URSI GASS 2020, Rome, Italy, 29 August - 5 September 2020

An Investigation of Nonlinear Adaptive Space-Time Processing for the Real-Time Detection of Extraterrestrial Transients

Alexander Faustmann and Jacki Gilmore
Stellenbosch University, Stellenbosch, South Africa, 7599, http://www.sun.ac.za/

Abstract

The detection and classification of extraterrestrial transients has become increasingly important to the scientific community. Current methods for achieving these goals are effective but computationally expensive. This paper investigates the potential of the augmented form of the least mean squared adaptive processor for real time beamforming in order to assist in the detection of transient signals. The extension of the adaptive least mean squared processor to the wideband non-linear form is detailed and a brief overview of bispectral analysis is given. The augmented LMS processor is found to be capable of detecting a desired signal based on its nonlinear properties.

1 Introduction

Much of the historical development of radio astronomy has been focused on the task of imaging the sky with increasing accuracy, such that images can be made of objects ever further away. The signals emitted by such distant sources require extremely sensitive instruments in order to detect. While the construction of larger and more sophisticated telescopes aids in achieving the desired sensitivity, recovering signals which are orders of magnitude below the noise floor of the receiving instruments requires the application of sophisticated signal processing. Many of the signal processing techniques used in synthesis imaging are highly effective at capturing the slowly varying sky, but are detrimental to transient signals. Recently, an increased interest in the detection and study of extraterrestrial transients such as pulsar emissions and fast radio bursts (FRBs) has led to a demand for better transient detection and classification algorithms.

Many approaches to the detection of transients make use of blind searches, some of which use aperture arrays to form multiple beams which are capable of searching different regions of the sky simultaneously. After detection, the classifications of radio transients typically involves processing several features which are extracted from the raw data, including signal to noise ratio (SNR), dispersion measure (DM) and bandwidth. The feature extraction and classification of radio transients is described in detail by [1]. The computational load of these detection and classification algorithms often require powerful servers and have been making increasing use of graphics processing unit (GPU) technology [2]. While parametric search algorithms and machine learning approaches work well, they are computationally expensive. The increasing data rates of modern radio telescopes leads to a demand for effective algorithms which are highly parallelizable and relatively simple.

FRBs are of particular interest currently as their origin as well as the underlying processes which generate them are largely unknown [3]. There is great interest in developing means of detecting FRBs in real time and synthesizing images of the regions of the sky from which they originate in order to determine their source. FRBs have been shown to have complex time frequency characteristics [4] and as a result of their extragalactic origin are also distinct from many other extraterrestrial transients in terms of their DMs, as shown by [5]. Given that radio transients often contain non-linear characteristics, an interesting prospect is whether the use non-linear adaptive filters would be capable of detecting unique nonlinearities contained within certain classes of signal, such as FRB or pulsar emissions.

The least mean squared error (LMS) adaptive array processor is well known for improving the performance of communication systems when provided with an adequate reference signal [6]. It is capable of adapting to its environment by nulling out unwanted interference and tracking desired signals. This paper investigates the extension of the classical LMS algorithm to the wideband non-linear case in order to detect signals with underlying known nonlinearities and to automatically form beams directed at these signals.

2 The LMS Adaptive Processor

Given a collection of N sensor elements the signal vector at the system input can be expressed as,

$$\boldsymbol{x}(t) = \begin{bmatrix} s(t-\tau_1) + n_1(t) \\ \vdots \\ s(t-\tau_N) + n_N(t) \end{bmatrix} = \boldsymbol{s}(t) + \boldsymbol{n}(t). \qquad (1)$$

In the above equation, $s(t)$ is some signal of interest, τ_i is the delay in the desired signal at the i_{th} element and $n_i(t)$ is the noise signal at the corresponding element. The noise signal may contain both externally generated interference signals as well as noise generated within the sensor elements. The output of the system is,

$$y(t) = [w_1 \quad \cdots \quad w_N]\boldsymbol{x}(t) = \boldsymbol{w}^T\boldsymbol{x}(t), \qquad (2)$$

where $\boldsymbol{w}$ is a column vector containing the real valued weights applied to the sensor elements and the superscript T denotes the transpose operation. The mean square of the difference between the output signal and the desired signal is therefore,

$$\boldsymbol{E}\{\epsilon^2\} = \boldsymbol{E}\left\{\left(\boldsymbol{s}(t) - \boldsymbol{w}^T\boldsymbol{x}(t)\right)^2\right\} \\ = \boldsymbol{S} - 2\boldsymbol{w}^T\boldsymbol{r}_{xs} + \boldsymbol{w}^T\boldsymbol{R}_{xx}\boldsymbol{w} \tag{3}$$

where $\boldsymbol{S}$ is the desired signal power, $\boldsymbol{r}_{sx}$ is a vector containing the correlation of the desired signal with each input signal and $\boldsymbol{R}_{xx}$ is the correlation matrix of the system input vector. Since the mean squared error is a quadratic equation, the optimal weights are calculated by setting its gradient to zero,

$$\boldsymbol{w}_{opt} = \boldsymbol{R}_{xx}^{-1}\boldsymbol{r}_{xs}, \tag{4}$$

which is the well-known Wiener equation. Eq. (4) represents the optimal steady state solution for the adaptive system. This choice of weights is known to produce an optimal signal to noise ratio given the second order statistics of a desired linear process and undesired linear noise. Of course, due to errors in the approximation of the signal statistics as well as many other factors, including perturbations in sensor element positions, it is unlikely that an open loop system will achieve this optimal state. The error between the output of the system and the desired signal is therefore calculated and used as feedback for weight adjustment. In a discrete time system, the control law for the system weights is give as,

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \mu\epsilon(k)\boldsymbol{x}(k). \tag{5}$$

A standard LMS system with 15 uniformly spaced sensor elements is simulated. An amplitude modulated (AM) signal acts as the desired signal and is incident on the array from a bearing of $\theta = {}^{\pi}/_{8}$ while a source of Gaussian interference lies at a bearing of $\theta = {}^{\pi}/_{3}$. The carrier frequency is provided as a reference for the processor. The resulting beam pattern is shown in dB in Fig. 1.

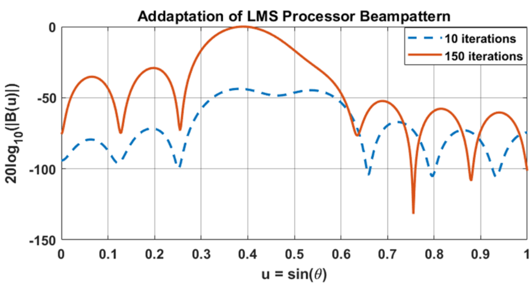

Figure 1. The adaptation of the LMS processor beam pattern. The desired signal is located at u = 0.382 while the interference signal location is u = 0.886.

A detailed description of the LMS adaptive processor as well as its transient behavior is provided by [6].

3 The Wideband LMS Adaptive Processor

In order to be used for transient detection, the system introduced in the previous section must be modified for wideband applications. A common approach to wideband array processing is the use of tapped delay lines (TDL). These lines can be used to implement temporal finite impulse response (FIR) filters which change the weights of the spatial filter in a manner which is frequency dependent, combined spatial and temporal filtering results in a space-time processor. The LMS algorithm can be extended to the TDL system in a straightforward manner. An in depth discussion of the wideband implementation of the LMS adaptive processor is given by [7]. Given a TLD adaptive processor with L delay steps, a N by L vector of signals is defined,

$$\boldsymbol{X}(t) = \begin{bmatrix} \boldsymbol{x}_1(t) \\ \vdots \\ \boldsymbol{x}_L(t) \end{bmatrix}, \\ \boldsymbol{x}_l(t) = \begin{bmatrix} x_{1l}(t) \\ \vdots \\ x_{Nl}(t) \end{bmatrix} = \begin{bmatrix} s(t-\tau_1 - l\Delta t) \\ \vdots \\ s(t-\tau_N - l\Delta t) \end{bmatrix}, \tag{6}$$

where Δt is the time step between each tap delay section. Each section in the TDL will now have its own weight vector, so that the output of the augmented system is given as,

$$y(t) = [\boldsymbol{w}_1 \quad \cdots \quad \boldsymbol{w}_L]\begin{bmatrix} \boldsymbol{x}_1(t) \\ \vdots \\ \boldsymbol{x}_L(t) \end{bmatrix} = \boldsymbol{W}^T\boldsymbol{X}. \tag{7}$$

The optimal weight vector can then be calculated using Eq. (4), except using the new signal and weight vector definitions. The weights can be updating according to Eq. (5). A tapped delay line LMS processor with 21 elements and 64 delay lines is simulated. The array is uniform and standard to a simulated frequency of 1 GHz. The desired signal is an analogue modulated signal with a bandwidth of 120 MHz, a center frequency of 500 MHz and has a bearing of $\theta = {}^{\pi}/_{6}$. The carrier frequency is provided as a reference signal. The resulting space-time beam is shown in Fig 2. The processor is found to have no main beam drift over the signal bandwidth.

4 Generalisation of the LMS Processor by Means of Volterra Series Expansion

The classical approach to systems analysis using impulse responses and single dimensional convolutions is only effective for the characterising linear systems. The general form of the output of a causal discrete time system can be expanded using the Volterra series as described by [8],

$$y(k) = \boldsymbol{h_0} + \sum_{m=0}^{m=\infty} \boldsymbol{h_1}(m)\boldsymbol{x}(k-m) + \sum_{m_1=0}^{m_1=\infty}\sum_{m_2=0}^{m_2=\infty} \boldsymbol{h_2}(m_1, m_2)\,\boldsymbol{x}(k - m_1)\boldsymbol{x}(k-m_2) + \cdots \quad (8)$$

So that system is characterised by an infinite sum of convolutions, where the nth convolution is n dimensional and represents the nth order response of the system. This method of representing nonlinear systems has the advantage that the response of the system is linearly dependent with respect to its convolution coefficients, which allows for straight forward application of gradient based optimisations. In practice, when using the Volterra expansion for system design or analysis, the highest desired order of the system response must be chosen and the Volterra series must be truncated accordingly. In this work, a quadratic LMS algorithm is used, so that the output of the space-time filter can be expressed using the first two terms of Eq. (8) and is given as,

$$y(k) = \boldsymbol{W_1X}(k) + \boldsymbol{X}(k)^{\boldsymbol{T}}\boldsymbol{W_2X}(k). \quad (9)$$

In the above equation the vector $\boldsymbol{W}_1$ determines the linear response of the system while the matrix $\boldsymbol{W}_2$ determines the quadratic response. The total error in the output of the system is now the sum of the errors caused by the linear and second order responses. The simplest method for augmenting the wideband LMS algorithm from the previous section in order to achieve the nonlinear case is by grouping all of the weights into single weight vector. The linear and second order system inputs can then also be grouped into a single input vector, so that the system output is,

$$y(k) = \boldsymbol{w^T x}(k). \quad (10)$$

This allows for the direct application of the LMS algorithm, as described in Sec. 2, except using the weight and input signal vectors which include the nonlinear coefficients.

5 Nonlinear Signal Analysis Using Higher-Order Spectra

Characterisation of nonlinear systems has been discussed in the previous section; it is now necessary to consider the analysis and characterisation of nonlinear signals. Linear random processes can be entirely characterized by their power spectral densities, provided that they meet certain stationarity criteria. In the case that a signal of interest is a stationary linear process the knowledge of its power spectrum can be used to design systems which operate upon the process. In the case of non-linear random processes, not all of the information regarding the nature of the process is contained within the power spectrum, and higher order spectra are require in order to fully characterise it. These higher-order spectra have found use

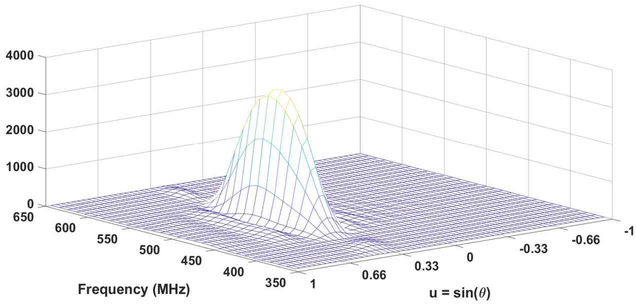

Figure 2 The space-time response beam of the wideband LMS processor. The desired signal is located at u = 0.5 and has a center frequency 500 MHz.

in many applications requiring the study and classification of signals which originate from highly non-linear processes, such as electroencephalography. A more complete description of higher-order spectra and nonlinear signal analysis is given by [9]. Just as the power spectral density of a signal is defined as the Fourier transform of its autocorrelation, the Nth order spectra is defined as the Fourier transform of the Nth order cumulant function,

$$F^n(\omega_1 \dots \omega_{N-1}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} E\,\{x(t)x(t-\tau_1) \dots x(t-\tau_{N-1})\}e^{-j\omega_1\tau_1} \dots e^{-j\omega_{N-1}\tau_{N-1}}d\tau_1 \dots d\tau_{N-1}. \quad (11)$$

In the case of a stationary process, the Nth order spectra has N-1 dimensions. Higher order spectra have symmetrical properties similar to those of the first order Fourier spectrum. A particularly advantageous property of higher order spectra is that they are high SNR domains, as spectra higher than order two are identically zero for Gaussian processes. A quadratic system is used in this work and so the spectrum of interest in this case is that of order three, which is commonly referred to as the bispectrum and is defined as,

$$B(\omega_1, \omega_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\,\{x(t)x(t-\tau_1)x(t-\tau_2)\}e^{-j\omega_1\tau_1}e^{-j\omega_2\tau_2}d\tau_1 d\tau_2 = F(\omega_1)F(\omega_2)F^*(\omega_1+\omega_2). \quad (12)$$

Its relation to the third order moment function implies that, statistically, the bispectrum is a measure of the average skew of a process. From a deterministic view, the primary nonlinearity represented by the bispectrum is that of frequency mixing.

6 Detection of Signals with Known Non-linear Characteristics

It is already known that adaptive non-linear filters are capable of non-linear channel equalization, the objective of the experiment was to determine whether the processor is capable of detecting a desired signal when the provided reference possesses the same nonlinear characteristics as the desired signal. In the case of the linear LMS processor, an AM signal can be recovered by providing the carrier frequency as a reference. It was stated in Sec. 5 that frequency mixing is fundamental to the bispectrum. A modulation triplet was therefore provided as the reference signal for a quadratic LMS processor.

The LMS algorithm described in Sec. 4 was simulated for a system with 15 sensor elements and 25 delay sections. The provided reference signal was,

$$r(t) = \cos(2\pi f_1 t) + \cos(2\pi f_2 t) + \cos(2\pi(f_1 + f_2)t), \tag{20}$$

where f_1 = 150 MHz and f_2 = 400 MHz. A random signal with a bandwidth of 60 MHz and a center frequency of 150 MHz was mixed with a sinusoid of frequency of 400 MHz. The mixed signal was then added to the original signal along with the modulated sinusoid in order to construct the desired signal. An interfering Gaussian noise source was introduced at a bearing which was different to that of the desired signal. It was found that the processor is capable of identifying the desired signal and attempted to block the interferer, as shown in Fig. 3. The quality of the reconstructed signal was found to be influenced by the number of delay sections, receiving elements and feedback gain. The system also took an average of 1 ns before it converged to track the desired signal. Lastly, it was found that many of the weights decayed to zero, potentially indicating that the processor can be pruned in order to improve its computational efficiency.

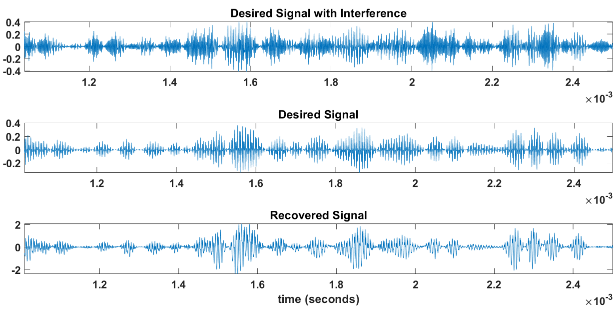

Figure 3. The nonlinear process is recovered from the interference

7 Conclusion

Simulations show that the TDL non-linear LMS processor is capable of detecting a desired signal with known non-linear properties while suppressing linear interference sources. The performance of the system when confronted with non-linear terrestrial signals is still to be determined. While one of the primary disadvantages of the higher order LMS processor is the very large number of weights it possesses, the algorithm is highly parallelizable, making it suitable for FPGA implementation in radio frequency front ends. It may also be possible to prune weights which are found to decay to negligible values. The LMS algorithm is a time domain algorithm and so it avoids the need for FFTs and IFFTs, as well as avoiding complex multiplications on account of using only real weight values.. Dedispersion is known to be very computationally expensive, an interesting prospect for investigation in future work may be the ability of adaptive processors to simultaneously detect and didesperse transient signals.

8 Acknowledgements

I would like to acknowledge my supervisor, Dr J. Gilmore, for permitting me great freedom in my research.

9 References

1. W. Farah et al., "Five new real-time detections of fast radio bursts with UTMOST", *Monthly Notices of the Royal Astronomical Society*, vol. 488, no. 3, pp. 2989-3002, 2019. Available: 10.1093/mnras/stz1748.

2. A. Magro, A. Karastergiou, S. Salvini, B. Mort, F. Dulwich and K. Zarb Adami, "Real-time, fast radio transient searches with GPU de-dispersion", *Monthly Notices of the Royal Astronomical Society*, vol. 417, no. 4, pp. 2642-2650, 2011. Available: 10.1111/j.1365-2966.2011.19426.x.

3. J. Cordes and S. Chatterjee, "Fast Radio Bursts: An Extragalactic Enigma", *Annual Review of Astronomy and Astrophysics*, vol. 57, no. 1, pp. 417-465, 2019. Available: 10.1146/annurev-astro-091918-104501.

4. J. Hessels et al., "FRB 121102 Bursts Show Complex Time–Frequency Structure", *The Astrophysical Journal*, vol. 876, no. 2, p. L23, 2019. Available: 10.3847/2041-8213/ab13ae.

5. L. Spitler et al., "Fast Radio Burst Discovered In The Arecibo Pulsar Alfa Survey", *The Astrophysical Journal*, vol. 790, no. 2, p. 101, 2014. Available: 10.1088/0004-637x/790/2/101.

6. R. Monzingo, R. Haupt and T. Miller, *Introduction to adaptive arrays*. Raleigh, NC: SciTech Pub., 2011.

7. Z. Ahmad, S. Yaoliang and Q. Du, "Adaptive Wideband Beamforming Based on Digital Delay Filter", *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, vol. 15, no. 3, pp. 261-274, 2016. Available: 10.1590/2179-10742016v15i3630.

8. A. Zaknich, *Principles of adaptive filters and self-learning systems*. [London]: Springer London, 2005.

9. C.L. Nikias and A.P Petropulu, *Higher-Order Spectra Analysis A Nonlinear Signal Processing Framework.* New Jersey: PTR Prentice-Hall, Inc., 1993.