Extending the dynamic range of RF receivers using nonlinear equalization

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Extending the Dynamic Range of RF Receivers using Nonlinear Equalization

Joel Goodman, Benjamin Miller, Matthew Herman, Michael Vai and Paul Monticiolo
MIT Lincoln Laboratory
Lexington, Massachusetts 02420
Email: {jgoodman,bamiller,nherman,mvai,umonticciolo}@ll.mit.edu

Abstract—Systems currently being developed to operate across wide bandwidths with high sensitivity requirements are limited by the inherent dynamic range of a receiver's analog and mixed-signal components. To increase a receiver’s overall linearity, we have developed a digital NonLinear Equalization (NLEQ) processor which is capable of extending a receiver's dynamic range from one to three orders of magnitude. In this paper we describe the NLEQ architecture and present measurements of its performance.

I. INTRODUCTION

In radar systems as well as many other applications (e.g., emerging Ultra-WideBand (UWB) communications), there is a need for a large dynamic range relative to the background noise and all other distortions. The need for dynamic range in radar returns is driven by the ratio of strong signals (such as clutter) versus weaker signals (returns from small targets). Systems currently being developed to operate across wide bandwidths with high sensitivity requirements are limited by the inherent dynamic range of the receiver's analog and mixed-signal components. Among these components (e.g., LNA, mixer), the ADC commonly has the lowest dynamic range [1]. A receiver’s deviation from its ideal ‘linear’ performance is commonly characterized by its spurious- and/or intermodulation-free dynamic range (SFDR and IFDR), which is a frequency-domain measurement that determines the minimum signal level that can be distinguished from distortion components. The SFDR and IFDR of an ADC are typically dominated by circuit-based (e.g., buffer amplifier, sample-and-hold) nonlinearities that are distinct from the nonlinear process of ideal quantization, which in principle can be circumvented with processing gain [2].

In this paper we develop an NLEQ processor composed of nonlinear polynomial filters to reduce polynomial distortions generated by the LNA, mixer and ADC. We also address practical problems in identifying an equalization architecture, such as separating the distortions that are induced by the analog signal generator (i.e., the excitation) from those that are generated by the receiver itself.

The uniqueness of NLEQ is that it suppresses wide band receiver nonlinear distortions in a computationally efficient fashion. Existing approaches to achieving computationally efficient polynomial filter architectures for RF compensation, principally developed to mitigate distortions generated by power amplifiers in transmitters, limit the multidimensional signal space over which the architecture can suppress spectral regrowth and in-band spurs [3], [4]. In this paper, we develop a technique to construct a polynomial filter architecture that searches over an unrestricted multidimensional signal space to select polynomial components that yield the highest equalization performance for a given computational complexity. In particular, we develop a coordinate system representation for polynomial filters, and leverage compressed sensing techniques to identify a sparse polynomial representation of an inverse nonlinearity.

The rest of this paper is organized as follows. In Section II, we develop a coordinate system representation used to construct an NLEQ architecture, and present an optimization procedure for selecting processing elements in that coordinate system. In Section III, we demonstrate the performance of NLEQ on both Maxim and Analog Devices ADCs, and in Section IV we conclude with a brief summary.

II. NONLINEAR EQUALIZER CONSTRUCTION

To construct an NLEQ architecture, we will first develop a horizontal coordinate system representation in a polynomial basis, and then describe two separate optimization techniques to identify a computationally efficient NLEQ architecture using components from that basis.

A. Polynomial Basis

The nonlinear system response of an ADC can often be described with the Pth-order (truncated) polynomial series expansion [5]

\[ y_{NL}(n) = \sum_{p=2}^{P} h_p[x(n)], \]

where

\[ \frac{h_p[x(n)]}{y_p(n)} = \sum_{m_1=0}^{N_p-1} \cdots \sum_{m_{p-1}=0}^{N_p-1} h_p(m_1, \ldots, m_{p}) \prod_{l=1}^{p} x(n - m_l) \]  

(1)

where \( N_p \) is the memory depth in each dimension of the \( p \)th-order Volterra kernel and \( h_p(m_1, \ldots, m_{p}) \) are the \( p \)th-order kernel coefficients. The number of non-redundant terms

This work is sponsored by DARPA under Air Force contract FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the United States Government.
in \( y_{NL}(n) \) is \( \sum_{p=2}^{P} \binom{N_p + p - 1}{p} \). Using \( N \) samples of \( x(n) \), equation (1) can be rewritten in matrix form as

\[
y_p = X^p(x)h_p
\] (2)

with

\[
y_p = \begin{bmatrix} y_p(n) & \ldots & y_p(n-N+1) \end{bmatrix}^T
\] (3)

where each of the \( \binom{N_p + p - 1}{p} \) columns of \( X^p(x) \) can be represented by

\[
\begin{bmatrix} \prod_{i=1}^p x(n-m_i) & \ldots & \prod_{i=1}^p x(n-m_i-N+1) \end{bmatrix}^T.
\]

The full series can now be expressed in matrix form as

\[
y_{NL} = X(x)h
\] (4)

with nonlinear convolution matrix \( X(x) = \begin{bmatrix} X^2(x) & \ldots & X^p(x) \end{bmatrix} \) and \( h = \begin{bmatrix} h^T_2 & \ldots & h^T_p \end{bmatrix}^T \).

It was shown in [6] that a large class of nonlinear systems can be approximated with arbitrarily small error using the polynomial representation in (1). However, this comes with the disadvantage that a relatively large number of parameters (factorial in \( N \) and \( p \)) are needed to represent systems with modest polynomial order and memory. To reduce the computational complexity when using (1) to model ADC nonlinearities, we develop an efficient representation of (1) in a new coordinate system.

The kernels in equation (1) can be rewritten in terms of elements of a horizontal coordinate system (HCS) in which a \( p \)th-order processing element (PE) is formulated as

\[
y_p^{HCS}(n, \alpha_2, \ldots, \alpha_p) = \sum_{m=1}^{N_p+i_1-1} h(m, \alpha_2, \ldots, \alpha_p)x(n-m) \prod_{l=2}^p x(n-\alpha_l),
\] (5)

where \( i_1 \) is used to center the data over the taps of the filter. Hence, it is possible to sum all the \( p \)th-order HCS PEs to obtain the \( p \)th order kernel

\[
y_p(n) = \sum_{\alpha_2=i_2}^{N_p+i_2-1} \sum_{\alpha_p=i_p}^{N_p+i_p-1} y_p^{HCS}(n, \alpha_2, \ldots, \alpha_p)
\] (6)

where the variables \( i_k \) in (5) are used to center the data over the taps of the multidimensional filter. Equation (5) geometrically corresponds to coefficients selected along a single horizontal (\( m \)) dimension while the other dimensions of the \( p \)th-order kernel \( (\alpha_2, \ldots, \alpha_p) \) remain fixed. The HCS representation has a regular structure for hardware implementation.

Using (5), we can derive an equalizer architecture for mitigating harmonic and intermodulation (nonlinear) distortion generated by an ADC, as illustrated in Fig. 1, where the nonlinear response of the ADC is estimated and subtracted from an appropriately delayed version of the ADC output. The objective of this section is to present techniques that enable the construction of a polynomial equalizer using a minimum set of PEs by efficiently searching over the multidimensional signal space.

1) Forward-Backward Sequential Estimation in Architecture Identification: In this section we develop a sequential algorithm for selecting PEs that minimize the mean square error, \( \varepsilon \). Let \( \varepsilon(y; \mathcal{P}) = \min_{\mathcal{H}} \| y - X_P h \|_2 \) denote the modeling error, where \( X_P = [X^{PE_1}, \ldots, X^{PE_L}] \), \( \mathcal{P} \) is the set of PEs in the architecture with cardinality \( |\mathcal{P}| = L \), and \( y \) is the vector of ADC output samples. Let \( \mathcal{P} \subset \mathcal{X} \), where \( \mathcal{X} \) is the set of user-defined candidate processing elements with \( |\mathcal{X}| \gg L \), then the pseudo-code for sequentially selecting processing elements to construct an NLEQ architecture is shown in Fig. 2. The total number of processing elements, \( |\mathcal{X}| \), can be adjusted according to computational considerations; the PEs that comprise a set are unique in their polynomial order, delay values and filter...
initialize $P = \emptyset$
while $\varepsilon(y; P') \geq \delta$
  // Add Processing Elements
  while $|P| \leq k$
    $p \leftarrow \arg \min_{p \in X} \varepsilon(y; P \cup \{p\})$
    $P \leftarrow P \cup \{p\}$,
    $X' \leftarrow X \setminus \{p\}$
  // Remove Processing Elements
  while $|P| \geq m$
    $p \leftarrow \arg \min_{p \in P} \varepsilon(y; P \setminus \{p\})$
    $P \leftarrow P \setminus \{p\}$,
    $X \leftarrow X \cup \{p\}$

Fig. 2. Pseudo-code for forward-backward sequential estimation. Each new PE $j$ from the set $X$ that yields the best minimum mean square error (MMSE) performance in combination with the previous $j-1$ PEs is added to the set $P$ in the forward stage. In the backward stage, a PE is removed one at a time from $P$ and placed back into $X'$ such that the PE removed has the least impact on MMSE performance.

coefficients. The parameter $\delta$ in the outer loop of the pseudo-code is a threshold for the MMSE; alternatively a fixed number of iterations can be used. In either case, the parameters $k$ and $m$, where $m \leq k$, are user defined and can change from iteration to iteration.

2) Global Estimation in Architecture Identification: An alternative to forward-backward sequential estimation is to formulate the problem of NLEQ architecture identification as a constrained optimization, in which the constraints are imposed to insure a computationally efficient solution. In [7], basis pursuit was used to find a sparse set of coefficients in a linear system identification problem. However, unlike linear systems, the size of the convolution matrix $X(x)$ representing the non-linear combinations of the input data can grow prohibitively large. One approach to reducing the dimensionality leverages the following theorem:

Theorem 1: Let $X \in \mathbb{R}^{N \times M}$, with $N \gg M$, be a non-singular matrix whose columns correspond to the $p$-fold product of the input, and let $y$ correspond to some vector in $\mathbb{R}^N$. Then there exists a projection matrix $P \in \mathbb{R}^{M \times N}$ with $\tilde{X} = PX \in \mathbb{R}^{M \times M}$, and hence an orthogonal projection matrix $P_{XX}^T = (I - \tilde{X}(\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T)$, such that $y^T P_{XX}^T \tilde{y} = y^T P_{XX}^T y$, where $\tilde{y} = Py$.

Proof: Consider a matrix $X$ with full column rank having singular value decomposition

$$X = \begin{bmatrix} U & \mathcal{U} \end{bmatrix} \begin{bmatrix} D & 0 \end{bmatrix} V^T \in \mathbb{R}^{N \times M},$$

with $U \in \mathbb{R}^{N \times M}$, and $N \gg M$. Further, consider the solutions to the least squares problems $h = \arg \min_h \|y - Xh\|$ and $h^* = \arg \min_h \|Py - PXh\|$, where $P$ is a $M \times N$ matrix that projects any $N$-dimensional vector $v$ down to an $M$-dimensional space. Then if $P = UT$, $h^* = h$.

It is possible to use the left singular vectors to reduce the dimensionality without any loss in performance; however, the computational cost of computing the SVD (singular value decomposition) of a large matrix, coupled with the fact that Theorem 1 is only valid for $N \geq M$, is of very little practical value. We propose two techniques for reducing the dimensionality in a computationally efficient fashion. To reduce the dimensionality of the rows, we leverage the following lemma [8].

Lemma 1 (Johnson-Lindenstrauss): For any $0 < \epsilon < 1$ and any integer $n$, let $K$ be a positive integer such that $K \geq (4(\epsilon^2/2 - \epsilon^3/3))^{-1}lnn$. Then for any set $V$ of $n$ points in $\mathbb{R}^N$, there is a map $f: \mathbb{R}^N \rightarrow \mathbb{R}^K$ such that for $u, v \in V$

$$(1 - \epsilon) \|u - v\| \leq \|f(u) - f(v)\| \leq (1 + \epsilon) \|u - v\|.$$  (9)

Extensions to the Johnson-Lindenstrauss lemma [9] have used concentration measure theory to show that if points in a vector space are projected onto a randomly selected subspace of suitably high dimension then the distances between the points (in a Euclidean sense) are approximately preserved. Therefore, we can construct a matrix $\Phi \in \mathbb{R}^{K \times N}$ whose elements are are randomly drawn from a Gaussian distribution ($\mathcal{N}(0, \frac{1}{\sqrt{K}})$), so that the projected matrix given by $\Phi X \in \mathbb{R}^{K \times M}$ where $K < N$ does not impart a significant error during architecture identification.

To further reduce the size of the matrix $X$, its columns can be pruned by projection and subtraction. First, the $M$ columns of $\Phi X$ are unit normalized so that each column can be considered as the coordinate of a point on the surface of a unit hypersphere. From the set of $M$ vectors, $L$ vectors are selected at one at a time, such that the $m$th vector chosen has the highest correlation with $\Phi y^{(m-1)}$ after the $m-1$ projections of the previously selected columns have been subtracted off, that is

$$\Phi y^{(1)} = \Phi y - \langle (\Phi X)_{i(1)}, \Phi y \rangle (\Phi X)_{i(1)}$$
$$\Phi y^{(2)} = \Phi y^{(1)} - \langle (\Phi X)_{i(2)}, \Phi y^{(1)} \rangle (\Phi X)_{i(2)}$$
$$\vdots$$
$$\Phi y^{(L)} = \Phi y^{(L-1)} - \langle (\Phi X)_{i(L)}, \Phi y^{(L-1)} \rangle (\Phi X)_{i(L)}.$$  (10)

In (10), the symbol $(\Phi X)_{i(m)}$ is the $m$th column of $\Phi X$ that has the highest correlation with $\Phi y^{(m-1)}$, whose column index is given by $i(m)$, and $(x, y)$ is the dot product of $x$ and $y$. The goal of projection and subtraction is to find the points on the unit hypersphere with widest angular separation that have a significant projected component on the received data. We use projection and subtraction to reduce the dimensionality of the data so that the subsequent constrained optimization is computationally tractable. Defining the matrix $\Theta \in \mathbb{R}^{M \times L}$ such that $X\Theta$ keeps only the columns of $X$ we get from projection and subtraction in (10), the constrained optimization to find a sparse NLEQ architecture is given by

$$\hat{h} = \arg \min_h e^T h$$
$$s.t. \|\Phi y - \Phi [X\Theta \Theta]h\|_2 \leq \epsilon$$
$$h \geq 0,$$  (11)

where $e = [1 \ 1 \ \ldots \ 1]^T \in \mathbb{R}^L$ and $\hat{h}_+, \hat{h}_- \in \mathbb{R}^L$, with $\hat{h} = [\hat{h}_+^T \ \hat{h}_-^T]^T$ and $h_{opt} = \hat{h}_+ - \hat{h}_-$. Equation (11) is easily solved using second order cone programming (SOCP) [10]. The basis pursuit (BP) cost function, along with the constraint
that $h \geq 0$, in (11) is an $\ell_1$ norm that favors sparse solutions [7]. The scalar regularization parameter $\varepsilon$ in the constraint balances the residual $\ell_2$ reconstruction error, that is, it ensures that the sparse solution $h_{opt}$ whose nonzero entries span the columns of $\Phi \Theta$ is in the cone of feasible solutions. Note that unlike forward-backward sequential estimation, the columns selected (non-zero entries of $h_{opt}$) correspond to single tap processing elements, where individual PEs are not connected with any specific coordinate system (e.g., horizontal, etc.).

### III. MEASURED RESULTS

#### A. Test Setup

To evaluate NLEQ performance, we used the MIT Lincoln Laboratory NLEQ testbed depicted in Fig. 3. The analog outputs of three Agilent E8257D tone generators were combined, filtered, and injected into an ADC that was seated in a temperature chamber. The testbed consists of three Agilent E8257D tone generators connected to an analog power combiner, an Agilent 16702B Logic Analyzer used for data capture, a temperature chamber where the time-interleaved ADCs are seated, and a Windows-based PC running Matlab to control the instrumentation for excitation and data capture.

![Simplified view of the the Lincoln Laboratory pNLEQ testbed.](Image)

**Fig. 3.** A simplified view of the the Lincoln Laboratory pNLEQ testbed. The testbed consists of three Agilent E8257D tone generators connected to an analog power combiner, an Agilent 16702B Logic Analyzer used for data capture, a temperature chamber where the time-interleaved ADCs are seated, and a Windows-based PC running Matlab to control the instrumentation for excitation and data capture.

#### B. Training and Verification

We trained both devices with a sequence of verification tone sets. These one-, two-, and three-tone sets were entirely different from the training sets used to derive the coefficients. We quantify dynamic range by taking the mean of the individual SFDR/IFDR measurements for each of the verification tone sets. This is the mean dynamic range (MDR) performance metric, which we measure in dBFS.

The individual tone generators did not impart intermodulation distortion, however, each tone generator did impart harmonic distortion. This makes it difficult to separate the distortions generated by the RF receiver from those of the excitation used during training. By operating in the second Nyquist zone (IF sampling), harmonic distortions fall out of band and were filtered off by the anti-aliasing filter preceding the ADC. Alternatively, if first Nyquist zone (base band) sampling were required, ignoring harmonic distortions in training and instead spacing two tones very close to one another to mimic harmonic distortion yielded satisfactory results.

#### C. Performance

Table II and Figures 4 and 5 illustrate the measured performance of NLEQ operating on data from the MAXIM MAX108 8-bit 1.5 GSps ADC, and the Analog Device AD6645 14-bit 105 MSps ADC. The MDR improvement of the MAX108 after NLEQ is roughly 22 dB, with an equalizer complexity of 217 operations per sample. The equalizer was constructed using forward/backward sequential estimation which selected 8 HCS PEs on the forward pass. Only marginal improvement (< 1 dB) using backward iterations was achieved. Basis pursuit yielded the same performance/computational complexity results as forward sequential estimation, indicating the robustness of the forward sequential estimation process in selecting HCS PEs. The MDR improvement of the AD6645 is only 10 dB, however, as is evident from Fig. 5, the distortions are virtually pushed down to the noise floor. The NLEQ architecture for the AD6645 required 9 HCS processing elements.

#### TABLE I

**SOURCE EXCITATION AND VERIFICATION PARAMETERS.**

<table>
<thead>
<tr>
<th>ADC</th>
<th>Excitation</th>
<th>Veriﬁcation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maxim MAX108</td>
<td>39 one-tone sets</td>
<td>350 one-tone sets</td>
</tr>
<tr>
<td></td>
<td>65 two-tone sets</td>
<td>280 two-tone sets</td>
</tr>
<tr>
<td></td>
<td>262 three-tone sets</td>
<td>2340 three-tone sets</td>
</tr>
<tr>
<td>Analog Device AD6645</td>
<td>39 one-tone sets</td>
<td>350 one-tone sets</td>
</tr>
<tr>
<td></td>
<td>60 two-tone sets</td>
<td>280 two-tone sets</td>
</tr>
<tr>
<td></td>
<td>262 three-tone sets</td>
<td>2340 three-tone sets</td>
</tr>
</tbody>
</table>

#### TABLE II

**EQUALIZATION RESULTS**

<table>
<thead>
<tr>
<th>Device</th>
<th>MDR (dBFS)</th>
<th>Improvement (dB)</th>
<th>Operations per Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAX108</td>
<td>51</td>
<td>22</td>
<td>217</td>
</tr>
<tr>
<td>AD6645</td>
<td>89</td>
<td>10</td>
<td>147</td>
</tr>
</tbody>
</table>
Fig. 4. Representative response of the Maxim 8-bit MAX108 to a two tone stimulus. The plot to the left represents the unequalized output of the ADC, with both intermodulation and harmonic distortions present. The plot to the right represents the equalized response of the MAX108 after running NLEQ, which demonstrates roughly 25 dB improvement in dynamic range.

Fig. 5. Representative response of the 14-bit AD6645 to a three tone stimulus. The plot to the left represents the unequalized output of the ADC, with both intermodulation and harmonic distortions present. The plot to the right represents the equalized response of the AD6645 after running NLEQ, which demonstrates roughly 10 dB improvement in dynamic range.

IV. SUMMARY

In this paper we developed a computationally efficient architecture for nonlinear equalization (NLEQ) in a new coordinate system to improve the dynamic range of RF receivers. We demonstrated a dynamic range improvement of over two orders of magnitude on the 8-bit, 1.5 MSps Maxim MAX108, and an order of magnitude improvement in the dynamic range of the 14-bit, 105 MSps Analog Devices AD6645. In both cases, the computational complexity of NLEQ was on the order of 200 operations per sample. We also demonstrated that NLEQ was capable of significantly improving the linearity of a device (AD6645) that already started out with a very high (90 dB) dynamic range.

REFERENCES