A Gradient Descent Implementation of Adaptive Pulse Compression

Patrick M. McCormick[1], Shannon D. Blunt[1], and Thomas Higgins[2]
[1]Radar Systems Lab, University of Kansas, Lawrence, KS
[2]Radar Division, Naval Research Laboratory, Washington, DC

***Abstract*—Gradient descent is an iterative method of determining the minima or maxima of a function. The algorithm can be used to solve a linear system of equations when the computational cost of a matrix inverse is too expensive for an application. Here, gradient descent is applied to Adaptive Pulse Compression (APC), yielding the GraD-APC algorithm. Specifically, a unit-gain constrained version of GraD-APC with optimal step size is derived for use with frequency modulated (FM) waveforms, particularly for cases in which the waveform time-bandwidth product is large enough to prohibit practical use of the original matrix inverse based APC. The range-profile estimation of GraD-APC is compared to that of fully adaptive APC using both simulated and experimentally measured data.**

Keywords—adaptive filtering, gradient descent

I. Introduction

The Adaptive Pulse Compression (APC) algorithm [1,2] was developed to mitigate the sidelobe interference that arises during pulse compression by generating an adaptive filter for each range cell. Since the inception of APC, it has branched in to many variants from joint-domain adaptive filtering [3,4] to applicability for FM waveforms [5,6]. APC employs re-iterative minimum mean-square error (RMMSE) estimation to obtain a range-adaptive filter for each range cell and requires at least one matrix inversion (of a structured covariance matrix) for each filter. The need for a matrix inverse is problematic when considering waveforms with high time-bandwidth product and/or the need for real-time implementation.

An early attempt to avoid full matrix inversion relied on the matrix inversion lemma [2] which performed reliably in low dynamic range scenarios. However, modifications to the algorithm were required to address high dynamic range scenarios due to numerical imprecision when using the matrix inversion lemma leading to error propagation for the APC filter update as a function of range. Dimensionality reduction techniques were also investigated to trade-off adaptive degrees of freedom for lower computational cost [7]. These reduction techniques split the full dimension covariance matrix into multiple lower dimension covariance matrices, thus reducing the overall computation cost. Interestingly, in [5] it was shown that this dimensionality reduction is also useful to compensate for the "oversampling" (relative to waveform 3 dB bandwidth) that is needed to represent FM waveforms with sufficient fidelity for application of APC. Here, the FM-amenable version of APC from [5] is formulated using a gradient descent implementation (yielding what is denoted as GraD-APC) as a means to avoid the matrix inverse altogether with minimal degradation in performance.

Descent methods are in most cases non-optimal (with less than infinite iterations), yet they have the desirable qualities of reduced computational complexity, which is usually $O(M)$ for filter length M, and better numerical stability [8]. Because of the reduction in computational cost, GraD-APC allows for the application of APC and its variants to problems with high dimensionality (e.g. high time-bandwidth product waveforms) and a lower computational threshold to achieve real-time operation.

An analog to APC and GraD-APC is the application of gradient descent to the deterministic minimum mean-square error (MMSE) beamformer or, if the problem is constrained, the linearly-constrained minimum variance (LCMV) beamformer [8]. The difference between these two sets of algorithms is that for APC/GraD-APC the desired signal and the covariance matrix are unknown (albeit structured), and therefore estimation of the covariance matrix and desired signal are determined using an alternating bootstrapping approach based on the initial matched filter (or mismatched filter) response.

With the incorporation of gradient descent, GraD-APC has two distinct iterative elements: an "inner loop" to estimate the filter (via gradient descent) and the original "outer loop" to estimate the range cell complex amplitude (the APC structure). This formulation could be directly extended to also incorporate joint adaptivity in the spatial [3], fast-time Doppler [9], slow-time Doppler [4], and/or polarization [6] domains.

Section II summarizes the signal model and FM-based APC filter derivation from [5]. Section III then introduces the gradient descent formulation and how it is applied to APC to form GraD-APC. In Section IV, the performance of GraD-APC is demonstrated in both a simulated environment and using free space measurements made with an LFM waveform.

II. Adaptive Pulse Compression

The received signal captured at discretized delay ℓ can be modeled as

$$y(\ell) = \mathbf{x}^T(\ell)\mathbf{s} + u(\ell) \tag{1}$$

where $\mathbf{s}$ is a length-M discretized version of the transmitted waveform $s(t)$, $u(\ell)$ is a sample of additive noise present at the receiver, $\mathbf{x}(\ell) = [x(\ell) \quad x(\ell-1) \quad \cdots \quad x(\ell-M+1)]^{\mathrm{T}}$ are M contiguous complex samples of the range profile ground truth (not known), and $(\bullet)^T$ is the transpose operation. The range profile vector $\mathbf{x}(\ell)$ contains the complex scaling that includes

This work was supported by the Radar Division of the Naval Research Laboratory.

transmitted power, antenna gain, spherical spreading losses, and radar cross section of the illuminated scatterers.

The signal-to-noise ratio (SNR) maximizing matched filter is denoted as

$$\hat{x}_{\mathrm{MF}}(\ell) = \mathbf{w}_{\mathrm{MF}}^{H}\mathbf{y}(\ell)\,, \tag{2}$$

where $\mathbf{y}(\ell) = [\,y(\ell) \quad y(\ell+1) \quad \cdots \quad y(\ell+M-1)]^{\mathrm{T}}$ is the collection of M contiguous discrete received samples that all contain the scatterer response $x(\ell)$ and $(\bullet)^{H}$ is the Hermitian operation. Here it is assumed that the matched filter, which is delay independent, is normalized as

$$\mathbf{w}_{\mathrm{MF}} = \frac{\mathbf{s}}{\mathbf{s}^{H}\mathbf{s}}\,. \tag{3}$$

Now define the delay dependent filtering as

$$\hat{x}(\ell) = \mathbf{w}^{H}(\ell)\mathbf{y}(\ell)\,, \tag{4}$$

where the delay dependent filter is found by minimizing the generalized APC cost function [5]

$$\begin{aligned} J(\ell;\mathbf{w}(\ell)) \;=\; & \sum_{k=0}^{K-1} E\left[\left|\frac{1}{K}x(\ell) - \mathbf{w}_k^{\mathrm{H}}(\ell)\mathbf{y}_k(\ell)\right|^2\right] \\ & +\mathrm{Re}\left\{\lambda\left(\sum_{k=0}^{K-1}\left(\mathbf{w}_k^{\mathrm{H}}(\ell)\mathbf{s}_k\right)-1\right)\right\} \end{aligned} \tag{5}$$

for use with FM waveforms, where K is the degree of "oversampling" on receive with respect to the 3-dB bandwidth of the transmitted waveform, which is needed to represent the FM waveform with sufficient fidelity. Thus $\mathbf{w}_k(\ell)$, $\mathbf{y}_k(\ell)$ and $\mathbf{s}_k$ are the kth length M/K polyphase-decomposed versions of the APC filter, receive vector, and transmitted waveform, respectively. The decomposition is necessary when applying the APC formulation for FM waveforms to avoid a noise enhancement effect that otherwise occurs in the inversion of the full $M \times M$ covariance matrix. Here it is assumed the data is resampled such that K is an integer. The bottom term in (5) is a unit gain constraint on the overall (non-decomposed) APC filter where λ is the Lagrange multiplier. Minimizing (5) with respect to $\mathbf{w}_k^{*}(\ell)$ yields the APC filter [5]

$$\mathbf{w}_k(\ell) = \frac{\left(\mathbf{C}_k(\ell)+\mathbf{R}_k\right)^{-1}\mathbf{s}_k}{\sum_{i=0}^{K-1}\mathbf{s}_i^{\mathrm{H}}\left(\mathbf{C}_i(\ell)+\mathbf{R}_i\right)^{-1}\mathbf{s}_i} \tag{6}$$

where $\mathbf{R}_k$ is the kth decomposed $M/K \times M/K$ noise covariance matrix, and $\mathbf{C}_k(\ell)$ is the kth decomposed $M/K \times M/K$ structured covariance matrix.

Fig. 1 shows a flowchart of the APC process. The blue box represents the range profile initialization using the matched filter, the green boxes represent polyphase decomposition or recombining and the red boxes represent steps in the APC procedure. The APC loop is enclosed in the dashed box and flows according to the bold arrows. The box labeled $\mathbf{w}_k(\ell)$ is found using (6). The flowchart in Fig. 1 is for a single range cell and it is assumed that estimation updates for all the range cells of interest are performed before the next iteration of filter updates is performed. It typically takes 2-5 iterations for this instantiation of APC to converge depending on the density of significant scatterers in the illuminated environment.

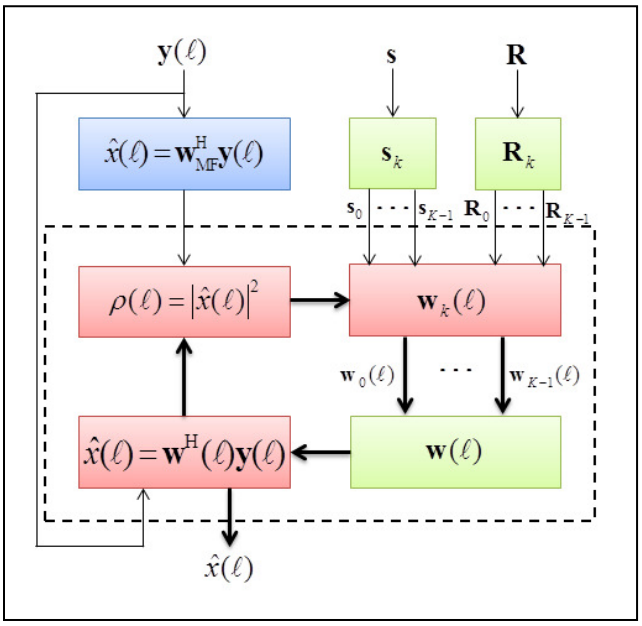

Fig. 1. Flowchart of APC operation

III. Gradient Descent - Adaptive Pulse Compression

Given the gradient of cost function J as $\nabla_{\mathbf{w}^*}J$, the gradient descent (GD) structure is [10]

$$\mathbf{w}_n = \mathbf{w}_{n-1} - \mu_n\left(\nabla_{\mathbf{w}^*}J(\mathbf{w})\right), \tag{7}$$

where n is the GD iteration index and μ_n is a non-negative step size that may be iteration dependent.

The APC cost function for use with FM waveforms in (5) employed a polyphase decomposition to avoid noise enhancement effects that arose during matrix inversion (due to the need to "over-sample" the FM waveform for sufficient fidelity). Because the GD structure in (7) avoids the need for matrix inversion, this polyphase decomposition is also no longer required. As such, the gain-constrained (and non-decomposed) APC cost function can be expressed as

$$J(\ell;\mathbf{w}(\ell)) = E\left[\left|x(\ell) - \mathbf{w}^{\mathrm{H}}(\ell)\mathbf{y}(\ell)\right|^2\right] + \mathrm{Re}\left\{\lambda\left(\mathbf{w}^{\mathrm{H}}(\ell)\mathbf{s}-1\right)\right\}. \tag{8}$$

Taking the gradient of (8) and assuming each range cell is uncorrelated with neighboring range cells and the noise yields

$$\nabla_{\mathbf{w}^*}J(\ell;\mathbf{w}(\ell)) = \left(\mathbf{C}(\ell)+\mathbf{R}\right)\mathbf{w}(\ell) - x^{*}(\ell)\mathbf{s} + \lambda\mathbf{s}\,, \tag{9}$$

where $\mathbf{R}$ is the $M \times M$ full-dimension noise covariance (for uncorrelated noise $\mathbf{R} = \sigma_u^2\mathbf{I}_M$, where σ_u^2 is the noise power and $\mathbf{I}_M$ is an $M \times M$ identity matrix). The $M \times M$ full-dimension structured covariance matrix $\mathbf{C}(\ell)$ is defined as

$$\mathbf{C}(\ell) = \sum_{\tau=-M+1}^{M-1} \rho(\ell+\tau)\, \mathbf{s}_\tau \mathbf{s}_\tau^H , \tag{10}$$

where $\rho(\ell) = \left|\hat{x}(\ell)\right|^2$ is the power of the current estimate of range cell $x(\ell)$ and

$$\mathbf{s}_\tau = \begin{cases} [s_{|\tau|} \quad s_{|\tau|+1} \quad \cdots \quad s_{M-1} \quad \mathbf{0}_{1\times|\tau|}]^T & \text{for} \quad \tau \le 0 \\ [\mathbf{0}_{1\times\tau} \quad s_0 \quad s_1 \quad \cdots \quad s_{M-1-\tau}]^T & \text{for} \quad \tau > 0 \end{cases} \tag{11}$$

are delay-shifted versions of the discretized waveform.

As in [5], the $K-1$ range cells on either side of the current range index ℓ are zeroed in (11) as

$$\rho(\ell \pm k) = 0 \tag{12}$$

for $k = 0,1,\ldots,K-1$. This zeroing has the effect of widening the beam in range to that of the matched filter resolution, which serves to minimize straddling loss [5] and focuses the adaptive degrees of freedom on sidelobe suppression instead of narrowing the mainbeam for super-resolution, which yields significant SNR loss and greatly increases convergence time.

Inserting (9) into the GD structure of (7) produces

$$\mathbf{w}_n(\ell) = \mathbf{w}_{n-1}(\ell) - \mu_n(\ell)\left(\left(\mathbf{C}(\ell)+\mathbf{R}\right)\mathbf{w}(\ell) - x^*(\ell)\mathbf{s} + \lambda\mathbf{s}\right). \tag{13}$$

The Lagrange multiplier can be determined by solving for λ in

$$\begin{aligned} &\mathbf{w}_n^H(\ell)\mathbf{s} = 1 \\ &= \left(\mathbf{w}_{n-1}(\ell) - \mu_n(\ell)\left(\left(\mathbf{C}(\ell)+\mathbf{R}\right)\mathbf{w}(\ell) - x^*(\ell)\mathbf{s} + \lambda\mathbf{s}\right)\right)^H \mathbf{s} \end{aligned}, \tag{14}$$

resulting in

$$\begin{aligned} \lambda = &-\mu_n^{-1}(\ell)\left(\mathbf{s}^H\mathbf{s}\right)^{-1} + \mu_n^{-1}(\ell)\left(\mathbf{s}^H\mathbf{s}\right)^{-1}\mathbf{s}^H\mathbf{w}_n(\ell) \\ &-\left(\mathbf{s}^H\mathbf{s}\right)^{-1}\mathbf{s}^H\left(\mathbf{C}(\ell)+\mathbf{R}\right)\mathbf{w}_n(\ell) + x^*(\ell) \end{aligned}. \tag{15}$$

Subsequently inserting (15) into (13) and simplifying yields

$$\mathbf{w}_n(\ell) = \mathbf{P}_\mathbf{s}^\perp\left(\mathbf{I}_M - \mu_n(\ell)\left(\mathbf{C}(\ell)+\mathbf{R}\right)\right)\mathbf{w}_{n-1}(\ell) + \mathbf{w}_\mathrm{q}, \tag{16}$$

where $\mathbf{P}_\mathbf{s}^\perp$ is the orthogonal projection matrix

$$\mathbf{P}_\mathbf{s}^\perp = \mathbf{I}_M - \mathbf{s}\left(\mathbf{s}^H\mathbf{s}\right)^{-1}\mathbf{s}^H \tag{17}$$

and $\mathbf{w}_\mathrm{q}$ is the quiescent filter

$$\mathbf{w}_\mathrm{q} = \mathbf{s}\left(\mathbf{s}^H\mathbf{s}\right)^{-1} \tag{18}$$

that is identical to the normalized matched filter of (3).

The optimal step size $\mu_{n,\mathrm{opt}}(\ell)$ at the nth iteration can be determined by solving

$$\min_{\mu_n(\ell)} J\left(\ell; \mathbf{w}_n(\ell)\right), \tag{19}$$

where $\mathbf{w}_n(\ell)$ is defined in (16). Plugging (16) into (8) and minimizing with respect to $\mu_n(\ell)$ yields the optimal step size

$$\mu_{n,\mathrm{opt}}(\ell) = \frac{\mathrm{Re}\left[\mathbf{g}_n^H(\ell)\left(\mathbf{C}(\ell)+\mathbf{R}\right)\left(\mathbf{P}_\mathbf{s}^\perp\mathbf{w}_{n-1}(\ell) + \mathbf{w}_\mathrm{q}\right)\right]}{\mathbf{g}_n^H(\ell)\left(\mathbf{C}(\ell)+\mathbf{R}\right)\mathbf{g}_n(\ell)}, \tag{20}$$

where

$$\mathbf{g}_n(\ell) = \mathbf{P}_\mathbf{s}^\perp\left(\mathbf{C}(\ell)+\mathbf{R}\right)\mathbf{w}_{n-1}(\ell). \tag{21}$$

While calculation of the optimal step size at each iteration may be computationally prohibitive, the formulation in (20) and (21) may provide a good guide to establish practical rules for selecting the step-size in practice (such as with LMS [10]).

We have now defined the two iterative processes for GraD-APC: the APC-based "outer loop" to converge to an estimate of $x(\ell)$; and the GD-based "inner loop" to converge to an estimate of the APC filter for a given range cell. Table I provides the details of the GraD-APC algorithm. Denote N as the total number of inner loop (GD) iterations. The number of outer loop (APC) iterations depends on the desired sidelobe suppression. A convergence criterion for the inner loop could also be implemented in lieu of a fixed number of iterations.

TABLE I: IMPLEMENTATION OF GRAD-APC ALGORITHM

1.	Collect M range samples corresponding to range index ℓ into the vector $\mathbf{y}(\ell)$.
2.	Obtain the initial range profile estimate $\hat{x}(\ell) = \hat{x}_{\mathrm{MF}}(\ell)$ via (2) and (3) and initialize the APC filter to the quiescent filter via (18) $\mathbf{w}_0(\ell) = \mathbf{w}_\mathrm{q}$.
3.	Compute the power estimates $\rho(\ell) = \lvert\hat{x}(\ell)\rvert^2$ and use to calculate the structured covariance matrice $\mathbf{C}(\ell)$ from (10) while implementing the zero-filling constraint for $\rho(\ell)$ from (12).
4.	Initialize inner loop iteration index to $n = 1$.
5.	Calculate the optimal step size $\mu_{n,\mathrm{opt}}(\ell)$ via (20) and (21).
6.	Calculate $\mathbf{w}_n(\ell)$ via (16) – (18).
7.	If $n < N$, go to step 5 and increment inner loop index $n = n+1$. Otherwise, go to step 8.
8.	Apply each GraD-APC filter to the associated data vector to obtain the updated range profile estimate as $\hat{x}(\ell) = \mathbf{w}_N^{\mathrm{H}}(\ell)\mathbf{y}(\ell)$.
9.	Initialize filters for next outer loop iteration as $\mathbf{w}_0(\ell) = \mathbf{w}_N(\ell)$ using current final filter estimates.
10.	Go to step 3 unless convergence or desired suppression is achieved.

The outer loop comprises steps 3 thru 8 in Table I while the inner loop is steps 5 thru 7. Note the initialization of the inner loop filter is dependent on the outer loop iteration. For the first iteration of the outer loop, the initial filter of the inner loop $\mathbf{w}_0(\ell)$ is set to the quiescent or matched filter. For the remaining outer loop iterations, $\mathbf{w}_0(\ell)$ is set to the final filter of the previous inner loop (step 9). Per outer loop iteration per range cell, APC from (6) requires $O(M^3/K^3)$ complex operations, while GraD-APC requires $O(M^2N)$ and is readily amenable for further reduction via parallel processing.

The nested loop structure of GraD-APC presents an interesting trade-off between filter convergence and

convergence of the range-profile estimate (or covariance matrix). It is advantageous in terms of time to not fully converge to a filter solution before updating the range profile estimate, but to find a "good enough" solution. This compromise permits a more rapid update of the gradient descent direction. These choices give flexibility to the desired amount of adaptivity/suppression. A comprehensive convergence analysis on the nested loop structure of GraD-APC is needed to understand fully the "optimal" choice in parameters.

IV. Simulation and Experimental Results

To judge the effectiveness of GraD-APC in estimating the range profile, it will be compared to the range profiles generated by the matched filter and by APC from (6). First, a high dynamic range profile containing two target responses is simulated using a linear frequency modulated waveform with time-bandwidth product of 100. Second, measured data captured in an open air environment using an LFM of time-bandwidth product of 64 is examined.

A. Simulation Results

Consider a range profile containing two point scatterers located at range indices 150 and 160. The magnitudes of these targets are 70 dB and 20 dB, respectively, relative to the noise power that is normalized to 0 dB. The phases of the two scatterers are randomly chosen from a uniform distribution on $U[0,2\pi]$. The illuminating waveform is an LFM with time-bandwidth product of 100. The sampling frequency is chosen to be K=3 times that of the 3-dB bandwidth of the waveform. Therefore $K-1=2$ range cells are zeroed on either side of the "current" range index ℓ for both GraD-APC and APC.

To start, the number of outer loop iterations is chosen to be 5 for GraD-APC and 2 for APC as it takes more iterations for GraD-APC to converge. GraD-APC is tested with two different amounts of inner loop iterations: $N=5$ and $N=10$.

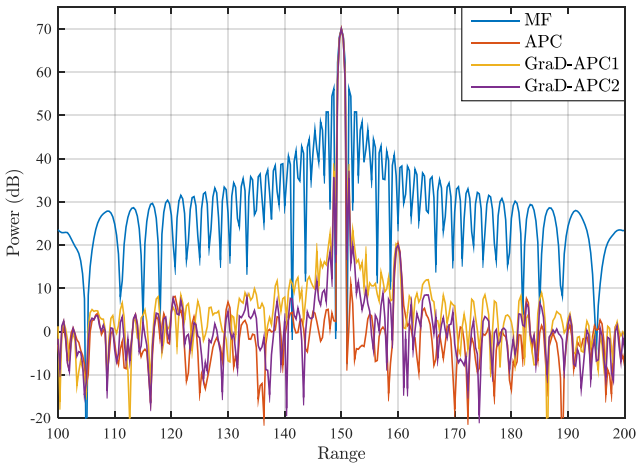

Fig. 2. Two-target range profile estimation using matched filter (blue), APC (red), GraD-APC (N=5) (yellow) and GraD-APC (N=10) (purple)

Figure 2 shows the simulation results for the two-target scenario. The yellow trace denoting GraD-APC1 is the scenario with $N=5$ inner loop iterations and the purple trace denoting GraD-APC2 has $N=10$ inner loop iterations. All APC variants exposed the 20 dB target after processing. The best overall sidelobe suppression is provided by the APC implementation of (7). The additional 5 inner loop iterations in GraD-APC2 reduce the range sidelobes an additional 5 dB as compared to GraD-APC1. Compared to APC, both GraD-APC variants have higher close-in sidelobes near the mainlobe of the larger target. These sidelobes are due to the GraD-APC solutions not fully converging during the given outer and inner loop iterations. Although the sidelobes for GraD-APC are not suppressed as much as with APC, the GraD-APC suppression is still close to 30 dB better than the matched filter response.

B. Open-Air Experimental Results

The dataset used here is the same used in [5] that was captured in part to test the modifications to APC to make it amenable to FM waveforms. Figure 3 shows the field of view for the experiment obtained from Google Maps. The main scatterers in the scene are indicated by the orange triangles and the radar location is marked by the orange circle. Figure 4 shows the quasi-monostatic setup of two quad-ridge antennas for simultaneous transmit and receive. An LFM with an approximate time-bandwidth of product of 64 occupies 80 MHz of bandwidth. The center frequency was 2.3 GHz and the transmit power was approximately 24 dBm. The baseband data were resampled to a sampling frequency $K=3$ times the 3-dB bandwidth of the LFM waveform.

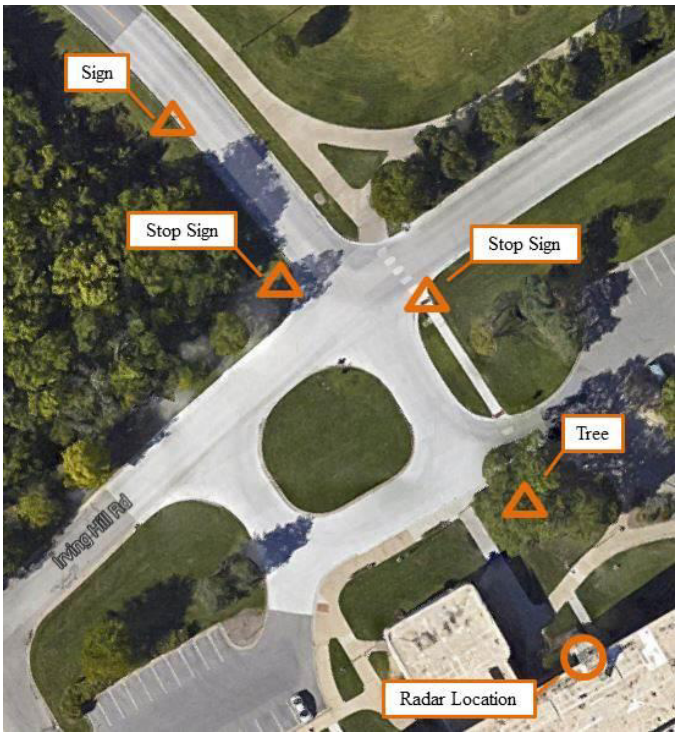

Fig. 3. Field of view for measured results

Figure 5 shows the range profile estimates using a matched filter (blue), APC (red), GraD-APC for $N=5$ inner loop iterations (yellow) and GraD-APC for $N=10$ inner loop iterations (purple). The number of outer loop iterations was the same as for the simulated results with 2 for APC and 5 for GraD-APC.

Fig. 4. Test setup for experimental measurements

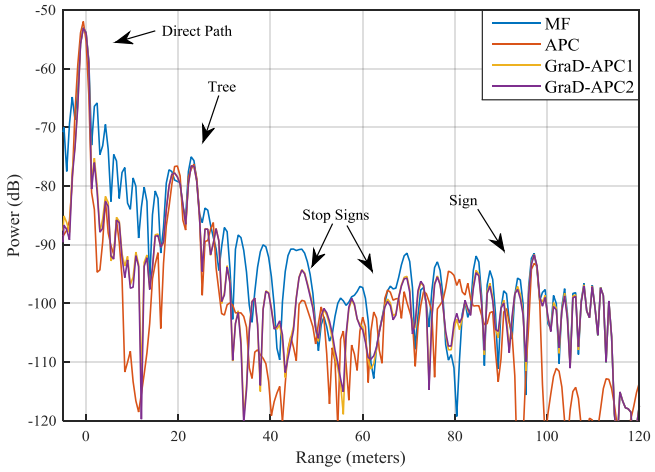

Fig. 5. Experimental range profile estimation using matched filter (blue), APC (red), GraD-APC (N=5) (yellow) and GraD-APC (N=10) (purple)

The difference between the two GraD-APC range-profiles is negligible, with the maximum difference between them being ~1 dB. APC again outperforms the two GraD-APC implementations, which is to be expected considering the simulation results. However all three of these APC-based schemes outperform the matched filter, with sidelobe suppression improvement ranging from 15 to 20 dB for APC and 10 to 15 dB for GraD-APC.

V. Conclusions

A new method for implementing adaptive pulse compression (APC) using gradient descent has been developed (denoted GraD-APC). The resulting absence of a matrix inverse opens the door for application of GraD-APC to higher dimensionality problems and may enable real-time operation. Further investigation is needed as to the "optimal" number of iterations, for both the inner and outer loops. Simulation and experimental results show that GraD-APC is a viable practical alternative to APC for suppression of range sidelobes. Ongoing work is exploring methods to reduce the eigenspread of the structured covariance matrix to accelerate convergence.

References

[1] S. D. Blunt and K. Gerlach, "A novel pulse compression scheme based on minimum mean-square error reiteration," *IEEE Intl. Radar Conf.*, Adelaide, Australia, Sept. 2003.

[2] S.D. Blunt and K. Gerlach, "Adaptive pulse compression via MMSE estimation," *IEEE Trans. Aerospace & Electronic Systems*, vol. 42, no. 2, pp. 572-584, Apr. 2006.

[3] P. McCormick, T. Higgins, S.D. Blunt, and M. Rangaswamy, "Adaptive receive processing of spatially modulated physical radar emissions," *IEEE Journal of Special Topics in Signal Processing*, vol. 9, no. 8, pp. 1415-1426, Dec. 2015.

[4] T. Higgins, S. Blunt, and A. Shackelford, "Time-range adaptive processing for pulse agile radar," *Intl. Waveform Diversity and Design Conf.*, Niagara Falls, Canada, Aug. 2010.

[5] D. Henke, P. McCormick, S.D. Blunt, and T. Higgins, "Practical aspects of optimal mismatch filtering and adaptive pulse compression for FM waveforms," *IEEE Intl. Radar Conf.*, Washington, DC, May 2015.

[6] P. McCormick, J. Jakabosky, S.D. Blunt, C. Allen, and B. Himed, "Joint polarization/waveform design and adaptive receive processing", *IEEE Intl. Radar Conf.*, Washington, DC, May 2015.

[7] S.D. Blunt, T. Higgins, and K. Gerlach, "Dimensionality reduction techniques for efficient adaptive pulse compression," *IEEE Trans. Aerospace and Electronic Systems*, vol. 46, no. 1, pp. 349-362, Jan. 2010.

[8] Van Trees, H.L., *Optimum Array Processing*, John Wiley & Sons, 2002, Chap. 2.

[9] S.D. Blunt, A. Shackelford, K. Gerlach, and K.J. Smith, "Doppler compensation & single pulse imaging via adaptive pulse compression," *IEEE Trans. Aerospace & Electronic Systems*, vol. 45, no. 2, pp. 647-659, Apr. 2009.

[10] S. Haykin, *Adaptive Filter Theory*, 5th *ed.*, Prentice Hall, 2013.