THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Compressive Sensing Technique for Mitigating Nonlinear Memory Effects in Radar Receivers

Compressive Sensing Technique for Mitigating Nonlinear Memory Effects in Radar Receivers

Euan Ward, *Student Member, IEEE,* Shahzad Gishkori, *Senior Member, IEEE* and Bernard Mulgrew, *Fellow, IEEE*

Abstract—Receiver nonlinearities pose a serious risk to the functionality of modern radar as they can compromise the sensor's immunity to interfering signals. With the radio frequency (RF) spectrum becoming increasingly crowded, it is now more important than ever that the sensor can maintain system performance when exposed to mutual interference. In this paper, we present a nonlinear compressive sensing (NCS) solution which, unlike the standard nonlinear equalization (NLEQ) techniques, is designed around the forward nonlinearity rather than the inverse. Importantly, in this study the NCS theory is extended to include nonlinear memory. Furthermore, a radar specific formalisation is derived which allows the nonlinear optimization to exploit the unique properties of pulsed-Doppler radar processing. As a result, the NCS solution can successfully restore system sensitivity back to the linear case when in-band interference drives the radar receiver into its nonlinear regime. Additionally, it is shown that the technique can consistently mitigate complex nonlinear memory effects generated in the RF receiver. This is a significant result as it proves that forward modelling techniques are a viable alternative to NLEQ. This is of particular importance to radar systems as they provide a far more explicit formalisation to mitigate nonlinear memory effects.

Index Terms—Radar, Mutual Interference, Receiver Nonlinearity, Nonlinear Equalization, Forward Modelling, Nonlinear Optimization, Nonlinear Compressive Sensing

I. INTRODUCTION

THE multi-sensor mutual interference problem is well-known in the radar community with the development of an intelligent radar network being the most widely accepted proposed solution [1]. The future of modern radar will undoubtedly involve a network of cognitive sensors that are frequency agile so that mutual interference is minimised [2]. However, some self-sufficient protection operating within each sensor is required if performance is to be maintained even in the worst interference scenario. In the case where an interferer is present in the radar's desired channel, the linearity of the radio frequency (RF) amplifiers in the receiver front-end effectively defines the system's immunity to the unwanted signal [3]–[5]. As long as the active components in the RF receiver remain in the linear regime, while the desired signal and the interference signal pass through the chain simultaneously, then the principle of superposition holds and the task of removing the interference effects can be seen as a linear signal separation problem [6].

E. Ward is a PhD Student at The University of Edinburgh supervised by B. Mulgrew, Institute for Digital Communications, Joint Research Institute in Signal and Image Processing, School of Engineering and Electronics, Edinburgh, EH9 3JL, UK, Email: euan.ward@ed.ac.uk; B.Mulgrew@ed.ac.uk

Shahzad Gishkori is with the King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia, Email: shahzad.gishkori@kaust.edu.sa
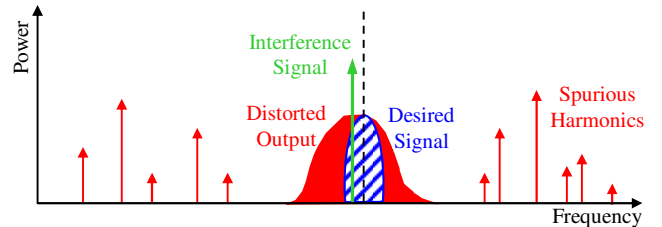

Fig. 1: Illustration of nonlinear distortion effects caused by an in-band interferer. Figure adapted from [9].

In an effort to save costs, modern radar systems are being designed from commercial-off-the-shelf RF components which have significantly poorer linearity specifications than their bespoke counterparts [5], [7], [8]. These modern sensors must therefore operate at the top of their linear regions in order to achieve acceptable performance. Crucially, this leaves no room to *back-off* the receiver in the presence of interference. Operating at the top of the linear region poses a serious risk to the functionality of these sensors as interference from other nearby radars can force the RF amplifiers in its receiver into a nonlinear state. With the use of consumer radars set to increase dramatically over the next decade modern mass-produced radar systems must evolve so that they can operate effectively beyond a purely linear regime.

The diagram in Fig. 1 illustrates the potential nonlinear effects an in-band interferer can cause if a radar's receiver linearity cannot be preserved. In contrast to the linear receiver, where the interferer and the desired signal are amplified separately, the nonlinearity causes the desired signal to become distorted as well as generating spurious harmonics across the spectrum [3]–[5], [10]–[12]. This paper presents a novel digital signal processing solution that is capable of mitigating the *in-channel* nonlinear distortion caused by the presence of an in-band interferer. It is assumed for the purposes of this paper that any spurious harmonics generated by the nonlinearity are successfully filtered out further down the RF receiver chain and are therefore omitted from this study.

While little work has been published on the mitigation of nonlinear receiver effects in the radar literature [4], [5], [10], [11], [13]–[15] it is an area of active research in the related field of communications. To save costs it is common for communications systems to operate their receivers in the weakly nonlinear regime and compensate for the nonlinear distortion digitally [5], [7], [8], [16], [17]. By far the most popular of these techniques is nonlinear equalization (NLEQ) where the corrupted signal is passed through an inverse

of the forward nonlinearity so that the dominant nonlinear effects are suppressed. There have been several studies into to the use of NLEQ techniques in radar [5], [10], [13]–[15]. In principle, they offer an attractive solution to the nonlinear receiver problem as they are mode independent and computationally inexpensive to implement. However, they are notoriously difficult to identify for systems with memory [15], [18] and their effectiveness is severely limited by the power of the input signals [15], [19].

The memory of a system is intrinsically linked to its bandwidth [8], [20] and with modern radars moving to increasingly wider receiver bandwidths they can no longer be approximated as strictly memoryless systems. It was shown in [11] that as nonlinear memory effects are introduced to the radar receiver, memoryless NLEQ techniques become increasingly decorrelated significantly reducing their effectiveness. In order for NLEQ techniques to be effective in mitigating nonlinear memory effects in the modern radar receiver, nonlinear inverse models with memory must be used. However, the functional form of these memory rich nonlinear equalisers is an open question in the radar literature as usually more memory and complexity is required in the inverse than the forward nonlinearity. Equally the nature of the signal required to train such a model is unclear [15], [18]. It is therefore thought that an alternative more robust solution must be sought when memory effects become significant in the RF receiver.

Forward modelling offers an alternative solution to standard NLEQ techniques. The technique removes the need for a complex nonlinear inverse structure as instead the forward nonlinear model is integrated into the signal processing itself. The extension of nonlinear behavioural models to included memory effects is well defined for the forward model but not so for the corresponding inverse structure. Designing a nonlinear mitigation solution around the forward nonlinear model rather than the inverse allows the technique to be extended to include memory effects in a well-defined manner. The problem can now be formulated as a nonlinear signal processing one where the signal of interest is observed through a nonlinear system.

The fact that the observations are intrinsically nonlinear means that traditional transform techniques cannot be used to infer the various signal components. Sparse signal processing offers a forward modelling solution. By setting the problem up as a sparse nonlinear optimisation problem the linear input signal can be estimated from the measured nonlinear output signal. Like most sparse signal processing problems, iterative solutions which are signal dependent must be sought. In the case of radar, this signal dependence results in the technique being mode specific. Nonlinear sparse signal processing is in its infancy with only one major paper published on the topic [21]. In [21] Blumensath showed that a sparse or structured signal observed through a nonlinear function could be recovered by the iterative hard thresholding (IHT) algorithm with constraints similar to those required in the linear setting. Importantly, Blumensath's proof holds under the condition that the system is not too nonlinear and therefore the error introduced in the linearisation is not too large.

Since [21] was published in 2013, Blumensath's IHT al-

gorithm for nonlinear compressive sensing (NCS) has been applied to a wide variety of different applications. In [22] Chen *et al.* apply the technique to a communications system for the first time with the aim of digitally correcting the nonlinear distortion from a front-end power amplifier leading to improved efficiency. The nonlinearity was described by means of a Rapp model and the NCS algorithm was shown to converge for both simulated and experimental data. This is a significant result as it proves that the NCS algorithm can be successfully employed to digitally mitigate nonlinear effects in a communications-based system. While Chen *et al.* employ Blumensath's NCS technique to mitigate a transmit side nonlinearity, Zhu *et al.* develop the technique further in [23] by applying it to the receive side of the system. The NCS technique in [23] focuses on a front-end LNA driven into its weakly nonlinear regime by the presence of an unwanted interferer in the spectrum. The LNA was modelled by means of a cubic order Taylor series and Blumensath's IHT technique was tested alongside a novel gradient pursuit method. Crucially, Zhu *et al.* prove in [23] that the linear input signal can be recovered by the NCS technique if the LNA is not too nonlinear and the Taylor series model parameters are known. While [22], [23] provides some encouragement, Blumensath's NCS technique [21] must be developed specifically to exploit the unique behaviour of radar systems if it is to be successful in mitigating nonlinear effects generated in the radar receiver.

In this paper, the NCS algorithm is extended and applied to mitigate nonlinear effects in the radar receiver caused by the presence of an in-band interferer. In addition to introducing this novel digital mitigation technique to the field of radar, the following key contributions are presented:

1) Input and output noise are addressed as part of the complex baseband nonlinear model, see section III-B.
2) The forward nonlinear function in the NCS algorithm is extended to include nonlinear memory effects, see section III-C.
3) Unlike previous work that deals with the purely real problem, the NCS algorithm is extended to the complex case, see section III-D.
4) A unique CS problem that exploits signal sparsity in slow-time and tackles a complex baseband nonlinearity with memory in fast-time is addressed, see section III-E.
5) The NCS algorithm is shown to mitigate nonlinear memory effects in the radar receiver for the first time, see section V.
6) The NCS algorithm is tested against real radar data for the first time, see section VI.

In section II, the interference scenario and the system model used to describe the nonlinear RF receiver are presented. In section III, the novel NCS algorithm is derived in full before the simulation used to test the performance of the technique are discussed in Section IV. Finally, simulated and experimental results are displayed in Sections V and VI respectively with the performance of the NCS algorithm discussed in detail.

II. SYSTEM MODEL

The Volterra series is a powerful black-box behavioural model capable of modelling complex nonlinear memory ef-

fects. It is often described as a Taylor series with memory and has a unique linear-in-the-parameters formalisation that makes it ideal for modelling weakly nonlinear RF amplifiers [8]. Mathematically it is described as a sum of multidimensional convolutions and in this paper we employed it in its complex baseband (BB) form [24] for computational efficiency. Importantly, when deriving the BB equivalent form of the Volterra series (BBVS) [9], [11], [24] the narrowband assumption is invoked which restricts its modelling capabilities to a small bandwidth around the carrier frequency. Referring to the nonlinear scenario illustrated in Fig. 1, this mean that only the *in-channel* distortion effects can be captured by the model. We must therefore assume that any spurious harmonics generated by the passband nonlinearity are successfully filtered out further down the RF receive chain. The full form of the BBVS model employed in this paper is displayed in (1) below:

$$y_{\{n\}} =$$
$$\sum_{p=0}^{P} \left\{ \sum_{i_1=0}^{L-1} \cdots \sum_{i_{2p+1}=0}^{L-1} \right\} h_{\{2p+1,\{i_1,\ldots,i_{2p+1}\}\}} \prod_{s=1}^{p} x_{\{n-i_s\}}^* \prod_{d=p+1}^{2p+1} x_{\{n-i_d\}} \tag{1}$$

where $y_{\{n\}}$ denotes the current BB output sample, $x_{\{n\}}$ represents the current BB input sample and P specifies the nonlinear model order. The BBVS model [9], [11], [24] described above takes a purely odd order form as the distortion generated from the even order terms fall outside the desired channel [24] and can therefore be ignored. Importantly, each power term in (1) describes a sum over Volterra taps. We use the phrase Volterra tap here to describe a single term from the multidimensional sum over $\{i_1, \ldots, i_{2p+1}\}$. In essence, the memory length L defines the amount of memory in the system with the terms $\{i_1, \ldots, i_{2p+1}\}$ defining each subsequent component's BB sample delay. The specific behaviour of the nonlinear model is defined by a set of nonlinear coefficients $h_{\{2p+1,\{i_1,\ldots,i_{2p+1}\}\}}$ referred to as the Volterra kernel coefficients in the case of the BBVS model. Importantly, in the memoryless case where $L = 1$ the BBVS model reproduces the baseband Taylor series (BBTS) model exactly[1]. The complexity of the BBVS model grows exponentially with nonlinear order P and memory length L, however even for small values of P and L the BBVS model can capture complex nonlinear memory effects. For the purposes of this paper, we simplify the entire radar RF receiver to be a single amplifier stage and choose to model its behaviour using one of the nonlinear black box behavioural models discussed above.

As mentioned previously, Fig. 1 illustrates the scenario of interest where the presence of an in-band interferer drives the radar receiver into the weakly nonlinear region causing the desired signal to be distorted. We therefore describe the BB input signal $x(t)$ as consisting of the desired radar signal $v(t)$ and the in-band interference signal $s(t)$.

$$x(t) = v(t) + s(t) \tag{2}$$

Importantly, the desired radar signal $v(t)$ consists of a linear combination of target, clutter and noise signals while we

choose signal $s(t)$ to be a continuous-wave (CW) interferer. For simplicity we assume that the radar transmits a single-tone pulsed signal and the CW interferer is ever-present in the received pulse repetition interval (PRI). The discrete-time input for the black-box receiver model (1) is therefore formed by sampling the continuous-time signal $x(t)$ through the Nyquist-Shannon sampling theorem [25] and stacking the fast-time signal samples $x_{\{n\}}$ into a column vector $\mathbf{x}$ of length N. Each received PRI signal is indexed with the subscript $\mathbf{x}_q$, for $q = 1 \ldots Q$, and is concatenated into a received signal matrix of size $N \times Q$ labelled $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_Q \end{bmatrix}$ in what is referred to as the slow-time dimension. The corresponding output matrix $\mathbf{Y}$ is constructed similarly by concatenating each subsequent PRI output $\mathbf{y}_q$ from the black-box receiver model into an $N \times Q$ matrix as follows $\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \ldots & \mathbf{y}_Q \end{bmatrix}$. The nonlinearity is therefore applied down the columns of matrix $\mathbf{X}$ to produce matrix $\mathbf{Y}$ acting in what is referred to as the fast-time dimension[2].

In this paper we focus on airborne pulsed Doppler radar and more specifically the medium pulse repetition frequency (MPRF) mode [6], [25], [26] for several key reasons:

1) The MPRF mode is one of the most popular modes of operation for modern radar as it provides both range and velocity measurements of the scene.
2) In a typical MPRF mode the radar is tasked with detecting low signal to noise ratio (SNR) targets in the presence of a strong clutter signal which can leave it susceptible to nonlinear effects.
3) The MPRF mode is intrinsically narrowband which as mentioned previously is the main assumption that must be satisfied when employing BB behavioural models to simulate the RF receiver.
4) The unique characteristics of the clutter in the MPRF mode makes it well-suited to the NCS mitigation technique.

Typically, in an MPRF radar mode pulse-Doppler processing is employed to detect targets in the scene. In standard pulse-Doppler radar processing [6], [25], a matched filter is applied down the columns of $\mathbf{Y}$ to convert the raw fast-time signals into range before a discrete-time Fourier transform (DFT) is applied across the received PRIs to reveal the Doppler behaviour of the artifacts in the scene. By performing this simple processing on the received data, the weak targets can be separated from the clutter in Doppler and due to their coherent nature can also be pulled above the noise floor. It is in this range-Doppler (RD) space that target detection analysis is typically performed. An example RD plot for the in-band interference scenario in the case where the RF receiver operates in its linear regime is displayed in Fig. 2.

In the case where an interferer drives a nonlinearity in the RF receiver, the resulting distortion generated in the RD

---

[1]The coefficients for the BBTS model are distinguished from those for the full BBVS model by omitting the delay-tap indices $\{i_1, \ldots, i_{2p+1}\}$ from the notation.

[2]A single matrix element is denoted by a non-bold symbol with dual subscripts, $x_{\{n,q\}}$. Furthermore, specific rows and columns of a matrix are denoted by a bold lowercase symbol with a single subscript that indicates the matrix dimension that has been *pulled out*. Importantly we use a tilde accent to distinguish row vectors from column vectors in the notation. Thus, $\tilde{\mathbf{x}}_n$ denotes a vector corresponding to row n of matrix $\mathbf{X}$ of size $1 \times Q$ and $\mathbf{x}_q$ denotes a vector corresponding to column q of matrix $\mathbf{X}$ of size $N \times 1$.
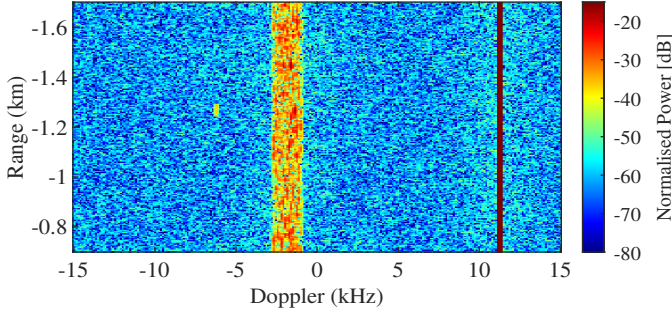
Fig. 2: Example RD plot for the linear in-band interference scenario. The target $\{-6.2kHz, -1.3km\}$ is clearly visible in the noise limited region away from the clutter at $-3.0kHz \rightarrow -0.7kHz$ and the CW interferer at $11.2kHz$ Doppler.

domain can have devastating effects on the radar's target detection performance. The most common in-band nonlinear distortion effect observed in the RD space involves the *broadening* of the main-beam clutter [14]. This clutter *broadening* effect can result in weak targets being masked by the nonlinear distortion. Importantly, the introduction of nonlinear memory effects is not expected to change the gross nonlinear effect described above but rather subtly change the structure and phase of the nonlinear distortion observed. While these subtle differences might not be enough to alter the target detection performance from the memoryless case, they can decorrelate the memoryless mitigation techniques [11].

III. NONLINEAR COMPRESSIVE SENSING

A. Signal Model

In this section we derive the novel NCS algorithm in full. We start by considering the input signal vector $\mathbf{x}_q$ for the qth PRI which consists of a linear combination of received pulses from both the radar scene and the interference source. As mentioned previously, $\mathbf{x}_q$ denotes the signal corresponding to column q of matrix $\mathbf{X}$ and contains a series of fast-time signal samples n. In standard pulse-Doppler radar processing a matched filter is applied down each received PRI in order to convert fast-time, n, into range, r. This is illustrated formally for a single PRI signal in (3).

$$\mathbf{x}_q = \sum_{r=1}^{R} \boldsymbol{\psi}_r \, x'_{\{r,q\}} = \boldsymbol{\Psi} \mathbf{x}'_q \qquad (3)$$

where $\boldsymbol{\psi}_r$ represents column vector r of the fast-time filter matrix $\boldsymbol{\Psi}$ size $N \times R$ which can take the form of a Toeplitz convolution matrix; and $\mathbf{x}'_q$ denotes the range converted signal vector for PRI q size $R \times 1$. Now that the received PRI signals have been converted from fast-time into range, their underlying structure can be studied in the slow-time dimension which acts along the rows of matrix $\mathbf{X}'$ where $\mathbf{X}' = \begin{bmatrix} \mathbf{x}'_1 & \mathbf{x}'_2 & \dots & \mathbf{x}'_Q \end{bmatrix}$. Note in this paper we adopt the convention whereby row vectors are denoted by the accent tilde e.g. $\tilde{\mathbf{x}}'_r = [x'_{\{r,1\}} \; x'_{\{r,2\}} \; \dots \; x'_{\{r,Q\}}]$. Let us therefore examine the slow-time signal vector for a single range gate r. The observed signal vector $\tilde{\mathbf{x}}'_r$ can be represented as a linear combination of J frequency vectors $\{\tilde{\boldsymbol{\phi}}_j\}_{j=1}^{J}$ where each

frequency vector is weighted by the corresponding complex coefficient $\theta_{\{r,j\}}$. Therefore the row vector $\tilde{\boldsymbol{\theta}}_r$, where $\tilde{\boldsymbol{\theta}}_r = [\theta_{\{r,1\}} \; \theta_{\{r,2\}} \; \dots \; \theta_{\{r,J\}}]$, contains a set of Fourier coefficients that describes the slow-time behaviour of the scatterers at a particular range gate r. This is displayed formally in (4) below,

$$\tilde{\mathbf{x}}'_r = \frac{1}{\sqrt{J}} \sum_{j=1}^{J} \theta_{\{r,j\}} \, \tilde{\boldsymbol{\phi}}_j = \tilde{\boldsymbol{\theta}}_r \boldsymbol{\Phi} \qquad (4)$$

where $\boldsymbol{\Phi}$ is a frequency basis matrix of size $J \times Q$ comprising the row vectors $\tilde{\boldsymbol{\phi}}_j$. By combing (3) and (4) we can define the two-dimensional linear signal model (5) which effectively describes the simple radar processing operation applied in a pulse-Doppler radar. In (5), the Fourier vectors $\tilde{\boldsymbol{\theta}}_r$ have been concatenated in range to form the matrix $\boldsymbol{\Theta}$ of size $R \times J$, e.g. $\boldsymbol{\Theta} = \begin{bmatrix} \tilde{\boldsymbol{\theta}}_1^T & \tilde{\boldsymbol{\theta}}_2^T & \dots & \tilde{\boldsymbol{\theta}}_R^T \end{bmatrix}^T$ where ^{T} denotes the matrix transpose operation.

$$\mathbf{X} = \boldsymbol{\Psi} \boldsymbol{\Theta} \boldsymbol{\Phi} \qquad (5)$$

Henceforth the matrices in (5) are referred to as follows: $\boldsymbol{\Psi}$ represents the range projection matrix that maps range to fast-time, $\boldsymbol{\Phi}$ describes the slow-time Fourier basis matrix and $\boldsymbol{\Theta}$ denotes the corresponding signal model regressors. Importantly, in this paper we assume that the radar transmits a single-tone waveform resulting in the matched filter operation being redundant as fast-time is directly equivalent to range in our received signal matrix. This assumption simplifies the structure of the range projection matrix $\boldsymbol{\Psi}$ to be an identity matrix of size $N \times R$ where $N = R$. If the radar were to transmit a more complex waveform then the range projection matrix $\boldsymbol{\Psi}$ would need to be updated to represent the new mapping from range to fast-time. In contrast, the Fourier basis matrix has a size $J \times Q$ and is constructed so that each column of the orthonormal matrix $\boldsymbol{\Phi}$ represents an independent frequency vector. When dealing with the radar signal model (5), it is often convenient to adopt the elementwise formalisation stated in (6) below.

$$x_{\{n,q\}} = \sum_j \left\{ \sum_r \Psi_{\{n,r\}} \theta_{\{r,j\}} \right\} \Phi_{\{j,q\}} \qquad (6)$$

In (6), the subscripts n, q, r, j are used to denote the elements of the corresponding matrix[3]. Additionally, the subscripts are used to label the dimensions of each specific matrix with: n representing the fast-time dimension, q indicating the slow-time dimension, r denoting the range dimension and j representing the Doppler dimension. Importantly, it is clear from (6) that the model regressors themselves describe the entire range-Doppler detection space.

Finally, we define the nonlinear observation model (7) which relates the measured nonlinear output matrix $\mathbf{Y}$ to the linear regressor matrix $\boldsymbol{\Theta}$ by combining the radar signal model (5) with the specific black-box nonlinear receiver model denoted by function $\Gamma(.)$.

$$\mathbf{Y} = \Gamma(\mathbf{X}) = \Gamma(\boldsymbol{\Psi} \boldsymbol{\Theta} \boldsymbol{\Phi}) \qquad (7)$$

By integrating this nonlinear observation model into the radar processing, the problem of recovering the linear input signal

---

[3]When indexing elements from a matrix, a lowercase symbol is used if the variable represents a set of signals while an uppercase symbol is used if it represents a basis/projection matrix.

X is equivalent to estimating the model regressor matrix Θ. Importantly, this allows prior knowledge of the radar detection space to be incorporated into the estimation process. The problem can therefore be setup as a sparse nonlinear optimisation one where the linear regressor matrix Θ can be estimated from the measured nonlinear output matrix **Y**. Like most solutions to sparse signal processing problems, the NCS algorithm is signal dependent and has an iterative formalisation. The algorithm's sparsity constraint and resultant signal dependence are discussed in detail at the end of this section once the iterative step has been derived in full. We start however by considering the effect of introducing noise to the nonlinear observation model in the following section.

B. Dealing with Noise

When dealing with noise in the nonlinear observation model we consider two independent noise sources: input noise **U** and output noise **W**. The former is additive noise present at the input of the receiver. The latter is additive noise present at the output of the receiver. For linear systems these could be used interchangeably [25], [26], however due the nonlinearity in the observation model we must start by treating these noise sources separately. Therefore, **U** might be thought of as representing the background noise from the radar scene with **W** representing the thermal noise in the RF receiver. If we assume that the black-box behavioural model (1) accurately captures the forward nonlinearity, then the output noise **W** can be considered to be genuinely additive and uncorrelated from sample to sample. It is therefore straightforward to address by simply tagging the $N \times Q$ noise matrix **W** to the nonlinear output matrix **Y**. It is important to note that the output noise term **W** may also be viewed as accounting for un-modelled terms in the in-band nonlinear behavioural model (1) i.e. terms beyond the Pth one. In which case, the output noise will be signal dependent and most likely correlated from sample to sample in the fast-time dimension n.

The input noise is more challenging to deal with in the case of the nonlinear observation model as unlike the linear case, we cannot assume that each output sample is a linear combination of signal plus noise. To study this further we follow the input noise through the memoryless forward nonlinear model described by (1) where $L = 1$. The full matrix form of the linear signal model with input noise is shown below. The input noise is represented by **U** and is a matrix of size $N \times Q$.

$$\mathbf{X} = \mathbf{\Psi}\mathbf{\Theta}\mathbf{\Phi} + \mathbf{U} \tag{8}$$

For mathematical simplicity we continue with an elementwise formalisation where each dimension is labelled using the previously defined subscripts. Thus, the elementwise signal model takes the following form:

$$x_{\{n,q\}} = \sum_j \left\{ \sum_r \Psi_{\{n,r\}} \theta_{\{r,j\}} \right\} \Phi_{\{j,q\}} + u_{\{n,q\}} \tag{9}$$

The memoryless observation model is formed by combining the elementwise formalisation of (7) with the BBTS black-box behavioural model from (1) where $L = 1$. Thus, the

elementwise form of the memoryless nonlinear observation model is as follows,

$$y_{\{n,q\}} = \sum_{p=0}^{P} h_{2p+1} |x_{\{n,q\}}|^{2p} x_{\{n,q\}} + w_{\{n,q\}} \tag{10}$$

Substituting (9) into (10) leads to,

$$y_{\{n,q\}} = \sum_{p=0}^{P} h_{2p+1} \left| \sum_j \left\{ \sum_r \Psi_{\{n,r\}} \theta_{\{r,j\}} \right\} \Phi_{\{j,q\}} + u_{\{n,q\}} \right|^{2p}$$
$$\left\{ \sum_j \left\{ \sum_r \Psi_{\{n,r\}} \theta_{\{r,j\}} \right\} \Phi_{\{j,q\}} + u_{\{n,q\}} \right\} + w_{\{n,q\}} \tag{11}$$

For what might be called the *signal limited case*, $\left| \sum_j \left\{ \sum_r \Psi_{\{n,r\}} \theta_{\{r,j\}} \right\} \Phi_{\{j,q\}} \right| >> |u_{\{n,q\}}|$, we find that $|x_{\{n,q\}}|^{2p} \approx \left| \sum_j \left\{ \sum_r \Psi_{\{n,r\}} \theta_{\{r,j\}} \right\} \Phi_{\{j,q\}} \right|^{2p}$. Thus, (16) can be reduced to the following form:

$$y_{\{n,q\}} = \sum_{p=0}^{P} h_{2p+1} \left| \sum_j \left\{ \sum_r \Psi_{\{n,r\}} \theta_{\{r,j\}} \right\} \Phi_{\{j,q\}} \right|^{2p}$$
$$\left\{ \sum_j \left\{ \sum_r \Psi_{\{n,r\}} \theta_{\{r,j\}} \right\} \Phi_{\{j,q\}} + u_{\{n,q\}} \right\} + w_{\{n,q\}} \tag{12}$$

and expanding out the brackets,

$$y_{\{n,q\}} =$$
$$\sum_{p=0}^{P} h_{2p+1} \left\{ \left| \sum_j \left\{ \sum_r \Psi_{\{n,r\}} \theta_{\{r,j\}} \right\} \Phi_{\{j,q\}} \right|^{2p} \right.$$
$$\left. \sum_j \left\{ \sum_r \Psi_{\{n,r\}} \theta_{\{r,j\}} \right\} \Phi_{\{j,q\}} \right\}$$
$$+ u_{\{n,q\}} \sum_{p=0}^{P} h_{2p+1} \left| \sum_j \left\{ \sum_r \Psi_{\{n,r\}} \theta_{\{r,j\}} \right\} \Phi_{\{j,q\}} \right|^{2p} + w_{\{n,q\}} \tag{13}$$

If we now re-define the signal model to remove the noise term $u_{\{n,q\}}$ as in (6), the associated memoryless nonlinear observation model becomes:

$$y_{\{n,q\}} = \sum_{p=0}^{P} h_{2p+1} |x_{\{n,q\}}|^{2p} x_{\{n,q\}}$$
$$+ u_{\{n,q\}} \sum_{p=0}^{P} h_{2p+1} |x_{\{n,q\}}|^{2p} + w_{\{n,q\}} \tag{14}$$

Examining (14) it is clear that in the *signal limited case* the associated input noise term is separate from the dominant nonlinear response described by (10). This strictly additive nature allows the input noise term to be absorbed by the output noise term $w_{\{n,q\}}$. We rename $w_{\{n,q\}}$ as system modelling error as it now encompasses both output and input noise terms. In essence, the system modelling error captures any deviation, both noise or otherwise, from the black-box behavioural model output and the true nonlinear output. Importantly, this novel approach for treating the noise in the nonlinear observation model will be validated in section IV of this paper.

C. IHT Algorithm for NCS

Now that the linear signal model has been defined in (5) we derive the IHT algorithm's iterative step in full. As with the input noise analysis above, we employ an elementwise formalisation for simplicity. By combining the BBVS expression (1) with the nonlinear observation model (7) we rewrite the nonlinear output signal $y_{\{n,q\}}$ in terms of the elementwise nonlinear function $F_{\{n,q,p,\{i_1,...,i_{2p+1}\}\}}$:

$$y_{\{n,q\}} = \sum_{p=0}^{P}\left\{\sum_{i_1=0}^{L-1}\cdots\sum_{i_{2p+1}=0}^{L-1}\right\}h_{\{2p+1,\{i_1,...,i_{2p+1}\}\}}$$
$$F_{\{n,q,p,\{i_1,...,i_{2p+1}\}\}} + w_{\{n,q\}} \tag{15}$$

where $F_{\{n,q,p,\{i_1,...,i_{2p+1}\}\}}$ takes the generalised BBVS form,

$$F_{\{n,q,p,\{i_1,...,i_{2p+1}\}\}} = \prod_{s=1}^{p} x^*_{\{n-i_s,q\}}\prod_{d=p+1}^{2p+1} x_{\{n-i_d,q\}} \tag{16}$$

In (15) and (16), i denotes the baseband sample delay as previously defined. As with any compressive sensing (CS) problem the task for the NCS IHT algorithm is to recover the desired signal $x_{\{n,q\}}$ from the corrupted output signal $y_{\{n,q\}}$. If we assume that the modelling error $w_{\{n,q\}}$ is small then we may define the observation error $e_{\{n,q\}}$ as follows:

$$e_{\{n,q\}} = y_{\{n,q\}} -$$
$$\sum_{p=0}^{P}\left\{\sum_{i_1=0}^{L-1}\cdots\sum_{i_{2p+1}=0}^{L-1}\right\}h_{\{2p+1,\{i_1,...,i_{2p+1}\}\}}F_{\{n,q,p,\{i_1,...,i_{2p+1}\}\}} \tag{17}$$

with the corresponding least squares (LS) cost function, C, defined in the usual manner (18).

$$C = \sum_{q}\sum_{n} e^*_{\{n,q\}}e_{\{n,q\}} \tag{18}$$

By seeking an iterative gradient-based solution the above cost function can by minimised with respect to the unknown linear signal model parameters Θ subject to any constraints on Θ e.g. sparsity. Unlike Blumensath's IHT algorithm for NCS [21] which dealt with the strictly real case, the cost function in (18) has been derived in the baseband domain and is therefore complex in nature. This is potentially problematic for the IHT algorithm as the usual mathematical concept of derivation to obtain a gradient is only defined for real numbers [27], [28]. An operational solution for the optimization of real cost functions as a function of a complex vector was developed by Brandwood in [29] and is applied here to deduce the gradient step. This will require a $1 \times J$ *row vector* of *co-factors* for each range gate regressor vector $\tilde{\boldsymbol{\theta}}_r$.

$$\frac{dC}{d\tilde{\boldsymbol{\theta}}_r} = \begin{bmatrix} \frac{\partial C}{\partial\theta_{\{r,1\}}} & \frac{\partial C}{\partial\theta_{\{r,2\}}} & \cdots & \frac{\partial C}{\partial\theta_{\{r,J\}}} \end{bmatrix} \tag{19}$$

and associated gradient vector,

$$\nabla_C(\tilde{\boldsymbol{\theta}}_r) = \left\{\frac{dC}{d\tilde{\boldsymbol{\theta}}_r}\right\}^* \tag{20}$$

Two equations define the gradient step in the iterative learning procedure:

$$\tilde{\boldsymbol{\theta}}_r^{\{k\}} = \tilde{\boldsymbol{\theta}}_r^{\{k-1\}} - \mu\nabla_C(\tilde{\boldsymbol{\theta}}_r^{\{k-1\}}) \tag{21}$$

and a constraint step which is applied to the entire $R \times J$ regressor matrix $\Theta^{\{k\}}$.

$$\Theta^{\{k\}} = \begin{bmatrix} \tilde{\boldsymbol{\theta}}_1^{\{k\}^T} & \tilde{\boldsymbol{\theta}}_2^{\{k\}^T} & \cdots & \tilde{\boldsymbol{\theta}}_R^{\{k\}^T} \end{bmatrix}^T$$
$$\Theta^{\{k\}} \leftarrow P_{\mathcal{A}}(\Theta^{\{k\}}) \tag{22}$$

In (21) and (22); $P_{\mathcal{A}}(.)$ denotes the projection operator that enforces the radar sparsity constraint, μ represents the step parameter and k denotes the IHT iteration number. These steps are applied repeatedly until the LS error C is suitably small at which point the algorithm can be considered to have converged. Obviously in the full formalisation of the IHT algorithm for NCS, the regressor matrix $\Theta^{\{k\}}$ must be updated as part of a loop over range gate r. This is illustrated in Algorithm 1 where the full algorithm is presented. Importantly, the function $vec(.)$ on line 7 of Algorithm 1 denotes the vectorisation operation which stacks matrices in a column-wise manner and ^{H} denotes the matrix Hermitian/conjugate transpose.

Algorithm 1 IHT algorithm for NCS. The algorithm runs for a maximum of k_{max} iterations or at least until it has converged to a minimum-error solution, see line 10. The model regressors Θ are initialized during the first iteration, see line 3, before being updated in all subsequent iterations. The error matrix $\mathbf{E}$ is calculated in line 6 from the current estimate of the input signal matrix $\mathbf{X}$ before being used to compute the LS error in line 7. The NCS gradient is calculated in lines 13-18 before the radar constraint $P_{\mathcal{A}}(.)$ is applied in line 20.

Input: k_{max}, $\boldsymbol{\Psi}$, $\boldsymbol{Y}$, $\boldsymbol{\Phi}$, Γ, R, μ, $P_{\mathcal{A}}$

1: **for** $k = 1 : k_{max}$ **do**
2:     **if** $(k = 1)$ **then**
3:        $\Theta = \boldsymbol{\Psi}^H\boldsymbol{Y}\boldsymbol{\Phi}^H$ Initialization
4:     **else**
5:        $\mathbf{X} = \boldsymbol{\Psi}\Theta\boldsymbol{\Phi}$
6:        $\mathbf{E} = \mathbf{Y} - \Gamma(\mathbf{X})$ Calculate error matrix
7:        $C = vec(\mathbf{E})^H vec(\mathbf{E})$ Compute LS error
8:        **if** $(k > 2)$ **then**
9:           **if** $(C^{\{k\}} > C^{\{k-1\}})$ **then**
10:             **break**   Convergence condition
11:           **end if**
12:        **end if**
13:        **for** $r = 1 : R$ **do**
14:           $\frac{dC}{d\tilde{\boldsymbol{\theta}}_r} = \begin{bmatrix} \frac{\partial C}{\partial\theta_{\{r,1\}}} & \frac{\partial C}{\partial\theta_{\{r,2\}}} & \cdots & \frac{\partial C}{\partial\theta_{\{r,J\}}} \end{bmatrix}$
15:           $\nabla_C(\tilde{\boldsymbol{\theta}}_r) = \left\{\frac{dC}{d\tilde{\boldsymbol{\theta}}_r}\right\}^*$
16:           $\tilde{\boldsymbol{\theta}}_r = \tilde{\boldsymbol{\theta}}_r - \mu\nabla_C(\tilde{\boldsymbol{\theta}}_r)$ Calculate NCS gradient
17:        **end for**
18:        $\Theta = \begin{bmatrix} \tilde{\boldsymbol{\theta}}_1^T & \tilde{\boldsymbol{\theta}}_2^T & \cdots & \tilde{\boldsymbol{\theta}}_R^T \end{bmatrix}^T$
19:     **end if**
20:     $\Theta = P_{\mathcal{A}}(\Theta)$ Apply radar constraint
21: **end for**
Output: Estimated $\hat{\Theta}$

Blumensath showed in [21] that the IHT algorithm can theoretically recover a sparse signal observed through a non-linear function under similar assumptions to those made in the linear CS case. However, Blumensath assumes that the nonlinear observation has a proper linear approximation which effectively projects it back into a linear function. In [22] and [23], the NCS algorithm is extended to exploit knowledge of the nonlinear function through which the input signal is observed. While the nonlinear functions in [22] and [23] are more complicated than that in [21] both are still strictly real and have a purely memoryless formalisation. For the NCS algorithm to be extended to the complex BBVS case we must consider the complex gradient calculation more carefully.

D. Calculating the Gradient

In order for the IHT algorithm to converge to the correct solution, the LS cost function C must be minimised with respect to the unknown model regressor matrix Θ subject to the chosen radar constraint $P_{\mathcal{A}}(.)$. The gradient step (21) therefore plays a crucial role in the iterative learning procedure as it constantly *steers* the algorithm towards the minimum of the cost function. The calculation of the gradient is non-trivial in the NCS algorithm as the desired signal has been observed through a complex nonlinear function. Let us start by considering a single term $\frac{\partial C}{\partial \theta_{\{r,j\}}}$ from the vector of *co-factors* described by (19). As with the previous sections in this paper we adopt an elementwise formalisation for mathematical simplicity.

$$\frac{\partial C}{\partial \theta_{\{r,j\}}} = \frac{\partial}{\partial \theta_{\{r,j\}}} \left\{ \sum_q \sum_n e^*_{\{n,q\}} e_{\{n,q\}} \right\}$$
$$= \sum_q \sum_n \left\{ e^*_{\{n,q\}} \frac{\partial}{\partial \theta_{\{r,j\}}} \{e_{\{n,q\}}\} + \frac{\partial}{\partial \theta_{\{r,j\}}} \{e^*_{\{n,q\}}\} e_{\{n,q\}} \right\}$$
$$(23)$$

Substituting (17) in for $e_{\{n,q\}}$ we find,

$$\frac{\partial C}{\partial \theta_{\{r,j\}}} =$$
$$- \sum_q \sum_n \left\{ e^*_{\{n,q\}} \sum_{p=0}^P \left\{ \sum_{i_1=0}^{L-1} \cdots \sum_{i_{2p+1}=0}^{L-1} \right\} h_{\{2p+1,\{i_1,\dots,i_{2p+1}\}\}} \right.$$
$$\frac{\partial}{\partial \theta_{\{r,j\}}} \{F_{\{n,q,p,\{i_1,\dots,i_{2p+1}\}\}}\}$$
$$\left. + e_{\{n,q\}} \sum_{p=0}^P \left\{ \sum_{i_1=0}^{L-1} \cdots \sum_{i_{2p+1}=0}^{L-1} \right\} h^*_{\{2p+1,\{i_1,\dots,i_{2p+1}\}\}} \right.$$
$$\left. \frac{\partial}{\partial \theta_{\{r,j\}}} \{F^*_{\{n,q,p,\{i_1,\dots,i_{2p+1}\}\}}\} \right\}$$
$$(24)$$

By applying the chain rule of differentiation to (24) we can expand the expression to the following form,

$$\frac{\partial C}{\partial \theta_{\{r,j\}}} =$$
$$- \sum_q \sum_n \left\{ e^*_{\{n,q\}} \sum_{p=0}^P \left\{ \sum_{i_1=0}^{L-1} \cdots \sum_{i_{2p+1}=0}^{L-1} \right\} h_{\{2p+1,\{i_1,\dots,i_{2p+1}\}\}} \right.$$
$$\left\{ \left\{ \frac{\partial F_{\{n,q,p,\{i_1,\dots,i_{2p+1}\}\}}}{\partial x_{\{n-i_1,q\}}} \frac{\partial x_{\{n-i_1,q\}}}{\partial \theta_{\{r,j\}}} + \frac{\partial F_{\{n,q,p,\{i_1,\dots,i_{2p+1}\}\}}}{\partial x^*_{\{n-i_1,q\}}} \right. \right.$$
$$\frac{\partial x^*_{\{n-i_1,q\}}}{\partial \theta_{\{r,j\}}} \right\} + \dots + \left\{ \frac{\partial F_{\{n,q,p,\{i_1,\dots,i_{2p+1}\}\}}}{\partial x_{\{n-i_{2p+1},q\}}} \frac{\partial x_{\{n-i_{2p+1},q\}}}{\partial \theta_{\{r,j\}}} \right.$$
$$\left. \left. + \frac{\partial F_{\{n,q,p,\{i_1,\dots,i_{2p+1}\}\}}}{\partial x^*_{\{n-i_{2p+1},q\}}} \frac{\partial x^*_{\{n-i_{2p+1},q\}}}{\partial \theta_{\{r,j\}}} \right\} \right\}$$
$$+ e_{\{n,q\}} \sum_{p=0}^P \left\{ \sum_{i_1=0}^{L-1} \cdots \sum_{i_{2p+1}=0}^{L-1} \right\} h^*_{\{2p+1,\{i_1,\dots,i_{2p+1}\}\}}$$
$$\left\{ \left\{ \frac{\partial F^*_{\{n,q,p,\{i_1,\dots,i_{2p+1}\}\}}}{\partial x_{\{n-i_1,q\}}} \frac{\partial x_{\{n-i_1,q\}}}{\partial \theta_{\{r,j\}}} + \frac{\partial F^*_{\{n,q,p,\{i_1,\dots,i_{2p+1}\}\}}}{\partial x^*_{\{n-i_1,q\}}} \right. \right.$$
$$\frac{\partial x^*_{\{n-i_1,q\}}}{\partial \theta_{\{r,j\}}} \right\} + \dots + \left\{ \frac{\partial F^*_{\{n,q,p,\{i_1,\dots,i_{2p+1}\}\}}}{\partial x_{\{n-i_{2p+1},q\}}} \frac{\partial x_{\{n-i_{2p+1},q\}}}{\partial \theta_{\{r,j\}}} \right.$$
$$\left. \left. \left. + \frac{\partial F^*_{\{n,q,p,\{i_1,\dots,i_{2p+1}\}\}}}{\partial x^*_{\{n-i_{2p+1},q\}}} \frac{\partial x^*_{\{n-i_{2p+1},q\}}}{\partial \theta_{\{r,j\}}} \right\} \right\} \right\}$$
$$(25)$$

As discussed previously, the derivative of a complex function of a complex variable does not exist in general. However, using the theory developed in [29] it can be defined to give a gradient for a real function C of a complex variable $\theta_{\{r,j\}}$ if $\theta^*_{\{r,j\}}$ is treated as a constant and $\frac{\partial \theta^*_{\{r,j\}}}{\partial \theta_{\{r,j\}}} = 0$. Therefore, in the context of the NCS algorithm:

$$x_{\{n-i_{2p+1},q\}} = \sum_j \left\{ \sum_r \Psi_{\{n-i_{2p+1},r\}} \theta_{\{r,j\}} \right\} \Phi_{\{q,j\}};$$
$$\frac{\partial x_{\{n-i_{2p+1},q\}}}{\partial \theta_{\{r,j\}}} = \sum_j \sum_r \Psi_{\{n-i_{2p+1},r\}} \Phi_{\{q,j\}};$$
$$x^*_{\{n-i_{2p+1},q\}} = \sum_j \left\{ \sum_r \Psi^*_{\{n-i_{2p+1},r\}} \theta^*_{\{r,j\}} \right\} \Phi^*_{\{q,j\}};$$
$$\frac{\partial x^*_{\{n-i_{2p+1},q\}}}{\partial \theta_{\{r,j\}}} = 0;$$
$$(26)$$

Additionally, since $\frac{\partial x^*_{\{n-i_{2p+1},q\}}}{\partial \theta_{\{r,j\}}} = 0$ then $x^*_{\{n-i_{2p+1},q\}}$ can also be treated as a constant for other partial derivatives when applying the chain rule. By applying these simple rules to (25) the form of the gradient can be drastically simplified. Importantly, the Fourier basis matrix Φ is not dependent on range and can therefore be removed as a common factor from the summation over p. Thus, the final expression for

the gradient takes the following form,

$$
\frac{\partial C}{\partial \theta_{\{r,j\}}} = -\sum_q \Phi_{\{q,j\}} \sum_n \Bigg\{
$$

$$
e^*_{\{n,q\}} \sum_{p=0}^{P} \left\{ \sum_{i_1=0}^{L-1} \cdots \sum_{i_{2p+1}=0}^{L-1} \right\} h_{\{2p+1,\{i_1,\ldots,i_{2p+1}\}\}}
$$

$$
\left\{ \left\{ \frac{\partial F_{\{n,q,p,\{i_1,\ldots,i_{2p+1}\}\}}}{\partial x_{\{n-i_1,q\}}} \right\} \Psi_{\{n-i_1,r\}} + \cdots \right.
$$

$$
\left. + \left\{ \frac{\partial F_{\{n,q,p,\{i_1,\ldots,i_{2p+1}\}\}}}{\partial x_{\{n-i_{2p+1},q\}}} \right\} \Psi_{\{n-i_{2p+1},r\}} \right\} \quad (27)
$$

$$
+ e_{\{n,q\}} \sum_{p=0}^{P} \left\{ \sum_{i_1=0}^{L-1} \cdots \sum_{i_{2p+1}=0}^{L-1} \right\} h^*_{\{2p+1,\{i_1,\ldots,i_{2p+1}\}\}}
$$

$$
\left\{ \left\{ \frac{\partial F^*_{\{n,q,p,\{i_1,\ldots,i_{2p+1}\}\}}}{\partial x_{\{n-i_1,q\}}} \right\} \Psi_{\{n-i_1,r\}} + \cdots \right.
$$

$$
\left. \left. + \left\{ \frac{\partial F^*_{\{n,q,p,\{i_1,\ldots,i_{2p+1}\}\}}}{\partial x_{\{n-i_{2p+1},q\}}} \right\} \Psi_{\{n-i_{2p+1},r\}} \right\} \right\}
$$

Integrating the above gradient expression with the rest of the IHT algorithm described in Algorithm 1 completes the iterative step of the NCS technique for the complex BBVS case. However, for the IHT algorithm to minimize the cost function with respect to the unknown linear signal model parameters Θ, the constraint step (22) must be applied effectively at each iteration. The nature of this radar constraint is discussed in detail in the next section.

E. Radar Constraint Step

The choice of the constraint applied at step (22) will be signal dependent and therefore unique to the radar's mode of operation. For the purposes of this paper we focus on the MPRF mode but the algorithm is not limited to this case. As mentioned previously, our unique formalisation of the NCS algorithm means that vector $\tilde{\theta}_r$ represents the Doppler spectrum at range gate r. We choose to exploit the sparsity constraint here as for the MPRF mode the dominant clutter is typically limited to a small region of the Doppler spectrum [26]. Furthermore, we expect the spectral behaviour of the clutter to be largely consistent across multiple range gates [26]. Before invoking the sparse threshold constraint $P_{\mathcal{A}}(.)$ we must first consider the unique formalisation of the NCS algorithm.

In the NCS algorithm, the memory of the BBVS model acts down the fast-time dimension while the sparse nature of the MPRF radar signal, which we wish to exploit, exists in the slow-time dimension. In the case where the nonlinearity has memory, the estimate $\tilde{\theta}_r^{\{k\}}$ not only depends on $\tilde{\theta}_r^{\{k\}}$ but also $\tilde{\theta}_{r-1}^{\{k\}}$ to $\tilde{\theta}_{r-\{L-1\}}^{\{k\}}$. Therefore, if the gradient (27) is to be calculated correctly in the memory case then every range regressor vector $\tilde{\theta}_r$ must be updated simultaneously at each iteration. This is displayed formally in (22) where the threshold constraint $P_{\mathcal{A}}(.)$ is applied to the full regressor matrix $\Theta^{\{k\}}$. Importantly, this is where we exploit our knowledge of the radar mode in the choice of sparsity level set. The

dependency between the range regressor vectors explains why the problem's complexity is intrinsically linked to the memory of the fast-time nonlinearity. This complexity ultimately manifests itself in the constraint step of the algorithm and can lead to the NCS algorithm requiring more iterations to converge to the correct solution.

There are a multitude of different ways to apply the sparse threshold holding constraint step in (22), however for this analysis we choose to employ Blumensath's hard thresholding method [21], [30]. This is illustrated in Algorithm 2 where the hard thresholding constraint implemented in this paper is presented in full. In short, the hard thresholding constraint $P_{\mathcal{A}}(\Theta^{\{k\}})$ represents a nonlinear operation where all but the $\mathcal{A}$ largest (in magnitude) elements for each range regressor vector $\tilde{\theta}_r$ are set equal to zero. While the number of retained regressors is predefined by the specified Doppler sparsity level for the MPRF radar scene, the locations of these $\mathcal{A}$ regressors are independently determined for every range regressor vector $\tilde{\theta}_r$ at each iteration k. Crucially, this allows the hard thresholding algorithm to identify artifacts in the scene such as targets, inhomogeneous clutter and interference that do not exist across all range gates. For best results the Doppler sparsity level should be set to be slightly lower than the true sparsity level of the MPRF radar scene to provide the hard thresholding algorithm with some leeway when determining the locations of the strongest regressors in each range gate vector $\tilde{\theta}_r$.

Algorithm 2 Hard thresholding algorithm that implements the radar constraint step $P_{\mathcal{A}}(\Theta^{\{k\}})$ at every iteration k of the NCS algorithm. Where the $sort(.)$ function outputs a vector of size $1 \times J$ that contains the indices of $\tilde{\theta}_r$ arranged in ascending order by the magnitude of the elements.

Input: Θ, R, J, $\mathcal{A}$
1: **for** $r = 1 : R$ **do**
2:    $\tilde{\alpha} = sort(|\tilde{\theta}_r|)$ Order regressor indices by magnitude
3:    **for** $j = 1 : \{J - \mathcal{A}\}$ **do**
4:      $\beta = \alpha_{\{j\}}$
5:      $\theta_{\{r,\beta\}} = 0$ Zero weakest $\mathcal{A}$ regressors
6:    **end for**
7: **end for**
Output: Θ

IV. SIMULATION ARCHITECTURE

In order to test the performance of the novel NCS algorithm derived above, an MPRF radar simulator capable of simulating in-band nonlinear distortion effects with memory had to be developed. In short, the comprehensive radar simulator simulated the pre-defined radar scene for a standard MPRF mode before passing the raw time-domain signals through the black-box nonlinear receiver model described by (15) and (16). The MPRF simulator was set-up specifically to generate the problematic in-band nonlinear effects detailed in section II and was configured to operate entirely in the BB domain. Once the radar scene and CW interferer characteristics were defined, the structure of the in-band nonlinear distortion effects produced in the simulation were entirely governed by the particular

nonlinear transfer function chosen. Furthermore, the gross level of the nonlinear distortion generated in the simulated RF receiver was determined by the magnitude of the specific nonlinear coefficients chosen for the memoryless case; h_1 and h_3. The magnitude of the linear coefficient h_1 was set equal to the RF receiver gain, while the magnitude of the third order coefficient h_3 was subsequently selected so that the unwanted nonlinear effects sat just below the level of the noise floor when the CW interferer was not present in the scene. In specifying the forward nonlinearity in this manner, we can assume that the simulated radar operates within its linear region when the interference is not present in the scene. Additionally, in order to ensure that the same strength of nonlinear effect was being generated in the radar receiver when memory effects were introduced to the problem, each BBVS kernel was normalised so that the respective power contributed from each group of terms in the BBVS model (15) was equivalent to that from the strictly memoryless case.

For this analysis, we assume that the dominant nonlinear receiver effects can be captured by terms up to the cubic order of the BBVS model [10], [12]. For the *in-band* nonlinear scenario, this corresponds to the case where the black-box receiver model, (15) and (16), has a nonlinear order of $P = 1$. We can therefore simplify the expression for the forward nonlinearity to the following form,

$$y_{\{n,q\}} = \sum_{i_1=0}^{L-1} h_{\{1,\{i_1\}\}} F_{\{n,q,0,\{i_1\}\}}$$
$$+ \sum_{i_1=0}^{L-1}\sum_{i_2=0}^{L-1}\sum_{i_3=0}^{L-1} h_{\{3,\{i_1,i_2,i_3\}\}} F_{\{n,q,1,\{i_1,i_2,i_3\}\}} \quad (28)$$

where,

$$F_{\{n,q,0,\{i_1\}\}} = x_{\{n-i_1,q\}}$$
$$F_{\{n,q,1,\{i_1,i_2,i_3\}\}} = x^*_{\{n-i_1,q\}} x_{\{n-i_2,q\}} x_{\{n-i_3,q\}} \quad (29)$$

Furthermore, the generalised expression for the NCS gradient can be drastically simplified for this specific case by: substituting (29) into (27), setting $P = 1$, performing the partial differentiation and collecting terms. The final expression for the NCS gradient implemented in the MPRF radar simulator therefore takes the following form,

$$\frac{\partial C}{\partial \theta_{\{r,j\}}} = -\sum_q \Phi_{\{q,j\}} \sum_n \left\{ e^*_{\{n,q\}} \sum_{i_1=0}^{L-1} h_{\{1,\{i_1\}\}} \Psi_{\{n-i_1,r\}} \right.$$
$$+ e^*_{\{n,q\}} \sum_{\{i_1,i_2,i_3\}=0}^{L-1} h_{\{3,\{i_1,i_2,i_3\}\}} \left\{ \left\{ x^*_{\{n-i_1,q\}} x_{\{n-i_3,q\}} \right\} \Psi_{\{n-i_2,r\}} \right.$$
$$\left. + \left\{ x^*_{\{n-i_1,q\}} x_{\{n-i_2,q\}} \right\} \Psi_{\{n-i_3,r\}} \right\}$$
$$\left. + e_{\{n,q\}} \sum_{\{i_1,i_2,i_3\}=0}^{L-1} h^*_{\{3,\{i_1,i_2,i_3\}\}} \left\{ x^*_{\{n-i_2,q\}} x^*_{\{n-i_3,q\}} \right\} \Psi_{\{n-i_1,r\}} \right\} \quad (30)$$

Importantly, the matrix rotations denoted by $\Psi_{\{n-i_{2p+1},r\}}$ in the above expression for the NCS gradient were performed in the simulations via a linear shift rather than a circular one. This simple extension meant that the simulations were more realistic than the theoretical NCS algorithm which is why it was implemented for all of the results presented in this paper. In essence, the linear shift introduces edge effects to the problem as not all of the range regressors vectors, $\tilde{\theta}_r$, that affect the gradient are updated as part of the IHT iteration.

While the magnitude of the linear gain and the nonlinear distortion generated in the forward black-box receiver model were fixed in the simulation, the particular choice of complex BB kernel coefficients were subject to change. It is understood that introducing nonlinear memory effects is not expected to change the gross level of the distortion observed at the output of the radar receiver but rather subtly change its structure and phase. The NCS algorithm was therefore tested against a wide variety of different nonlinear transfer functions with varying degrees of memory in order to gain a comprehensive understanding of its performance. Crucially, the nonlinear models tested in the radar simulation had to fit into the Volterra series framework otherwise the above expression for the NCS gradient (30) becomes invalid. The specific nonlinear models tested in the simulation were therefore: the BBTS model, the BBVS model, the BB Hammerstein model and the BB parallel Hammerstein model [8], [31]. For each nonlinear model listed above, the radar simulator was configured to perform statistical convergence analysis on the NCS algorithm whereby the algorithm's mitigation performance was tested against randomly generated sets of kernel coefficients.

While this broad convergence analysis is important, it is also important to determine the true performance the NCS algorithm for a MPRF radar mode with a realistic black-box nonlinear receiver model. Unfortunately, there is a severe lack of published data in the available literature on the memory behaviour of front-end receiver amplifiers. One of the very few papers that has published such data is [8] by Vansebrouck *et al.* In [8], the authors' employ a quintic, $P = 2$, parallel Hammerstein model with memory length $L = 5$ to describe a wideband communications receiver centred on 250MHz. It is clear from the data published in [8] that the third-order nonlinear effects are far more dominant than the fifth order distortion generated which allows us to truncate Vansebrouck's nonlinear model to order $P = 1$. Furthermore, to implement the nonlinear model described in [8] in our MPRF radar simulator we must translate the passband nonlinear impulse response from a centre frequency of 250MHz to baseband. In performing this translation, we convert the real passband coefficients to complex BB coefficients and disregard the even order terms in the model as they fall outside the bandwidth of the desired receiver channel. The final forward nonlinear model implemented in the MPRF radar simulator was therefore a BB parallel Hammerstein model with coefficients equal to those displayed in Table I. In terms of the full Volterra formalisation, (15), these coefficients correspond to the *on-diagonal* elements of the respective Volterra kernels, i.e. where $i_1 = i_2 = i_3$, with all other *off-diagonal* elements set equal to zero. Note, in order to distinguish this specific nonlinear model from other BBVS models employed in this paper we refer to it as the "BBVS-[8]" model as it was derived from the parallel Hammerstein model in [8]. An example RD plot outputted from the MPRF radar simulator that corresponds to

the above parallel Hammerstein model is displayed in Fig. 3 with the corresponding desired linear output displayed in Fig. 2.

i	Linear Kernel: $h_{\{1,\{i\}\}}$	Cubic Kernel: $h_{\{3,\{i,i,i\}\}}$
0	$\{1 + 0j\}$	$\{0.1886 + 0j\}$
1	$\{0\}$	$\{0 + 0.1766j\}$
2	$\{0\}$	$\{-0.0881 + 0j\}$
3	$\{0\}$	$\{0 + 0.0072j\}$
4	$\{0\}$	$\{0.1625 + 0j\}$

TABLE I: List of the dominant BBVS kernel coefficients implemented in the MPRF radar simulator for the NCS performance analysis. All coefficients not listed in Table I were set equal to zero in the BBVS-[8] model.
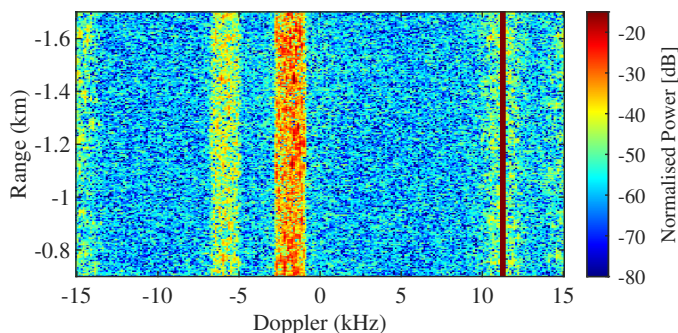

Fig. 3: Simulated RD plot for the BBVS-[8] nonlinear in-band interference scenario. Target $\{-6.2kHz, -1.3km\}$ masked by clutter *broadening* effect at $-8.0kHz \rightarrow -5kHz$ Doppler.

In order to study the performance of the novel NCS algorithm for a typical MPRF radar mode in detail, statistical probability of detection (PD) analysis was performed in the RD domain on a single target. On each burst, the MPRF radar simulator generated input receive data matrices, $\mathbf{X}$, consisting of 256 received pulses based on a pre-defined radar scene. The simulation then generated three output matrices, $\mathbf{Y}$, by passing the input data through the following black-box receiver models: the BBVS-[8] model, the corresponding BBTS model and finally the linear gain model. In all cases, both input and output noise were added to the received signals with the SNR level referenced to the input so that the output noise level was not biased by the specific nonlinear transfer function chosen, $\text{SNR}_{noise} = 35dB$. A further three data sets were then generated by applying the memory rich NCS algorithm to the BBVS-[8] receiver output and by applying the memoryless NCS algorithm to both the BBVS-[8] and BBTS outputs respectively. Simulation results for a single burst illustrating the nonlinear distortion generated from the BBVS-[8] model and the subsequent correction by the memory rich NCS algorithm are displayed in Fig. 3 and Fig. 4 respectively. It is clear from Fig. 4 that the memory rich NCS algorithm has performed well in restoring the simulated radar's performance back to the desired linear case. However, the performance and robustness of the algorithm can be studied in much more detail through the stochastic based PD analysis. It is important

to recognise at this point that the noise assumptions invoked during the derivation of the NCS algorithm have not been implemented in the radar simulator. Therefore, the fact that the NCS algorithm has recovered the correct solution in Fig. 4 clearly validates the noise approach detailed in section III-B.

As well as performing the PD analysis on the scenarios discussed above, further results were generated to probe the robustness of the algorithm for varying levels of modelling error, $w_{\{n,q\}}$. While all of the simulations conducted had some degree of modelling error due to the input and output noise added to the receive data matrices, more in-depth analysis was required to fully understand the limitations of the algorithm. The NCS algorithm is fundamentally designed around the forward black-box receiver model and therefore its performance is limited by how well the nonlinear behaviour of the RF receiver has been characterised. For a real radar system, it is highly likely that this nonlinear system identification will be performed offline and it is therefore reasonable to assume in the simulations that the forward nonlinearity has been identified to a high degree of accuracy. However, it is important to understand how much modelling error the algorithm can tolerate before it fails. To perform this modelling error analysis as part of the simulation, the MPRF radar simulator was configured to identify the forward nonlinearity for each burst through varying levels of output noise. The accuracy of the forward nonlinear coefficients was therefore dependent on the level of the output noise through which they were learnt. The linear least squares (LLS) algorithm was used to learn the nonlinear coefficients by means of a noise identification procedure [9] with the residual error value providing an accurate measure of the modelling error in the simulation.
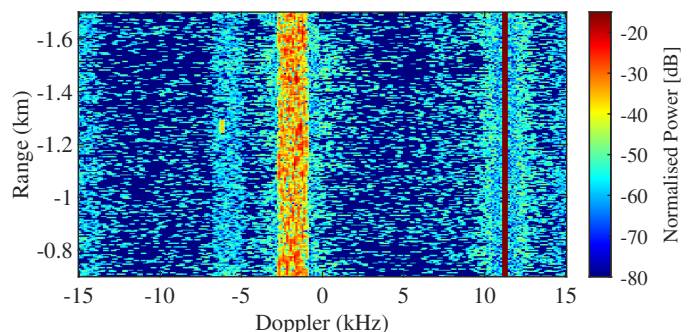

Fig. 4: RD map for the BBVS-[8] in-band interference scenario mitigated by the memory rich NCS algorithm. The algorithm has succeeded in recovering the target $\{-6.2kHz, -1.3km\}$ from the nonlinear clutter *broadening* effect.

V. SIMULATION RESULTS

A. Convergence Analysis

For the NCS convergence analysis, the mitigation algorithm was tested against different classes of randomly generated nonlinear transfer functions in order to study its overall convergence properties both in the memoryless and memory rich case. While this stochastic based convergence analysis will not

Nonlinear Model	Monotonicity	Successful Convergence %
BBTS	Strictly Monotonic	100
BBTS	Non-monotonic	49
BB Hammerstein	Strictly Monotonic	100
BB Hammerstein	Non-monotonic	53
BB Parallel Hammerstein	Strictly Monotonic	100
BB Parallel Hammerstein	Non-monotonic	89
Full BBVS	N/A	83

TABLE II: List of NCS convergence results. 100 individual nonlinear transfer functions were used to generate the successful convergence value in each case.

inform us about the effectiveness of the NCS algorithm for all nonlinear radar receivers, as there are infinitely many, it can provide us with a better understanding of the specific nonlinear characteristics that will prove problematic for convergence. We use the percentage of successful convergence as a measure of the algorithm's efficacy, with a 100% convergence defining the situation where the NCS algorithm always recovers the weak target from the corrupted RD map, see Fig. 3 and Fig. 4. The success of the mitigation technique was determined for each realisation by means of local area average (LAA) target thresholding algorithm combined with a residual error thresholding technique. The NCS modelling error was strictly limited by the receiver output noise in this case, $\text{SNR}_{noise} = 50dB$, as there was no input noise and the nonlinear coefficients were set equal to the true forward model coefficients in the NCS algorithm. The convergence results for the NCS algorithm are displayed in Table II with every convergence percentage value determined from 100 realisations of the simulation each with a unique nonlinear transfer function.

When dealing with nonlinear transfer functions, the concept of monotonicity is a fundamental one. In the case of a static nonlinearity, the transfer function is considered strictly monotonic if there exists a one-to-one mapping between the inputs and the outputs of the nonlinear model [32]. Importantly, if a function is strictly monotonic then there exists a unique inverse function that maps the outputs back onto the inputs. If however this one-to-one mapping does not exist, then the nonlinear transfer function is described as being non-monotonic and is therefore non-invertible. In the context of the NCS algorithm, non-monotonicity of the forward nonlinearity manifests itself in the LS cost function whereby the ambiguity generates multiple local minima for the algorithm to converge to. In other words, there is more than one solution to the problem. This is most easily observed for the BBTS result in Table II. We must be very careful when interpreting the convergence results in Table II as the nonlinear transfer functions generated in the simulation are not representative of a real RF receiver. However, it is clear from the results that if the memoryless forward nonlinearity is strictly monotonic then

the NCS algorithm will converge to the correct solution. In the case where the memoryless nonlinearity is non-monotonic, the NCS algorithm's ability to recover the correct solution is dependent on the particular forward nonlinear transfer function through which the desired input signal was observed.

In the case where memory is introduced to the forward nonlinearity, we have to be careful with how we interpret the concept of monotonicity which is only defined for one-dimensional functions. We start with the Hammerstein nonlinear model as it the simplest extension from the BBTS model to the memory case. The Hammerstein model is described by a memoryless nonlinearity followed by a FIR filter and therefore its monotonicity is entirely defined by its static nonlinearity. Examining the convergence results for the BB Hammerstein model we find that they reflect that of the BBTS case. This makes sense as the monotonicity of the forward nonlinearity is defined in exactly the same way as the BBTS model. Importantly, the NCS algorithm has been shown to converge to the correct solution consistently in the case where the forward nonlinearity exhibits nonlinear memory effects. Taking this analysis one step further, the Hammerstein model is extended to the parallel Hammerstein model by applying an individual FIR filter to each term in the static nonlinearity [8]. In essence, the parallel Hammerstein model consists of L static nonlinear functions that can be thought of as acting on individual taps of a FIR filter. Therefore, the nonlinear transfer function is considered strictly monotonic if all of the L static nonlinearities are themselves strictly monotonic. In this case, the convergence of the NCS algorithm is guaranteed which is indicated by the successful convergence percentage of 100% in Table II. Furthermore, we define the non-monotonic case for the parallel Hammerstein model to be when any one of the L static nonlinearities stop being strictly monotonic. Interestingly, the successful convergence percentage is much higher in this case than in the standard Hammerstein and BBTS cases suggesting that the NCS algorithm can tolerate some degree of non-monotonicity in its nonlinear memory terms. Finally, we consider the full BBVS model where unfortunately the nonlinear cross-terms mean that we cannot define the concept of monotonicity in this case. Examining the successful convergence percentage for the full BBVS case we observe that the NCS algorithm can successfully recover the desired input from the corrupted output even when the forward nonlinearity exhibits complex nonlinear memory effects. While the general convergence of the NCS algorithm cannot be guaranteed in this case, the final result highlights the capabilities of this novel mitigation technique.

B. Radar PD Analysis

For the PD performance analysis of the NCS algorithm, the MPRF radar simulator employed an adaptive cell-averaging constant false alarm rate (CA-CFAR) thresholding technique to study the simulated radar's detection performance in the different scenarios. The CA-CFAR technique is discussed in detail in [25] but in short, the algorithm was configured to maintain a specified probability of false alarm rate (PFAR) which it achieved by studying the power and statistical behaviour of

the RD cells in the neighbourhood of the target cell. The simulated radar therefore exhibits maximum sensitivity when its receiver operates in the linear regime as in this case the cells that surround the target cell in the RD map are strictly noise limited. For the nonlinear scenarios, the level of the unwanted distortion was set by the memoryless coefficients as discussed previously with the BBVS-[8] model coefficients described by those in Table I. Importantly, for this analysis both the BBTS and BBVS-[8] nonlinearities had strictly monotonic forward transfer functions. When the CW interferer is introduced to the radar scene, the unwanted clutter *broadening* effect drives the target detection threshold up and consequently reduces the sensitivity of the sensor. Thus, for the NCS algorithm to restore the system sensitivity of the radar back to the linear case it must reduce the level of distortion in the RD detection space without removing the potential targets from the scene. To study the sensitivity of the radar in the different scenarios the input SNR of the radar test target was varied with each SNR value forming a single data point in the respective PD curves. The simulation results for the PD analysis are displayed in Fig. 5 with the CA-CFAR algorithm set to maintain a PFAR of 0.2. For each data point in Fig. 5, RD maps for 400 bursts were used to estimate the PD value.
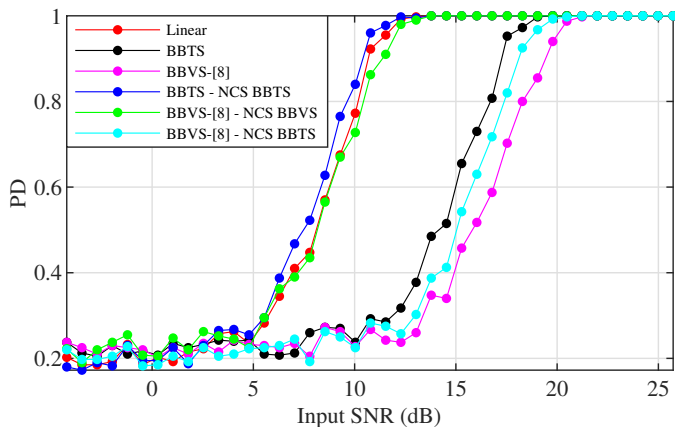

Fig. 5: Performance PD results for the NCS algorithm. Ideal system sensitivity illustrated by linear receiver curve, red; Memoryless nonlinear receiver curve, black; Memory rich nonlinear receiver curve, magenta; Memoryless nonlinear receiver output corrected by memoryless NCS algorithm, blue; Memory rich nonlinear receiver output correct by memory rich NCS algorithm, green; Memory rich nonlinear receiver output corrected by memoryless NCS algorithm, cyan.

Examining the simulation results in Fig. 5, the first thing that we observe is that the PD curves for the BBTS and BBVS-[8] receiver outputs fall off at a much higher input SNR value than that for the linear gain case. This reflects the loss of system sensitivity experienced by the simulated radar in the case where its RF receiver operates in the nonlinear regime. Somewhat surprisingly, the PD curves for the BBTS and BBVS-[8] cases do not lie on top of each other in Fig. 5 despite the strength of the nonlinearities being matched in the MPRF radar simulator. This slight deviation between the two curves is not due to any variation in the strength of the clutter

broadening effect but rather reflects subtle phase changes introduced by the nonlinear memory that alters the relationship between the nonlinear distortion and the test target. Examining the blue curve in Fig. 5, it is clear that the memoryless NCS algorithm has succeeded in fully restoring the sensitivity of the simulated radar in the case where the nonlinear RF receiver is strictly memoryless. This is a significant result as it confirms that the NCS technique can be successfully employed to mitigate nonlinear receiver effects in modern radar systems. Interestingly, when the memoryless NCS technique is applied to the memory rich BBVS-[8] output the mitigation algorithm fails to restore any of the simulated radar's lost performance. Much like the memoryless NLEQ techniques discussed in [11], the memoryless NCS algorithm is effectively decorrelated by the subtle amplitude and phase effects introduced by nonlinear memory terms in the forward receiver model. This confirms that if memory effects are significant in the RF receiver then memory rich mitigation techniques must be employed if system sensitivity is to be restored.

In the case of the NCS mitigation technique, the result of introducing complex memory effects to the algorithm's formalisation is displayed in Fig. 5, see the green curve. Clearly, introducing the memory terms to the NCS algorithm has drastically improved the mitigation technique's performance in the BBVS-[8] scenario with the simulated radar's system sensitivity more or less restored back to the linear gain case. This was fundamentally due to the expression for the NCS gradient (30) which incorporated complex nonlinear memory effects so that each iteration of the IHT algorithm was correctly pointed towards the minimum of the cost function. This is a significant result as it shows that the novel NCS algorithm provides a forward modelling framework capable of compensating for complex nonlinear memory effects generated in the modern radar receiver.

C. Modelling Error Analysis

In this section of the simulation results we examine the robustness of the NCS algorithm to varying levels of modelling error. As discussed previously, the amount of modelling error in the PD simulations was governed by the input/output noise of the simulated radar receiver as well as the error in the forward model coefficients implemented in the NCS algorithm. While the SNR ratio was fixed at 35dB for both the input and output noise in all simulations, the level of uncertainty in the NCS forward model coefficients was varied. The PD analysis performed by the MPRF radar simulator was therefore exactly the same as that described in the previous section, with the exception that the NCS algorithm was applied multiple times on each burst using different sets of forward model coefficients. We focus on the scenario where the nonlinearity is described by the BB parallel Hammerstein model in Table I, BBVS-[8], and is mitigated by the corresponding memory rich NCS algorithm. The results are displayed in Fig. 6 with each PD data point estimated from 400 bursts and the CA-CFAR set to maintain a PFAR of 0.2 as before. As discussed previously, the modelling error (ME) is given by the residual error for the LLS noise identification procedure.
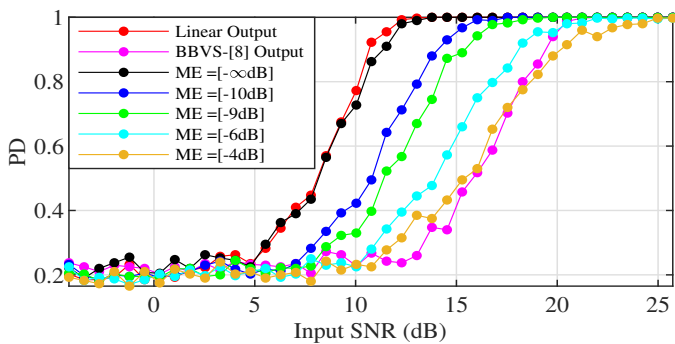
Fig. 6: Modelling error PD analysis for the NCS algorithm. Desired linear gain PD curve shown in red with the corresponding BBVS-[8] nonlinear PD curve shown in magenta. PD curves corresponding to the NCS mitigated outputs are reference by their ME level from low to high respectively. A ME value of $-\infty dB$ describes the case where the true nonlinear coefficients are employed in the NCS algorithm.

Examining the PD result in Fig. 6, it is clear that modelling error in the forward black-box receiver model can have quite a profound effect on the mitigation performance of the NCS algorithm. This makes sense as the mitigation technique is fundamentally designed around the forward nonlinearity. Encouragingly, the NCS algorithm can tolerate a reasonable amount of modelling error in the forward model coefficients before a significant drop in performance is observed. This drop-off in performance of the NCS algorithm appears to occur quite sharply around $ME = -10dB$, which suggests that beyond this point the iterative learning procedure struggles to consistently identify the minimum of the cost function and therefore the desired solution. Despite this, the NCS algorithm can still provide noticeable performance improvements from the unmitigated BBVS-[8] case even when the modelling error is significantly distorting the IHT estimate on each iteration.

VI. TRIALS DATA RESULTS

In order to study the performance and robustness of the NCS algorithm further, the novel mitigation technique was tested against real radar trials data that had been digitally corrupted by a nonlinearity and a synthetically generated frequency modulated CW (FMCW) interferer. The FMCW interference scenario was chosen for this analysis as it is an extremely popular waveform for modern radar systems [2] meaning it is also one of the most likely signals to cause mutual interference in the crowded RF environment. Importantly, unlike the simple CW interferer the FMCW waveform has a bandwidth which for this analysis was chosen to be equal to the simulated bandwidth of the radar receiver. Performing this type of analysis will further test the robustness of the NCS algorithm as the high bandwidth nature of the FMCW interferer will generate more sophisticated nonlinear memory effects in the simulated nonlinearity. This is due to the fact that the memory of a system is intrinsically linked to its bandwidth [8], [20]. The trials data used for this analysis was collected by Leonardo's experimental AEXAR system as part of a flight trial where the radar was setup to operate in an MPRF mode

with a downward looking configuration. The range, Doppler and power values of the input trials data were normalised for this analysis so that they matched those from the MPRF radar simulator. By performing this normalisation, the same nonlinear scenario studied in the previous results section could be digitally stimulated for the trials data case. Thus, the forward nonlinearity took the form of the BBVS-[8] model with coefficients equal to those listed in Table I. The final results from the trials data analysis are displayed in Fig. 7 with mitigation outputs from both the memoryless and memory rich NCS algorithm presented. In both cases, the forward model coefficients employed in the NCS algorithms were learnt by means of a noise identification procedure which importantly allowed additional modelling error to be introduced to the memory rich case. The residual LLS forward modelling error attributed to the memory rich case was therefore $-14dB$ which compared to $-6dB$ for the memoryless NCS case.

Examining the trials data results from Fig. 7 we firstly observe that the ground clutter in the linear RD map, between $-4kHz \rightarrow 5kHz$ Doppler, is noticeably more complex than the simulated clutter generated in the MPRF radar simulator. This is useful as it allows the performance of NCS algorithm to be tested against a less uniform clutter spectrum where the sparsity may not be at exactly the same Doppler locations in each range gate. Importantly, there are three distinct radar targets in the trials data RD map located around $-9kHz$ Doppler which can be used as a marker to gauge the performance of the NCS mitigation technique. In Fig. 7b, these targets are no longer distinguishable from the radar background as they have become *entangled* with the nonlinear clutter *broadening* effect stimulated by the presence of the synthetic FMCW interferer in the scene, $-4.5kHz$ Doppler. It is clear from Fig. 7c that the memoryless NCS algorithm is incapable of correcting for the nonlinear memory effects in the radar receiver resulting in the radar targets still being indiscernible from the background. Examining the final result in Fig. 7d, we observe that introducing memory terms to the NCS algorithm allows it to *unscramble* the complex nonlinear memory effects generated in the RF receiver and ultimately restore the sensor's performance almost back to the linear case. Crucially, the test targets previously lost in the nonlinear clutter *broadening* effect are now clearly identifiable above the RD background. This is an extremely encouraging result as is highlights the effectiveness of the NCS algorithm in mitigating complex nonlinear receiver effects generated by modern radars in sophisticated clutter and interference scenarios.

VII. CONCLUSION

In this paper, we presented a novel forward modelling technique designed to digitally compensate for complex nonlinear memory effects in the radar receiver. The technique builds on the work previously presented by Blumensath in [21] and was specifically targeted at the in-band interference scenario in radar. In order for the forward modelling technique to be successful, Blumensath's nonlinear compressive sensing framework had to be extended to represent the unique processing applied in modern radar systems. Furthermore,
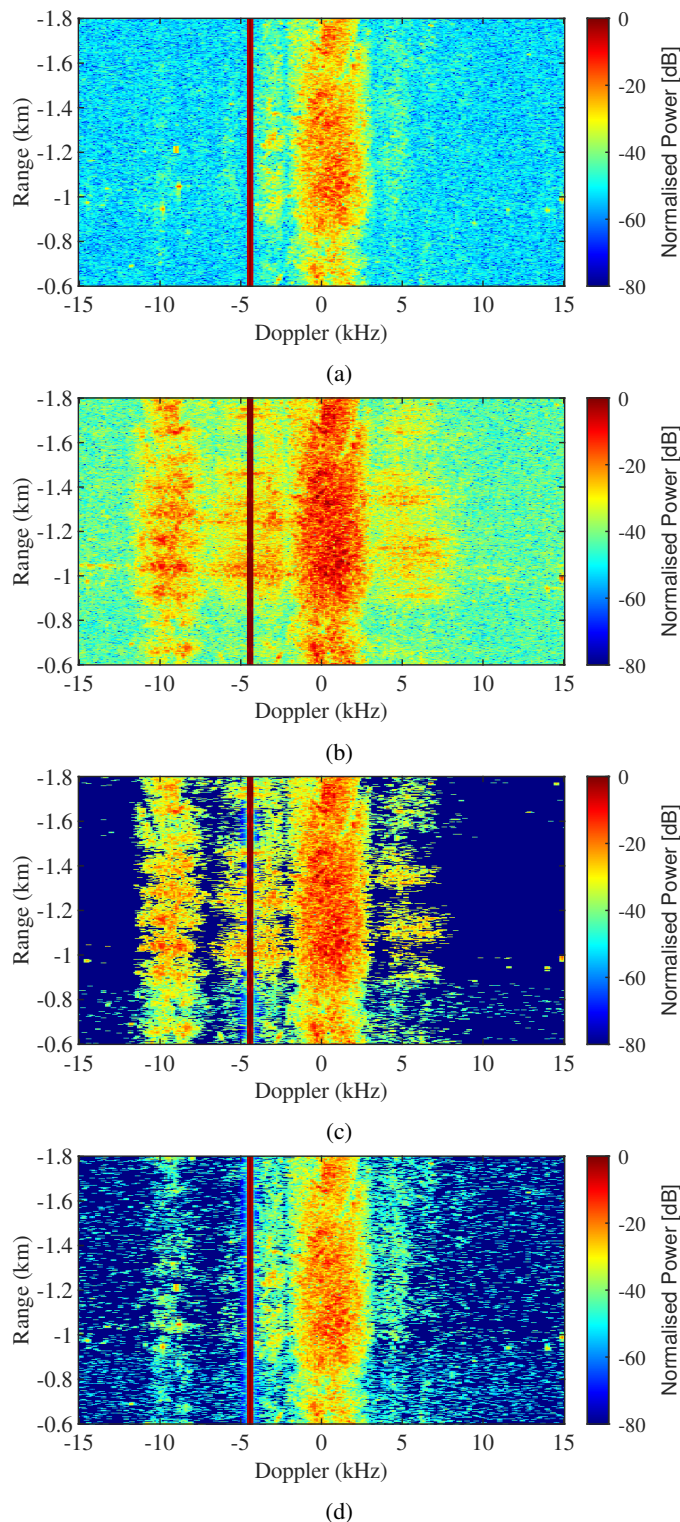
Fig. 7: Trials data RD plots; (a) Linear gain output; (b) Nonlinear BBVS-[8] output; (c) BBVS-[8] output mitigated by memoryless NCS; (d) BBVS-[8] output mitigated by memory rich NCS. Ground clutter located between $-4kHz \rightarrow 5kHz$ Doppler with the synthetically added FMCW interferer located at $-4.5kHz$ Doppler. The mitigation technique successfully recovers the strongest targets located around $-9kHz$ Doppler only when memory terms are introduced to the NCS algorithm.

the NCS algorithm had to be expanded beyond those results published in [22] and [23] so that complex nonlinear memory effects could be mitigated by the NCS technique for the first time. The novel NCS algorithm was tested extensively by means of a MPRF radar simulator that was capable of simulating sophisticated nonlinear receiver effects. The convergence properties of the NCS algorithm were studied for a wide variety of nonlinear transfer functions with the convergence of the algorithm guaranteed in the case where the forward nonlinearity was strictly monotonic. Importantly, more work needs to be done to fully understand the convergence behaviour of the NCS algorithm in the non-monotonic case however a full analysis of this problem is beyond the scope of this paper.

In addition to the convergence analysis, the NCS algorithm was shown to successfully mitigate deleterious effects from both memoryless and memory rich receiver nonlinearities in a MPRF radar mode through in-depth PD analysis. This impressive mitigation performance was shown to hold for real-world MPRF radar data and for significant levels of modelling error built into the forward model of the NCS algorithm. If nonlinear memory effects prove to be significant in the modern radar receiver then forward modelling techniques offer a digital signal processing solution that is not built around the inverse nonlinear transform. Crucially, the unique formalisation of the NCS algorithm provides a more explicit framework to compensate for complex nonlinear memory effects than the standard NLEQ mitigation techniques.

REFERENCES

[1] H. Griffiths, "The concept of the intelligent radar network," in *Novel Radar Techniques and Applications, Volume 2: Waveform Diversity and Cognitive Radar, and Target Tracking and Data Fusion*, R. Klemm, Ed. London, UK: IET, 2017.

[2] P. Liu, Y. Liu, T. Huang, Y. Lu, and X. Wang, "Decentralized automotive radar spectrum allocation to avoid mutual interference using reinforcement learning," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 1, pp. 190–205, 2021.

[3] C. Fischer and W. Gruener, "Quantifying cross modulation effects from pulsed interferers on radar signals," in *2016 European Radar Conference (EuRAD)*, Oct 2016, pp. 1–4.

[4] I. M. Mellor, F. J. Adams, and P. G. Richardson, "Cross modulation cancellation for airborne phased array radar," in *2007 IET International Conference on Radar Systems*, Oct 2007, pp. 1–4.

[5] M. Allén, J. Marttila, M. Valkama, S. Singh, M. Epp, and W. Schlecker, "Digital full-band linearization of wideband direct-conversion receiver for radar and communications applications," in *2015 49th Asilomar Conference on Signals, Systems and Computers*, Nov 2015, pp. 1361–1368.

[6] W. L. Melvin, Ed., *Principles of Modern Radar: Advanced techniques*, ser. Radar, Sonar & Navigation. Institution of Engineering and Technology, 2012.

[7] Q. Zou, M. Mikhemar, and A. H. Sayed, "Digital compensation of cross-modulation distortion in software-defined radios," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 3, pp. 348–361, June 2009.

[8] R. Vansebrouck, C. Jabbour, O. Jamin, and P. Desgreys, "Fully-digital blind compensation of non-linear distortions in wideband receivers," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 8, pp. 2112–2123, Aug 2017.

[9] E. Ward and B. Mulgrew, "Baseband equivalent volterra series for modelling cross-channel nonlinear distortion," in *2019 IEEE Radar Conference (RadarConf)*, April 2019, pp. 1–4.

[10] D. J. Rabideau, "Hybrid mitigation of distortion in digital arrays," in *IEEE International Radar Conference, 2005.*, May 2005, pp. 236–241.

[11] E. Ward and B. Mulgrew, "Mitigation of cross-modulation effects in radar receivers with memory," in *2020 IEEE International Radar Conference (RADAR)*, April 2020, pp. 856–859.

[12] G. F. Earl, "The influence of receiver cross-modulation on attainable hf radar dynamic range," *IEEE Transactions on Instrumentation and Measurement*, vol. IM-36, no. 3, pp. 776–782, Sep. 1987.

[13] B. J. Harker, Z. Dobrosavljevic, E. P. Craney, C. M. Tubb, and G. L. Harris, "Dynamic range improvements and measurements in radar systems," *IET Radar, Sonar Navigation*, vol. 1, no. 6, pp. 398–406, Dec 2007.

[14] N. Peccarelli, Z. Peck, and J. L. Garry, "Analysis and mitigation of receiver induced nonlinearities on pulse-doppler radars," in *2020 IEEE International Radar Conference (RADAR)*, April 2020, pp. 333–338.

[15] E. Ward and B. Mulgrew, "Memory nleq techniques to mitigate cross-modulation effects in radar," in *2021 IEEE Radar Conference (Radar-Conf21)*, 2021, pp. 1–6.

[16] Y. Ma, Y. Yamao, K. Ishibashi, and Y. Akaiwa, "Adaptive compensation of inter-band modulation distortion for tunable concurrent dual-band receivers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 12, pp. 4209–4219, 2013.

[17] M. Valkama, A. Shahed hagh ghadam, L. Anttila, and M. Renfors, "Advanced digital signal processing techniques for compensation of nonlinear distortion in wideband multicarrier radio receivers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 6, pp. 2356–2366, June 2006.

[18] S. A. Billings, *Nonlinear system identification : NARMAX methods in the time, frequency, and spatio-temporal domains*. Chichester, West Sussex, United Kingdom: John Wiley & Sons, Inc., 2013.

[19] M. Schetzen, *The Volterra & Wiener Theories of Nonlinear Systems*. Malabar, Florida: Krieger Publishing Company, 1989.

[20] Z. Dunn, M. Yeary, C. Fulton, and N. Goodman, "Wideband digital predistortion of solid-state radar amplifiers," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 5, pp. 2452–2466, 2016.

[21] T. Blumensath, "Compressed sensing with nonlinear observations and related nonlinear optimization problems," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3466–3474, June 2013.

[22] H. Chen, H. T. Kung, and M. Comiter, "Nonlinear compressive sensing for distorted measurements and application to improving efficiency of power amplifiers," in *2017 IEEE International Conference on Communications (ICC)*, 2017, pp. 1–7.

[23] F. Zhu, A. Liu, and V. K. N. Lau, "Joint interference mitigation and data recovery for massive carrier aggregation via non-linear compressive sensing," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1389–1404, 2018.

[24] S. Benedetto, E. Biglieri, and R. Daffara, "Modeling and performance evaluation of nonlinear satellite links-a volterra series approach," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-15, no. 4, pp. 494–507, July 1979.

[25] M. A. Richards, Ed., *Principles of Modern Radar: Basic principles*, ser. Radar, Sonar & Navigation. Institution of Engineering and Technology, 2010.

[26] C. Alabaster, *Pulse Doppler Radar: Principles, technology, applications*, ser. Radar, Sonar & Navigation. Institution of Engineering and Technology, 2012.

[27] S. Haykin, Ed., *Adaptive Filter Theory*. Prentice Hall, 2002.

[28] K. Kreutz-Delgado, "The complex gradient operator and the cr-calculus," *ArXiv e-print*, no. "2009arXiv0906.4835K", June 2009. [Online]. Available: https://arxiv.org/abs/0906.4835

[29] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proceedings F - Communications, Radar and Signal Processing*, vol. 130, no. 1, pp. 11–16, February 1983.

[30] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265 – 274, 2009.

[31] A. Y. Kibangou and G. Favier, "Wiener-hammerstein systems modeling using diagonal volterra kernels coefficients," *IEEE Signal Processing Letters*, vol. 13, no. 6, pp. 381–384, 2006.

[32] J. Tsimbinos, "Identification and Compensation of Nonlinear Distortion," Ph.D. dissertation, Institute for Telecommunications Research, School of Electronic Engineering, University of South Australia, South Australia, Australia, 1995.

Euan Ward (Student Member, IEEE) received the MPhys degree in Physics (First Class) from the University of St Andrews, St Andrews, Scotland, in 2016. After graduation, he worked for a year as a Systems Engineer with the Radar Modelling Team at Leonardo, Edinburgh, before joining The University of Edinburgh's Institute for Digital Communications to pursue a Ph.D. degree in 2017. During his second year of study (2018), he was awarded a three-year Industrial Fellowship by The Royal Commission for the Exhibition of 1851 to drive his PhD research towards industrial application. His research interests include adaptive signal processing techniques and in particular their application to modern radar systems.

Shahzad Gishkori (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Engineering and Technology Lahore, Lahore, Pakistan, in 2002, and the M.Sc. (cum Laude) and Ph.D. degrees both in electrical engineering from the Delft University of Technology, Delft, The Netherlands, in 2009 and 2014, respectively.
From 2014 to 2015, he was a Postdoctoral Research Associate with Imperial College London, London, U.K. From 2016 to 2020, he was a Postdoctoral Research Associate with the University of Edinburgh, Edinburgh, U.K. Since March 2020, he is working as a Research Fellow at King Abdullah University of Science and Technology, Saudi Arabia. His research interests include signal & image processing and machine learning.

Bernard (Bernie) Mulgrew (Fellow, IEEE) received the B.Sc. degree from Queen's University Belfast, Belfast, U.K, in 1979 and the Ph.D. degree from University of Edinburgh, Edinburgh, U.K., in 1987. After graduation, he worked for four years as a Development Engineer with the Radar Systems Department, Ferranti, Edinburgh. From 1983 to 1986, he was a Research Associate with the Department of Electrical Engineering, University of Edinburgh. He was appointed to lectureship, in 1986, promoted to Senior Lecturer, in 1994 and became a Reader, in 1996. The University of Edinburgh appointed him as a Personal Chair in October 1999 (Professor of Signals and Systems). He has co-author of three books on signal processing. His research interests include adaptive signal processing and estimation theory and in their application to radar and sensor systems. Prof. Mulgrew is a fellow of Royal Academy of Engineering, Royal Society of Edinburgh, and Institution of Engineering and Technology.