MULTI-STAGE REDUCED-RANK ADAPTIVE FILTER WITH FLEXIBLE STRUCTURE

Sergueï Burykh and Karim Abed-Meraim

ENST PARIS, Signal and Image Processing Dept.,
46, rue Barrault,
75634 Paris Cedex 13, FRANCE
e-mail: {burykh,abed}@tsi.enst.fr

ABSTRACT

An adaptive reduced-rank filter based on Conjugate-Gradient Algorithm is presented - Conjugate-Gradient Reduced-Rank Filter (CGRRF). The proposed method computes reduced-rank Wiener solutions iteratively starting from rank 1 (matched filter). It is shown that any of these filters is equivalent to the Multi-Stage Wiener Filter proposed by Goldstein and Reed of a particular rank. Contrary to other techniques of comparable performance, the rank of CGRRF (number of stages) can be easily adapted to achieve required performance/complexity trade-off. This property is illustrated numerically using adaptive rank selection technique along with other examples showing good transient and asymptotic performance of the method.

1 INTRODUCTION

Adaptive filtering is widely used in signal processing applications such as equalization, array signal processing, multi-user detection, to name a few. The frequent problem which arises when designing an adaptive filtering system is that large observation size, and, therefore, large filter length, means inevitably high computational cost, slow convergence and poor tracking performance. However, this situation corresponds to many important practical cases such as Direct-Sequence Code-Division-Multiple-Access (DS-CDMA) systems which use high spreading factors, radar or Global Positioning System (GPS) array processing.

Reduced-rank adaptive filters provide a way out of this dilemma. The basic idea behind the rank reduction is to project the observation onto a lower-dimensional subspace. The adaptation is then performed within this subspace with a low-order filter resulting in computational savings and improving convergence and tracking characteristics. Various reduced-rank filters were proposed [4]. Among the most promising ones are Multi-Stage Wiener Filter (MSWF) [2] and Auxiliary-Vector Filters (AVF) [5, 6]. It can be shown that MSWF and AVF use the same projection subspace - Krylov subspace. This fact was further studied in [1] where a number of reduced-rank techniques were related to Krylov subspace methods for linear systems [3].

This work is developed in the direction opposite to [1]. We start from a Krylov subspace method known as Conjugate-Gradient Algorithm (CGA) [3] in order to design the corresponding reduced-rank filter (CGRRF). Being mathematically equivalent to MSWF, CGRRF offers additional structural flexibility in the sense that its rank (D) can be easily adjusted to attain required performance/complexity trade-off. This flexibility is of major practical importance in the context of limited and/or distributed processing power, for example, mobile-side multi-user detection for CDMA systems or detection and tracking of multiple targets in radars.

2 DATA MODEL

Throughout the paper, the notations *, ^{T}, and ^{H} are used to denote the conjugate, transpose, and conjugate transpose operations, respectively.

Let $\mathbf{r}(k) = [r_1(k)\ r_2(k)\ \ldots\ r_N(k)]^T$ be the $N \times 1$ vector consisting of N data samples observed at time instant k, which is modeled as

$$\mathbf{r}(k) = \mathbf{H}\mathbf{s}(k) + \mathbf{n}(k), \tag{1}$$

where $\mathbf{s}(k)$ denotes $M \times 1$ vector of source signals $s_1(k), s_2(k) \ldots s_M(k)$, $\mathbf{H}$ is a $N \times M$ channel matrix and $\mathbf{n}(k)$ stands for a $N \times 1$ noise vector. In the sequel, $\mathbf{s}(k)$ and $\mathbf{n}(k)$ are supposed to be zero-mean and wide-sense stationary with respective covariance matrices $E[\mathbf{s}(k)\mathbf{s}^H(k)] = \mathrm{diag}(\epsilon_1, \epsilon_2, \ldots\ \epsilon_M)$ and $E[\mathbf{n}(k)\mathbf{n}^H(k)] = \mathbf{R}_n$.

The model (1) can be used, for example, to represent M narrowband sources impinging on a N-element antenna array, or in a context of a synchronous DS-CDMA system. In the latter case, $\mathbf{s}(k)$ is the vector of signals transmitted by M system users and the ith column of channel matrix $\mathbf{H}$ models channel signature of user i, i.e., ith spreading code convolved with ith channel impulse response.

3 FILTER RANK REDUCTION

Let us consider the problem of estimation of source signal $s_1(k)$ given the observation (1). General linear estimator can be written as

$$\hat{s}_1(k) = \mathbf{w}^H \mathbf{r}(k), \tag{2}$$

where $\mathbf{w}$ is a $N \times 1$ vector (filter). The well-known *full-rank* Wiener filter is the solution of the following linear system (normal equations):

$$\mathbf{R}\mathbf{w}_{opt}^{fr} = \mathbf{c}_1, \tag{3}$$

where $\mathbf{c}_1 = E[\mathbf{r}(k)s_1^*(k)]$ is the desired signal-data cross-correlation vector and $\mathbf{R} = E[\mathbf{r}(k)\mathbf{r}^H(k)]$ is the covariance matrix of $\mathbf{r}(k)$. The important property of the Wiener filter is that it is the only filter that minimizes the mean-squared estimation error (MSE), or, in other words, average error energy. In our notations, MSE can be written as

$$J(\mathbf{w}) = E[\|\hat{s}_1(k) - s_1(k)\|^2] = \epsilon_1 + \mathbf{w}^H\mathbf{R}\mathbf{w} - \mathbf{w}^H\mathbf{c}_1 - \mathbf{c}_1^H\mathbf{w}. \tag{4}$$

Let $\{\mathcal{S}^i\}$, $i = 1, 2, \ldots$ be a sequence of subspaces in $\mathcal{C}^N$. The *reduced-rank Wiener filter in subspace* $\mathcal{S}^i$ is defined as

$$\mathbf{w}_{opt}^i \stackrel{\text{def}}{=} \arg\min_{\mathbf{w}\in\mathcal{S}^i} J(\mathbf{w}). \tag{5}$$

The above definition includes full-rank Wiener filter as a particular case when $\mathcal{S}^{\mathbf{i}} = \mathcal{C}^N$. Let $D \stackrel{\text{def}}{=} \dim(\mathcal{S}^i)$ and let $\{\mathbf{q}_j\}$, $j = 1, \ldots, D$ be an orthonormal basis of $\mathcal{S}^i$. It can be shown that reduced-rank Wiener filter is the solution of the following linear system:

$$\left(\mathbf{Q}^H\mathbf{R}\mathbf{Q}\right)\mathbf{w}_{opt}^i = \mathbf{Q}^H\mathbf{c}_1, \tag{6}$$

where $\mathbf{Q} \stackrel{\text{def}}{=} [\mathbf{q}_1\, \mathbf{q}_2\, \ldots\, \mathbf{q}_D]$. Note that (6) is a system of $D \le N$ equations. Therefore, choosing $D \ll N$ may lead to substantial gain in complexity compared with conventional system (3). This complexity gain is one of the benefits of reduced-rank techniques. Another advantage of rank reduction is faster convergence and better tracking properties [4]. On the other hand, reduced-rank Wiener filter performs only local MSE optimization, hence $J(\mathbf{w}_{opt}^i) \ge J(\mathbf{w}_{opt}^{fr})$.

4 CONJUGATE-GRADIENT ALGORITHM

Here, we summarize briefly the basic ideas behind the family of conjugate-gradient methods. More detailed presentation can be found in standard textbooks on computational linear algebra [3].

Consider the following general iterative algorithm:

$$\mathbf{w}^0 = \mathbf{0} \tag{7}$$

$$\mathbf{w}^i = \mathbf{w}^{i-1} + c_i\mathbf{u}_i, \quad i = 1, 2, \ldots, D \tag{8}$$

with the sequences of complex coefficients c_i and of unit-norm vectors $\mathbf{u}_i$ chosen according to some optimization criterion.

The criterion considered here is $J(\mathbf{w}^i)$ (see (4)), so it is natural to require that $J(\mathbf{w}^i) \le J(\mathbf{w}^{i-1})$. Note also from (8) that $\mathbf{w}^i$ is always in $\mathcal{U}^i = span\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_i\}$. The question is: whether it is possible to choose c_i and $\mathbf{u}_i$ to give the reduced-rank Wiener filter in $\mathcal{U}^i$? In other words, we require that

$$\mathbf{w}^i = \mathbf{w}_{opt}^i = \arg\min_{\mathbf{w}\in\mathcal{U}^i} J(\mathbf{w}). \tag{9}$$

The following lemma helps to answer this question.

Lemma 1 *For the requirement (9) to be satisfied, it is sufficient that*

1. $\mathbf{u}_i$ *are mutually* $\mathbf{R}$-conjugate, *that is,*

$$\mathbf{u}_i^H\mathbf{R}\mathbf{u}_j = 0, \quad i \ne j \tag{10}$$

2. c_i *is given by*

$$c_i = \mathbf{u}_i^H\mathbf{e}_{i-1}/\mathbf{u}_i^H\mathbf{R}\mathbf{u}_i, \tag{11}$$

where $\mathbf{e}_i \stackrel{\text{def}}{=} \mathbf{c}_1 - \mathbf{R}\mathbf{w}^i$.

Proof 1 *See [3].*

Different versions of conjugate-gradient algorithm result from different ways to compute the sequence of $\mathbf{R}$-conjugate vectors $\mathbf{u}_i$ [3]. The version of our choice, shown in Table I, requires only one matrix-by-vector multiplication per iteration. D iterations of the algorithm result in a sequence $\{\mathbf{w}_{opt}^i\}$ of D reduced-rank Wiener filters.

Initialization:
$\mathbf{w}_{opt}^0 = \mathbf{0}$
$\beta_1 = 0$
$\mathbf{u}_1 = \mathbf{e}_0 = \mathbf{c}_1$
For $i = 1, 2, \ldots, D$
  if $i > 1$
    $\beta_i = \|\mathbf{e}_{i-1}\|^2/\|\mathbf{e}_{i-2}\|^2$
    $\mathbf{u}_i = \mathbf{e}_{i-1} + \beta_i\mathbf{u}_{i-1}$
  End
  $\mathbf{z}_i = \mathbf{R}\mathbf{u}_i$
  $c_i = \|\mathbf{e}_{i-1}\|^2/\mathbf{u}_i^H\mathbf{z}_i$
  $\mathbf{e}_i = \mathbf{e}_{i-1} - c_i\mathbf{z}_i$
  $\mathbf{w}_{opt}^i = \mathbf{w}_{opt}^{i-1} + c_i\mathbf{u}_i$

Table I: Summary of CGRRF

5 PROPOSED REDUCED-RANK FILTER

CGRRF has a multi-stage structure (see Fig. 1). Stage i of CGRRF computes the reduced-rank Wiener filter in $\mathcal{U}^i$. The received signal is then filtered giving the estimate $\hat{s}_1^i(k)$. The complexity order of rank D CGRRF is $O(DN^2)$ flops per filter update (compared to $O(N^3)$ for the full-rank filter using direct matrix inversion)[1].

[1] $O(ND)$ approximate adaptive implementations of CGRRF can be developed using the approach explained in [1].

CGRRF can be linked to other reduced-rank methods by means of Krylov subspaces. More precisely, define *Krylov subspace*

$$\mathcal{K}^i(\mathbf{R}, \mathbf{c}_1) \stackrel{\text{def}}{=} \text{span}\{\mathbf{c}_1, \mathbf{R}\mathbf{c}_1, \ldots, \mathbf{R}^{i-1}\mathbf{c}_1\}. \quad (12)$$

It is an established fact [1] that MSWF [2] minimizes $J(\mathbf{w})$ in Krylov subspace $\mathcal{K}^D(\mathbf{R}, \mathbf{c}_1)$. The proposed filter also has this property, as stated by the following lemma.

Lemma 2 *For all* $1 \leq i \leq D$, $\mathcal{U}^i = \mathcal{K}^i(\mathbf{R}, \mathbf{c}_1)$.

Proof 2 *See [3].*

This means that MSWF and CGRRF of the same rank are mathematically equivalent, as they minimize $J(\mathbf{w})$ over the same subspace. However, CGRRF has the advantage of computing Wiener filters of *all* ranks ranging from 1 to D, and filter outputs (symbol estimates) of different ranks $\hat{s}_1^i(k)$, $i = 1 \ldots D$ are simultaneously available. This property makes possible, for example, real-time filter rank selection by measuring the SINR at the output of each stage and adapting filter rank D to achieve a given target SINR. Moreover, CGRRF of any rank $i < D$ is always at hand while for MSWF the system (6) has to be re-solved for each value of i[2].

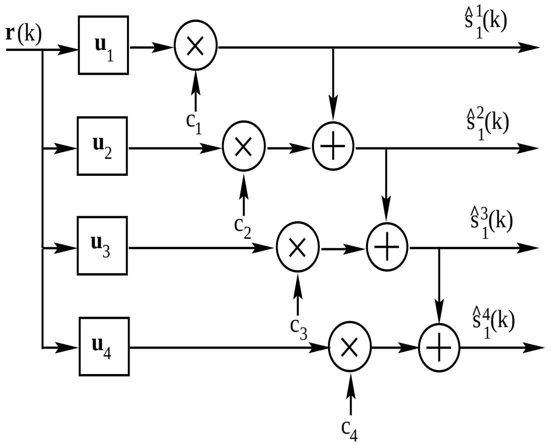

Figure 1: Conjugate-Gradient Reduced-Rank Filter (rank $D = 4$).

6 COMPUTER SIMULATIONS

In the first experiment, low sample size performance of the proposed method is studied. A synchronous single-path DS-CDMA system with $K = 4$ active users is considered. Each user signal is spread with a Gold sequence of length $N = 7$. The receiver employs uniform linear antenna array of 3 elements (half-wavelength element spacing was used) so that the received signal $\mathbf{r}(k) \stackrel{\text{def}}{=} \left[\mathbf{r}_1^T(k)\, \mathbf{r}_2^T(k)\, \mathbf{r}_3^T(k)\right]^T$ is of dimension $3N = 21$. Angles of arrival are chosen randomly in $[0\ 2\pi]$. In Figure 2, BER of the desired user vs. interferers' SNR is evaluated for CGRRF of ranks 2, 3 and 4 and for the full-rank filter. The covariance matrix $\mathbf{R}$ is estimated using a block of 50 observations of $\mathbf{r}(k)$. The desired user SNR is fixed at 8 dB. Single-user bound and RAKE (matched filter) performance are given for reference. The figure shows that CGRRF performs better than full-rank filter over the whole range of interferers' SNR. Also note that for low MAI levels it is better to use lower ranks.

Next experiment studies steady-state performance of CGRRF. A synchronous DS-CDMA system with processing gain $N = 31$ and $K = 16$ active users is considered. Each user signal propagates through a non-varying multipath (four propagation paths per user) channel with path delays chosen randomly in the range $0 \ldots 3T_c$ (T_c stands for the chip period) and path fadings chosen as i.i.d. Gaussian variables. Desired user's SNR was fixed at 8 dB while interfering users were at 14 dB. In Figure 3, performance of CGRRF is compared with that of standard methods, such as Recursive-Least-Squares (RLS) algorithm [4] and two variants of Auxiliary-Vector Filter (AVF) [5, 6] denoted here as AVF-1 and AVF-2. CGRRF, AVF-1 and AVF-2 are of the same rank $D = 4$. All algorithms estimate covariance matrix $\mathbf{R}$ as

$$\mathbf{R}(k) = \gamma\, \mathbf{R}(k-1) + (1-\gamma)\, \mathbf{r}(k)\mathbf{r}^H(k), \quad (13)$$

with forgetting factor $\gamma = 0.99$. Output Signal-to-Interference-Plus-Noise Ratio (SINR), averaged over 250 Monte-Carlo trials, is measured. It can be observed that CGRRF of rank 4 shows SINR within 1 dB from RLS and within 1.5 dB from SINR of optimal Wiener filter. Also, CGRRF outperforms AVF of the same rank. This gain is explained by the fact that AVF only approximates reduced-rank Wiener filter in $\mathcal{K}^D(\mathbf{R}, \mathbf{c}_1)$ [1].

Third experiment demonstrates structural flexibility of the proposed filter. Each 250 samples we estimate instantaneous SINR at the output of the Dth stage which is then time-averaged using the forgetting factor of 0.95. The value of rank (D) is then either increased or decreased by 1 in order to keep time-averaged SINR within the range 3 ± 1 dB. In Figures 4 and 5 one realisation of time-averaged SINR vs. time and of rank D vs. time is shown. Initially, the simulation set-up of the previous experiment is used. Starting rank value is 2. Over the first 1000 samples rank can be seen to converge and then stabilize at $D = 5$. At time $k = 4000$, six interfering users quit the system and D decreases to 3. Finally, at $k = 6000$ three of interferers re-enter the communication and filter rank again grows to 4. Figure 4 shows that the reception quality is kept reasonably well within the required limits. With this or similar technique available processing power is used in an efficient way.

[2]More exactly, weighting coefficients w_i which result from backward recursion of MSWF have to be re-computed.

7 CONCLUSIONS

Adaptive reduced-rank filter based on conjugate-gradient technique has been presented. The proposed algorithm (CGRRF) combines flexible structure with fast convergence and near-optimal steady-state performance achieved with low filter ranks. CGRRF can be conveniently implemented in high-rate communications systems with strong constraints on number of training symbols as well as on available processing power.

References

[1] S. Burykh and K. Abed-Meraim, "Reduced-rank Adaptive Filtering Using Krylov Subspace Basis: A Comparative Study," *in Proc. of 39th Allerton Conference on Communications, Control and Computing*, October 2001.

[2] J. S. Goldstein, I. S. Reed and L. L. Scharf, "A Multistage Representation of the Wiener Filter Based on Orthogonal Projections," *IEEE Trans. on Inf. Theory*, pp. 2943-2959, November 1998.

[3] G. H. Golub and C. F. Van Loan, *Matrix Computations*, John Hopkins Univ. Press, 1989.

[4] M. L. Honig and M. K. Tsatsanis, "Adaptive Techniques for Multiuser CDMA Receivers", *IEEE Signal Proc. Magazine*, pp. 49-61, May 2000.

[5] D. A. Pados and S. N. Batalama, "Joint Space-Time Auxiliary-Vector Filtering for DS/CDMA Systems with Antenna Arrays," *IEEE Trans. on Comm.*, pp. 1406-1415, September 1999.

[6] D. A. Pados and G. N. Karystinos, "An Iterative Algorithm for the Computation of the MVDR Filter", *IEEE Trans. on Signal Proc.*, pp. 290-300, February 2001.

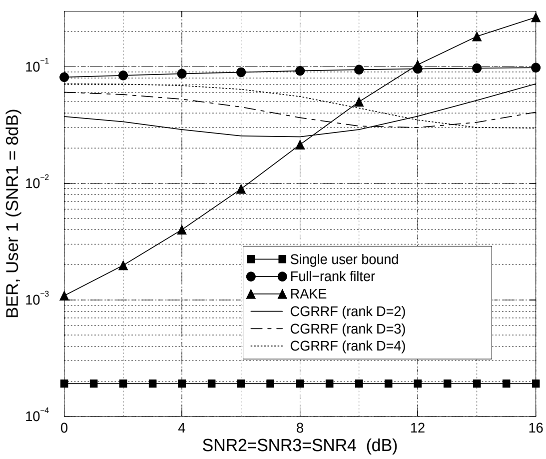

Figure 2: BER vs. interferer's SNR for CGRRF and full-rank filter.

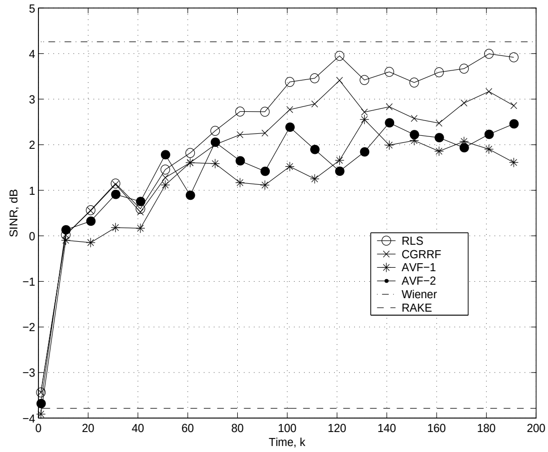

Figure 3: Output SINR vs. time for CGRRF, AVF-1, AVF-2 and RLS algorithms. Rank $D = 4$.

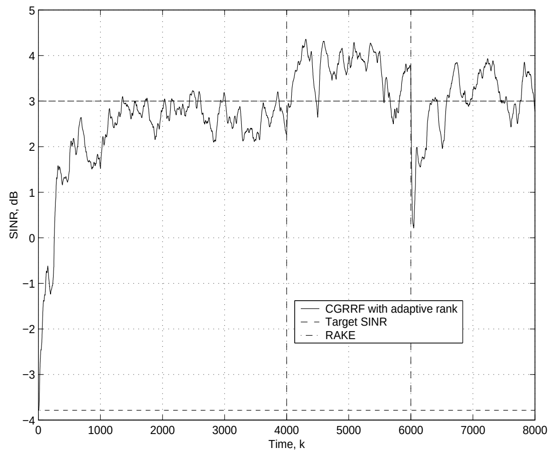

Figure 4: Averaged SINR vs. time for CGRRF with adaptive rank selection.

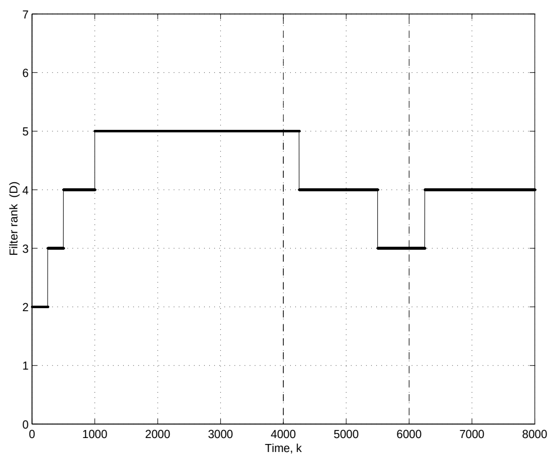

Figure 5: Rank D vs. time for CGRRF with adaptive rank selection.