

A Spectrum Sensing Technique for Cognitive Radios in the Presence of Harmonic Images

N.A. Moseley, E.A.M Klumperink and B. Nauta
IC Design Group, CTIT Research Institute, University of Twente
PO Box 217, 7500 AE Enschede
The Netherlands
n.a.moseley@ewi.utwente.nl

Abstract—Harmonic downmixing is an important effect that must be taken into account when performing sensitive spectrum sensing using direct-conversion receivers. When the local oscillator waveform contains harmonics of the fundamental frequency, the quadrature mixer in the receiver will downconvert RF signals found at these harmonics, termed harmonic images, in addition to the RF signals around the fundamental frequency. The harmonic images will be detected by power spectral density based sensing algorithms and will cause certain parts of the desired spectrum to be mistakenly flagged as occupied. Although harmonic downmixing is important to consider, it is an often neglected effect.

This paper presents a harmonic rejection spectral sensing technique, that exploits two quadrature mixers. The mixers work with different LO frequencies, which decorrelate the harmonic images so that cross-correlation of their outputs renders an improved spectral estimate. In addition to rejecting the harmonic images, spurious signals entering the receiver through the analog baseband inputs will also be rejected. The frequency resolution of the spectral estimate is scalable and the rejection of the harmonic images increases 15 dB per 1000-fold increase of the correlation time. The complexity of the algorithm is analyzed and its performance is shown by means of simulations. The effect of I/Q imbalance is also taken into account.

I. INTRODUCTION

Cognitive radios ideally support a large frequency range, one that spans multiple octaves. Generating the local oscillator (LO) signals for such a wide range of frequencies is predominantly achieved using digital circuits[1]. For instance, digital frequency dividers are used to divide the output of a high-frequency oscillator. In this way, a frequency synthesizer that spans a multiple-octave frequency range can be realized. However, such a frequency synthesizer generates a digital square-wave LO signal, which contains many harmonics in addition to the fundamental frequency.

The direct-conversion receiver has a low external component count compared to other receiver architectures, such as the superheterodyne. It is therefore the architecture of choice when high-integration and low-cost are desired. Such features are important in consumer applications, such as PDAs, mobile phones and laptops. However, when receiver's mixers are driven by digital circuits, they suffer from

*harmonic downmixing*¹.

Harmonic downmixing is an undesirable effect whereby RF signals found at multiples of the local oscillator frequency, termed *harmonic images*, are downmixed to baseband, in addition to the desired signal. The effect is caused by harmonics in the local oscillator signal, which have significant strengths for a square-wave LO. The harmonic images cause interference to the desired signals and will degrade the bit-error rate (BER) at the decoder, or even make the desired signals undecodable. Harmonic downmixing also causes problems in spectrum sensing applications, such as found in cognitive radio networks. The spectrum sensing algorithm, which is often based on energy detection or power spectral density measurements, will mistake a harmonic image signal, which is outside the band of interest, for a signal that is within the band of interest. As a result, certain parts of the spectrum are erroneously flagged as occupied while they are empty in reality.

The effect of harmonic downmixing is neglected in most papers as they assume perfect sinusoidal LO signals and perfect multiplying mixers, or sufficient RF filtering. However, sufficient RF filtering is expensive, bulky and can be difficult to implement. A different approach could be to remove the harmonics from the LO signal by filtering. However, this leads to the same problems of bulky, expensive filters as the RF filtering case. Clearly, these solutions are not desirable when aiming for low-cost or embedded applications.

Also note that a perfectly sinusoidal LO signal does not always solve the harmonic downmixing problem because many practical mixer circuits, such as switching mixers, cause nonlinear distortion of the LO signal. These distortions add harmonics to the LO signal, which again leads to harmonic downmixing.

Disentangling the harmonic images from the desired signal by wideband sampling of the mixer output, which contains multiple versions of the harmonic images and the desired signal, is possible in theory. The sampling bandwidth required would be at least $x + 1$ times the LO frequency, where x is the number of harmonic images to be removed. Unfortunately,

¹All radio architectures can suffer from harmonic downmixing. Here we limit ourselves to direct-conversion receivers as the harmonic downmixing problem is most severe in such receivers.

such wideband analog-to-digital converters (ADCs) are beyond the current state of the art, except for very low LO frequencies.

Given the above, it is clear that a different way of dealing with harmonic downmixing is needed.

In this paper, we present an algorithm based on cross-correlation of two baseband signals, r_1 and r_2 . Both contain signals from the desired spectrum and from the harmonic images. A frequency offset Δf is used in the analog frontend to frequency shift the harmonic images and the desired signal by different amounts. After a frequency shift, $-\Delta f$, in the digital domain, the desired signal in r_1 and r_2 cross correlate. This principle allows correct spectrum sensing under conditions where harmonic images are present.

This paper is organized as follows. In Section II, a mixer model is developed that shows how multiple RF signals are downmixed to baseband and how they relate to the Fourier series of the LO waveform. Then, the model is extended to quadrature mixers. In Section III, the principle of using a frequency offset to decorrelate the desired signals from the harmonic images is presented. In addition, the sensing algorithm based on cross-correlation is explained and is formulated in terms of the quadrature mixer model from Section II. In Section IV, the performance of the algorithm is analyzed and evaluated using simulations, while the final section offers the conclusions.

II. A MIXER MODEL INCORPORATING HARMONIC DOWNMIXING

When a multiplying mixer's LO port is driven by a pulse-like signal, such as a digitally generated square wave, multiple frequencies present at the RF port are downconverted to baseband. In addition to the desired mixing product, additional mixing products appear at the mixer's baseband output. While the desired mixing product is governed by the fundamental frequency of the LO signal, the other products are governed by the higher harmonics present in the LO signal. If a switching mixer is used instead of a multiplying mixer, the same argument holds as the *effective* LO waveform is pulse-like regardless of the actual LO waveform.

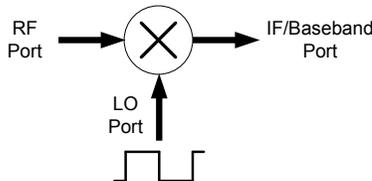


Fig. 1. A mixer driven by a square wave at the LO port.

The mixing products found at the IF, or baseband, output of the mixer, depend on the Fourier series of the effective LO waveform. Given an RF signal $x_{RF}(t)$ and an effective LO waveform $s_{LO}(t)$, the signal at the output, $y(t)$ of a mixer is given by the following equation:

$$y(t) = x_{RF}(t) \cdot s_{LO}(t) \quad (1)$$

The RF signal can be written as a sum of baseband-equivalent signals, where each baseband-equivalent signal represents a piece of RF spectrum centered around the harmonics of the LO waveform. In this way, each mixing product can be described separately. Given a set of band-limited baseband-equivalent signals $z_p(t)$, where $p \in [1, 2, 3, \dots, \infty)$, the RF signal is expressed as the following sum:

$$x_{RF}(t) = \Re \left\{ \sum_{p=1}^{\infty} z_p(t) \cdot e^{j \cdot 2\pi \cdot p \cdot f_{LO} \cdot t} \right\}, \quad (2)$$

where \Re denotes taking the real part of its complex argument and f_{LO} is the LO frequency in Hertz. The signal $z_p(t)$ represents a part of the RF spectrum, namely RF signals found around a frequency of $p \cdot f_{LO}$ Hz.

The \Re operator may be removed by realizing that $\Re\{x\} = \frac{1}{2}(x + x^*)$, where $(\cdot)^*$ denotes the complex conjugate. Therefore, (2) can be simplified to give:

$$x_{RF}(t) = \frac{1}{2} \sum_{p=1}^{\infty} (z_p(t) \cdot e^{j \cdot 2\pi \cdot p \cdot f_{LO} \cdot t} + z_p^*(t) \cdot e^{-j \cdot 2\pi \cdot p \cdot f_{LO} \cdot t}) \quad (3)$$

Given that the LO waveform is periodic, it is fully described by its complex-valued Fourier series coefficients $\{c_0, c_1, \dots, c_\infty\}$:

$$s_{LO}(t) = c_0 + \sum_{p=1}^{\infty} (c_p \cdot e^{j \cdot 2\pi \cdot f_{LO} \cdot p \cdot t} + c_p^* \cdot e^{-j \cdot 2\pi \cdot f_{LO} \cdot p \cdot t}) \quad (4)$$

By combining (3) and (4), an expression is obtained that shows the relation of the downmixed products to the baseband-equivalent signals $z_p(t)$ and the Fourier coefficients of the mixer's LO waveform:

$$y(t) = \frac{1}{2} \sum_{p=1}^{\infty} (c_p^* \cdot z_p(t) + c_p \cdot z_p^*(t)) \quad (5)$$

+ high frequency mixing products.

In a direct-conversion receiver, the high frequency mixing products in (5) are removed by a post-mixer filter and/or the anti-aliasing lowpass filters before the A/D converters. In receiver applications, (5) may thus be simplified to:

$$y(t) = \frac{1}{2} \sum_{p=1}^{\infty} (c_p^* \cdot z_p(t) + c_p \cdot z_p^*(t)) \quad (6)$$

The purpose of a receiver is to recover $z_1(t)$, which is the desired signal spectrum. Note that, in the general case, $z_1(t)$ is a complex-valued signal, while $y(t)$ is real-valued. It is therefore not possible to fully recover $z_1(t)$ from $y(t)$ alone. Recovery of $z_1(t)$ is only possible by employing quadrature techniques.

A. Quadrature downconversion

A quadrature receiver uses two mixers to recover $z_1(t)$, where each mixer is driven by a different LO waveform. The two LO waveforms have the same shape but they are phase-shifted by 90 degrees with respect to each other. The phase-shift allows $z_1^*(t)$ to be removed while retaining the desired signal spectrum $z_1(t)$.

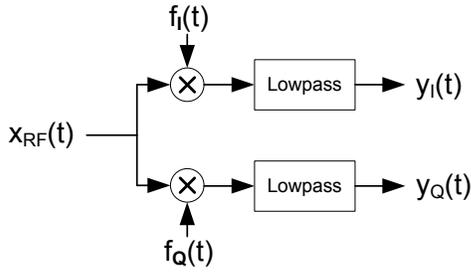


Fig. 2. A quadrature downmixer employing two mixers. The LO signals $f_I(t)$ and $f_Q(t)$ are 90-degree out of phase with respect to each other.

We will now apply the mixer model to a quadrature receiver architecture and show the dependence of the received baseband signal $r(t)$, given the baseband-equivalent signals $z_p(t)$, and the Fourier series coefficients of the two LO waveforms. For generality, we use different variables for the Fourier coefficients of the two LO waveforms, which allows the LO waveforms to differ in shape and not only in phase.

Figure 2 shows the quadrature downconverter principle. Given the mixer model (6), the in-phase output $y_I(t)$ and the quadrature phase output $y_Q(t)$ can be written as:

$$\begin{aligned} y_I(t) &= \frac{1}{2} \sum_{p=1} (c_p^* \cdot z_p(t) + c_p \cdot z_p^*(t)) \\ y_Q(t) &= \frac{1}{2} \sum_{p=1} (d_p^* \cdot z_p(t) + d_p \cdot z_p^*(t)), \end{aligned} \quad (7)$$

where c_p and d_p are the Fourier series coefficients of the LO signals $f_I(t)$ and $f_Q(t)$, respectively. The two real-valued outputs $y_I(t)$ and $y_Q(t)$ form a complex-valued baseband signal $r(t)$, where $r(t)$ may be expressed as:

$$\begin{aligned} r(t) &= y_I(t) + j \cdot y_Q(t) \\ &= \frac{1}{2} \sum_{p=1} ((c_p - j \cdot d_p)^* \cdot z_p(t) + (c_p + j \cdot d_p) \cdot z_p^*(t)) \end{aligned} \quad (8)$$

As stated previously, the purpose of a receiver is to recover the baseband signal, or spectrum, $z_1(t)$. By examining (8), it is clear that this is possible when $c_1 = -j \cdot d_1$ so that $z_1^*(t)$ is removed and $c_p = 0, d_p = 0$ for any $p > 1$ so that the harmonic images are removed. Note that the conjugated terms $z_p^*(t)$ constitute interference and are caused by I/Q imbalance. We refer to them as *I/Q imbalance images*. Appendix B gives the I/Q imbalance image suppression as function of $c_p - j \cdot d_p$ and $c_p + j \cdot d_p$.

Unfortunately, the above requirements are only met when the effective LO waveform is a perfect sine wave. A mixer based on switches cannot be used as this always leads to a pulse-like effective LO waveform. In fact, only a perfect multiplying mixer driven by a perfect sine wave will achieve the conditions. Unfortunately, both are impossible to realize exactly and very difficult to approximate well in practice.

Traditionally, receivers employ bandpass filters at the RF port of the mixer to remove the harmonic image signals.

Direct-conversion receivers with a wide frequency range, such as desired for cognitive radios, would require a tunable RF filter [2] or a bank of selectable bandpass filters. However, a highly-selective tunable filter with a wide tuning range is very challenging, if not impossible, to make, owing to the many tunable components that are required. In addition, the frequency response of such filters tends to deviate from the optimum shape as they are tuned across the band, thereby reducing their selectivity. Also, a bank of selectable bandpass filters is bulky, expensive and therefore unattractive in low-cost or highly integrated systems.

Given the above, it would seem that switching mixers or pulse-like LO waveforms cannot be used effectively in direct-conversion receivers. However, analog multi-path techniques exist that employ multiple mixers in parallel to obtain an effective LO waveform that lacks certain harmonics. Therefore, certain harmonic images, most notably the 3rd, 5th and all even-order harmonic images, are suppressed, thereby reducing the demands on the RF filter. This technique was demonstrated by [3] for transmitters and by [4] for receivers. The multi-path technique offers a typical suppression of around 40 dB. This figure is limited by mismatches within the analog circuits. To further enhance the harmonic image suppression of the analog frontend, we have proposed a digital compensator based on blind adaptive signal processing which increases the suppression of the strongest harmonic image [5], typically by 40 dB.

As the multi-path technique does not completely remove the targeted harmonic images, there will be residual energy from the harmonic images present in the baseband signal $r(t)$. Performing spectrum sensing using $r(t)$, based on the power spectral density, leads to the detection of signals related to the harmonic images, which are false positives as they these signals do not come from the band of interest, $z_1(t)$.

In this paper we describe a method for performing spectrum sensing based on cross-correlation of two baseband signals. The method detects the presence of signals in the desired spectrum² while it rejects signals from the harmonic images.

III. SPECTRUM SENSING BASED ON TWO BASEBAND SIGNALS

The signal detection scheme is based on obtaining two discrete-time observations of the RF spectrum. The observation signals, $r_1(n)$ and $r_2(n)$ are produced by two quadrature mixers, as shown in Fig. 3. The local oscillators LO1 and LO2 are tuned to f_{LO} and $f_{LO} + \Delta f$ Hz, respectively. The A/D converters use a sampling rate of F_s samples per second. Therefore, the bandwidth, W_{bb} , of each complex-valued baseband signal is equal to F_s .

Owing to harmonic downmixing, $r_1(n)$ represents multiple parts of the RF spectrum simultaneously. In addition to representing an RF part centered around f_{LO} Hz, it will also

²In normal receiver applications we wish to obtain the desired signal. In spectrum sensing, however, we wish to obtain a desired spectrum, regardless whether there is an RF signal present or not. Note that the terms refer to the same baseband-equivalent signal $z_1(t)$.

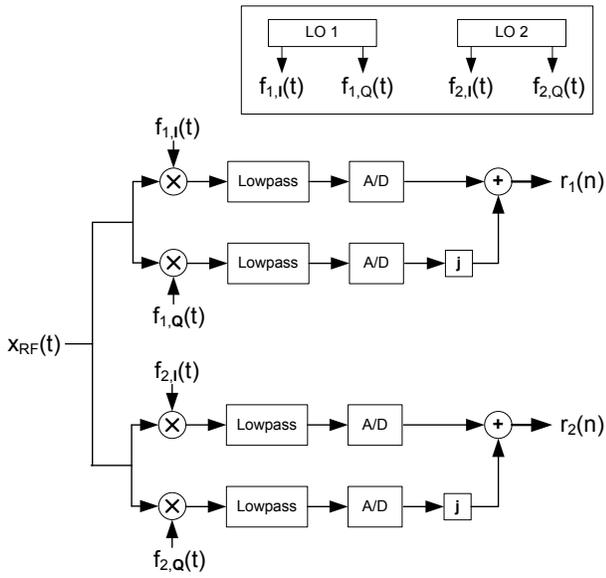


Fig. 3. The proposed spectrum sensing frontend, which produces two discrete-time complex-valued baseband outputs $r_1(n)$ and $r_2(n)$.

represent RF signals centered around $k \cdot f_{LO}$ Hz, where k is any positive integer. Or more precisely, $r_1(n)$ contains signals from the following parts of the RF spectrum:

$$\left\{ k \cdot f_{LO} + x \mid |x| < \frac{1}{2}W_{bb}, k \in [0, 1, 2, \dots, \infty) \right\} \quad (9)$$

The second baseband signal, $r_2(n)$ is generated using an LO frequency of $f_{LO} + \Delta f$ Hz, where Δf is a frequency offset. In effect, $r_2(n)$ represents a different part of the spectrum, namely:

$$\left\{ k \cdot (f_{LO} + \Delta f) + x \mid |x| < \frac{1}{2}W_{bb}, k \in [0, 1, 2, \dots, \infty) \right\} \quad (10)$$

By comparing (9) to (10), it can be seen that the harmonic images ($k > 1$) experience a greater frequency shift, namely $(k - 1) \cdot \Delta f$ more, than the desired signal spectrum ($k = 1$). This effect is exploited by our algorithm to discern the desired signal spectrum from the harmonic images.

By choosing Δf to be much smaller than W_{bb} , a large part of the desired signal spectrum in $r_2(n)$ overlaps with that in $r_1(n)$. The overlapping part is $W_{bb} - 2\Delta f$ Hz wide. Figure 4 illustrates the different shifting amounts of each harmonic image caused by the frequency offset. The desired signals, shown in black, shift by Δf , while the second harmonic image signal, shown in white, shifts by twice that amount.

We will now restate the above in a mathematical sense.

A. Describing $r_1(n)$ and $r_2(n)$

The baseband signal $r_1(n)$ can be written as a weighted sum of the desired signal spectrum $z_1(n)$ and the harmonic

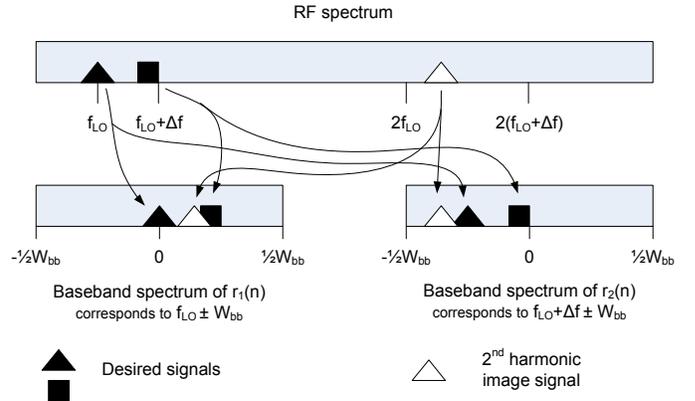


Fig. 4. Position of $r_1(n)$ and $r_2(n)$ in the RF spectrum and their corresponding baseband spectra.

images $\{z_p(n) \mid p > 1\}$:

$$r_1(n) = \sum_{p=1}^{\infty} \alpha_p \cdot z_p(n) + \beta_p \cdot z_p^*(n), \quad (11)$$

where $\alpha_p = (c_p - j \cdot d_p)^*$ and $\beta_p = (c_p + j \cdot d_p)$. Furthermore, c_p and d_p are the Fourier series coefficients of the quadrature mixer associated with LO1.

The second baseband signal, $r_2(n)$, can be written as a weighted sum of different baseband-equivalent signals, \tilde{z}_p .

$$r_2(n) = \sum_{p=1}^{\infty} \epsilon_p \cdot \tilde{z}_p(n) + \gamma_p \cdot \tilde{z}_p^*(n), \quad (12)$$

where $\epsilon_p = (\tilde{c}_p - j \cdot \tilde{d}_p)^*$ and $\gamma_p = (\tilde{c}_p + j \cdot \tilde{d}_p)$. In addition, \tilde{c}_p and \tilde{d}_p are the Fourier series coefficients of the quadrature mixer associated with LO2.

As $z_p(n)$ and \tilde{z}_p must partially overlap in the RF frequency domain, \tilde{z}_p is also partially defined by $z_p(n)$ and their relation can be described as follows:

$$\tilde{z}_p(n) = F_p \{ z_p(n) \cdot e^{-j2\pi \cdot f_d \cdot p \cdot n} \} + \text{additional image components}, \quad (13)$$

where $f_d = \frac{\Delta f}{F_s}$ and the $F_p\{\cdot\}$ operator filters any components that are not shared by both p^{th} -order baseband signals z_p and \tilde{z}_p . It is not necessary to define the additional image components in detail, as they will not correlate with $z_p(n)$. Only their combined variance, σ_{add}^2 , is of interest.

Given the signal descriptions presented above, we proceed to apply our algorithm, which is based on cross-correlation.

B. Cross-correlation based signal detection

Before applying a cross-correlation algorithm to $r_1(n)$ and $r_2(n)$, which detects the signal power of common signals in each subband, the desired spectrum within $r_2(n)$ must be frequency aligned. As shown by (13), the desired signal spectrum $z_1(n)$ is offset by Δf with respect to $\tilde{z}_1(n)$. By

applying a frequency translation to $r_2(n)$, a frequency-aligned version, $s(n)$ is obtained:

$$s(n) = r_2(n) \cdot e^{j \cdot 2\pi \cdot f_d \cdot n} + \text{additional image components} \quad (14)$$

Note that the frequency shift caused by (14) is common for all signals within $r_2(n)$. Equation (14) can be written in terms of the baseband signals $z_p(n)$ using (13):

$$s(n) = \left\{ \sum_{p=1}^{\infty} \epsilon_p \cdot \tilde{z}_p(n) + \gamma_p \cdot \tilde{z}_p^*(n) \right\} \cdot e^{j \cdot 2\pi \cdot f_d \cdot n} + \text{additional image components}, \quad (15)$$

which can be further expanded to:

$$s(n) = \left\{ \sum_{p=1}^{\infty} \epsilon_p \cdot F_p \left\{ z_p(n) \cdot e^{-j \cdot 2\pi \cdot f_d \cdot p \cdot n} \right\} \right\} \cdot e^{j \cdot 2\pi \cdot f_d \cdot n} + \left\{ \sum_{p=1}^{\infty} \gamma_p \cdot F_p \left\{ z_p^*(n) \cdot e^{j \cdot 2\pi \cdot f_d \cdot p \cdot n} \right\} \right\} \cdot e^{j \cdot 2\pi \cdot f_d \cdot n} + \text{additional image components} \quad (16)$$

Moving the frequency translation into the sum leads to:

$$s(n) = \sum_{p=1}^{\infty} \epsilon_p \cdot F_p \left\{ z_p(n) \cdot e^{-j \cdot 2\pi \cdot f_d \cdot (p-1) \cdot n} \right\} + \sum_{p=1}^{\infty} \gamma_p \cdot F_p \left\{ z_p^*(n) \cdot e^{j \cdot 2\pi \cdot f_d \cdot (p+1) \cdot n} \right\} + \text{additional image components} \quad (17)$$

Note that for $p = 1$, (17) shows that $z_1(n)$ no longer has a frequency translation. However, owing to the fact that $z_1(n)$ is not represented in its entirety by $r_2(n)$ (and therefore $s(n)$), the $F_p\{\cdot\}$ operator is still needed. As stated previously, $r_2(n)$ and $r_1(n)$ overlap by $W_{bb} - 2\Delta f$ Hz. Also note that the I/Q imbalance image, $z_1^*(n)$, in $s(n)$ does not align with itself in $r_1(n)$. The same holds for the harmonic images ($p > 1$).

The signals $r_1(n)$ and $s(n)$ are decomposed into subbands by a filterbank so that each subband can be cross correlated separately thereby producing frequency dependent cross-correlation information. The decomposition can be done efficiently using the Fast Fourier Transform (FFT). The FFT can be viewed as having N parallel subband channels, which share a common input $x(n)$. A single channel is shown in Fig. 5.

Each channel consists of a complex multiplier, an N -point FIR filter H and a decimation-by- N stage. A new subband output, $X(k)$ is produced for every N input samples of $x(n)$. The impulse response of the filter H is an N -point all-one vector, $[1, 1, 1, \dots]$, known as a rectangular window. It is well-known that the bin-to-bin isolation of the rectangular window is poor as the first sidelobe is only 13 dB down, and that applying a non-rectangular window to the FFT input vector, $[x(n) \dots x(n+N-1)]$, allows trading frequency resolution for better isolation.

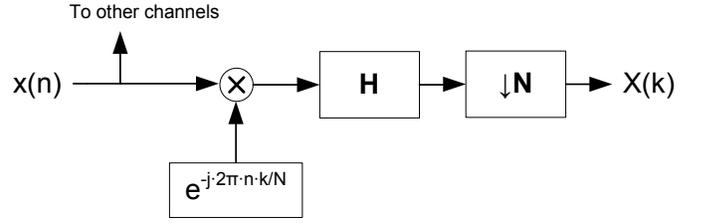


Fig. 5. A time domain description of the k^{th} bin, or channel, (out of N bins/channels) of an N -point FFT. Each channel consists of a complex mixing operation, a filter H and a decimation-by- N stage.

As the main focus of this paper is not on filter banks, we will not go into the window design in great detail. We refer to [6] and for a thorough treatment of filter banks and to [7] for their application in cognitive radios.

Here, we limit ourselves to mention that we apply a Nuttall window [8] given by (18), which has the first sidelobe at -93.3 dB. Unfortunately, it offers a frequency resolution which is four times lower. Figure 6 shows the spectrum of the rectangular window and the Nuttall window, where the differences in mainlobe width (which determines the frequency resolution) and sidelobe level (which determines the bin-to-bin isolation) become evident.

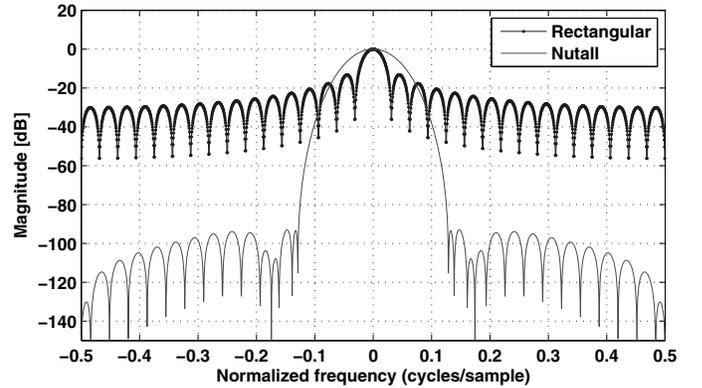


Fig. 6. The spectrum of the rectangular window and a Nuttall window, both are $N=32$ points long.

$$h_{nuttall}(n) = a_0 - a_1 \cdot \cos\left(\frac{2\pi \cdot n}{N-1}\right) + a_2 \cdot \cos\left(\frac{4\pi \cdot n}{N-1}\right) - a_3 \cdot \cos\left(\frac{6\pi \cdot n}{N-1}\right) \quad (18)$$

$$a_0 = 0.355768 \quad a_1 = 0.487396$$

$$a_2 = 0.144232 \quad a_3 = 0.012604$$

Given an N -point FFT and the Nuttall window, each subband is approximately $\frac{8 \cdot W_{bb}}{N}$ Hz wide.

We proceed by cross correlating the k^{th} subband of $r_1(n)$, which is denoted by $x_k(i)$, with the k^{th} subband of $s(n)$, which is denoted by $y_k(i)$. As there are N subbands, the valid range

of k is $[0, \dots, N-1]$. Note that the sampling rate of $x_k(i)$ and $y_k(i)$ has been reduced to $\frac{F_s}{N}$ because of the decimation-by- N of the FFT, hence the need for the new time index, i .

The zero-lag cross-correlation of the complex signals x_k and y_k is defined as:

$$R_{x_k, y_k} = E\{x_k \cdot y_k^*\}, \quad (19)$$

where $E\{\}$ denotes the expectation operator.

Assuming that x_k and y_k are stationary ergodic processes, their cross-correlation can be estimated based on a limited number of samples. Given N samples of x_k and y_k , the following equation is used to obtain an unbiased estimate \widehat{R}_{x_k, y_k} of R_{x_k, y_k} :

$$\widehat{R}_{x_k, y_k} = \frac{1}{N} \sum_{i=0}^{N-1} x_k(i) \cdot y_k(i)^* \quad (20)$$

See Appendix A for proof relating to the unbiasedness of the estimator.

Let x_k denote the signals in the k^{th} subband of r_1 , y_k denote the signals in the k^{th} subband of s , and let $z_{p,k}$ be the signals in the k^{th} subband of the p^{th} harmonic image. By choosing Δf to be a multiple of the subband spacing, $\frac{W_{bb}}{N}$, we can write

$$x_k(i) = \sum_{p=1}^{\infty} \alpha_p \cdot z_{p,k}(i \cdot N) + \beta_p \cdot z_{p, N-k-1}^*(i \cdot N) \quad (21)$$

and

$$\begin{aligned} y_k(i) = & \sum_{p=1}^{\infty} \epsilon_p \cdot F_p \{z_{p, k+(p-1) \cdot o}(i \cdot N)\} \\ & + \sum_{p=1}^{\infty} \gamma_p \cdot F_p \{z_{p, N-k-1-(p+1) \cdot o}^*(i \cdot N)\} \\ & + \text{additional image components,} \end{aligned} \quad (22)$$

where o is Δf expressed in the number of subband channels:

$$o = \frac{N \cdot \Delta f}{W_{bb}} \quad (23)$$

Note that the subband index can go below zero or above $N-1$ in (21) and (22). In that case, the baseband-equivalent signal is outside the baseband bandwidth and is equal to zero. Also note that decimation-by- N takes place, owing to the use of the FFT.

Assuming that the baseband signals $z_{p,k}$ are wide-sense stationary, the cross-correlation expression for each subband (20) can be written as:

$$\begin{aligned} R_{x_k, y_k} &= E\{x_k(i) \cdot y_k(i)^*\} \\ &= \alpha_p \cdot \epsilon_p^* \cdot |z_{p,k}|^2 \end{aligned} \quad (24)$$

The assumptions made in (24) are that subband signals $z_{p,k}$ and $z_{p, k+(p-1) \cdot o}$ do not correlate unless $p = 1$ and that $z_{p, N-k-1}$ and $z_{p, N-k-1-(p+1) \cdot o}^*$ never correlate, regardless of p . The assumptions will hold true if Δf is large enough that the subbands are independent, i.e. larger than the inter-channel spacing of the wireless standards being monitored.

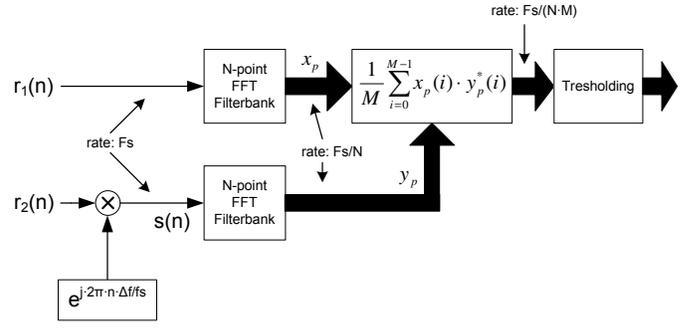


Fig. 7. A block diagram of the baseband processing algorithm. The signals $r_1(n)$ and $r_2(n)$ are produced by the downmixer shown in Fig. 3.

Whether this offset can be reduced to smaller values is a topic for further research.³

Note that the correlation (24) depends on ϵ_p and α_p , which are related to the mixers in the frontend. This fact can be used as a way to perform system identification of the mixers, which may be useful for calibration purposes. Also consider that changing the discrete-time frequency shift, f_d , in (14) allows a different harmonic image to be selected instead of $z_1(n)$. The correlation (24) will provide information on the power of that harmonic image. Such information is useful in adaptive harmonic rejection schemes, such as [5].

After obtaining an estimate of the cross-correlation (24) for each subband using the estimator (20), the results are thresholded to decide whether the subband can be considered occupied or empty. The integration time given by the number of samples, M , should be chosen large.

A block diagram of the entire algorithm is given in Fig. 7. The time it takes the algorithm to produce an occupancy vector is equal to $\frac{N \cdot M}{F_s}$ seconds.

C. Implementation aspects

The algorithm from Fig. 7 need not be implemented as shown. For instance, when the LO offset Δf is chosen to be an integer multiple of the subband spacing of the FFT, then the downmixing action performed on $r_2(n)$ can be achieved by shifting the elements in the vector y_p .

A further reduction in complexity is achieved by realizing that only the overlapping part of $r_1(n)$ and $r_2(n)$ provides correlative information about the desired signal spectrum. Therefore, not all elements from x_p and y_p need be taken into account when performing the cross-correlation. In effect, the number of correlators can be reduced from N to $N_{corr} = \lceil N \frac{W_{bb} - 2\Delta f}{W_{bb}} \rceil$, where $\lceil \cdot \rceil$ denotes a ceiling operation.

A radix-2 FFT requires $2N \cdot \log_2(N)$ real multiplies and $3N \cdot \log_2(N)$ real adds. In addition to the FFT, a window must be applied. The windowing operation consists of $2N$ real multiplies. There are two FFTs and therefore two window operations necessary, bringing the count to $4N \cdot \log_2(N) + 4N$

³Consider that an OFDM signal consists of multiple carriers, of which most can be considered independently modulated. In such a case, a frequency offset smaller than the OFDM channel spacing will not result in significant correlation.

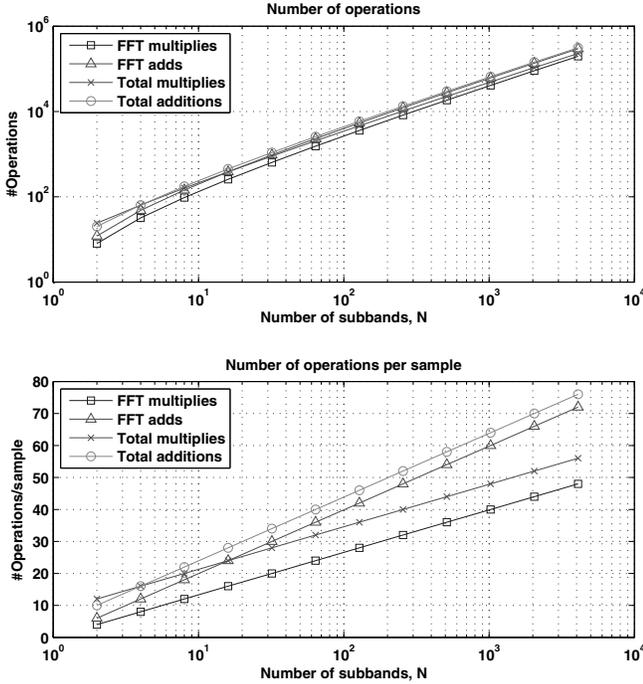


Fig. 8. Algorithmic complexity as a function of the number of subbands, N . The number of correlators, N_{corr} , is equal to N .

multiplies and $6N \cdot \log_2(N)$ real additions. Each correlator performs one complex multiply, one complex add and a conjugation for each new x_p (or y_p) vector. The conjugation of y_p can be performed by the complex multiply. Therefore, a correlator equates to 4 real multiplies and 4 real adds. There are N_{corr} correlators, so the total complexity becomes $4N \cdot \log_2(N) + 4N_{corr} + 4N$ multiplies and $6N \cdot \log_2(N) + 4N_{corr}$ adds which must be performed $\frac{F_s}{N}$ times a second.

Figure 8 shows the total number of operations and the number of operations per sample as a function of the number of subbands, N . The number of correlators N_{corr} was chosen equal to N . The lower graphs shows that the number of operations per sample does not become prohibitively large for a large number of subbands. For every doubling of the number of subbands, the number of multiplies grows by four and the number of additions grows by six. The increase is solely due to the FFT.

IV. SIMULATIONS

We performed simulations to demonstrate the effectiveness of the detection algorithm. The aim is to compare a traditional detector based on the power spectral density with our detection algorithm. The baseband-equivalent mixer model (11) was used to generate the baseband signals $r_1(n)$ and $r_2(n)$. The α_p and β_p coefficients were chosen to represent a switching

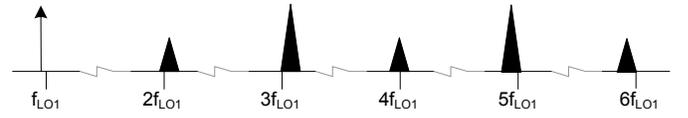


Fig. 9. Graphical representation of the RF signals and their frequencies as given by Table I.

mixer with a 50% duty cycle square wave:

$$\alpha_p = \begin{cases} 0 & \text{if } p \text{ is even,} \\ \frac{1}{p} & \text{if } p \text{ is odd} \end{cases} \quad (25)$$

$$\beta_p = 0 \quad (26)$$

The fact that $\beta_p = 0$ means that there is no I/Q imbalance.

The desired signal is a sinusoid and the harmonic image signals are all 64-QAM modulated using random bitstreams. The 64-QAM signals occupy approximately 1/16 of the baseband bandwidth. Their strength (standard deviation σ) and frequency are shown in Table I and in Fig. 9. The signals were deliberately placed so that they overlap after downmixing. Furthermore, two of the image signals are 34 dB stronger (at the antenna) than the desired signal.

harmonic	frequency [cycles/sample]	strength [σ]
1	-0.015	1
2	0.01	1
3	0.02	50
4	-0.005	1
5	-0.015	50
6	0.025	1

TABLE I
FREQUENCY AND STRENGTH OF SIMULATED SIGNALS

The number of subbands was set to $N = 1024$. Taking the spectral spreading of the Nutall window into account, this leads to a resolution of $1024/8 = 128$ distinct frequency bins. The frequency offset was set to $d_f = \frac{60}{1024}$, which equates to a distance of 60 FFT bins.

Figure 10 shows the power spectral density (PSD) of $r_1(n)$ and $s(n)$. The PSD of $r_1(n)$ shows that almost all harmonic images are on top of each other and that the sinusoidal desired signal can be seen peaking just above them. It is clear that the different harmonic images present in $s(n)$ undergo different frequency shifts compared to $r_1(n)$. However, the desired signal undergoes no frequency shift. A detector based on the PSD will detect the presence of the harmonic images, as can be seen from Fig.10. Therefore, it is not suited for spectrum sensing when harmonic images are present.

Our detector should be able to distinguish the harmonic images from the desired signal. That this is indeed the case, is shown by Fig. 11, which gives the power spectrum based on the cross-correlation for correlation length $M=1,64$ and 1024. As a reference, the PSD of $r_1(n)$ is also plotted.

Figure 12 is a zoomed in version of Fig. 11 which shows the spectrum around the sinusoidal signal present in the desired spectrum. The PSD curve shows that harmonic images are

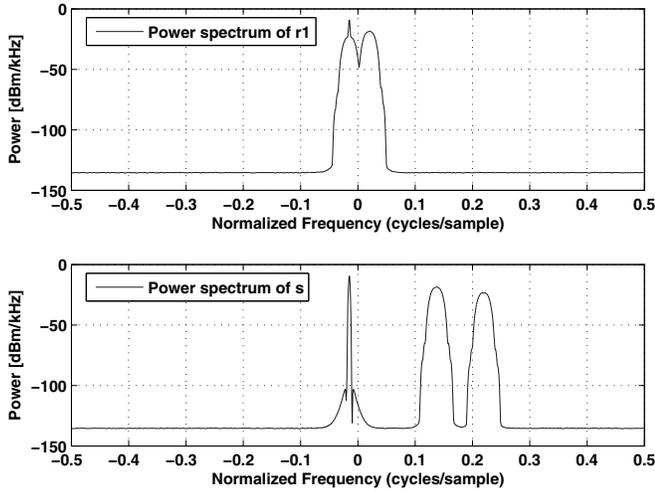


Fig. 10. The power spectral density of $r_1(n)$ and $s(n)$. The quadrature downmixers do not suffer from I/Q imbalance.

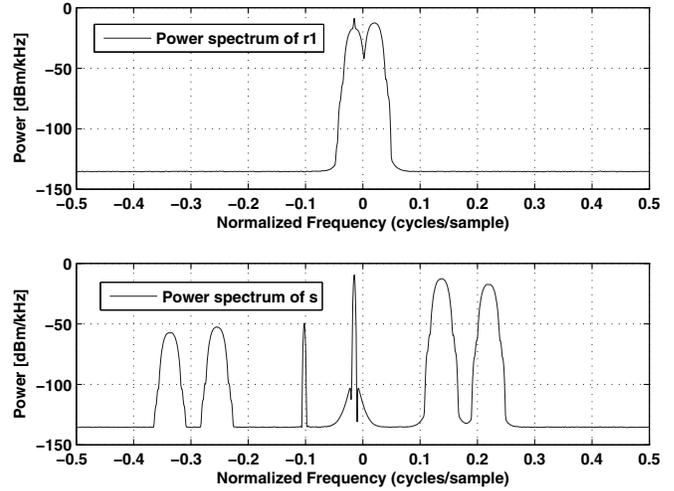


Fig. 13. The power spectral density of $r_1(n)$ and $s(n)$. The quadrature downmixers have I/Q imbalance. The I/Q imbalance image rejection is 40 dB.

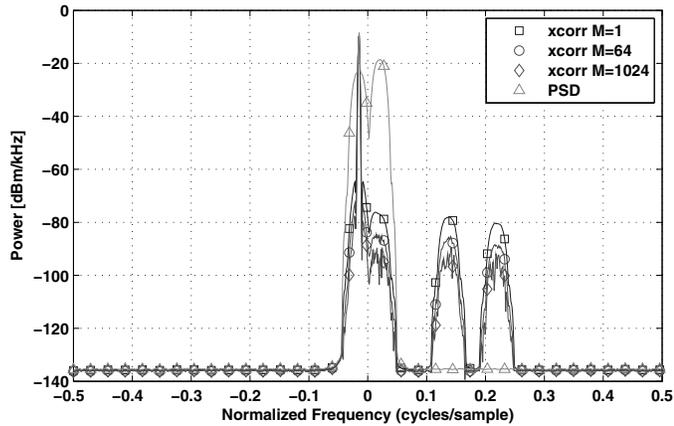


Fig. 11. Power spectral density of $r_1(n)$ and the cross-correlation vectors for $M=1, 16$ and 1024 . The quadrature downmixers do not suffer from I/Q imbalance.

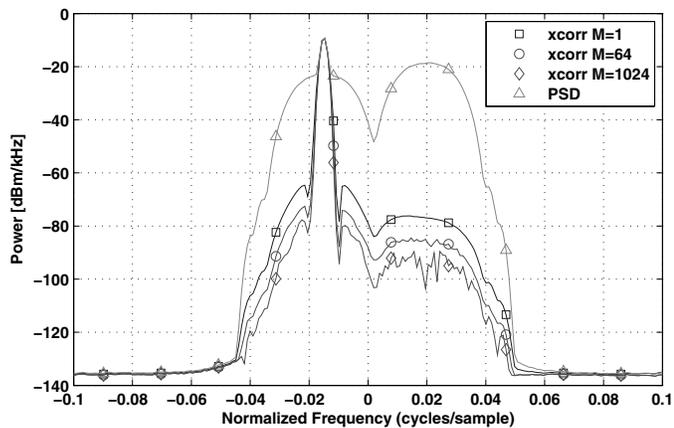


Fig. 12. A zoomed in version of Fig. 11. The quadrature downmixers do not suffer from I/Q imbalance.

being detected, as the PSD response is only 9.3 dB (at 0.02 cycles/sample) below that of the sinusoid. Examining the cross-correlation curves shows that the harmonic images have been rejected by more than 61.5 dB for $M = 1$ to more than 80 dB for $M = 1024$ at 0.02 cycles/sample. Note that the large rejection for $M = 1$ dB is owing to the fact that there is no signal present at 0.02 cycles/sample in $s(n)$.

The harmonic images are suppressed by 61.5 dB for $M = 1$ and every 1000-fold increase of M will add 15 dB of suppression. As the averaging is done in the power domain, not the amplitude domain, we cannot expect more. Given enough samples, the non-correlating peaks will disappear in the noise and our cross-correlation algorithm will be able to separate the harmonic images from the desired spectrum. However, this may take an unrealistic number of samples.

The previous simulation did not take into account the fact that imperfections in the analog frontend, such as an amplitude or phase mismatch between the I and Q paths of r_1 or r_2 , cause I/Q imbalance. When I/Q imbalance is present, the complex conjugated terms in (21) and (22) come into play. We performed the above simulations again, this time with $\beta_p = 0.01 \cdot e^{j \cdot 2\pi \cdot \phi}$, where the phase angle ϕ was arbitrarily chosen to equal 0.1. The I/Q imbalance image rejection of this quadrature downmixer is 40 dB.

The I/Q imbalance images can clearly be seen in the PSD plot of $s(n)$, see Fig. 13. However, the I/Q imbalance images in $r(n)$ are disguised by the harmonic images and the desired signal. By comparing Fig. 10 to Fig. 13, the rejection of the harmonic images remains the same. This is easily explained as the I/Q imbalance images manifest themselves in subbands different from the harmonic images. Therefore, they are not seen by the same correlators. The I/Q imbalance images experience the same rejection of 15 dB per 1000-fold increase in the number of correlator samples, M .

While the number of correlation samples, M , needed to fully reject the harmonic images may prove prohibitively large,

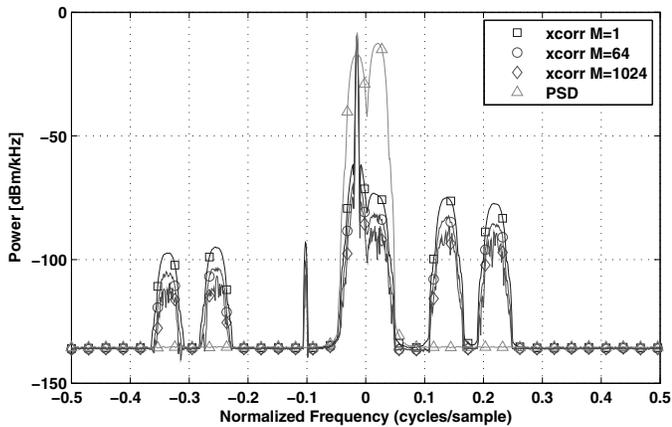


Fig. 14. Power spectral density of $r_1(n)$ and the cross-correlation vectors for $M=1, 16$ and 1024 . The quadrature downmixers have I/Q imbalance. The I/Q imbalance image rejection is 40 dB.

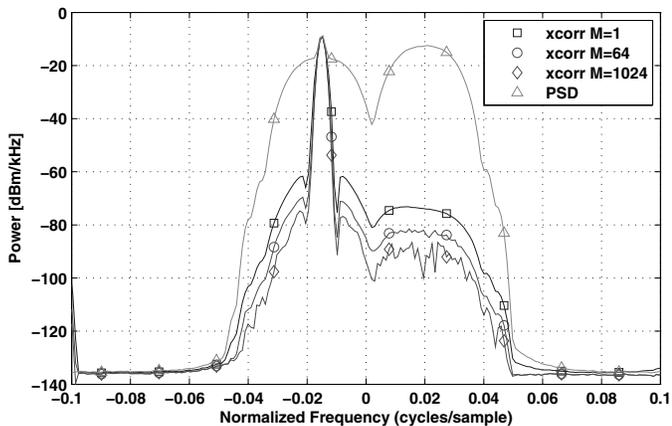


Fig. 15. A zoomed in version of Fig. 14. The quadrature downmixers have I/Q imbalance. The I/Q imbalance image rejection is 40 dB.

our method may still be useful. Consider the fact that if the PSD, which is obtainable from less samples, indicates no signal presence in a frequency bin, the cross-correlation information of that bin can be disregarded. However, when the PSD indicates the presence of a signal in a bin, the signal's origin must be checked. To check the origin of the signal, we can compare two cross-correlation vectors. One obtained using $M = 64$ and averaged in magnitude 16 times and the other obtained using $M = 1024$. If the power in the bin reduces by approximately $10 \cdot \log_{10} \left(\frac{1024}{64} \right) = 6$ dB, it must be occupied by a harmonic image. If the attenuation is less, there is correlative energy and a signal from the desired spectrum is also present. Whether this solution is feasible, is a topic for further investigation.

Another way to make use of our cross-correlation method, is to apply a multi-path harmonic-rejection mixer [4], which offers about 40 dB of harmonic image rejection without the use of an RF filter. As the harmonic images are already rejected by a fair amount, the cross-correlation algorithm need not provide as much rejection as without a harmonic-rejection mixer. A rejection of 40 dB reduces the required number of

correlation samples by a factor of 10^8 , which is an extremely large reduction. Note that the 40 dB of harmonic rejection offered by harmonic rejection mixers is not enough to solve the problem of erroneous detection of the harmonic images as these images can be more than 100 dB above the noise floor.

In addition to rejecting the harmonic images, the cross-correlation method will also reject spurious signals that find their way into r_1 and r_2 . These spurious signals can come from other parts of the receiver circuitry, such as the frequency synthesizer, the clock generators and the digital baseband processor. Spurious signals can also enter the receiver circuitry from the power supply terminals of the circuit. The rejection occurs because the digital-domain frequency shift, (14), will decorrelate them.

In conclusion, note that because the algorithm is based on cross-correlation, its performance is independent of the modulation schemes of the desired signal and the harmonic images. Only the energy contained within each FFT subchannel, x_k or y_k , and the number of averaged cross-correlation samples, denoted by M , determine the suppression of the harmonic images.

V. CONCLUSIONS

We have introduced the important problem of harmonic downmixing in the context of normal receiver operation and spectrum sensing applications, such as cognitive radios. Harmonic downmixing is the result of harmonics in the effective LO waveform, which cause RF signals found at multiples of the LO's fundamental frequency to be downconverted to baseband.

In cognitive radio, power spectral density (PSD) measurements are used to perform spectrum sensing. However, given the presence of harmonic images in the baseband signal of the sensing receiver, a PSD-based algorithm will mistakenly identify the frequency as occupied. In this paper we presented an algorithm based on cross-correlation, which is able to reject the harmonic images making correct occupancy identification possible.

The algorithm uses two baseband observations, $r_1(n)$ and $r_2(n)$, which are generated by two quadrature mixers. The LO frequencies of the mixers are offset by Δf Hz. Because of this, the harmonic images undergo different frequency shifts with respect to the desired signal spectrum and each other. This causes a decorrelation of the harmonic images, while the desired signal spectrum remains correlated. A subband cross-correlator is used to identify which frequency bins are occupied by a signal within the desired spectrum.

The complexity of the algorithm was analyzed. A 1024-band measurement needs 48 real multiplications and 64 real additions per sample. At positions where the harmonic images are present and detected by the PSD method, the cross-correlation method is able to reject these signals by more than 80 dB. This figure is not affected by I/Q imbalance.

In addition to rejecting the harmonic images, any spurious signals entering the receiver through the analog baseband inputs will also be rejected.

ACKNOWLEDGMENTS

The authors would like to thank Zhiyu Ru and Mark Oude Alink for the many fruitful discussions.

APPENDIX A

PROOF OF THE UNBIASEDNESS OF THE CROSS-CORRELATION ESTIMATOR

We will show that the estimator (20) is unbiased when x and y are jointly wide-sense stationary. To prove this, we must show that the expected value of the estimate \widehat{R}_{xy} is equal to R_{xy} .

We begin by writing out the expectation of the estimator:

$$E\{\widehat{R}_{xy}\} = E\left\{\frac{1}{N}\sum_{n=0}^{N-1}x(n)\cdot y(n)^*\right\} \quad (27)$$

We proceed by moving the expectation operator into the sum:

$$E\{\widehat{R}_{xy}\} = \frac{1}{N}\sum_{n=0}^{N-1}E\{x(n)\cdot y(n)^*\} \quad (28)$$

As the expectation of the terms $E\{x(n)\cdot y(n)^*\}$ are all equal to R_{xy} when x and y are jointly wide-sense stationary, we may simplify the above to:

$$E\{\widehat{R}_{xy}\} = \frac{1}{N}\sum_{n=0}^{N-1}R_{xy} \quad (29)$$

,which can be reduced to:

$$E\{\widehat{R}_{xy}\} = R_{xy} \quad (30)$$

Hence, the estimator (20) is unbiased.

APPENDIX B

I/Q IMBALANCE AND HARMONIC REJECTION PARAMETERS

The relation between the complex-valued baseband output, the baseband-equivalent signals $z_p(t)$ and the Fourier coefficients of the quadrature LO waveforms, $\{c_0, c_1, \dots, c_\infty\}$ and $\{d_0, d_1, \dots, d_\infty\}$, see (8), may be restated as:

$$r(t) = \frac{1}{2}\sum_{p=1}(\alpha_p \cdot z_p(t) + \beta_p \cdot z_p^*(t)), \quad (31)$$

where $\alpha_p = (c_p - j \cdot d_p)^*$ and $\beta_p = (c_p + j \cdot d_p)$.

The magnitudes $|\beta_p|$ and $|\alpha_p|$ give the amount of harmonic image suppression.

The suppression of the p^{th} harmonic image with respect to the desired spectrum can be defined as

$$L_{\text{harmonic},p} = \frac{|\alpha_1|^2}{|\alpha_p|^2} \quad (32)$$

,or in decibel:

$$L_{\text{harmonic},p} = 10 \cdot \log_{10}\left(\frac{|\alpha_1|^2}{|\alpha_p|^2}\right) \text{ dB} \quad (33)$$

Each harmonic image has an associated I/Q imbalance image, owing to I/Q imbalance in the receiver. The I/Q imbalance image suppression of the p^{th} harmonic image is defined by (34).

$$L_{\text{quad},p} = \frac{|\alpha_p|^2}{|\beta_p|^2} \quad (34)$$

In decibel it is given by:

$$L_{\text{quad},p} = 10 \cdot \log_{10}\left(\frac{|\alpha_p|^2}{|\beta_p|^2}\right) \text{ dB} \quad (35)$$

This is an extension of the formulation of L_{quad} given by [9].

REFERENCES

- [1] E. Klumperink, R. Shrestha, E. Mensink, V. Arkesteijn, and B. Nauta, "Cognitive radios for dynamic spectrum access - polyphase multipath radio circuits for dynamic spectrum access," *Communications Magazine, IEEE*, vol. 45, no. 5, pp. 104–112, May 2007.
- [2] J.R. Tourret et al., "Sip tuner with integrated LC tracking filter for both cable and terrestrial TV reception," *Solid-State Circuits, IEEE Journal of*, vol. 42, no. 12, pp. 2809–2821, Dec. 2007.
- [3] J. A. Weldon, R. S. Narayanaswami, J. C. Rudell, L. Lin, M. Otsuka, S. Dedieu, L. Tee, K.-C. Tsai, C.-W. Lee, and F. Paul R. Gray, "A 1.75-GHz highly integrated narrow-band CMOS transmitter with harmonic-rejection mixers," *IEEE J. Solid-State Circuits*, vol. 36, no. 12, pp. 2003–2015, Dec. 2001.
- [4] Z. Ru, E. A. M. Klumperink, and B. Nauta, "A discrete-time mixing receiver architecture with wideband harmonic rejection," in *ISSCC Digest of technical papers*, Feb. 2008.
- [5] N. A. Moseley, E. A. M. Klumperink, and B. Nauta, "A two-stage approach to harmonic rejection mixing using blind interference cancellation," *IEEE Trans. Circuits Syst. II*, Accepted for publication.
- [6] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Prentice Hall, 1993.
- [7] B. Farhang-Boroujeny, "Filter bank spectrum sensing for cognitive radios," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1801–1811, May 2008.
- [8] A. Nuttall, "Some windows with very good sidelobe behavior," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 1, pp. 84–91, Feb 1981.
- [9] M. Valkama, M. Renfors, and V. Koivunen, "Advanced methods for I/Q imbalance compensation in communication receivers," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2355–2344, Oct. 2001.