Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 1999-52

A Comparison between Iterative block-LMS and Partial Rank Algorithm [1]

Koen Eneman, Simon Doclo, Marc Moonen [2]

March 1999
Internal report ESAT-SISTA

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/eneman/reports/99-52.ps.gz*

[2]ESAT (SISTA) - Katholieke Universiteit Leuven, Kardinaal Mercierlaan 94, 3001 Leuven (Heverlee), Belgium, Tel. 32/16/321809, Fax 32/16/321970, WWW: *http://www.esat.kuleuven.ac.be/sista*. E-mail: *koen.eneman@esat.kuleuven.ac.be*, *simon.doclo@esat.kuleuven.ac.be*. Simon Doclo is a Research Assistant supported by the I.W.T. Marc Moonen is a Research Associate with the F.W.O. Vlaanderen. This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of the concerted research action MIPS ('Model–based Information Processing Systems') of the Flemish Government, the F.W.O. research project nr. G.0295.97, the IUAP program P4–02 (1997–2001) ('Modeling, Identification, Simulation and Control of Complex Systems'), the IT–project *Multimicrophone Signal Enhancement Techniques for Hands–free Telephony and Voice Controlled Systems* (AUT/970517/Philips ITCL) of the I.W.T. and was partially sponsored by Philips–ITCL. The scientific responsibility is assumed by its authors.

A Comparison between Iterative block-LMS and Partial Rank Algorithm

Koen Eneman, Simon Doclo, Marc Moonen

March 17, 1999

1 Introduction

In this report, we examine the relation between two adaptive filtering techniques : iterative block-LMS [1] and PRA (Partial Rank Algorithm) [2][3]. It is shown that iterative block-LMS approaches PRA with stepsize 1, if the number of iterations goes to infinity. This is a simple block-extension of the fact that the normalised LMS algorithm can be considered as the limiting result of reiterating the LMS algorithm [4]. In each iteration of block-LMS, the a-priori (and a-posteriori) error decreases, and the difference between the actual and the calculated filter coefficients decreases. The computational complexity for both algorithms is compared.

2 Notation

- The size of each block is L and the filter length is N.
- The input vector at time k is denoted by $\mathbf{x}_k = [x_k \;\; x_{k-1} \;\; \dots \;\; x_{k-N+1}]^T$.
- In block-LMS as well as in PRA, we consider L consecutive input vectors, such that the input data matrix $\mathbf{X}_n \in \mathbb{R}^{N \times L}$ for block n is a Toeplitz-matrix,

$$\mathbf{X}_n = [\mathbf{x}_{nL} \;\; \mathbf{x}_{nL+1} \;\; \dots \;\; \mathbf{x}_{nL+L-1}] = \begin{bmatrix} x_{nL} & x_{nL+1} & \dots & x_{nL+L-1} \\ x_{nL-1} & x_{nL} & \dots & x_{nL+L-2} \\ \vdots & \vdots & & \vdots \\ x_{nL-N+1} & x_{nL-N+2} & \dots & x_{nL-N+L} \end{bmatrix}. \tag{1}$$

- For block-LMS, the adaptive filter of length N for block n at iteration step r is denoted by $\mathbf{w}_n^{(r)}$.
- For PRA, the adaptive filter of length N for block n is denoted by $\mathbf{w}_n$.
- The desired signal for block n is denoted by $\mathbf{d}_n = [d_{nL} \;\; d_{nL+1} \;\; \dots \;\; d_{nL+L-1}]^T$.
- The error signal for block n is denoted by $\mathbf{e}_n = [e_{nL} \;\; e_{nL+1} \;\; \dots \;\; e_{nL+L-1}]^T$.

3 Block-LMS and PRA

Block-LMS and PRA (as many adaptive filtering algorithms) consist of two parts :

1. calculation of the a-priori error
2. weight update using this a-priori error

3.1 Calculation of the a-priori error

The calculation of the a-priori error is the same for both algorithms. For each element in $\mathbf{e}_n^{pri}$, we can write

$$e_k^{pri} = d_k - \mathbf{x}_k^T \mathbf{w}_n,\ k = nL \ldots nL + L - 1,$$

such that

$$\mathbf{e}_n^{pri} = \mathbf{d}_n - \mathbf{X}_n^T \mathbf{w}_n. \tag{2}$$

3.2 Weight update

The weight update formula for block-LMS has the same form as the (well-known) weight update formula for LMS, except for the fact that the mean gradient over a block of size L is calculated and that the filters are updated every L samples :

$$\begin{aligned} \mathbf{w}_{n+1} &= \mathbf{w}_n + \mu \sum_{k=nL}^{nL+L-1} \mathbf{x}_k e_k^{pri} \\ &= \mathbf{w}_n + \mu \mathbf{X}_n \mathbf{e}_n^{pri} \\ &= \mathbf{w}_n + \mu \mathbf{X}_n \left(\mathbf{d}_n - \mathbf{X}_n^T \mathbf{w}_n \right) \end{aligned} \tag{3}$$

For the same block of L samples, we can as well iterate this formula :

$$\mathbf{w}_n^{(0)} = \mathbf{w}_n \tag{4}$$

$$\mathbf{w}_n^{(r+1)} = \mathbf{w}_n^{(r)} + \mu \mathbf{X}_n \left(\mathbf{d}_n - \mathbf{X}_n^T \mathbf{w}_n^{(r)} \right) \tag{5}$$

Block-LMS for which equation 5 is iterated is called iterative block-LMS or block row action projection (BRAP)[5]. The different steps of the algorithm are summarised in table 1.
The partial rank algorithm (PRA) is a block version of the affine projection algorithm (APA) [5]. The weight update equation is an extension of equation 3. A normalisation step is added, resulting in the following weight update equation :

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \alpha \mathbf{X}_n \left(\mathbf{X}_n^T \mathbf{X}_n + \delta \mathbf{I}_L \right)^{-1} \left(\mathbf{d}_n - \mathbf{X}_n^T \mathbf{w}_n \right) \tag{6}$$

If $L = 1$, PRA reduces to NLMS. The steps of the PRA-algorithm are summarised in table 2.

4 The Relation between iterative block-LMS and PRA

4.1 Weight update

Theorem 1 *For the rth iteration, the weights $\mathbf{w}_n^{(r)}$ can be written in function of $\mathbf{w}_n^{(0)}$ as*

$$\mathbf{w}_n^{(r)} = \mathbf{w}_n^{(0)} + \mu \mathbf{X}_n \left[\sum_{l=0}^{r-1} \left(\mathbf{I}_L - \mu \mathbf{X}_n^T \mathbf{X}_n \right)^l \right] \left(\mathbf{d}_n - \mathbf{X}_n^T \mathbf{w}_n^{(0)} \right). \tag{7}$$

Proof : For $r = 1$, the proof is trivial and follows from equation 3.
Suppose equation (7) holds for $\mathbf{w}_n^{(r)}$, then we can prove that

$$\begin{aligned}
\mathbf{w}_n^{(r+1)} &= \mathbf{w}_n^{(r)} + \mu\mathbf{X}_n\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(r)}\right) \\
&= \mathbf{w}_n^{(0)} + \mu\mathbf{X}_n\left[\sum_{l=0}^{r-1}\left(\mathbf{I}_L - \mu\mathbf{X}_n^T\mathbf{X}_n\right)^l\right]\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right) + \\
&\quad \mu\mathbf{X}_n\left[\mathbf{d}_n - \mathbf{X}_n^T\left(\mathbf{w}_n^{(0)} + \mu\mathbf{X}_n\left[\sum_{l=0}^{r-1}\left(\mathbf{I}_L - \mu\mathbf{X}_n^T\mathbf{X}_n\right)^l\right]\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right)\right)\right] \\
&= \mathbf{w}_n^{(0)} + \mu\mathbf{X}_n\left[\sum_{l=0}^{r-1}\left(\mathbf{I}_L - \mu\mathbf{X}_n^T\mathbf{X}_n\right)^l + \mathbf{I}_L - \mu\mathbf{X}_n^T\mathbf{X}_n\sum_{l=0}^{r-1}\left(\mathbf{I}_L - \mu\mathbf{X}_n^T\mathbf{X}_n\right)^l\right]\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right) \\
&= \mathbf{w}_n^{(0)} + \mu\mathbf{X}_n\left[\mathbf{I}_L + \left(\mathbf{I}_L - \mu\mathbf{X}_n^T\mathbf{X}_n\right)\sum_{l=0}^{r-1}\left(\mathbf{I}_L - \mu\mathbf{X}_n^T\mathbf{X}_n\right)^l\right]\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right) \\
&= \mathbf{w}_n^{(0)} + \mu\mathbf{X}_n\left[\mathbf{I}_L + \sum_{l=1}^{r}\left(\mathbf{I}_L - \mu\mathbf{X}_n^T\mathbf{X}_n\right)^l\right]\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right) \\
&= \mathbf{w}_n^{(0)} + \mu\mathbf{X}_n\left[\sum_{l=0}^{r}\left(\mathbf{I}_L - \mu\mathbf{X}_n^T\mathbf{X}_n\right)^l\right]\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right)
\end{aligned}$$

Theorem 2 *The iteration described in equations (4) and (5) converges to*

$$\mathbf{w}_n^{(\infty)} = \mathbf{w}_n^{(0)} + \mathbf{X}_n\left(\mathbf{X}_n^T\mathbf{X}_n\right)^{-1}\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right), \tag{8}$$

if μ is chosen such that $\mu \leq \frac{2}{\lambda_{max}}$, with λ_{max} the maximum eigenvalue of $\mathbf{X}_n^T\mathbf{X}_n$.

Proof : If we take the eigendecomposition of $\mathbf{X}_n^T\mathbf{X}_n$

$$\mathbf{X}_n^T\mathbf{X}_n = \mathbf{Q}^T\Lambda\mathbf{Q},$$

then we can write

$$\begin{aligned}
\sum_{l=0}^{r-1}\left(\mathbf{I}_L - \mu\mathbf{X}_n^T\mathbf{X}_n\right)^l &= \sum_{l=0}^{r-1}\left(\mathbf{Q}^T\mathbf{Q} - \mu\mathbf{Q}^T\Lambda\mathbf{Q}\right)^l \\
&= \mathbf{Q}^T\left[\sum_{l=0}^{r-1}\left(\mathbf{I}_L - \mu\Lambda\right)^l\right]\mathbf{Q} \\
&= \mathbf{Q}^T\mathrm{diag}\left(\sum_{l=0}^{r-1}\left(1 - \mu\lambda_j\right)^l\right)\mathbf{Q} \\
&\overset{\mu,\lambda_j\neq 0}{=} \mathbf{Q}^T\mathrm{diag}\left(\frac{1 - \left(1 - \mu\lambda_j\right)^r}{\mu\lambda_j}\right)\mathbf{Q},
\end{aligned} \tag{9}$$

such that equation (7) becomes

$$\mathbf{w}_n^{(r)} = \mathbf{w}_n^{(0)} + \mathbf{X}_n\mathbf{Q}^T\mathrm{diag}\left(\frac{1 - \left(1 - \mu\lambda_j\right)^r}{\lambda_j}\right)\mathbf{Q}\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right). \tag{10}$$

For r going to ∞, equation (10) becomes

$$\begin{aligned}\lim_{r\to\infty}\mathbf{w}_n^{(r)} &= \mathbf{w}_n^{(0)} + \mathbf{X}_n\mathbf{Q}^T\mathrm{diag}\left(\lim_{r\to\infty}\frac{1-(1-\mu\lambda_j)^r}{\lambda_j}\right)\mathbf{Q}\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right)\\ &= \mathbf{w}_n^{(0)} + \mathbf{X}_n\mathbf{Q}^T\mathrm{diag}\left(\frac{1}{\lambda_j}\right)\mathbf{Q}\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right) \qquad (11)\end{aligned}$$

under the condition that $|1-\mu\lambda_j| \leq 1 \quad \forall j$, which is satisfied if $(1-\mu\lambda_{max}) \geq -1$ (assuming $\mu > 0$) or $\mu \leq \frac{2}{\lambda_{max}}$. We can rewrite equation (11) as

$$\begin{aligned}\mathbf{w}_n^{(\infty)} &= \mathbf{w}_n^{(0)} + \mathbf{X}_n\mathbf{Q}^T\Lambda^{-1}\mathbf{Q}\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right)\\ &= \mathbf{w}_n^{(0)} + \mathbf{X}_n\left(\mathbf{X}_n^T\mathbf{X}_n\right)^{-1}\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right). \qquad (12)\end{aligned}$$

which is the same weight update formula as for PRA with $\alpha = 1$ and $\delta = 0$ (see equation 6).

4.2 A-posteriori error

The a-posteriori error is defined as :

$$\mathbf{e}_n^{post} = \mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_{n+1} = \mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(R-1)}. \qquad (13)$$

After R iterations with block-LMS the a-posteriori error is

$$\begin{aligned}\mathbf{e}_n^{post} &= \mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(R-1)}\\ &= \mathbf{d}_n - \mathbf{X}_n^T\left[\mathbf{w}_n^{(0)} + \mathbf{X}_n\mathbf{Q}^T\mathrm{diag}\left(\frac{1-(1-\mu\lambda_j)^{R-1}}{\lambda_j}\right)\mathbf{Q}\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right)\right]\\ &= \left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right) - \mathbf{X}_n^T\mathbf{X}_n\mathbf{Q}^T\mathrm{diag}\left(\frac{1-(1-\mu\lambda_j)^{R-1}}{\lambda_j}\right)\mathbf{Q}\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right)\\ &= \left[\mathbf{I}_L - \mathbf{Q}^T\Lambda\mathbf{Q}\mathbf{Q}^T\mathrm{diag}\left(\frac{1-(1-\mu\lambda_j)^{R-1}}{\lambda_j}\right)\mathbf{Q}\right]\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right) \qquad (14)\\ &= \left[\mathbf{I}_L - \mathbf{Q}^T\mathrm{diag}\left(1-(1-\mu\lambda_j)^{R-1}\right)\mathbf{Q}\right]\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right). \qquad (15)\end{aligned}$$

For R going to ∞, the a-posteriori error becomes zero :

$$\lim_{R\to\infty}\mathbf{e}_n^{post} = \left[\mathbf{I}_L - \mathbf{Q}^T\mathbf{Q}\right]\left(\mathbf{d}_n - \mathbf{X}_n^T\mathbf{w}_n^{(0)}\right) = 0, \qquad (16)$$

as could be expected. Iterative block-LMS approaches PRA with $\alpha = 1$ and $\delta = 0$ and PRA is defined such that the a-posteriori error is zero.

5 Computational complexity

5.1 Iterative block-LMS

The formulas defining iterative block-LMS (BRAP) and the corresponding cost are given in table 1.

for each block of L samples do

$$\mathbf{X}_n = \begin{bmatrix} x_{nL} & x_{nL+1} & \dots & x_{nL+L-1} \\ x_{nL-1} & x_{nL} & \dots & x_{nL+L-2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{nL-N+1} & x_{nL-N+2} & \dots & x_{nL-N+L} \end{bmatrix}$$

$$\mathbf{w}_n^{(0)} = \mathbf{w}_n$$

$$\mathbf{d}_n = \begin{bmatrix} d_{nL} \\ \vdots \\ d_{nL+L-1} \end{bmatrix}$$

for $r = 0$ to $R-1$ do

$\mathbf{y}_n^{(r)} = \mathbf{X}_n^T \mathbf{w}_n^{(r)}$	$2LN - L$
$\mathbf{e}_n^{(r)} = \mathbf{d}_n - \mathbf{y}_n^{(r)}$	L
$\mathbf{w}_n^{(r+1)} = \mathbf{w}_n^{(r)} + \mu \mathbf{X}_n \mathbf{e}_n^{(r)}$	$2LN + L$

}

$$\mathbf{w}_{n+1} = \mathbf{w}_n^{(R-1)}$$

}

Table 1: Block Row Action Projection (BRAP)

The total cost per sample for iterative block-LMS (BRAP) (see table 1) is

$$4NR + R \tag{17}$$

5.2 Partial Rank Algorithm

The formulas defining PRA are given in table 2. In order to obtain vector $\mathbf{p}_n$ the inverse of a regularised symmetric matrix $\mathbf{M}_n + \delta \mathbf{I}_L = \mathbf{X}_n^T \mathbf{X}_n + \delta \mathbf{I}_L$ has to be computed. As data is shifted block by block efficient QR-based updating of $\mathbf{X}_n^T \mathbf{X}_n$ is not possible for $L > 1$. However, $\mathbf{X}_n^T \mathbf{X}_n$ can be updated[1] :

$$\mathbf{M}_n = \mathbf{X}_n^T \mathbf{X}_n = \left[\begin{array}{c|c|c|c} {\mathbf{X}_n^0}^T & {\mathbf{X}_n^1}^T & \dots & {\mathbf{X}_n^{N/L-1}}^T \end{array} \right] \left[\begin{array}{c} \mathbf{X}_n^0 \\ \hline \mathbf{X}_n^1 \\ \hline \vdots \\ \hline \mathbf{X}_n^{N/L-1} \end{array} \right] \tag{18}$$

$$= \mathbf{M}_{n-1} + {\mathbf{X}_n^0}^T \mathbf{X}_n^0 - {\mathbf{X}_n^{N/L}}^T \mathbf{X}_n^{N/L} \tag{19}$$

[1] assuming N is multiple of L and defining $\mathbf{X}_n^k = \begin{bmatrix} x_{nL-kL} & \dots & x_{nL+L-1-kL} \\ \vdots & \ddots & \vdots \\ x_{nL-L+1-kL} & \dots & x_{nL-kL} \end{bmatrix}$

for each block of L samples do

$$\mathbf{X}_n^k \overset{\forall k}{=} \begin{bmatrix} x_{nL-kL} & \cdots & x_{nL+L-1-kL} \\ \vdots & \ddots & \vdots \\ x_{nL-L+1-kL} & \cdots & x_{nL-kL} \end{bmatrix}, \quad k = 0 \rightarrow \frac{N}{L}$$

$$\mathbf{X}_n = \begin{bmatrix} \mathbf{X}_n^0 \\ \mathbf{X}_n^1 \\ \vdots \\ \mathbf{X}_n^{N/L-1} \end{bmatrix}$$

$$\mathbf{d}_n = \begin{bmatrix} d_{nL} \\ \vdots \\ d_{nL+L-1} \end{bmatrix}$$

$\mathbf{y}_n = \mathbf{X}_n^T \mathbf{w}_n$	$2LN - L$
$\mathbf{e}_n = \mathbf{d}_n - \mathbf{y}_n$	L
$\mathbf{M}_n = \mathbf{M}_{n-1} + {\mathbf{X}_n^0}^T \mathbf{X}_n^0 - {\mathbf{X}_n^{N/L}}^T \mathbf{X}_n^{N/L}$	$(2L+1)L(L+1)$
$\mathbf{p}_n = \left(\mathbf{M}_n + \delta \mathbf{I}_L\right)^{-1} \mathbf{e}_n$	$\frac{2}{3}L^3 + L^2 + L$
$\mathbf{w}_{n+1} = \mathbf{w}_n + \alpha \mathbf{X}_n \mathbf{p}_n$	$2LN + L$

}

Table 2: Partial Rank Algorithm (PRA)

This reduces the cost substantially if N/L is large. $\mathbf{p}_n = (\mathbf{M}_n + \delta \mathbf{I}_L)^{-1} \mathbf{e}_n$ can be computed efficiently by solving the linear set of equations $(\mathbf{M}_n + \delta \mathbf{I}_L)\mathbf{p}_n = \mathbf{e}_n$ using Gauss elimination (cost : $\frac{2L^3}{3}$) and back substitution (cost : L^2). The total cost per sample for PRA (see table 2) is

$$4N + \frac{8L^2}{3} + 4L + 3 \tag{20}$$

5.3 BRAP vs. PRA

Whereas the cost of BRAP is independent of the block length, PRA becomes more expensive for large values of L. For small values of R, BRAP will be cheaper than PRA. The more iterations R are performed with BRAP, the more expensive it becomes, but the better it approaches the convergence behaviour of PRA (see section 6).

We now compute R_{crit}, the value for R for which BRAP is as expensive as PRA. By combining eq. 17 and 20, we obtain

$$R_{crit} = \frac{4N + \frac{8L^2}{3} + 4L + 3}{4N + 1}. \tag{21}$$

For long adaptive filters, $N >> L$. By taking the limit for N to ∞ we find

$$R_{crit}^{\infty} = 1 \tag{22}$$

As could be expected, for small block lengths L, R_{crit} is small and PRA outperforms BRAP both in terms of convergence and cost.
The effect of parameter L on the convergence of PRA is shown in figure 1 and 2. Both

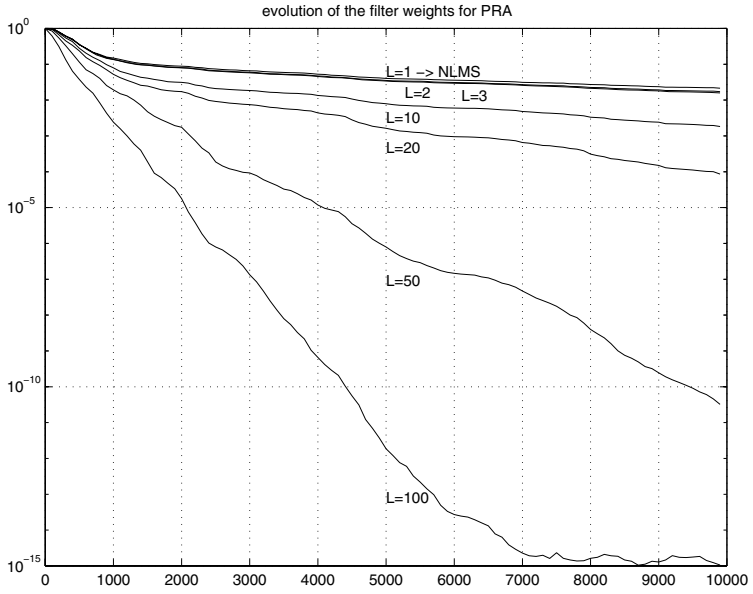

Figure 1: Error between the filter coefficients and the true path as a function of time for PRA. Coloured stationary noise was filtered with a 150-taps FIR filter. The adaptive filter length N is 150.

coloured noise and a music signal are filtered with an artificial path and identified with PRA. Block length L was changed from 1 (=NLMS) to 100. Large values of L lead to better steady-state suppression and initial convergence. For large values of L also R_{crit} will be considerable and BRAP with $R < R_{crit}$ might be employed to obtain a cheap adaptive algorithm with a performance comparable to that of PRA. As ever, there exists a trade-off between good performance at a high cost (PRA) and reduced performance at a lower cost (BRAP with $R < R_{crit}$). The optimal R can be determined using the following heuristic :

- increase R by making a trade-off between the expected improvement in performance and extra cost, taking in account
  - the available processor power
  - the additional cost and improvement in performance when choosing PRA
- if $R \geq R_{crit}$, choose PRA

Some performance tests are needed to estimate the expected improvement in performance for BRAP when R is increased.

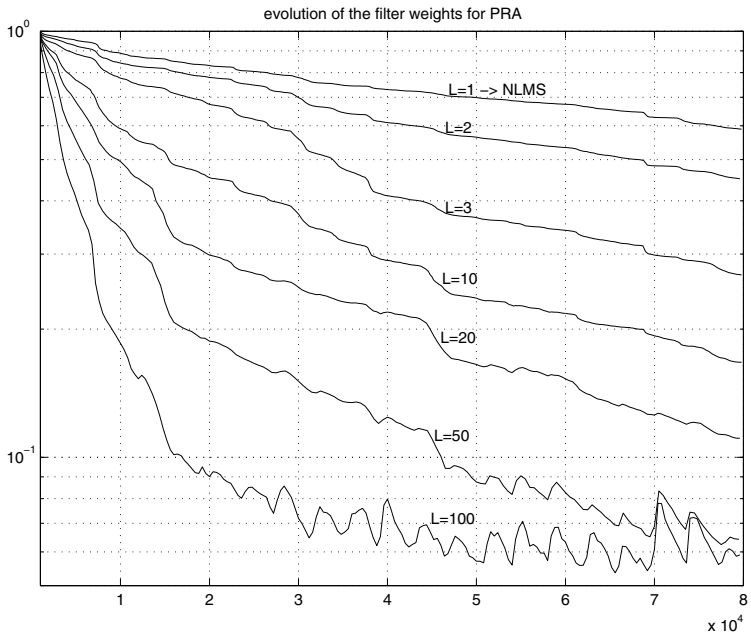

Figure 2: Error between the filter coefficients and the true path as a function of time for PRA. A music signal was filtered with a 1200-taps FIR filter. The adaptive filter length is 1000.

6 Simulations

Coloured stationary noise was filtered with a 100-taps FIR filter. PRA and BRAP with $R = 1, 2, 3, 10, 20, 50, 100$ are compared. The length of the adaptive filters N is 100. In a first experiment the block length L was chosen to be 20. For this value of L, $R_{crit} = 3.86$. Figure 3 shows the evolution of the filter coefficients. Figure 4 and 5 show the evolution of the a-priori and a-posteriori errors respectively. In a second experiment the block length L was chosen to be 3. Figure 6 shows the evolution of the filter coefficients. For this value of L, $R_{crit} = 1.09$. Figure 7 and 8 show the evolution of the a-priori and a-posteriori errors respectively.

7 Conclusion

In this report we have shown that the partial rank algorithm (PRA) can be considered as the limiting result of reiterating the block-LMS algorithm (BRAP) with the same data. As a consequence, for BRAP, the convergence of the adaptive weights and the a-priori error is inferior to that of PRA. For a small number of iterations the computational cost of BRAP is lower than that of PRA. This means that for large block sizes and limited processing power iterative block-LMS can be a viable alternative for PRA.

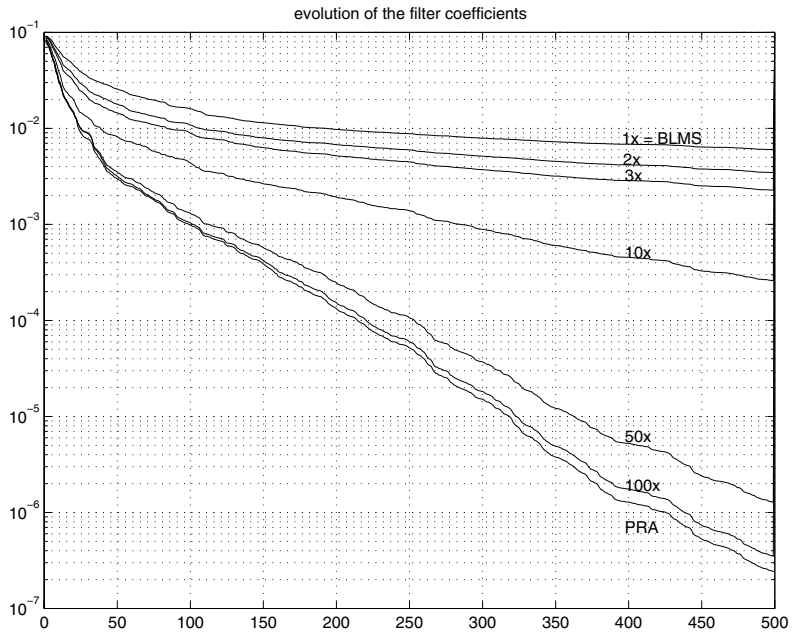

Figure 3: Error between the filter coefficients and the true path as a function of time for BRAP and PRA, $L = 20$. Coloured stationary noise was filtered with a 100-taps FIR filter. The adaptive filter length N is 100.

8 Acknowledgements

Simon Doclo is a Research Assistant supported by the I.W.T. Marc Moonen is a Research Associate with the F.W.O. Vlaanderen. This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of the concerted research action MIPS ('Model–based Information Processing Systems') of the Flemish Government, the F.W.O. research project nr. G.0295.97, the IUAP program P4–02 (1997–2001) ('Modeling, Identification, Simulation and Control of Complex Systems'), the IT–project *Multimicrophone Signal Enhancement Techniques for Hands–free Telephony and Voice Controlled Systems* (AUT/970517/Philips ITCL) of the I.W.T. and was partially sponsored by Philips–ITCL. The scientific responsibility is assumed by its authors.

References

[1] S. Haykin, *Adaptive Filter Theory*, Information and system sciences series. Prentice Hall, 3rd edition, 1996.

[2] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties", *Electron. Commun. Japan*, vol. 67–A, pp. 19–27, 1984.

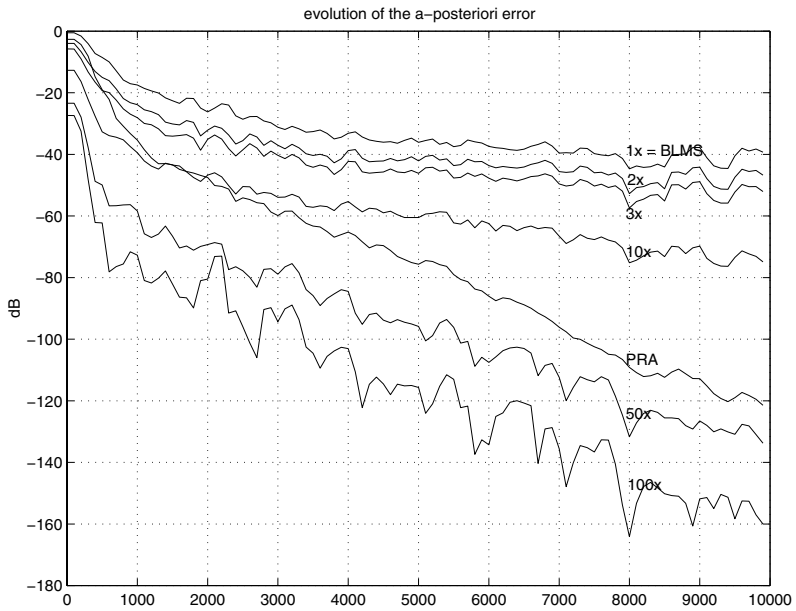

Figure 4: Evolution of the a-priori errors for BRAP and PRA, $L = 20$. Coloured stationary noise was filtered with a 100-taps FIR filter. The adaptive filter length N is 100.

[3] S. G. Kratzer and D. R. Morgan, "The partial-rank algorithm for adaptive beamforming", in *Proc. SPIE Int. Soc. Optic. Eng.*, 1985, vol. 564, pp. 9–14.

[4] D. R. Morgan and S. G. Kratzer, "On a Class of Computationally Efficient, Rapidly Converging, Generalized NLMS Algorithms", *IEEE Signal Processing Letters*, vol. 3, no. 8, pp. 245–247, aug 1996.

[5] S. Gay, *Fast Projection Algorithms with Application to Voice Echo Cancellation*, PhD thesis, Rutgers, The State University of New Jersey, New Brunswick, New Jersey, USA, October 1994.

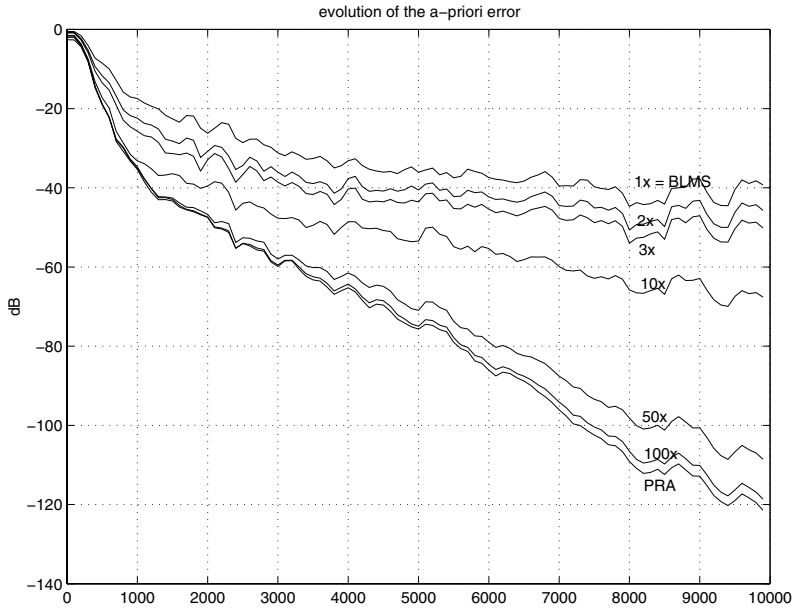

Figure 5: Evolution of the a-posteriori errors for BRAP and PRA, $L = 20$. Coloured stationary noise was filtered with a 100-taps FIR filter. The adaptive filter length N is 100.

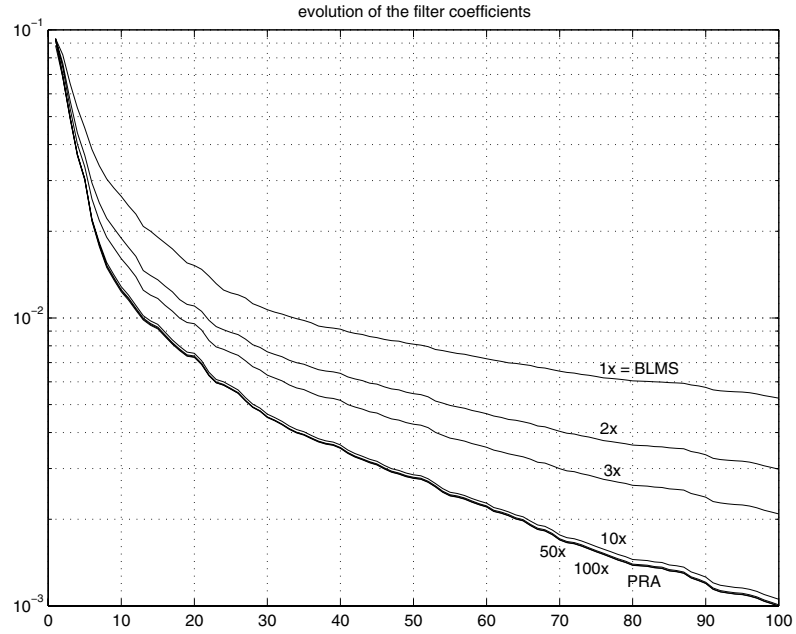

Figure 6: Error between the filter coefficients and the true path as a function of time for BRAP and PRA, $L = 3$. Coloured stationary noise was filtered with a 100-taps FIR filter. The adaptive filter length N is 100.

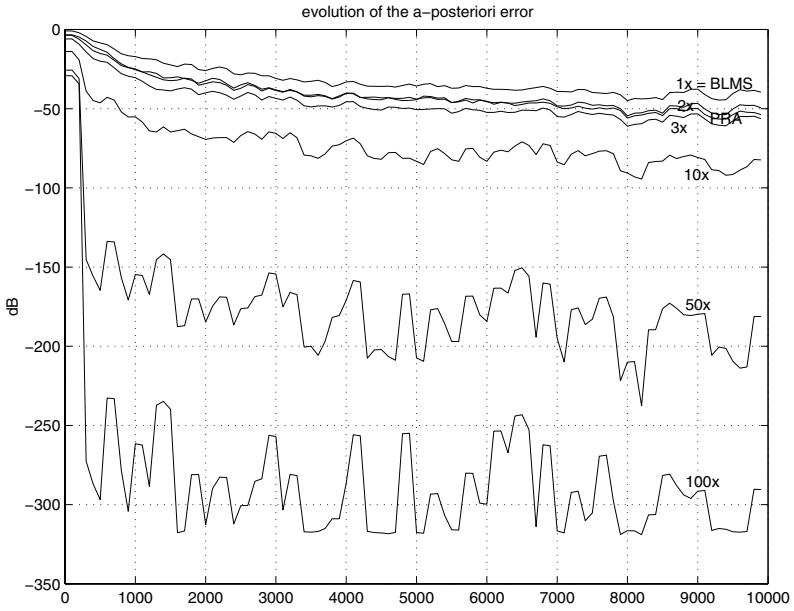

Figure 7: Evolution of the a-priori errors for BRAP and PRA, $L = 3$. Coloured stationary noise was filtered with a 100-taps FIR filter. The adaptive filter length N is 100.

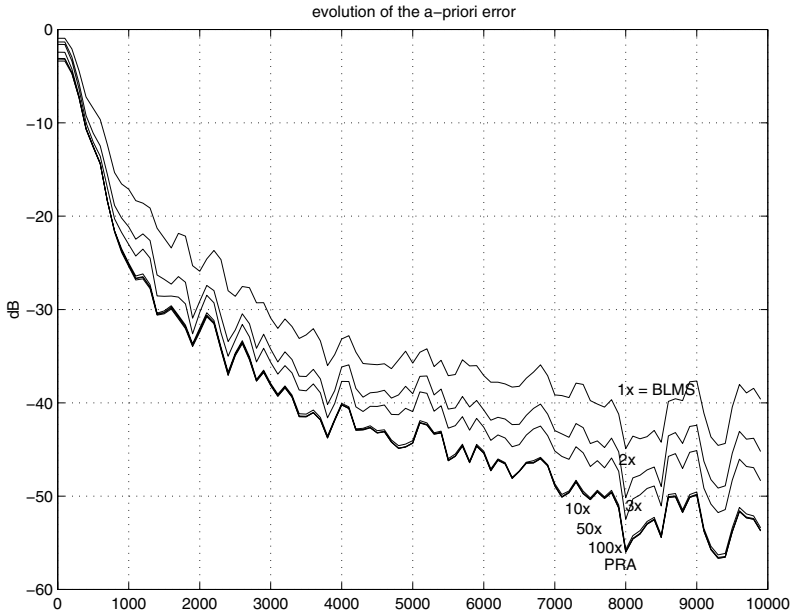

Figure 8: Evolution of the a-posteriori errors for BRAP and PRA, $L = 3$. Coloured stationary noise was filtered with a 100-taps FIR filter. The adaptive filter length N is 100.