A Nonlinearity Mitigation Method for a Broadband RF Front-End in a Sensor Based on Best Delay Searching

Wen Zhao, Hong Ma, Hua Zhang*, Jiang Jin, Gang Dai and Lin Hu

School of Electronic Information and Communications, Huazhong University of Science & Technology, 1037 Luoyu Road, Wuhan 430074, China; D201177667@hust.edu.cn (W.Z.); mahong@hust.edu.cn (H.M.); jinjiang@hust.edu.cn (J.J.); daigang99@126.com (G.D.); D201077465@hust.edu.cn (L.H.)

* Correspondence: zhanghua@hust.edu.cn; Tel.: +86-134-7679-2182

Received: 22 July 2017; Accepted: 14 September 2017; Published: 28 September 2017

Abstract: The cognitive radio wireless sensor network (CR-WSN) is experiencing more and more attention for its capacity to automatically extract broadband instantaneous radio environment information. Obtaining sufficient linearity and spurious-free dynamic range (SFDR) is a significant premise of guaranteeing sensing performance which, however, usually suffers from the nonlinear distortion coming from the broadband radio frequency (RF) front-end in the sensor node. Moreover, unlike other existing methods, the joint effect of non-constant group delay distortion and nonlinear distortion is discussed, and its corresponding solution is provided in this paper. After that, the nonlinearity mitigation architecture based on best delay searching is proposed. Finally, verification experiments, both on simulation signals and signals from real-world measurement, are conducted and discussed. The achieved results demonstrate that with best delay searching, nonlinear distortion can be alleviated significantly and, in this way, spectrum sensing performance is more reliable and accurate.

Keywords: broadband RF front-end; nonlinear distortion; non-constant group delay distortion; nonlinearity mitigation; cognitive radio wireless sensor network

1. Introduction

Being capable of gathering information from the environment with simple processing, low complexity, and low power consumption, wireless sensor networks (WSN) have attracted an increasing amount of attention and have seen rapid development, especially in micro-electro-mechanical systems (MEMS) technology [1]. Generally, since the Industrial Scientific Medical (ISM) band, where WSNs operate, is shared with various kinds of successful communication technologies [2], the detection performance in a complex electro-magnetic environment is affected by false alarms or missed detection, and inevitably degrades. Instead of the traditional fixed-spectrum arrangement method, cognitive radio (CR) has the ability to sense underutilized or unutilized bands [3] to increase the spectrum efficiency to some extent. Hence, CR-WSN is supposed to be a promising technique, taking into account both spectrum sensing accuracy and efficiency [4–6].

At present, most studies regarding CR-WSN in published papers focus on the development of sensing techniques [7–13]. The literature [7] provides a spectrum-sensing technique based on the autocorrelation of received samples and decides whether signals exist by means of a derived Euclidean distance. In Reference [8], a spectrum sensor scheduling algorithm and a data sensor resource allocation algorithm are proposed for a heterogeneous cognitive radio sensor network (HCRSN) to achieve the sustainability of spectrum sensors and the conservation of data sensors.
Except for the development on improving spectrum sensing performance, the sensing performance evaluation of CR networks is studied in Reference [9].

By contrast, the limitation of radios, as another essential reason causing vacant spectrum inaccessibility and deteriorated sensing performance, should be discussed further. Normally, the broadband RF front-end of the sensor in CRs offers the secondary users (SUs) [14] unoccupied spectrum opportunistically and, in the meantime, protects licensed primary users (PUs) from interference. If there is vacant spectrum available for SUs during a certain period, the sub-band of interest can be used for signal reception by a flexible change of the filter in the digital domain. However, strong signals from PUs in adjacent channels may generate unwanted intermodulation to interfere with SUs’ signals or occupy vacant spectrum. In this case, detection mechanisms in a CR-WSN are very likely to make the wrong judgment on the service condition of the spectrum, resulting in spectrum efficiency degradation. Therefore, for the sake of facilitating the development of CR-WSN, a broadband RF front-end with sufficient spurious-free dynamic range (SFDR) is an important guarantee.

Distortions, namely the RF front-end impairments, contain nonlinear distortion, in-phase/quadrature (I/Q) imbalance, in-band group delay distortion, carrier frequency offset, and phase noise [15]. I/Q imbalance usually takes place at the direct-conversion receiver (DCR) whose I/Q mixer fails to provide an exactly 90-degree phase shift as desired. This can be avoided through digital down-conversion (DDC) after the analog to digital converter (ADC). Additionally, de-noise methods [16,17] have been widely studied and have achieved a certain degree of noise reduction effect.

Loosely speaking, compared with other kinds of distortions in the RF front-end, nonlinear distortion is more difficult to mitigate [18], especially the spectral hyperplasia generated by strong signals inside or outside the sub-band of interest which occupies free bandwidth or hides weaker target signals away. In Reference [19], circuits were designed to receive the I/Q signal and calculate its different kinds of nonlinear distortions to compensate for the distorted baseband signals. In Reference [20], intermodulation products were generated by a tunable nonlinear circuit and depressed nonlinearities in the original signals after adaptive filtering.

In addition to the above nonlinearity suppression methods by analog circuits, a great many methods [21–27] based on more flexible digital signal processing (DSP) have been published. In Reference [21], a modified post-distortion algorithm was proposed to mitigate nonlinearities in a broadband receiver which is far away from the original fundamental waves, as well as to corrupt the other sub-channels. This method succeeded in designing a broadband direct digitization RF receiver. However, it could hardly adapt to compensating intermodulation products since its signal extraction is unsuitable for the components inside or nearby the adjacent channel of strong signals. In Reference [22], a nonlinear behavior model of a baseband signal was designed to eliminate nonlinearities in an RF receiver. The objective signal was converted to a baseband and then centered at DC after rotation. Then, each distortion component and its frequency content function were calculated to form an output of a nonlinear model for compensation. Unfortunately, this method has an unwanted limitation in the number of input signals. In References [23,24], a finite impulse response (FIR) band-pass/band-stop filter was formed according to the strong signals’ frequency location and practical adaptive filtering-based learning algorithms were then proposed. Our previous work [27] suggested a blind identification criterion for minimizing the short-time energy of the nonlinear components in the frequency domain. However, suffering from the non-ideal frequency response characteristic of FIR, the methods mentioned above may hardly adapt to the situation where multiple unknown signals with different power levels exist simultaneously. Unfortunately, in most cases, this is exactly the electromagnetic environment in which the RF front-end of the sensor works.

On the other hand, in-band group delay distortion, which introduces inter-symbol interference to the output signal, inevitably degrades the overall CR-WSN system performance. Many references [28–33] have been presented up to now. Due to an analog circuit’s lack of flexibility and applicability, strategies in DSP have become a trend. In Reference [34], the cumulative distribution function (CDF) of a given modulation signal was calculated and a measure of distortion of the in-band group delay was then constructed in the statistical domain, which could be optimized to obtain
the equalizing filter coefficients. Nevertheless, this method cannot be utilized in the RF front-end in the sensor because the statistical characteristic of various modulation mode signals with different levels is hardly determined and, furthermore, the processing procedure of the proposed algorithm is too complicated to meet the requirement in the context of CRs.

In addition, it is worth noting that even though non-constant group delay belongs to linear distortion, when it joins with the memory nonlinear effect, the system modeling of group delay compensation and nonlinearity mitigation in series cannot achieve the desired improvement any longer. Even worse, the equalizer could increase the memory depth of the nonlinear behavior model. According to the best knowledge of the author, in the existing state-of-the-art literature, few methods involving nonlinearity mitigation take group delay distortion compensation into account.

In this paper, a simplified RF front-end model containing a low noise amplifier (LNA) and filter with a non-constant group delay in the passband is chosen as the object of analysis. Then, we emphasize how the group delay distortion brings a phase change to each frequency component at the output. Furthermore, in terms of the time domain, it can be interpreted as a delay between the strong signals producing nonlinearities and the weak signals including the nonlinear components and low signal-to-noise ratio (SNR) signals. Thereupon, the instantaneous power residual error can be regarded as a function of nonlinear model kernels and delay. Consequently, a blind identification criterion for nonlinearity mitigation with best delay searching is designed.

To eliminate the disadvantage of filter imperfections and be able to deal with multicarrier situations without too many extra operations added, strong signals and nonlinearities are extracted based on the threshold power in the frequency domain. Afterwards, their time domain waveforms are obtained by the inverse fast Fourier transform (IFFT) technique. Compared to the signal extraction realized by pass/stop-band filters, the proposed method can save not only the amplitudes of the original signals, but also their phase spectrum. This is a significant preparation step for making a reliable and precise best delay search. After that, an iterative formula based on the weighed iterative improvement (WII) [35] method is applied to minimize the ill-conditioned equation problem.

The outline of the remainder of this paper is as follows: Section 2 explains the joint effect of group delay distortion and nonlinear distortion through a simplified model from the broadband RF front-end and, based on it, the blind identification criterion is designed. Then the mitigation architecture and optimization solution are suggested in Section 3. In Section 4, simulation experiments and RF measurement results are provided to verify the performance of the proposed method. Finally, conclusions are drawn in Section 5.

2. System Model and Problem Analysis

As mentioned above, the broadband RF front-end with sufficient linearity is an important guarantee of a reliable spectrum sensing. In this section, first, a classical broadband RF front-end of a sensor is chosen as an example and then a simplified model is built to discuss its joint effect of group delay distortion and nonlinear distortion. Finally, an identification criterion for nonlinearity mitigation is suggested.

2.1. RF Front-End of a Sensor Node in CR-WSN

A broadband reconfigurable direct RF front-end architecture is shown in Figure 1. The input signals from the antenna are amplified by the LNA. After that, a tunable anti-aliasing filter configures the center frequency to determine the frequency location to be sampled. Then, by a series of pulses, the pulse sampler down-converts the signals to a certain intermediate frequency (IF). The second band-pass filter removes the components outside the IF bandwidth so that signals can be sampled and converted to digital signals by ADC. Finally, baseband signals are available through DDC.
It is worth noting that, except for the heterodyne receiver shown in Figure 1, zero-IF and low-IF receivers are also in general use in sensors. However, no matter which kind the receiver belongs to, the joint effect of the non-constant group delay distortion and nonlinear distortion could occur in the receiver. Strictly speaking, an ideal linearity condition is hardly met in a real RF front-end. Almost every part in a receiver, such as the LNA, mixer, BPF, and ADC, can generate nonlinear components. Moreover, as long as the RF front-end has a broad passband, group delay is supposed to vary in a large range that cannot be ignored and, especially, the BPF’s nonlinear phase-frequency response and the nonlinear device’s memory effect will both make it happen. Thus, in the actual receiver system, the joint effect certainly exists and becomes more complicated.

Besides these two kinds of distortions, in zero-IF or low-IF receivers, another significant effect resulting in decreased SFDR is I/Q imbalance, which is due to an imprecise 90-degree phase shift in the I/Q mixer. The main objective of this paper is to analyze the joint effect of non-constant group delay distortion and nonlinear distortion of a broadband RF front-end and provide its corresponding solution. In order to highlight the key point and make the result of the proposed method simple and explicit, a heterodyne receiver is taken as the object of study.

2.2. Joint Effect of Group Delay Distortion and Nonlinear Distortion

According to Figure 1, the proposed method is expected to focus on baseband signals after the RF front-end and DDC. In order to see how a non-constant group delay in the passband influences the signals with nonlinearity, a simplified grey-box model containing the LNA and band-pass filter (BPF) in series is built in Figure 2. When passing through the receive chain, signals are distorted. Even though all kinds of modules with non-ideal linear characteristics in the RF front-end could bring in nonlinearity and non-constant group delay, we should pay attention to the main influence and make the following assumptions that the main source where nonlinearities are generated is at the LNA stage, and non-constant group delay distortion in the passband is due to the non-ideal phase-frequency response of the BPF.

Let \( x(n) \) be the chosen signals with a fixed center frequency and bandwidth. After the LNA, \( y(n) \) are signals containing nonlinearity. \( z(n) \) means the signals passing through the BPF are distorted both by nonlinear products and non-constant group delay distortion. If no BPF exists, our previous method [27] presents a blind identification criterion of forming a nonlinear behavior model to compensate for nonlinear interference:

\[
 s_y(n) = y(n) - f\{y(n)\} 
\]  

(1)
where \( s_r(n) \) are the compensated signals and \( f \{ \cdot \} \) indicates the nonlinear behavior model, such as the Volterra model and Hammerstein model.

Generally speaking, a constant group delay in the passband of RF filter circuits is desirable. However, it is difficult to achieve in broadband systems, or even in narrowband communications, when tight filtering is required [15]. Since the hypothesis is that the BPF only brings the non-constant group delay to \( y(n) \), it can be regarded as a linear filter adding nonlinear phase increments to each frequency component in the passband. From the perspective of the frequency domain, to each frequency spectrum:

\[
Z(k) = Y(k) \cdot H(k) = Y(k) \cdot A(k)e^{j\varphi(k)}
\]  

(2)

where \( Y(k) \) and \( Z(k) \) are the frequency spectra of \( y(n) \) and \( z(n) \), respectively. Their frequency value is represented as \( k (k = 0,1,2,\ldots,N-1) \). \( N \) is the number of points carrying out the discrete Fourier transform (DFT). \( H(k) \) represents the frequency response function of the BPF whose amplitude is \( A(k) \) and phase is \( \varphi(k) \). Then, the Inverse Discrete Fourier Transform (IDFT) of \( Z(k) \) is expressed as:

\[
z(r) = \frac{1}{N} \sum_{k=0}^{N-1} Z(k) e^{\frac{2\pi jkr}{N}}
\]  

(3)

and the DFT of \( y(n) \) is:

\[
Y(k) = \sum_{n=0}^{N-1} y(n)e^{-\frac{2\pi jkn}{N}}
\]  

(4)

Substituting Equations (2) and (4) into Equation (3):

\[
z(r) = \frac{1}{N} \sum_{k=0}^{N-1} Y(k)A(k)e^{j\varphi(k)} e^{\frac{2\pi jkr}{N}}
\]

\[
= \frac{1}{N} \sum_{k=0}^{N-1} A(k)e^{j\varphi(k)} \sum_{n=0}^{N-1} y(n)e^{-\frac{2\pi jkn}{N}} e^{\frac{2\pi jkr}{N}}
\]

\[
= \frac{1}{N} \sum_{n=0}^{N-1} y(n) \sum_{k=0}^{N-1} A(k)e^{j\frac{2\pi j(k-n)\varphi(k)}}
\]  

(5)

In Reference [34], the phase characteristic of the BPF is modeled by a \( P \)th order polynomial function and, without loss of generality, we obtain:

\[
\varphi(k) = \frac{2\pi}{N} \sum_{p=1}^{P} c_p k^p
\]  

(6)

Next, we employ FIR to simulate which has the same amplitude response and phase response as the BPF in Figure 2. Since the phase response of the BPF is nonlinear in the passband, there exists no linear relationship between \( \varphi(k) \) and \( k \). Meanwhile, considering that the time domain response coefficients of FIR belong to the real number field, to meet to the symmetry condition, we suppose that \( A(k) \) has even symmetry and \( \varphi(k) \) has odd symmetry relative to \( k = \frac{N}{2} \) (\( N \) is an even number), i.e., \( A(k) = A(N-k) \), \( \varphi(k) = -\varphi(N-k) \), and \( A\left(\frac{N}{2}\right) = 0 \), \( \varphi\left(\frac{N}{2}\right) = 0 \). Then, Equation (5) can be written as:
\[
z(r) = \frac{1}{N} \sum_{n=0}^{N-1} y(n) \sum_{k=1}^{N-1} A(k) e^{\frac{2\pi j (r-n) \phi(k)}{N}}
= \frac{1}{N} \sum_{n=0}^{N-1} y(n) \left\{ A(0) + \sum_{k=1}^{N-1} A(k) e^{\frac{2\pi j (r-n) \phi(k)}{N}} + \sum_{k=-N-1}^{-1} A(k) e^{\frac{2\pi j (r-n) \phi(k)}{N}} \right\}
= \frac{1}{N} \sum_{n=0}^{N-1} y(n) \left\{ A(0) + \sum_{k=1}^{N-1} A(k) e^{\frac{2\pi j (r-n) \phi(k)}{N}} + \sum_{k=-N-1}^{-1} A(k) e^{\frac{2\pi j (r-n) \phi(k)}{N}} \right\}
= \frac{1}{N} \sum_{n=0}^{N-1} y(n) \left\{ A(0) + \sum_{k=1}^{N-1} A(k) e^{\frac{2\pi j (r-n) \phi(k)}{N}} + \sum_{k=-N-1}^{-1} A(k) e^{\frac{2\pi j (r-n) \phi(k)}{N}} \right\}
\]

(7)

Apparently, \( \phi(0) = \sum_{p=1}^{P} c_{p} \cdot 0^{p} = 0 \). Utilizing the transformation \( k = N - k' \), and since \( k \) and \( k' \) are operated in the same accumulation range, Equation (7) can be written as:

\[
z(r) = \frac{1}{N} \sum_{n=0}^{N-1} y(n) \left\{ A(0) + 2 \sum_{k=1}^{N-1} A(k) \cos\left[\frac{2\pi}{N} (r-n) + \phi(k)\right]\right\}
\]

(8)

If the BPF has a linear phase response, \( \phi_{\text{ideal}}(k) = \frac{2\pi}{N} c_{k} k \), the second term in Equation (8) becomes:

\[
\sum_{k=1}^{N-1} A(k) \cos\left[\frac{2\pi}{N} (r-n) + \phi(k)\right] = \sum_{k=1}^{N-1} A(k) \cos\left[\frac{2\pi}{N} (r-n + c_{k})\right]
\]

(9)

This means that for a fixed time \( r \), there exists \( n = r + c_{1} \) making:

\[
\sum_{k=1}^{N-1} A(k) \cos\left[\frac{2\pi}{N} (r-n + c_{k})\right] = \sum_{k=1}^{N-1} A(k)
\]

(10)

While \( n \neq r + c_{1} \), the odd symmetry of the cosine function in the range of 0 to \( \pi \) makes this accumulation term equal to 0. Hence, we obtain:

\[
z(r) = \frac{1}{N} A(0) \sum_{n=0}^{N-1} y(n) + 2 \sum_{n=0}^{N-1} A(k) y(n - c_{1})
\]

(11)

On account of \( A(0) \ll \sum_{k=1}^{N-1} A(k) \) and the direct component having no variation in phase, the final expression is:

\[
z(r) = \frac{2}{N} y(n - c_{1}) \sum_{k=1}^{N-1} A(k)
\]

(12)
products outputted at the same time $n$ in $y(n)$ are sampled at different moments of $z$. Taking this effect into account when designing the nonlinearity identification criterion, modeling is expected to be more accurate and distortion will be mitigated more thoroughly. Thus:

$$s_z(n) = z(n) - f[z(n), \tau]$$

To be exact, $\tau$ should be a function of frequency. However, to make the realization easy, we choose a simplified method, in that:

$$s_z(n) = z(n - \tau_{\text{best}}) - f[z(n)]$$

Here, $\tau_{\text{best}}$ is not a delay really existing in the RF front-end, but a sampling interval between the output of the nonlinear behavior model and the distortion component in $z$. When the best delay is found, the minimum of the power residuals after optimization is expected to be more precise. In addition, the searching range to which the best delay belongs can be set according to prior knowledge of the $S$ parameter of the RF front-end. Once the $S$ parameter is obtained, it can be determined by the differential of the phase angle of $S_{21}$ to the frequency.

2.3. Design of Blind Identification Criterion

The scheme for nonlinearity cancellation is presented in Figure 3, where $g(\cdot)$ is assumed as the transfer function of a broadband RF front-end including the linear term with memory $g_s(\cdot)$ and the nonlinear term with memory $g_S(\cdot)$. Hence:

$$z(n) = g[x(n)] = g_s[x(n)] + g_S[x(n)]$$

The input signals $x(n)$ could be divided into strong signals $x_s(n)$ which produce nonlinear interference when passing through the RF front-end, and weak signals $x_i(n)$ containing nonlinear distortions, low power signals, and noise $n_i(n)$, namely:

$$x(n) = x_s(n) + x_i(n) + n_i(n)$$

Substituting Equation (16) into (15):

$$z(n) = g_s[x_s(n)] + g_s[x_i(n)] + g_S[x_s(n)] + g_S[x_i(n)] + g_s[x_i(n), x_i(n)] + n_i(n)$$

$$\approx g_s[x_s(n)] + g_s[x_i(n)] + g_S[x_s(n)] + g_S[x_i(n)] + n_i(n)$$

Normally, as a result of $x_s(n) >> x_i(n)$, $g_S[x_i(n)]$ and $g_S[x_i(n), x_i(n)]$ are much less than $g_S[x_s(n)]$ and below the noise floor; thus, they can be neglected in $z(n)$. If $h(\cdot)$ is the nonlinear behavior model and meets the condition that $h[g_i(\cdot)] = g_S(\cdot)$, based on the discussion in Section 2.2, we obtain:

$$s_z(n) = z(n - \tau_{\text{best}}) - h[z(n)]$$

$$= g_s[x_s(n - \tau_{\text{best}})] + g_s[x_i(n - \tau_{\text{best}})] + g_S[x_s(n - \tau_{\text{best}})] + g_S[x_i(n - \tau_{\text{best}})] + n_i(n - \tau_{\text{best}})$$

$$- h[g_s[x_s(n)] + g_s[x_i(n)] + g_S[x_s(n)] + g_S[x_i(n)] + n_i(n)]$$

$$\approx g_s[x_s(n - \tau_{\text{best}})] + g_s[x_i(n - \tau_{\text{best}})] + g_S[x_s(n - \tau_{\text{best}})] - h[g_s[x_s(n)] + n_i(n)]$$

Since $x_s(n) >> x_i(n)$, there is an assumption that the output of the nonlinear behavior model $h[g_s[x_s(n)] + g_s[x_i(n)] + g_S[x_s(n)] + n_i(n)] \approx h[g_s[x_s(n)]$. Actually, Equation (18) provides the identification criterion of nonlinear mitigation. To each delay within the searching range, optimization can be employed to make the power residual between $g_S[x_s(n - \tau_{\text{best}})]$ and
\[ h[g_1(x_i(n))] \] achieve its minimal value. The minimal one among these values is acquired at the best delay and, hence, the corresponding coefficients of the nonlinear model are regarded as the optimal solution for nonlinearity mitigation.

\[
1(n) \xrightarrow{\text{Wideband RF-front end}} z(n) \xrightarrow{\text{Best delay}} z(n-\tau_{\text{best}}) + s_z(n)
\]

**Figure 3.** Principle of nonlinearity mitigation.

### 3. Proposed Mitigation Architecture for an RF Front-End

Based on the identification criterion above, in this section, mitigation architecture is proposed in detail and followed by WII optimization method for nonlinear model kernels calculation.

#### 3.1. Mitigation Architecture

As shown in Figure 4, mitigation architecture can be divided into three steps including signal extraction, best delay searching and power residuals optimization, and nonlinearity compensation. Since the reason and method of step 3 have been introduced in Section 2, step 1 and step 2 will be discussed in this section.

**Figure 4.** Mitigation architecture.

Let us begin with the process of step 1. Signals \( x(t) \) are first sampled and converted to digital baseband signals \( z(n) \). Subsequently, noise \( n_z(n) \) is obtained when the RF front-end is at the same working state of receiving \( z(n) \), such as the center frequency, passband bandwidth, system gain, and attenuation. Then, a fast Fourier transform (FFT) is carried out to obtain the frequency spectra of \( z(n) \) and \( n_z(n) \). By substituting the power spectrum of \( z(n) \) for the noise power spectrum...
according to the set threshold, strong signals $Z_s(k)$ and weak signals $Z_w(k)$ in the frequency domain are separated from $z(n)$, respectively. Finally, the waveforms of $z_1(n)$ and $z_2(n)$ in the time domain are available through IFFT.

It is worth mentioning that this type of signal extraction saves both amplitudes and phases of the original signal and, hence, has the following benefits. In the literature [23,27], FIRs are utilized. Suffering from imperfect performance of amplitude-frequency and phase-frequency response, there will be unwanted amplitude and phase changes in $z_1(n)$ and $z_2(n)$. This is inevitably harmful for best delay searching and nonlinear behavior model kernel calculation. In addition, compared with continually designing and changing multi-stopband/multi-passband digital filters with extremely high performance, in the proposed approach, only power threshold altering is needed. Therefore, this method is more likely to handle the situation of multiple signals with different power levels or wide ranges of bandwidth, which is apparently more flexible and convenient.

In the rest of Section 3, step 2 is explained in detail including the choice of nonlinear behavior model for complex baseband signal in Section 3.1, the objective function establishment in Section 3.1, and the optimization solution in Section 3.2.

The generalized Hammerstein model [24] is chosen to simulate the nonlinear behavior in the RF front-end. If the input is $x(n)$, the expression is:

$$\text{ham}(n) = \alpha_1(n) * x(n) + \alpha_2(n) * x^2(n) + \alpha_3(n) * x^3(n) + ...$$

(19)

where $\{\alpha_1(n), \alpha_2(n), \alpha_3(n)\}$ are impulse responses taking memory effect into account for each order, and $\text{ham}(n)$ is the output. Even-order nonlinear components are usually far away from others [18] in the spectrum, and even if the bandwidth of the RF front-end is extremely wide, proper circuit design methodologies [18,36] can provide a high second-order intercept point where harmful effects resulting from the even-order nonlinearity are kept to a minimum. On the contrary, the odd-order nonlinearity appearing at the frequency near or inside the total frequency band of interest is hardly eliminated simply by some analog filters and, hence, more complicated to deal with. In the baseband, most nonlinearities belong to odd-order intermodulation components, and the third-order and fifth-order spectrum hyperplasia are much stronger than others. Therefore, Equation (19) is simplified to:

$$\text{ham}(n) = \alpha_1(n) * z_1(n) + \alpha_3(n) * |z_1(n)|^2 z_1(n) + \alpha_5(n) * |z_1(n)|^4 z_1(n)$$

(20)

In Equation (20), "\(|\)\" represents the modulus operator, as $z_1(n)$ is the baseband complex signal and notation "\(*\)" denotes convolution. Let $M_1, M_3,$ and $M_5$ be the memory depth of each order. The total number of kernels is $M = M_1 + M_3 + M_5$. Equation (20) can be expressed in a matrix form: 

$$\text{ham}(n) = \omega^T u(n)$$

(21)

where:

$$\omega = \begin{bmatrix} \alpha_{0,1} & \alpha_{0,1} & ... & \alpha_{0,1} & \alpha_{0,3} & \alpha_{0,3} & ... & \alpha_{0,3} & \alpha_{0,5} & \alpha_{0,5} & ... & \alpha_{0,5} \\ \end{bmatrix}^T$$

(22)

and:

$$u(n) = \begin{bmatrix} z_1(n) & z_1(n-1) & ... & z_1(n-M_1+1) \\ |z_1(n)|^2 z_1(n) & |z_1(n-1)|^2 z_1(n-1) & ... & |z_1(n-M_3+1)|^2 z_1(n-M_3+1) \\ |z_1(n)|^4 z_1(n) & |z_1(n-1)|^4 z_1(n-1) & ... & |z_1(n-M_5+1)|^4 z_1(n-M_5+1) \end{bmatrix}^T$$

(23)
The next step is to calculate the error signals between the outputs of the Hammerstein model $ham(n)$ and the delayed weak signal $z_2(n+\tau)$, while the power residual $E(\omega, \tau)$ is regarded as the objective function to be minimized. Thus, $E(\omega, \tau)$ is:

$$E(\omega, \tau) = \sum_{n=1}^{N} |ham(n) - z_2(n+\tau)|^2 = [U\omega - Z_2(\tau)]^H [U\omega - Z_2(\tau)]$$  \hspace{1cm} (24)

where $U = [u(1) \ u(2) \ ... \ u(N)]$, $Z_2(\tau) = [z_2(1) \ z_2(2) \ ... \ z_2(N)]^T$ and the notation $H$ is the transposed-conjugate operator. Once the searching range of the best delay is determined, the optimization problem becomes:

$$\arg \min_{\omega, \tau} E(\omega, \tau) \hspace{1cm} (25)$$

Equation (25) reveals that for each delay $\tau$ within the searching range, $E(\omega, \tau)$ reaches its minimum value after optimization. Among these values, the minimal one is chosen as the optimum solution and its corresponding $\tau$ and $\omega$ are taken as the best delay and the coefficients of nonlinear behavior kernels. As $\omega$ in the objective function belonging to the complex number field, applying deviation directly to $E(\omega, \tau)$ is infeasible. Let:

$$\omega_k = \rho_k e^{i\phi_k}, \ d = U^H Z_2(\tau) = [d_1 \ d_2 \ ... \ d_M]^T, \ R = U^H U = \begin{bmatrix} r_{11} & r_{12} & ... & r_{1M} \\ r_{21} & r_{22} & ... & r_{2M} \\ ... & ... & ... & ... \\ r_{M1} & r_{M2} & ... & r_{MM} \end{bmatrix} \hspace{1cm} (26)$$

where $\rho_k$ and $\phi_k$ represent the amplitude and phase of $\omega_k$, respectively. From Equation (24), we obtain:

$$E(\omega, \tau) = \omega^H R\omega - \omega^H d - d^H \omega + Z_2^H(\tau) Z_2(\tau)$$ \hspace{1cm} (27)

Taking the derivative of $E(\omega, \tau)$ with respect to $\rho_k$ and $\phi_k$:

$$\frac{\partial E}{\partial \rho_k} = e^{-i\phi_k} \sum_{i=1}^{M} r_i \omega_i^* + e^{i\phi_k} \sum_{i=1}^{M} \omega_i^* r_i - d_i^* e^{i\phi_k} - e^{-i\phi_k} d_k = 0$$ \hspace{1cm} (28)

$$\frac{\partial E}{\partial \phi_k} = -j \rho_k e^{-i\phi_k} \sum_{i=1}^{M} r_i \omega_i + j \rho_k e^{i\phi_k} \sum_{i=1}^{M} \omega_i^* r_i - j d_i^* \rho_k e^{-i\phi_k} + j \rho_k e^{i\phi_k} d_k = 0$$ \hspace{1cm} (29)

Let: $\varphi = \begin{bmatrix} e^{i\phi_1} \\ e^{i\phi_2} \\ ... \\ e^{i\phi_M} \end{bmatrix}$, $\rho = \begin{bmatrix} \rho_1 \\ \rho_2 \\ ... \\ \rho_M \end{bmatrix}$, $\frac{\partial E}{\partial \rho} = \begin{bmatrix} \frac{\partial E}{\partial \rho_1} & \frac{\partial E}{\partial \rho_2} & ... & \frac{\partial E}{\partial \rho_M} \end{bmatrix}$, $\frac{\partial E}{\partial \phi} = \begin{bmatrix} \frac{\partial E}{\partial \phi_1} & \frac{\partial E}{\partial \phi_2} & ... & \frac{\partial E}{\partial \phi_M} \end{bmatrix}$ and substituting them into Equations (28) and (29):

$$\frac{\partial E}{\partial \rho} = \varphi^* R\omega + \varphi R^* \omega^* - \varphi^* d - \varphi^* d = 0$$ \hspace{1cm} (30)

$$\frac{\partial E}{\partial \phi} = -j \rho \varphi^* R\omega + j \rho \varphi R^* \omega^* - j \rho \varphi^* d + j \rho \varphi^* d = 0$$ \hspace{1cm} (31)

Combining Equations (30) and (31), the result is $R\omega = d$, i.e.,

$$U^H U\omega = U^H Z_2(\tau)$$ \hspace{1cm} (32)
3.2. Optimization Solution

To solve linear Equation (32), whose solutions \( \omega \) are a series of complex numbers, we make a transformation by assuming \( R = R_x + jR_m \), \( d = d_x + jd_m \), and \( \omega = \omega_x + j\omega_m \), where subscripts “re” and “im” mean the real and imaginary parts, respectively. Thus:

\[
\begin{bmatrix}
R_x - R_m \\
R_m \\
R_x
\end{bmatrix}
\begin{bmatrix}
\omega_x \\
\omega_m
\end{bmatrix} = \begin{bmatrix}
d_x \\
d_m
\end{bmatrix}
\quad (33)
\]

If \( A = \begin{bmatrix} R_x - R_m & R_m \\ R_m & R_x \end{bmatrix} \), \( \bar{\omega} = [\omega_x^T \omega_m^T]^T \), Equation (33) can be written as:

\[
A\bar{\omega} = \bar{d}
\quad (34)
\]

In Equation (32), the odd-order Hammerstein model is set from 1 to 5 so that the first diagonal element of \( U^H U \) is \( \sum_{n=1}^{N} |z(n)|^2 \) and the last one is \( \sum_{n=1}^{N} |z(n-M_s+1)|^2 \). As a result of too large a power difference, \( U^H U \) and \( A \) easily tend to be ill-conditioned matrices. In this paper, WII [35] is used to deal with this ill-conditioned problem. The prerequisite of adopting WII is that \( A \) is a real symmetry non-negative definite matrix. Since \( R \) is a Hamiltonian matrix \( R = R^H \) and \( \omega^H R \omega = (U \omega)^H (U \omega) = |U \omega|^2 \geq 0 \), it can easily be verified. The final iterative formula is:

\[
\omega^{(k+1)} = \omega^{(k)} + (A + \lambda I)^{-1} (\bar{d} - A\omega^{(k)})
\]

where \( \lambda \) is the weighting factor. Although there is a matrix inversion step appearing in Equation (35), \( (A + \lambda I)^{-1} \) does not update during the whole iteration procedure. Hence, unlike other methods, such as the recursive least squares method (RLS), \( (A + \lambda I)^{-1} \) can be computed in advance and stays unchanged until the entire delay searching is finished. Therefore, the recurrence formula (Equation (35)) only contains some simple multiply-add operations, making it suitable for efficient realization in software-defined radio (SDR).

4. Simulation Experiments and Results Analysis

In Section 4, the proposed method is utilized to evaluate its SFDR improvement, both on simulation signals and signals from a real RF front-end. To illustrate the advantages of the best delay-based nonlinear mitigation, the spectrum after nonlinear compensation at the best delay is compared with the one without delay, and the curves of optimized power residuals \( E(\omega, \tau) \) in Equation (24) at each delay in the searching range are also provided. Then, constellation diagrams are obtained to verify the demodulation improvement and SFDR enhancement. Finally, by the energy detection (ED) method [37], sensing performance promotion is offered.

4.1. The Verification of Simulation Signals

The simulation signals are comprised of two strong 16 Quadrature Amplitude modulation (16QAM) signals and a weak one. Both of the strong signals are generated with the same amplitude of −45 dBFS and the same symbol rate of 2 MHz. Their center frequencies are 125 MHz and 130 MHz, respectively. The weak one located in the polluted region is at 119 MHz with −97 dBFS, and its symbol rate is 0.5 MHz. Their baseband signals are obtained by carrying out DDC. Hence, from the distorted spectrum in Figure 5a, unwanted nonlinearities cover the weak 16QAM’s spectrum and distort it seriously so that its signal to interference plus noise ratio (SINR) inevitably decreases. In addition, the green region affected by intermodulation is originally supposed to be a vacant spectrum. Therefore, false-alarm probability in this region rises and spectrum efficiency degrades.
Figure 5. Nonlinearity mitigation results of simulation signals including two strong 16QAM signals and one weak 16QAM signal: (a) distorted signals; and (b) signals after nonlinearity mitigation.

Figure 5 provides the relationship between the delay and optimized power residuals in the full-spectrum band. When $\tau$ equals 10, the optimized residuals reach the minimum value compared to the others. Therefore, 10 is chosen as the best delay. In Figure 5b, the blue spectrum indicates the mitigation performance when best delay searching is employed, while the red spectrum shows the compensation result without the best delay. The nonlinear behavior model adopted here is the Hammerstein model in Equation (20) and the model orders are 1, 3, and 5 with the same memory depth of 20. By carrying out mitigation methods, SFDR in the full band rises by 10–15 dB.

In Figure 5b, it seems that both the proposed mitigation methods have achieved satisfactory results, while the apparent differences between them are only focused on the adjacent channels of strong 16QAMs. However, these differences are crucial evidence for explaining the advantage of the best delay-based method. In Figure 5a, there mainly exist two kinds of intermodulation products. The first kind is generated when one strong 16QAM is modulated by another. In Figure 5a, they are located at the center of 120 MHz and 135 MHz and have a 9 MHz bandwidth, since the roll-off factor is 0.5. The second is caused by the strong signal itself and is located in an adjacent channel to the strong 16QAM. In the first adjacent channels of the strong 16QAMs, different kinds of distortions are vector overlapped in the frequency domain. For the first type of intermodulation product, both mitigation methods have a significant improvement while, for the second one, the best delay-based mitigation method outperforms the mitigation-only method. In particular, in the marked false-alarm region ranging from 127 MHz to 128 MHz, the spectrum after mitigation at the best delay is about −110 dBFS, which is close to the background noise level and 5 dB lower than that without the best delay. If the black line of −108 dBFS is used as the ED threshold, the red spectrum will always be detected as occupied so that SUs would miss a transmission opportunity. This is also proved by Figure 6a, which shows the relationship between the delay and optimized power residuals in the false-alarm region. The best delay of 10 corresponds to its minimal optimized power residuals. Consequently, we conclude that when the best delay is determined, the nonlinear behavior model can provide better outputs to deal with the more complicated situation where different kinds of nonlinearities simultaneously exist.
Figure 6. Relationships between delay and power residuals after optimization: (a) optimized power residuals in false-alarm region; and (b) optimized power residuals in full spectrum band.

Figure 7 shows the constellations of the weak 16QAM in Figure 5 before and after mitigation. Due to the nonlinear components coming from the spectrum hyperplasia of adjacent strong 16QAM, the weak signal is interfered and its SINR is not large enough to directly carry on demodulation, thus its constellation is randomly distributed in Figure 5a. Meanwhile, in Figure 5b,c, benefiting from the mitigation method, the intermodulation in the polluted region is obviously depressed, finally obtaining a regularly-distributed constellation. This experiment illustrates that the proposed nonlinear distortion mitigation method can improve the ability of detecting weak signals, which is a crucial capacity of sensing performance.

Since the constellations in Figure 7b,c seem to be nearly the same, their error vector magnitudes (EVMs) and bit error rates (BERs) are calculated to specifically explain the advantage of the best delay-based nonlinearity mitigation method. After normalization, the result is obtained and shown in Table 1. Compared to the results without the best delay, when the best delay condition is met, the EVM is lower and the BER is also better, as expected. As mentioned above, the major components distorting the weak 16QAM belong to the first kind of intermodulation product. Since both the mitigation methods can effectively eliminate these nonlinearities and result in a considerable SFDR increase after nonlinearity elimination, it is rational that the EVM and BER achieved by the best delay-based approach do not have an additional increase in the order of magnitude.
Table 1. EVM and BER of weak 16QAM after nonlinearity mitigation with or without the best delay.

<table>
<thead>
<tr>
<th></th>
<th>Mitigation without the Best Delay</th>
<th>Mitigation with the Best Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVM</td>
<td>22.3%</td>
<td>15.1%</td>
</tr>
<tr>
<td>BER</td>
<td>$8.54 \times 10^{-3}$</td>
<td>$3.12 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

4.2. The Verification of Signals from a Real-World RF Device

In this part, signals passing through real-world RF are utilized for testing. For the sake of corresponding with the simulation signal experiments, the signals used for testing contain three strong 16QAM signals and one weak 16QAM signal. The sampling rate here is 112 MHz. The strong 16QAMs have the same symbol rate of 2 MHz and their center frequencies are 157 MHz, 162 MHz, and 168 MHz, respectively. The weak one, with a 0.5 MHz symbol rate, is located at 150 MHz. The center frequency of the intermediate frequency sampling is set as 160 MHz and the passband bandwidth is 40 MHz.

The order and memory depth of the Hammerstein model employed is set to be the same as those of the simulation signals. From Figure 8b, among the optimized power residuals of each delay, when delay $\tau$ equals 11, the minimal value is 88.3% of the optimized $E(\omega, \tau)$ without the best delay. This indicates that nearly 11.7% of power residuals, when mitigation is already finished without the best delay, can still be reduced if best delay searching is carried out. In Figure 8a, the best delay is also decided as 11, which reduces the optimized power residuals in the false-alarm region to almost one third of that without delay. A comparison of the spectrum before and after nonlinearity compensation is shown in Figure 9. As expected, SFDR increases at least 10 dB and the best delay approach can eliminate 3–5 dB more nonlinearities in most parts of the spectrum.

![Figure 8](image-url)

**Figure 8.** Relationships between delay and power residuals after optimization: (a) the optimized power residuals in the false-alarm region; and (b) the optimized power residuals in the full-spectrum band.

To verify the improvement of spectrum sensing performance of the proposed mitigation method, the green region in Figure 9 is chosen as the band of interest. Its receiver operating characteristics (ROC) for each case are calculated for the detection of spurious signals. The ED [37] method is utilized as a demonstration where the detection threshold is calculated by:

$$
\gamma = \sigma_w^2 \left( N + \sqrt{2NQ^{-1}(P_f)} \right)
$$

(36)

where $\sigma_w^2$ is the noise variance of the sub-band, $Q^{-1}(\cdot)$ is the inverse function of the normal distribution function, $N$ is the point used, and $P_f$ is the false-alarm probability which is used as a metric for CR-WSN failing to detect the absence of the primary signal.
Figure 9. Nonlinearity mitigation results of simulation signals including three strong 16QAM signals and one weak 16QAM signal: (a) distorted signals; and (b) signals after nonlinearity mitigation.

As shown in Figure 10, $P_d$ denotes the spurious signal detection probability. The curve in red represents the ROC of the distorted signal whose $P_d$ is always one when $P_f$ changes from 0 to 1. This means that, as a result of the existence of intermodulation, the ED method will always detect the green region as busy. This result is acceptable since the intermodulation component’s spectrum power is 10 dB higher than the noise floor in Figure 9a.

As for the comparison between the mitigation method with and without best delay searching, the pure noise’s ROC is treated as a standard whose characteristic conforms to $P_d \approx P_f$. Considering that the ROC curve here is in the false-alarm region, $P_d$ indicates the probability of detecting a nonexistent signal. For a fixed probability of $P_f$, the blue curve’s $P_d$ is always greater than that of the pure noise, while the purple’s curve is always less. This is to say that, by the proposed best delay-based mitigation method, the sensor is less likely to detect a false signal. Consequently, we can draw the conclusion that nonlinearities can be mitigated more effectively when best delay searching is employed and, in that way, the reliability of spectrum sensing is enhanced.

Figure 10. The receiver operating characteristics (ROC) before and after nonlinearity mitigation.

Finally, constellations of the distorted weak 16QAM by the intermodulation components are calculated in Figure 11. Randomly-distributed constellation points become regular after the nonlinearities are eliminated. This result is also the same as that obtained in the simulation signal.
Therefore, when handling real-world signals, the proposed nonlinearity mitigation method achieves the desired result.

Figure 11. The constellations of the 16QAM signals in the polluted regions: (a) distorted weak 16QAM; (b) weak 16QAM after mitigation without delay; and (c) weak 16QAM after mitigation at the best delay.

To sum up, the performance of the proposed method is demonstrated through the results of the simulation signals and signals from real-world measurements. The suppression of the nonlinear components is significantly better when best delay searching is performed. Thus, the proposed method provides a high-performance linearization solution for a broadband RF front-end in a sensor node.

5. Conclusions

In this paper, aiming to compensate for the intermodulation components brought by adjacent strong signals, we offer the best delay searching-based nonlinearity mitigation method to enhance the SFDR of the RF front-end and improve the reliability of spectrum sensing in CR-WSN. To summarize, the advantages of the proposed method are as follows:

1. There are plenty of non-ideal linear characteristic modules bringing in nonlinear distortions in the RF front-end of a sensor. Moreover, when the group delay distortion coming from the time delay electronic components joins with nonlinearities, it is necessary to design this blind identification criterion of nonlinearity mitigation at the best delay. Here, the best delay means the sampling interval between the output of the nonlinear behavior model and the distorted signal which corresponds to the minimal power residual after optimization.

2. Based on the proposed identification criterion, the mitigation architecture with best delay searching is presented. In this architecture, an odd-order generalized Hammerstein model is utilized as the nonlinear behavior model to compensate for the intermodulation components of the strong signals. Additionally, a flexible band-split method with low complexity is provided in the frequency domain to save both amplitudes and phases contained in the original signals. Finally, to solve the system of linear equations in the complex field which, in most cases, is ill-conditioned, the WII method is employed.

In conclusion, the proposed method achieves satisfactory results when tested both on the simulation signals and signals from real-world measurements. If applying it to the RF front-end in the sensor, the performance of spectrum sensing can be improved predictably. Our next stage is to find a better solution for modeling the group delay distortion instead of using only sampling interval changes. This further development is beyond of the main scope of this paper, but offers an interesting and relevant topic for future studies.

Author Contributions: Wen Zhao proposed the original algorithm and wrote this paper; Hong Ma, Jiang Jin and Hua Zhang revisited the paper and supervised the entire process; Lin Hu gave some valuable suggestions; Gang Dai offered the RF-front end for testing the proposed method performance of signals from real-world measurement and suggested some constructive advises during revision.
Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

- SFDR: Spurious-Free Dynamic Range
- RF: Radio Frequency
- WSN: Wireless Sensor Network
- MEMS: Micro-Electro-Mechanical Systems
- ISM: Industrial Scientific Medical
- CR: Cognitive Radio
- HCRSN: Heterogeneous Cognitive Radio Sensor Network
- SU: Secondary User
- PU: Primary User
- I/Q: In-Phase/Quadrature
- DCR: Direct-Conversion Receiver
- ADC: Analog-to-Digital Converter
- DDC: Digital Down-Conversion
- DSP: Digital Signal Processing
- FIR: Impulse Response Filter
- CDF: Cumulative Distribution Function
- LNA: Low Noise Amplifier
- SNR: Signal Noise Ratio
- IFFT: Inverse Fast Fourier Transform
- WII: Weighed Iterative Improvement method
- IF: Intermediate Frequency
- BPF: Band-Pass Filter
- DFT: Discrete Fourier Transform
- IDFT: Inverse Discrete Fourier Transform
- FFT: Fast Fourier Transform
- RLS: Recursive Least Squares method
- SDR: Software Defined Radio
- ED: Energy Detection
- SINR: Signal to Interference plus Noise Ratio
- EVM: Error Vector Magnitude
- BER: Bit Error Rate
- ROC: Receiver Operating Characteristics

References


© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).