In the following, we present three competing detectors [10], [11], [12] that are also based on second-order statistics of the received signal. To our knowledge, [10], [11], [12] represent the current state-of-the-art for the problem that we consider.

1) Autocorrelation-based detector of [10]: The method of [10] was called an autocorrelation-based detector and it uses the empirical mean of the sample value products r_i, normalized by the received power, as test statistic. More precisely, the test proposed in [10] is

$$\Lambda_{\text{AC}} = \frac{\sum_{i=0}^{(N_c+N_d)-1}\overline{R}_i}{\frac{N_c+N_d}{N}\sum_{i=0}^{N-1}|x_i|^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\text{AC}}. \tag{24}$$

The detector proposed in [10] does not require any knowledge about the noise variance σ_n^2.

Referring to Figure 3, the detector of [10] essentially uses the average of the 40 samples, and does not exploit the fact that only 8 of the samples have non-zero mean and the other 32 have zero mean. Thus, the detector of [10] ignores the fact that the received signal under H_1 is not stationary. The basic problem with this is that the samples x_i and x_{i+N_d} that correspond to signal samples that are repeated in the CP ($s_i = s_{i+N_d}$) are strongly correlated. On the other hand, the samples x_i and x_{i+N_d} that correspond to signal samples that are not repeated in the CP, are independent (because s_i and s_{i+N_d} are independent). Hence, taking the average of the sample value products as in (24) does not exploit all of the structure in the problem.

2) CP-detector of [11]: The detector of [11] is similar to the detector of [10] described in Section IV-D1, in the sense that it also uses the empirical mean of the sample value products r_i and therefore does not exploit the non-stationarity of the signal. The test proposed in [11] is

$$\Lambda_{\text{CP}} = \left|\frac{1}{N}\sum_{i=0}^{(N_c+N_d)-1} R_i + c\right|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\text{CP}}, \tag{25}$$

where

$$c \triangleq \frac{\frac{Nc}{Nc+Nd}\sigma_n^2}{\left(\left(1+2\left(\frac{Nc}{Nc+Nd}\right)^2\right)\frac{\sigma_s^2}{\sigma_n^2}+2\right)}.$$

It should be noted that this test statistic depends on σ_n^2 and σ_s^2. The work of [11] also proposed to use $c = 0$, to remove the required knowledge of these parameters. However, even if c is set to zero, the decision threshold depends on the noise variance σ_n^2.

3) Sliding-window detector of [12]: The detector of [12] uses a sliding window that sums over N_c consecutive samples, and takes the maximum. The test statistic is

$$\max_\tau \left|\sum_{i=\tau}^{\tau+N_c-1} r_i\right|.$$

The statistic (26) only takes one OFDM symbol at a time into account. A straightforward extension of this detector for K symbols, is to use the test

$$\Lambda_{\text{SW}} \triangleq \max_\tau \left|\sum_{i=\tau}^{\tau+N_c-1} R_i\right| \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\text{SW}}. \tag{26}$$

We will use the extended statistic (26) in our comparisons. The main drawback of the detector proposed in [12] is that it requires knowledge about σ_n^2 to set the decision threshold.

E. Detector complexity

In this section, we compare the proposed second-order GLRT detector with the benchmarks of Section IV-D in terms of complexity. We use the number of summations as an approximate measure of the detector complexity, because the number of summations is much greater than the number of divisions and multiplications. It should be noted that all these detectors require $(K-1)(N_c+N_d) = O(KN_d)$ additions to create the N_c+N_d averaged sample value products R_i from (13). Note that the CP length N_c is proportional to the data length N_d. The autocorrelation-based detector (24) and the CP-detector (25) require additionally $O(N_d)$ summations, given the averaged sample value products. That is, the total number of required additions is $O(KN_d)$ ($K \geq 1$). The sliding-window detector (26) and the proposed second-order GLRT detector (21) contain sums over S_τ, which have to be computed for each τ. However, since only two terms differ in the sums between two consecutive values of τ, the sums can be computed differentially. Then, both the sliding-window detector and the proposed GLRT detector require additionally $O(N_d)$ summations, and their total complexity is $O(KN_d)$. To conclude, the complexities of all the presented detectors based on second-order statistics are in the same order of magnitude. It should also be noted that the proposed GLRT detector and the sliding-window detector [12] exploit knowledge of N_c. Usually multiple values of N_c are possible. Then, these two

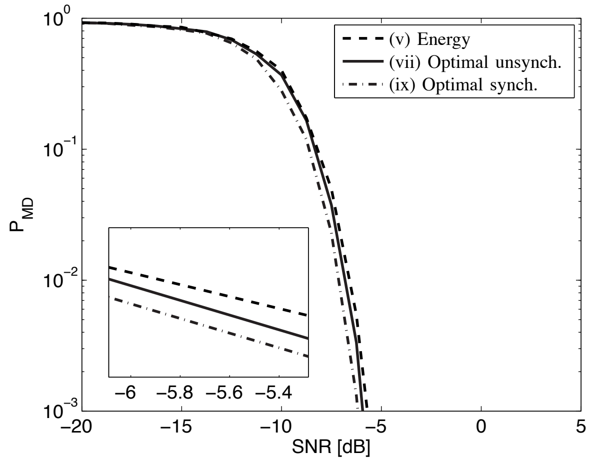

Fig. 4. Probability of missed detection P_{MD} versus SNR for different schemes with known parameters. $P_{\text{FA}} = 0.05$, $N_d = 32$, $N_c = 8$, $K = 10$.

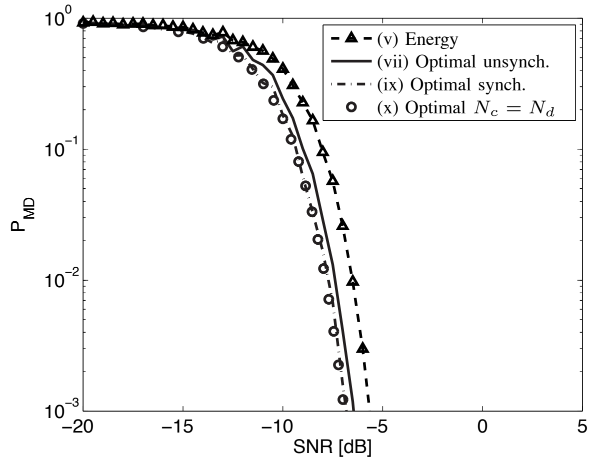

Fig. 5. Same as Figure 4, but with $N_d = 20$, $N_c = 20$.

detectors have to perform detection with several CP lengths in parallel, whereas the autocorrelation-based detector [10] and the CP-detector [11] need not do this.

V. Comparisons

We show some numerical results for the proposed detection schemes, obtained by Monte-Carlo simulation. All simulations are run until at least 100 detections (and missed detections) are observed. Performance is given as the probability of missed detection, P_{MD}, as a function of SNR. The SNR in dB is defined as $10\log_{10}(\sigma_s^2/\sigma_n^2)$. The noise variance was set to $\sigma_n^2 = 1$, and the SNR was varied from -20 dB to 5 dB. The data vector $\mathbf{q}$ was drawn randomly with the distribution $\mathbf{q} \sim \mathcal{C}N(\mathbf{0}, \sigma_s^2\mathbf{I})$. In the simulations, the probability of false alarm P_{FA} was fixed to find the detection threshold, η, and the probability of missed detection, P_{MD}. The thresholds have to be evaluated empirically, as there appears to be no closed form solution in most cases. The number of received OFDM symbols was set to $K = 10$. Choosing a larger K or larger N_c and N_d yield a larger number of received samples, and thus better performance. Note however, that scaling K is not equivalent to scaling N_c and N_d for most detectors, because the synchronization error τ depends on the absolute values of N_c and N_d. All detectors and their parameter knowledge requirements are summarized in Table II.

A. Optimal Neyman-Pearson detector

We start by comparing the optimal detector of Section III, with its special cases. We compare the following schemes, where the enumeration refers to Table II:

- (v) Energy, (5), σ_n^2 known, σ_s^2 and τ unknown.
- (vii) Optimal unsynchronized, (3), σ_n^2 and σ_s^2 known, τ unknown.
- (ix) Optimal synchronized, (4), σ_n^2, σ_s^2 and τ known.
- (x) Optimal longest CP ($N_c = N_d$), theory, (9), σ_n^2 known and $\tau = 0$.

Result 1: Comparison of Detectors (Figure 4). In this first scenario we study the optimal detector (vii) with known σ_n^2 and σ_s^2, and the special cases thereof. The IFFT size is set to $N_d = 32$ and the CP is chosen as $N_c = N_d/4 = 8$. The probability of false alarm is set to $P_{\text{FA}} = 0.05$. The results are shown in Figure 4. It is notable that the energy detector is near-optimal (within 0.2 dB SNR), even though the signal has a substantial correlation structure. This observation is also in line with [1], where the optimal detector for a BPSK modulated signal was derived, and it was shown that knowing the modulation format does not appreciably improve the detector performance over the energy detector. Notable is also that knowledge of σ_s^2 does not significantly improve the detection performance, since the energy detector only requires knowledge of σ_n^2 to set the decision threshold. We also note that perfect synchronization (knowing τ, scheme (viii)) does *not* substantially improve the detector performance.

Result 2: Longest possible CP (Figure 5). The purpose of the second scenario is to investigate what happens when the CP is as long as possible ($N_c = N_d$) and the OFDM signal therefore has as much correlation structure as possible. Here we choose $N_c = N_d = 20$, to get the same number of samples per OFDM symbol as in Figure 4. The results are shown in Figure 5. Here, the (unsynchronized) optimal detector outperforms the energy detector by about 1 dB SNR. Scheme (ix), where the synchronization mismatch τ is known (but not necessarily zero), performs almost as well as the repetition scheme (x). Some performance is lost, due to lost correlation when the received signal does not consist of K *complete* OFDM symbols (as when $\tau = 0$). Asymptotically, when the number of received OFDM symbols $K \to \infty$, these two schemes should have identical performance.

B. Detectors based on second-order statistics

In this section, we compare the detectors of Section IV, that are based on second-order statistics of the received signal. We also include the optimal detector with known σ_n^2 and σ_s^2, as a lower bound for the probability of missed detection. In this case we compare the following detectors, where again the enumeration refers to Table II:

- (i) Autocorrelation-based of [10], (24), σ_n^2 and σ_s^2 unknown.

TABLE II
SUMMARY OF DETECTORS, WHERE − MEANS UNKNOWN AND × MEANS KNOWN PARAMETER RESPECTIVELY.

ID	Ref.	Detector	Test	σ_n^2	σ_s^2	τ	N_d	N_c
i	[10]	Autocorrelation	(24)	−	−	−	×	−
ii	Proposed	2nd order, GLRT	(21)	−	−	−	×	×
iii	[12]	Sliding Window	(26)	×	−	−	×	×
iv	Proposed	2nd order, GLRT	(22)	×	−	−	×	×
v	[3]	Energy	(5)	×	−	−	−	−
vi	Proposed	2nd order, GLRT	(23)	×	×	−	×	×
vii	Proposed	Optimal unsynch.	(3)	×	×	−	×	×
viii	[11]	CP detection	(25)	×	×	−	×	×
ix	Proposed	Optimal synch.	(4)	×	×	×	×	×
x	Proposed	Optimal $N_c = N_d$	(9)	×	−	0	×	×

(ii) Proposed GLRT, (21), σ_n^2 and σ_s^2 unknown.
(iii) Sliding window of [12], (26), σ_n^2 known, σ_s^2 unknown.
(iv) Proposed GLRT with known σ_n^2, (22).
(vi) Proposed GLRT with known σ_n^2 and σ_s^2, (23).
(vii) Optimal unsynchronized, (3), σ_n^2 and σ_s^2 known.
(viii) CP detection of [11], (25), σ_n^2 and σ_s^2 known.

Result 3: Comparison of Detectors (Figure 6).
In this first scenario of detectors based on second-order statistics, the parameter values are the same as for Figure 4, except that the number of received symbols is increased to $K = 50$. The smaller complexity of the second-order statistics detectors compared to the optimal detector allows for a larger value of K. Figure 6 shows the results. The results show that knowledge of the parameters can improve the detector performance significantly, in the order of 5 dB SNR. We also note that the GLRT detector (ii), proposed in this paper, outperforms the autocorrelation-based detector (i) in the low P_{MD} region. Moreover, the improvement increases with decreasing P_{MD} (increasing SNR). In the IEEE 802.22 WRAN standard, a secondary user must be able to detect a primary user DVB-T signal with $P_{\text{MD}} \leq 10^{-1}$ [22]. At $P_{\text{MD}} = 10^{-1}$, the performance improvement of the GLRT detector (ii) over the autocorrelation-based detector (i) is in the order of 2.3 dB SNR. At lower P_{MD}, the improvement can be up to 5 dB SNR. The gain comes from exploiting the knowledge of the CP length, N_c, and the fact that the proposed detector exploits the non-stationarity of the OFDM signal. However, at high P_{MD} the autocorrelation-based detector (i) slightly outperforms the GLRT detector (ii). With these settings, this occurs approximately for $P_{\text{MD}} > 0.8$. We believe this effect appears owing to the suboptimality of GLRT, especially with respect to the synchronization error.

Result 4: Comparison of all detectors (Figure 7).
In this scenario, we show a comparison of all the presented unsynchronized detectors, using the same parameter values as in Figure 4. It is clear that the detector based on second-order statistics is suboptimal if σ_n^2 and σ_s^2 are known. In this scenario there is a $2-3$ dB gain in using the optimal detector (vii) based on the received data compared to the detector based on second order statistics (vi). Parts of the performance loss can also be attributed to the approximations made when deriving the second-order statistics detector. We note that the proposed detector based on second-order statistics (iv), and

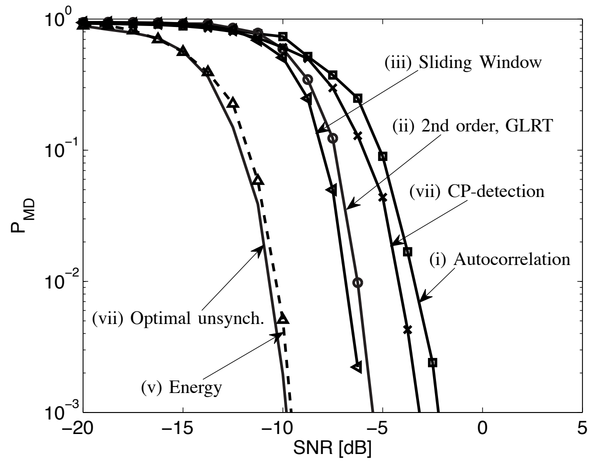

Fig. 6. Comparison of the correlation-based detection schemes. $P_{\text{FA}} = 0.05$, $N_d = 32$, $N_c = 8$, $K = 50$.

the sliding window detector (iii) have essentially the same performance when σ_n^2 is known. Worth noting is also that if σ_n^2 is known, the energy detector is near-optimal and outperforms the detectors based on second-order statistics. Thus, if σ_n^2 is known, no significant improvement over the energy detector can be achieved. However, if σ_n^2 is unknown, there can be a significant gain, as is shown in Result 5 below. The number of received symbols (samples) in this scenario is only a fifth compared to Figure 6. Moreover, the largest P_{MD} value where the GLRT detector (ii) outperforms the autocorrelation-based detector (i) is again approximately 0.8. This means also that the probability of detection is approximately 0.2. The introduction of cognitive radios in a primary network will require a larger probability of detection to avoid causing too much interference. Then, in most relevant cases, the GLRT detector (ii) is preferable over the autocorrelation-based detector (i).

Result 5: Noise uncertainty (Figure 8).
In this scenario, we consider noise uncertainty of 1 dB. That is, the noise variance supplied to the detectors deviates 1 dB from the true noise variance. The parameters are otherwise the same as in Figure 4. The results are shown in Figure 8. We note first of all that the performances of the detectors (i)-(ii), which

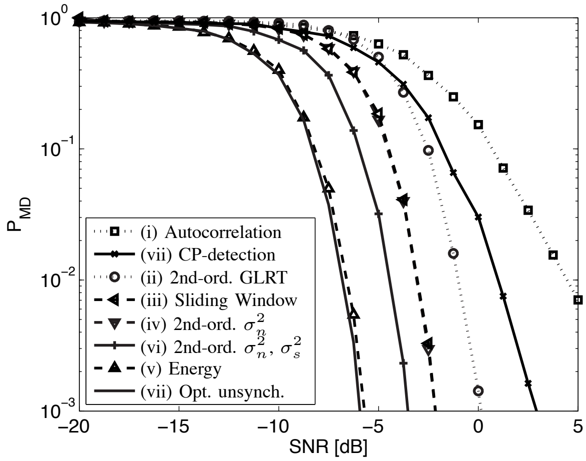

Fig. 7. Same parameters as Figure 4. Solid lines: known σ_n^2 and σ_s^2, dashed lines: known σ_n^2, unknown σ_s^2, and dotted lines: unknown σ_n^2 and σ_s^2.

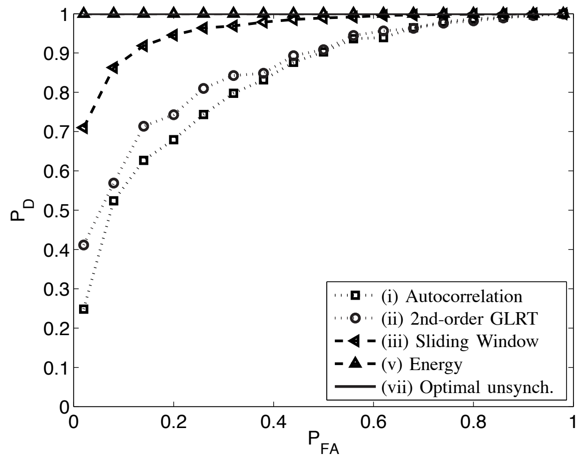

Fig. 9. ROC curve at -5 dB SNR, and otherwise the same parameters as Figure 4.

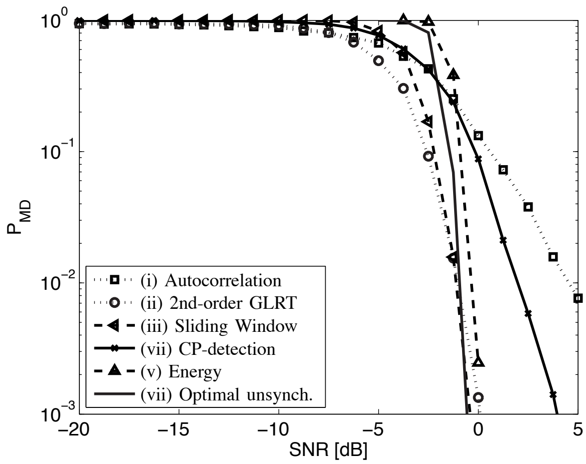

Fig. 8. Same parameters as Figure 4, and noise uncertainty 1 dB.

do not depend on the noise variance, are unaffected by the noise uncertainty. Furthermore, we note that the performances of both the optimal detector (vii) and the energy detector (v) have deteriorated with approximately 5 dB SNR, as compared to Figure 7. The sliding window detector (iii) and the CP-detector (vii) seem to be slightly more robust to the noise uncertainty. To conclude, even for small noise uncertainty levels, the proposed GLRT detector is superior to all compared detectors that assumes perfect knowledge of σ_n^2.

Result 6: Receiver operating characteristics (ROC) (Figure 9).
The receiver operating characteristics at -5 dB SNR, and otherwise the same parameters as in Figure 4, are shown in Figure 9. The results, in particular the order of the detectors, are similar to the previous results.

VI. Concluding Remarks

We derived the optimal Neyman-Pearson detector for an OFDM signal when the noise power and the signal power were known. Numerical comparisons showed that the energy detector is near-optimal (within 1 dB SNR) even if the cyclic prefix is relatively long, so that the signal has a substantial correlation structure.

We also proposed a detector based on the second-order statistics of the received OFDM signal, that does not require any knowledge of the noise variance or the signal variance. We showed numerically that the proposed detector can improve the detection performance with 5 dB SNR compared to state-of-the-art detectors such as the one in [10]. For simplicity, we made a few approximations in the derivation of the proposed GLRT detector. We assumed low SNR, and that the averaged sample value products are independent. We also used a Gaussian approximation via the central limit theorem. The detector performance might be further improved by not making these approximations. In this work we used a GLRT-approach, which is suboptimal. There are other ways of dealing with the unknown parameters, for example by marginalization. This is a topic for future research.

Appendix

A. *Efficient Computation of* $\mathbf{x}^H\left(\mathbf{Q}_\tau^{-1}-\frac{1}{\sigma_n^2}\mathbf{I}\right)\mathbf{x}$

The covariance matrix $\mathbf{Q}_\tau$ has the form

$$\begin{bmatrix} \ddots & & & \mathbf{0} & & & \\ & \sigma_n^2+\sigma_s^2 & 0 & \cdots & 0 & \sigma_s^2 & \\ & 0 & \sigma_n^2+\sigma_s^2 & \cdots & 0 & 0 & \\ \mathbf{0} & \vdots & & \ddots & & \vdots & \mathbf{0} \\ & 0 & \cdots & & \sigma_n^2+\sigma_s^2 & 0 & \\ & \sigma_s^2 & 0 & \cdots & 0 & \sigma_n^2+\sigma_s^2 & \\ & & & \mathbf{0} & & & \ddots \end{bmatrix},$$

where the nonzero off-diagonal elements correspond to samples of the OFDM signal that are equal. Because of the simple structure of $\mathbf{Q}_\tau$, its inverse has the form

$$\mathbf{Q}_\tau^{-1} = \begin{bmatrix} \ddots & & & \mathbf{0} & & & \\ & a & 0 & \cdots & 0 & b & \\ & 0 & c & \cdots & 0 & 0 & \\ \mathbf{0} & \vdots & & \ddots & & \vdots & \mathbf{0} \\ & 0 & \cdots & & c & 0 & \\ & b & 0 & \cdots & 0 & a & \\ & & & \mathbf{0} & & & \ddots \end{bmatrix}. \quad (27)$$

The elements a, b, c can be obtained from the equation $\mathbf{Q}_\tau \mathbf{Q}_\tau^{-1} = \mathbf{I}$, which yields the following linear system of equations

$$\begin{cases} a\left(\sigma_n^2+\sigma_s^2\right)+b\sigma_s^2=1 \\ a\sigma_s^2+b\left(\sigma_n^2+\sigma_s^2\right)=0 \\ c\left(\sigma_n^2+\sigma_s^2\right)=1 \end{cases} \Rightarrow \begin{cases} a=\frac{\sigma_n^2+\sigma_s^2}{2\sigma_n^2\sigma_s^2+\sigma_n^4} \\ b=-\frac{\sigma_s^2}{2\sigma_n^2\sigma_s^2+\sigma_n^4} \\ c=\frac{1}{\sigma_n^2+\sigma_s^2} \end{cases}.$$

For the computation of the likelihood function, we are not interested in the matrix inverse in itself, but in the quadratic form $\mathbf{x}^H\left(\mathbf{Q}_\tau^{-1}-\frac{1}{\sigma_n^2}\mathbf{I}\right)\mathbf{x}$. This quadratic form can be efficiently computed by using (27). First note that $\mathbf{Q}_\tau^{-1}-\frac{1}{\sigma_n^2}\mathbf{I}$ is of the form shown in (28). The product $\mathbf{y} \triangleq \left(\mathbf{Q}_\tau^{-1}-\frac{1}{\sigma_n^2}\mathbf{I}\right)\mathbf{x}$ is a vector where the elements y_i and y_{i+N_d} corresponding to a signal sample q_i that lies in the CP and have been observed twice (i.e. $i \in S_\tau$), are $y_i = y_{i+N_d} = \frac{-\sigma_s^2}{\sigma_n^2(\sigma_n^2+2\sigma_s^2)}\left(x_i+x_{i+N_d}\right)$. The element y_k corresponding to a signal sample q_k that has been observed only once, is $y_k = \frac{-\sigma_s^2}{\sigma_n^2(\sigma_n^2+\sigma_s^2)}x_k$. Then,

$$\mathbf{x}^H\left(\mathbf{Q}_\tau^{-1}-\frac{1}{\sigma_n^2}\mathbf{I}\right)\mathbf{x} = \sum_{i=0}^{K(N_c+N_d)-1}\alpha_i, \quad (29)$$

where

$$\alpha_i \triangleq \begin{cases} \frac{-\sigma_s^2}{\sigma_n^2(\sigma_n^2+2\sigma_s^2)}x_i^*\left(x_i+x_{i+N_d}\right), i \in S_\tau, \\ \frac{-\sigma_s^2}{\sigma_n^2(\sigma_n^2+2\sigma_s^2)}x_i^*\left(x_{i-N_d}+x_i\right), (i-N_d) \in S_\tau, \\ \frac{-\sigma_s^2}{\sigma_n^2(\sigma_n^2+\sigma_s^2)}|x_i|^2, \text{if } q_i \text{ is observed once}, \\ 0, \text{if } q_i \text{ is not observed}. \end{cases} \quad (30)$$

B. *Efficient Computation of* $\det(\mathbf{Q}_\tau)$

The computation of $\det(\mathbf{Q}_\tau)$ can also be simplified. We start by rewriting the determinant as follows:

$$\begin{aligned} \det(\mathbf{Q}_\tau) &= \det\left(\sigma_n^2\mathbf{I}_N+\sigma_s^2\mathbf{T}_\tau\mathbf{T}_\tau^T\right) \\ &= \det\left(\sigma_n^2\left(\mathbf{I}_N+\frac{\sigma_s^2}{\sigma_n^2}\mathbf{T}_\tau\mathbf{T}_\tau^T\right)\right) \\ &= \sigma_n^{2N}\det\left(\mathbf{I}_N+\frac{\sigma_s^2}{\sigma_n^2}\mathbf{T}_\tau\mathbf{T}_\tau^T\right). \end{aligned}$$

Using the identity $\det(\mathbf{I}_m+\mathbf{AB}) = \det(\mathbf{I}_n+\mathbf{BA})$ for matrices $\mathbf{A}$ and $\mathbf{B}$ of compatible dimensions, we can simplify the determinant further:

$$\det\left(\mathbf{I}_N+\frac{\sigma_s^2}{\sigma_n^2}\mathbf{T}_\tau\mathbf{T}_\tau^T\right) = \det\left(\mathbf{I}_{(K+1)N_d}+\frac{\sigma_s^2}{\sigma_n^2}\mathbf{T}_\tau^T\mathbf{T}_\tau\right).$$

The matrix $\mathbf{T}_\tau^T\mathbf{T}_\tau$ is diagonal with diagonal elements

$$d_i = \begin{cases} 0, \text{ if data } q_i \text{ is not observed}, \\ 1, \text{ if data } q_i \text{ is observed once}, \\ 2, \text{ if data } q_i \text{ is observed twice}. \end{cases}$$

That is, a diagonal element $d_i = 2$ corresponds to a data sample q_i which is repeated in the CP. Consider again the example shown in Figure 2, for $\tau = 1$. Then the data samples q_4, q_7, q_8 are observed twice, $q_1, q_2, q_3, q_5, q_6, q_{11}$ are observed once, and q_9, q_{10}, q_{12} are not observed at all. In general, the number of data samples that are received twice (number of $d_i = 2$) is $KN_c - \mu(\tau)$, where $\mu(\tau)$ is given by (1). The number of data samples that are not received at all (number of $d_i = 0$) is $N_d - \mu(\tau)$, and the number of samples that are received once (number of $d_i = 1$) is $K(N_d - N_c) + 2\mu(\tau)$.

Since the matrix $\mathbf{I}_{(K+1)N_d} + \frac{\sigma_s^2}{\sigma_n^2}\mathbf{T}_\tau^T\mathbf{T}_\tau$ is diagonal, the determinant is simply the product of the diagonal elements. That is

$$\begin{aligned} &\det\left(\mathbf{I}_{(K+1)N_d}+\frac{\sigma_s^2}{\sigma_n^2}\mathbf{T}_\tau^T\mathbf{T}_\tau\right) = \prod_{i=1}^{(K+1)N_d}\left(1+\frac{\sigma_s^2}{\sigma_n^2}d_i\right) \\ &= \left(1+\frac{\sigma_s^2}{\sigma_n^2}\right)^{K(N_d-N_c)+2\mu(\tau)}\left(1+2\frac{\sigma_s^2}{\sigma_n^2}\right)^{KN_c-\mu(\tau)}. \end{aligned}$$

To conclude,

$$\begin{aligned} &\det(\mathbf{Q}_\tau) \\ &= \sigma_n^{2N}\left(1+\frac{\sigma_s^2}{\sigma_n^2}\right)^{K(N_d-N_c)+2\mu(\tau)}\left(1+2\frac{\sigma_s^2}{\sigma_n^2}\right)^{KN_c-\mu(\tau)}, \end{aligned} \quad (31)$$

where $\mu(\tau)$ is given by (1).

C. *Moments of* $\overline{R}_i$

First we compute

$$E\left[\overline{R}_i\right] = E\left[\frac{1}{K}\sum_{k=0}^{K-1}\overline{r}_{i+k(N_c+N_d)}\right] = E\left[\overline{r}_i\right],$$

where

$$\overline{r}_i = \text{Re}\left(x_i^* x_{i+N_d}\right) = \overline{x}_i\overline{x}_{i+N_d} + \widetilde{x}_i\widetilde{x}_{i+N_d}.$$

In general, the expected value of $\overline{r}_i$ is

$$E\left[\overline{r}_i\right] = E\left[\overline{x}_i\overline{x}_{i+N_d}\right] + E\left[\widetilde{x}_i\widetilde{x}_{i+N_d}\right]. \quad (32)$$

The first term of (32) can be written as

$$\begin{aligned} &E\left[\overline{x}_i\overline{x}_{i+N_d}\right] = E\left[\left(\overline{s}_i+\overline{n}_i\right)\left(\overline{s}_{i+N_d}+\overline{n}_{i+N_d}\right)\right] \\ &= E\left[\overline{s}_i\overline{s}_{i+N_d}\right]+E\left[\overline{s}_i\overline{n}_{i+N_d}\right]+E\left[\overline{n}_i\overline{s}_{i+N_d}\right]+E\left[\overline{n}_i\overline{n}_{i+N_d}\right] \\ &= E\left[\overline{s}_i\overline{s}_{i+N_d}\right]. \end{aligned}$$

Similarly, the second term of (32) is

$$E\left[\widetilde{x}_i\widetilde{x}_{i+N_d}\right] = E\left[\widetilde{s}_i\widetilde{s}_{i+N_d}\right].$$

There are three different cases we need to consider: there is no signal (H_0), there is a signal and the signal samples are equal ($H_1, i \in S_\tau$), there is a signal but the samples are independent

$$\mathbf{Q}_\tau^{-1} - \frac{1}{\sigma_n^2}\mathbf{I} = \begin{bmatrix} \ddots & & & \mathbf{0} & & & \\ & \frac{-\sigma_s^2}{\sigma_n^2(\sigma_n^2+2\sigma_s^2)} & 0 & \cdots & 0 & \frac{-\sigma_s^2}{\sigma_n^2(\sigma_n^2+2\sigma_s^2)} & \\ & 0 & \frac{-\sigma_s^2}{\sigma_n^2(\sigma_n^2+\sigma_s^2)} & \cdots & 0 & 0 & \\ \mathbf{0} & \vdots & & \ddots & & \vdots & \mathbf{0} \\ & 0 & \cdots & & \frac{-\sigma_s^2}{\sigma_n^2(\sigma_n^2+\sigma_s^2)} & 0 & \\ & \frac{-\sigma_s^2}{\sigma_n^2(\sigma_n^2+2\sigma_s^2)} & 0 & \cdots & 0 & \frac{-\sigma_s^2}{\sigma_n^2(\sigma_n^2+2\sigma_s^2)} & \\ & & & \mathbf{0} & & & \ddots \end{bmatrix} \quad (28)$$

$(H_1, i \notin S_\tau)$. Then,

$$\begin{cases} E[\overline{r}_i|H_1, i \in S_\tau] = E[\overline{s}_i^2] + E[\widetilde{s}_i^2] = \frac{\sigma_s^2}{2} + \frac{\sigma_s^2}{2} = \sigma_s^2, \\ E[\overline{r}_i|H_1, i \notin S_\tau] = E[\overline{s}_i]E[\overline{s}_{i+N_d}] + E[\widetilde{s}_i]E[\widetilde{s}_{i+N_d}] \\ = 0, \\ E[\overline{r}_i|H_0] = 0. \end{cases}$$

Next, we compute the variance of $\overline{R}_i$:

$$\text{Var}[\overline{R}_i] = \text{Var}\left[\frac{1}{K}\sum_{k=0}^{K-1}\overline{r}_{i+k(N_c+N_d)}\right] = \frac{1}{K}\text{Var}[\overline{r}_i]. \quad (33)$$

The variance of $\overline{r}_i$ in (33) is

$$\text{Var}[\overline{r}_i] = E[\overline{r}_i^2] - E[\overline{r}_i]^2,$$

where

$$\begin{aligned} E[\overline{r}_i^2] &= E\left[(\overline{x}_i\overline{x}_{i+N_d} + \widetilde{x}_i\widetilde{x}_{i+N_d})^2\right] \\ &= E[\overline{x}_i^2\overline{x}_{i+N_d}^2] + E[\widetilde{x}_i^2\widetilde{x}_{i+N_d}^2] + 2E[\overline{x}_i\overline{x}_{i+N_d}]E[\widetilde{x}_i\widetilde{x}_{i+N_d}] \\ &= 2E[\overline{x}_i^2\overline{x}_{i+N_d}^2] + 2E[\overline{x}_i\overline{x}_{i+N_d}]^2 \\ &= 2E\left[(\overline{s}_i+\overline{n}_i)^2(\overline{s}_{i+N_d}+\overline{n}_{i+N_d})^2\right] + 2E[\overline{s}_i\overline{s}_{i+N_d}]^2 \\ &= 2E\left[(\overline{s}_i^2+\overline{n}_i^2+2\overline{s}_i\overline{n}_i)(\overline{s}_{i+N_d}^2+\overline{n}_{i+N_d}^2+2\overline{s}_{i+N_d}\overline{n}_{i+N_d})\right] \\ &\quad + 2E[\overline{s}_i\overline{s}_{i+N_d}]^2 \\ &= 2E\left[\overline{s}_i^2\overline{s}_{i+N_d}^2 + \overline{s}_i^2\overline{n}_{i+N_d}^2 + 2\overline{s}_i^2\overline{s}_{i+N_d}\overline{n}_{i+N_d} + \overline{n}_i^2\overline{s}_{i+N_d}^2\right. \\ &\quad + \overline{n}_i^2\overline{n}_{i+N_d}^2 + 2\overline{n}_i^2\overline{s}_{i+N_d}\overline{n}_{i+N_d} + 2\overline{s}_i\overline{n}_i\overline{s}_{i+N_d}^2 \\ &\quad \left.+ 2\overline{s}_i\overline{n}_i\overline{n}_{i+N_d}^2 + 4\overline{s}_i\overline{n}_i\overline{s}_{i+N_d}\overline{n}_{i+N_d}\right] + 2E[\overline{s}_i\overline{s}_{i+N_d}]^2 \\ &= 2E\left[\overline{s}_i^2\overline{s}_{i+N_d}^2 + \overline{s}_i^2\overline{n}_{i+N_d}^2 + \overline{n}_i^2\overline{s}_{i+N_d}^2 + \overline{n}_i^2\overline{n}_{i+N_d}^2\right] \\ &\quad + 2E[\overline{s}_i\overline{s}_{i+N_d}]^2. \end{aligned}$$

Firstly, consider the case $i \in S_\tau$. Then, the variance of $\overline{r}_i$ is

$$\begin{aligned} &\text{Var}[\overline{r}_i|H_1, i \in S_\tau] = E[\overline{r}_i^2|H_1, i \in S_\tau] - E[\overline{r}_i|H_1, i \in S_\tau]^2 \\ &= 2E[\overline{s}_i^4 + \overline{s}_i^2\overline{n}_{i+N_d}^2 + \overline{n}_i^2\overline{s}_i^2 + \overline{n}_i^2\overline{n}_{i+N_d}^2] + 2E[\overline{s}_i^2]^2 - \sigma_s^4 \\ &= 2\left(3\left(\frac{\sigma_s^2}{2}\right)^2 + \frac{\sigma_s^2}{2}\frac{\sigma_n^2}{2} + \frac{\sigma_n^2}{2}\frac{\sigma_s^2}{2} + \frac{\sigma_n^2}{2}\frac{\sigma_n^2}{2} + \left(\frac{\sigma_s^2}{2}\right)^2\right) \\ &\quad - \sigma_s^4 \\ &= 2\left(\sigma_s^4 + \frac{\sigma_s^2\sigma_n^2}{2} + \frac{\sigma_n^4}{4}\right) - \sigma_s^4 \\ &= \sigma_s^4 + \sigma_s^2\sigma_n^2 + \frac{\sigma_n^4}{2}. \end{aligned}$$

Secondly, consider the case $i \notin S_\tau$. Then, the variance of $\overline{r}_i$ is

$$\begin{aligned} &\text{Var}[\overline{r}_i|H_1, i \notin S_\tau] = E[\overline{r}_i^2|H_1, i \notin S_\tau] - E[\overline{r}_i|H_1, i \notin S_\tau]^2 \\ &= 2\left(E[\overline{s}_i^2]E[\overline{s}_{i+N_d}^2] + E[\overline{s}_i^2]E[\overline{n}_{i+N_d}^2]\right. \\ &\quad + E[\overline{n}_i^2]E[\overline{s}_{i+N_d}^2] + E[\overline{n}_i^2]E[\overline{n}_{i+N_d}^2] \\ &\quad \left.+ E[\overline{s}_i]^2E[\overline{s}_{i+N_d}]^2\right) \\ &= 2\left(\frac{\sigma_s^2}{2}\frac{\sigma_s^2}{2} + \frac{\sigma_s^2}{2}\frac{\sigma_n^2}{2} + \frac{\sigma_n^2}{2}\frac{\sigma_s^2}{2} + \frac{\sigma_n^2}{2}\frac{\sigma_n^2}{2}\right) \\ &= 2\left(\frac{\sigma_s^2}{2} + \frac{\sigma_n^2}{2}\right)^2. \end{aligned}$$

Finally, consider H_0 $(s_i = 0)$:

$$\begin{aligned} \text{Var}[\overline{r}_i|H_0] &= E[\overline{r}_i^2|H_0] - E[\overline{r}_i|H_0]^2 \\ &= 2E[\overline{n}_i^2\overline{n}_{i+N_d}^2] = \frac{\sigma_n^4}{2}. \end{aligned}$$

To conclude,

$$\begin{cases} E[\overline{R}_i|H_1, i \in S_\tau] = \sigma_s^2, \\ E[\overline{R}_i|H_1, i \notin S_\tau] = 0, \\ E[\overline{R}_i|H_0] = 0, \\ \text{Var}[\overline{R}_i|H_1, i \in S_\tau] = \frac{1}{K}\left(\sigma_s^4 + \sigma_s^2\sigma_n^2 + \frac{\sigma_n^4}{2}\right), \\ \text{Var}[\overline{R}_i|H_1, i \notin S_\tau] = \frac{2}{K}\left(\frac{\sigma_s^2}{2} + \frac{\sigma_n^2}{2}\right)^2, \\ \text{Var}[\overline{R}_i|H_0] = \frac{\sigma_n^4}{2K}. \end{cases}$$

D. Moments of $\widetilde{R}_i$

First we compute

$$E[\widetilde{R}_i] = E\left[\frac{1}{K}\sum_{k=0}^{K-1}\widetilde{r}_{i+k(N_c+N_d)}\right] = E[\widetilde{r}_i],$$

where

$$\widetilde{r}_i = \text{Im}(x_i^* x_{i+N_d}) = \overline{x}_i\widetilde{x}_{i+N_d} - \widetilde{x}_i\overline{x}_{i+N_d}.$$

The expected value of $\widetilde{r}_i$ is

$$\begin{aligned} E[\widetilde{r}_i] &= E[\overline{x}_i\widetilde{x}_{i+N_d} - \widetilde{x}_i\overline{x}_{i+N_d}] \\ &= E[\overline{x}_i]E[\widetilde{x}_{i+N_d}] - E[\widetilde{x}_i]E[\overline{x}_{i+N_d}] = 0. \end{aligned}$$

Next, we compute the variance of $\widetilde{R}_i$:

$$\text{Var}[\widetilde{R}_i] = \text{Var}\left[\frac{1}{K}\sum_{k=0}^{K-1}\widetilde{r}_{i+k(N_c+N_d)}\right] = \frac{1}{K}\text{Var}[\widetilde{r}_i]. \quad (34)$$

The variance of $\widetilde{r}_i$ in (34) is

$$\begin{aligned}
&\text{Var}\left[\widetilde{r}_i\right] = E\left[\widetilde{r}_i^2\right] - E\left[\widetilde{r}_i\right]^2 = E\left[\left(\overline{x}_i\widetilde{x}_{i+N_d} - \widetilde{x}_i\overline{x}_{i+N_d}\right)^2\right]\\
&= E\left[\overline{x}_i^2\widetilde{x}_{i+N_d}^2 + \widetilde{x}_i^2\overline{x}_{i+N_d}^2 - 2\overline{x}_i\widetilde{x}_{i+N_d}\widetilde{x}_i\overline{x}_{i+N_d}\right]\\
&= E\left[\overline{x}_i^2\right]E\left[\widetilde{x}_{i+N_d}^2\right] + E\left[\widetilde{x}_i^2\right]E\left[\overline{x}_{i+N_d}^2\right]\\
&\quad - 2E\left[\overline{x}_i\overline{x}_{i+N_d}\right]E\left[\widetilde{x}_{i+N_d}\widetilde{x}_i\right]\\
&= 2E\left[\overline{x}_i^2\right]^2 - 2E\left[\overline{s}_i\overline{s}_{i+N_d}\right]^2.
\end{aligned}$$

Firstly, consider the case $i \in S_\tau$. Then, the variance of $\widetilde{r}_i$ is

$$\begin{aligned}
&\text{Var}\left[\widetilde{r}_i|H_1, i\in S_\tau\right]\\
&= 2E\left[\overline{x}_i^2|H_1, i\in S_\tau\right]^2 - 2E\left[\overline{s}_i\overline{s}_{i+N_d}|H_1, i\in S_\tau\right]^2\\
&= 2\left(\frac{\sigma_s^2}{2}+\frac{\sigma_n^2}{2}\right)^2 - 2E\left[\overline{s}_i^2\right]^2\\
&= 2\left(\frac{\sigma_s^2}{2}+\frac{\sigma_n^2}{2}\right)^2 - 2\left(\frac{\sigma_s^2}{2}\right)^2\\
&= \frac{\sigma_s^4}{2}+\frac{\sigma_n^4}{2}+\sigma_s^2\sigma_n^2 - \frac{\sigma_s^4}{2} = \frac{\sigma_n^4}{2}+\sigma_s^2\sigma_n^2.
\end{aligned}$$

Secondly, consider the case $i \notin S_\tau$. Then, the variance of $\widetilde{r}_i$ is

$$\begin{aligned}
\text{Var}\left[\widetilde{r}_i|H_1, i\notin S_\tau\right] &= 2E\left[\overline{x}_i^2|H_1, i\notin S_\tau\right]^2\\
&\quad - 2E\left[\overline{s}_i\overline{s}_{i+N_d}|H_1, i\notin S_\tau\right]^2\\
&= 2\left(\frac{\sigma_s^2}{2}+\frac{\sigma_n^2}{2}\right)^2.
\end{aligned}$$

Finally, consider H_0:

$$\text{Var}\left[\widetilde{r}_i|H_0\right] = 2E\left[\overline{x}_i^2|H_0\right]^2 = 2E\left[\overline{n}_i^2\right]^2 = \frac{\sigma_n^4}{2}.$$

To conclude,

$$\begin{cases}
E\left[\widetilde{R}_i|H_1, i\in S_\tau\right] = 0,\\
E\left[\widetilde{R}_i|H_1, i\notin S_\tau\right] = 0,\\
E\left[\widetilde{R}_i|H_0\right] = 0,\\
\text{Var}\left[\widetilde{R}_i|H_1, i\in S_\tau\right] = \frac{1}{K}\left(\frac{\sigma_n^4}{2}+\sigma_s^2\sigma_n^2\right),\\
\text{Var}\left[\widetilde{R}_i|H_1, i\notin S_\tau\right] = \frac{2}{K}\left(\frac{\sigma_s^2}{2}+\frac{\sigma_n^2}{2}\right)^2,\\
\text{Var}\left[\widetilde{R}_i|H_0\right] = \frac{\sigma_n^4}{2K}.
\end{cases}$$

E. Derivation of Cov $\left[\overline{R}_i, \widetilde{R}_i\right]$

The covariance of $\overline{R}_i, \widetilde{R}_i$ is

$$\text{Cov}\left[\overline{R}_i, \widetilde{R}_i\right] = E\left[\overline{R}_i\widetilde{R}_i\right] - E\left[\overline{R}_i\right]E\left[\widetilde{R}_i\right].$$

Since $E\left[\widetilde{R}_i\right] = 0$, we have

$$\begin{aligned}
&\text{Cov}\left[\overline{R}_i, \widetilde{R}_i\right] = E\left[\overline{R}_i\widetilde{R}_i\right]\\
&= E\left[\frac{1}{K}\sum_{k=0}^{K-1}\overline{r}_{i+k(N_c+N_d)}\frac{1}{K}\sum_{l=0}^{K-1}\widetilde{r}_{i+l(N_c+N_d)}\right]\\
&= \frac{1}{K^2}\sum_{k=0}^{K-1}\sum_{l=0}^{K-1}E\left[\overline{r}_{i+k(N_c+N_d)}\widetilde{r}_{i+l(N_c+N_d)}\right].
\end{aligned}$$

The expected value of the product $\overline{r}_k\widetilde{r}_l$ is computed as follows

$$\begin{aligned}
&E\left[\overline{r}_k\widetilde{r}_l\right] = E\left[\left(\overline{x}_k\overline{x}_{k+N_d} + \widetilde{x}_k\widetilde{x}_{k+N_d}\right)\left(\overline{x}_l\widetilde{x}_{l+N_d} - \widetilde{x}_l\overline{x}_{l+N_d}\right)\right]\\
&= E\left[\overline{x}_k\overline{x}_{k+N_d}\overline{x}_l\widetilde{x}_{l+N_d} - \overline{x}_k\overline{x}_{k+N_d}\widetilde{x}_l\overline{x}_{l+N_d}\right.\\
&\quad\left. + \widetilde{x}_k\widetilde{x}_{k+N_d}\overline{x}_l\widetilde{x}_{l+N_d} - \widetilde{x}_k\widetilde{x}_{k+N_d}\widetilde{x}_l\overline{x}_{l+N_d}\right]\\
&= E\left[\overline{x}_k\overline{x}_{k+N_d}\overline{x}_l\right]E\left[\widetilde{x}_{l+N_d}\right] - E\left[\overline{x}_k\overline{x}_{k+N_d}\overline{x}_{l+N_d}\right]E\left[\widetilde{x}_l\right]\\
&\quad + E\left[\widetilde{x}_k\widetilde{x}_{k+N_d}\widetilde{x}_{l+N_d}\right]E\left[\overline{x}_l\right] - E\left[\widetilde{x}_k\widetilde{x}_{k+N_d}\widetilde{x}_l\right]E\left[\overline{x}_{l+N_d}\right]\\
&= 0.
\end{aligned}$$

Thus, Cov $\left[\overline{R}_i, \widetilde{R}_i\right] = 0$ for all cases (H_0, H_1, $i \in S_\tau$, $i \notin S_\tau$).

REFERENCES

[1] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in *Proc. Allerton Conference on Communication, Control, and Computing*, October 2004, pp. 1662–1671.

[2] E. G. Larsson and M. Skoglund, "Cognitive radio in a frequency-planned environment: some basic limits," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4800–4806, December 2008.

[3] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, no. 4, pp. 523–531, April 1967.

[4] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: half a century of research," *Signal Processing*, vol. 86, no. 4, pp. 639–697, 2006.

[5] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *IEEE Signal Processing Mag.*, vol. 8, no. 2, pp. 14–36, 1991.

[6] A. V. Dandawaté and G. B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Trans. Signal Process.*, vol. 42, no. 9, pp. 2355–2369, 1994.

[7] Motorola Inc., "Long term evolution (LTE): Overview of LTE air-interface," Technical White Paper, http://business.motorola.com/experiencelte/pdf/LTEAirInterfaceWhitePaper.pdf, 2007.

[8] DVB Project Office, "DVB fact sheet - digital terrestrial television," http://www.dvb.org/technology/fact_sheets/DVB-T-Fact-Sheet.0709.pdf, July 2009.

[9] M. Oner and F. Jondral, "Cyclostationarity based air interface recognition for software radio systems," in *Proc. IEEE Radio and Wireless Conference*, pp. 263–266, Sept. 2004.

[10] S. Chaudhari, V. Koivunen, and H. V. Poor, "Autocorrelation-based decentralized sequential detection of OFDM signals in cognitive radios," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2690–2700, 2009.

[11] Z. Lei and F. Chin, "OFDM signal sensing for cognitive radios," in *Proc. IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–5, Sept. 2008.

[12] Huawei Technologies and UESTC, "Sensing scheme for DVB-T," *IEEE Std.802.22-06/0127r1*, Jul. 2006.

[13] J. Lundén, V. Koivunen, A. Huttunen, and H. V. Poor, "Collaborative cyclostationary spectrum sensing for cognitive radio systems," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4182–4195, Nov. 2009.

[14] R. Tandra and A. Sahai, "Fundamental limits on detection in low SNR under noise uncertainty," in *Proc. IEEE International Conference on Wireless Networks, Communications and Mobile Computing*, June 13-16 2005, vol. 1, pp. 464–469.

[15] 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.

[16] S. Kraut, L.L. Scharf, and L.T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 1–16, Jan 2001.

[17] Z. Quan, S. Cui, H.V. Poor, and A. Sayed, "Collaborative wideband sensing for cognitive radios," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 60–73, November 2008.

[18] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I*, John Wiley and Sons, Inc., 1968.

[19] E. Axell and E. G. Larsson, "A Bayesian approach to spectrum sensing, denoising and anomaly detection," in *Proc. IEEE ICASSP 2009*, pp. 2333–2336, 19-24 Apr. 2009.

[20] C.R. Rao, *Linear Statistical Inference and Its Applications*, Wiley, 1965, Second corrected printing, Aug. 1968.

[21] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers, Scientists and Mathematicians*, Springer-Verlag New York, Inc., 2006.

[22] S. Shellhammer and G. Chouinard, *Spectrum Sensing Requirements Summary*, IEEE Std.802.22-06/0089r4, Jun. 2006.

Erik Axell is a Ph.D. student with the Communication Systems Division of the Department of Electrical Engineering, Linköping University, Sweden, since March 2006. He received his Lic. Eng. degree in december 2009 and his M.Sc. degree in 2005, from Linköping University, Sweden. From February 2005 to December 2005 he was with the Ericsson Research group in Linköping, Sweden. His main research interests are within the area of wireless communications and especially spectrum sensing and system aspects of cognitive radio.

Erik G. Larsson is Professor and Head of the Division for Communication Systems in the Department of Electrical Engineering (ISY) at Linköping University (LiU) in Linköping, Sweden. He joined LiU in September 2007. He has previously been Associate Professor (Docent) at the Royal Institute of Technology (KTH) in Stockholm, Sweden, and Assistant Professor at the University of Florida and the George Washington University, USA.

His main professional interests are within the areas of wireless communications and signal processing. He has published some 60 journal papers on these topics, he is co-author of the textbook *Space-Time Block Coding for Wireless Communications* (Cambridge Univ. Press, 2003) and he holds 10 patents on wireless technology.

He has been Associate Editor for the *IEEE Transactions on Signal Processing*, the *IEEE Signal Processing Letters* and the *IEEE Transactions on Vehicular Technology*. He is a member of the IEEE Signal Processing Society SAM and SPCOM technical committees. He is active in conference organization, most recently as the technical area chair for communication systems of the Asilomar Conference on Signals, Systems and Computers 2010, and as general co-chair of CrownCom 2010.