Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 2000-127

On Iterating the Partitioned Block Frequency–Domain Adaptive Filter [1]

Koen Eneman, Marc Moonen [2]

December 2000

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/eneman/reports/00-127.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*. This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of the Belgian State Interuniversity Poles of Attraction Programme – IUAP P4–02 (1997–2001) : Modeling, Identification, Simulation and Control of Complex Systems, Concerted Research Action GOA–MEFISTO–666 of the Flemish Government, Research Project F.W.O. nr. G.0295.97 ('Design and implementation of adaptive digital signal processing algorithms for broadband applications') and IWT project AUT/970517: MUSETTE ('Multimicrophone Signal Enhancement Techniques for handsfree telephony and voice controlled systems'). The scientific responsibility is assumed by its authors.

On Iterating the Partitioned Block Frequency–Domain Adaptive Filter

Koen Eneman Marc Moonen
ESAT – Katholieke Universiteit Leuven
Kardinaal Mercierlaan 94, B–3001 Leuven – Belgium
phone : +32/16321809 +32/16321060
fax : +32/16321970
email : koen.eneman@esat.kuleuven.ac.be
marc.moonen@esat.kuleuven.ac.be

Abstract

In this text the PBFDRAP algorithm is considered, an adaptive filtering technique combining the partitioned block frequency–domain adaptive filter (PBFDAF) and row action projection (RAP). The PBFDRAP is obtained by iterating the filter and weight updating part of the PBFDAF, leading to an improved error output and better estimated filter weights. In this paper the PBFDRAP algorithm is analysed and more in particular the behaviour of the algorithm if the number of iteration steps goes to infinity. In [2] the relation between two adaptive filtering techniques was examined : iterated block–LMS and the Partial Rank Algorithm (PRA). It was shown that iterated block–LMS approaches PRA with stepsize $\mu = 1$, if the number of iteration steps goes to infinity. Similar computations can be done for the PBFDRAP. It will be shown that the PBFDRAP algorithm approaches a normalised version of the PBFDAF that is based on projected subband energies with maximum stepsize $\mu = 1$, if the number of iteration steps goes to infinity. For certain parameter settings the infinitely iterated PBFDRAP reduces to well–known adaptive filter algorithms.

1 Introduction

Of all existing adaptive algorithms the Least Mean Squares algorithm may be best known [10]. LMS–based algorithms have a complexity that is linear in the filter length, but they suffer from a rather slow convergence for coloured input signals such as speech. More efficient implementations of the LMS algorithm are available, often relying on frequency–domain techniques [8]. The so–called Partitioned Block Frequency–Domain Adaptive Filter (PBFDAF) [1][7][9] is often used nowadays to cancel long acoustic echoes. In this paper a combination of the PBFDAF algorithm and row action projection (RAP) will be discussed. Row Action Projection was initially proposed as an improvement to the LMS algorithm. A good reference on Row Action Projection can

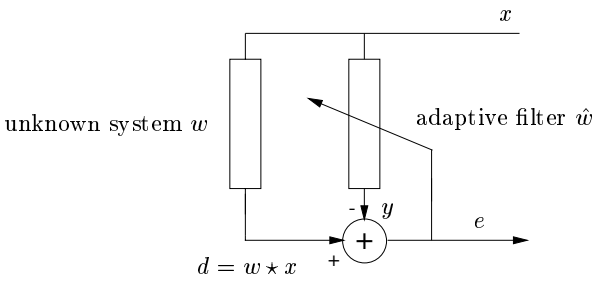


Figure 1: Adaptive filtering setup

be found in [5].

2 Frequency–Domain Adaptive Filter

As a cheaper alternative to the LMS algorithm, the frequency–domain adaptive filter (FDAF) was introduced, which is a direct translation of Block–LMS to the frequency domain [8]. Instead of linear convolutions and correlations circular operations are performed in the frequency domain. This requires some 'restore' operations which are based on overlap–save or overlap–add techniques. The overlap–save FDAF essentially comes down to the following procedure :

$$\mathbf{X}^{(n)} = \text{diag}\left\{ \mathbf{F} \begin{bmatrix} x[(n+1)L-M+1] \\ \vdots \\ x[(n+1)L] \end{bmatrix} \right\} \Bigg\uparrow M, \tag{1}$$

$$\mathbf{y}^{(n)} = \begin{bmatrix} \mathbf{0}_{N-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \mathbf{X}^{(n)} \mathbf{w}^{(n)} \tag{2}$$

$$\mathbf{d}^{(n)} = \begin{bmatrix} \mathbf{0} \\ \mathbf{d}_n \end{bmatrix} \begin{matrix} \updownarrow N-1+s \\ \updownarrow L \end{matrix}, \qquad \mathbf{d}_n = \begin{bmatrix} d[nL+1] \\ \vdots \\ d[(n+1)L] \end{bmatrix} \Bigg\uparrow L \tag{3}$$

$$\mathbf{e}^{(n)} = \mathbf{d}^{(n)} - \mathbf{y}^{(n)} \tag{4}$$

$$\mathbf{w}^{(n+1)} = \mathbf{w}^{(n)} + \mathbf{F} \begin{bmatrix} \mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1} \boldsymbol{\Delta} \mathbf{X}^{(n)*} \mathbf{F} \mathbf{e}^{(n)} \tag{5}$$

During each iteration n, L new x–samples are taken in, and L new filter output samples e are produced. Vector $\mathbf{d}_n$ has length L and contains the L most recent values of signal d (for signal conventions see figure 1). L is called the block length, and hence the corresponding input/output delay of the FDAF algorithm is $2L-1$. $\mathbf{F}$ is the $M \times M$ DFT matrix : element (p,q) of $\mathbf{F}$ is given by $\mathbf{F}(p,q) = e^{-j\frac{2\pi pq}{M}}$. Further, N is the adaptive filter length and $M = N+L-1+s$, $s \geqslant 0$. The standard version of the algorithm, for which $\boldsymbol{\Delta} = \mu \mathbf{I}_M$, corresponds to an exact translation of Block–LMS to the frequency domain. In the most general case however, $\boldsymbol{\Delta} = \text{diag}(\boldsymbol{\mu}_s^{(n)})$ contains

subband dependent stepsizes $\boldsymbol{\mu}_s^{(n)}(m)$ that are function of the block index n. The performance and convergence properties of the FDAF algorithm are comparable to those of the LMS algorithm. It appears that the FDAF algorithm is computationally attractive only if the block length L has the same order of magnitude as the filter length N. In practice however, this leads to unacceptable input/output delays.

3 Partitioned Block Frequency–Domain Adaptive Filter

By splitting the adaptive filter $\hat{w}$ (see figure 1) in equal parts and transforming to the frequency domain, a mixed time/frequency–domain adaptive filter can be obtained, called the Partitioned Block Frequency–Domain Adaptive Filter (PBFDAF) [1][7][9]. The N–taps fullband adaptive filter $\hat{w}^{(n)}[k]$ at block index n is partitioned in $\frac{N}{P}$ parts[1] $\mathbf{w}^{(n,p)}$ of length P each, which are transformed to the frequency domain :

$$\mathbf{w}^{(n,p)} \stackrel{\forall p}{=} \begin{bmatrix} \hat{w}^{(n)}[pP] \\ \vdots \\ \hat{w}^{(n)}[(p+1)P-1] \end{bmatrix}, \quad p = 0 : \frac{N}{P} - 1 \tag{6}$$

$$\mathbf{w}_p^{(n)} \stackrel{\forall p}{=} \mathbf{F} \begin{bmatrix} \mathbf{w}^{(n,p)} \\ \mathbf{0} \end{bmatrix} \begin{matrix} \updownarrow P \\ \updownarrow L-1+s \end{matrix} \tag{7}$$

The equations defining the PBFDAF are :

$$\mathbf{X}_p^{(n)} \stackrel{\forall p}{=} \operatorname{diag}\left\{ \mathbf{F} \begin{bmatrix} x[(n+1)L - pP - M + 1] \\ \vdots \\ x[(n+1)L - pP] \end{bmatrix} \right\} \Big\uparrow M, \quad p = 0 : \frac{N}{P} - 1 \tag{8}$$

$$\mathbf{y}^{(n)} = \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{w}_p^{(n)} \tag{9}$$

$$\mathbf{d}^{(n)} = \begin{bmatrix} \mathbf{0} \\ \mathbf{d}_n \end{bmatrix} \begin{matrix} \updownarrow P-1+s \\ \updownarrow L \end{matrix}, \qquad \mathbf{d}_n = \begin{bmatrix} d[nL+1] \\ \vdots \\ d[(n+1)L] \end{bmatrix} \Big\uparrow L \tag{10}$$

$$\mathbf{e}^{(n)} = \mathbf{d}^{(n)} - \mathbf{y}^{(n)} \tag{11}$$

$$\mathbf{w}_p^{(n+1)} \stackrel{\forall p}{=} \mathbf{w}_p^{(n)} + \mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1} \boldsymbol{\Delta} \mathbf{X}_p^{(n)*} \mathbf{F} \mathbf{e}^{(n)}, \quad p = 0 : \frac{N}{P} - 1 \tag{12}$$

Each time a PBFDAF iteration, described by Eq. 8–12, is performed L new x–samples are taken in, and L new filter output samples e are produced. Vector $\mathbf{d}_n$ contains the L most recent values of signal d (see figure 1 for signal conventions). L is called the block length, and hence the corresponding input/output delay is $2L-1$. $\mathbf{F}$ is the $M \times M$ DFT matrix, $\boldsymbol{\Delta} = \operatorname{diag}(\boldsymbol{\mu}_s^{(n)})$ contains the subband dependent stepsizes and

[1] It is assumed that $\frac{N}{P}$ is integer.

$M = P + L - 1 + s$, $s \geqslant 0$. Block length L of the PBFDAF can be adjusted, resulting in a cheap adaptive filter with acceptable processing delay.
There exist two variants of this algorithm, called the constrained and the unconstrained PBFDAF. The algorithm described by Eq. 8–12 is the constrained PBFDAF. For the unconstrained version the "projection" type operation

$$\mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1} \tag{13}$$

is left out of Eq. 12. The unconstrained updating requires 3 DFTs per iteration, provided P is divisible by L : equation 8 then requires only 1 DFT computation as $\mathbf{X}_p^{(n)}$ for $p > 0$ can be recuperated from previous iterations. The constrained PBFDAF is more expensive, having an extra $\frac{2N}{P}$ DFTs to compute. The latter on the other hand has better convergence properties. By introducing stepsize normalisation the convergence can be further improved.

4 Partitioned Block Frequency–Domain RAP

Stepping several times through the filter and weight updating part of the PBFDAF algorithm (Eq. 9, 11 and 12) with block index n kept constant leads to an improved update of the error output and the filter weights :

`for` $r = 1$ `to` R `do` $\{$

$$\mathbf{y}^{(n,r)} = \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{w}_p^{(n,r)} \tag{14}$$

$$\mathbf{d}^{(n)} = \begin{bmatrix} \mathbf{0} \\ \mathbf{d}_n \end{bmatrix} \begin{matrix} \updownarrow P-1+s \\ \updownarrow L \end{matrix}, \qquad \mathbf{d}_n = \begin{bmatrix} d[nL+1] \\ \vdots \\ d[(n+1)L] \end{bmatrix} \Big\updownarrow L \tag{15}$$

$$\mathbf{e}^{(n,r)} = \mathbf{d}^{(n)} - \mathbf{y}^{(n,r)} \tag{16}$$

$$\mathbf{w}_p^{(n,r+1)} \stackrel{\forall p}{=} \mathbf{w}_p^{(n,r)} + \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)*}\mathbf{F}\mathbf{e}^{(n,r)}, \quad p = 0 : \frac{N}{P} - 1 \tag{17}$$

$\}$

in which $\mathbf{w}_p^{(n,r)}$, $\mathbf{y}^{(n,r)}$, $\mathbf{e}^{(n,r)}$ are respectively the adaptive filter weights, an estimate of $w \star x$ and the error output at iteration step r. Further,

$$\mathbf{w}_p^{(n+1,1)} \stackrel{\forall p}{=} \mathbf{w}_p^{(n,R+1)}, \quad p = 0 : \frac{N}{P} - 1. \tag{18}$$

This extension of the PBFDAF algorithm which iterates several times on the same block of data $(\mathbf{X}_p^{(n)}, \mathbf{d}_n)$ will be called the partitioned block frequency–domain RAP (PBFDRAP) adaptive filter. RAP stands for Row Action Projection and was initially proposed as an improvement to the LMS algorithm. A good reference on Row Action Projection can be found in [5].

Matrix $\mathbf{G}$ in equation 17 defines the type of updating : for the unconstrained PBFDRAP, $\mathbf{G} = \mathbf{I}_M$, for the constrained PBFDRAP

$$\mathbf{G} = \mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1}. \tag{19}$$

Remark that the PBFDRAP reduces to the PBFDAF if $R = 1$. In [4] the PBFDRAP is discussed more in detail and some experimental results are given showing how the PBFDRAP can be used in an acoustic echo cancellation setup.

5 On Iterating the PBFDRAP

In [2] the relation between two adaptive filtering techniques was examined : iterated block–LMS and PRA (Partial Rank Algorithm), which is a block version of the APA–algorithm [5]. It was shown that iterated block–LMS approaches PRA with stepsize $\mu = 1$, if the number of iteration steps 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 [6]. Stepsize $\mu = 1$ is the maximum stepsize allowed with PRA. This maximum stepsize guarantees optimal convergence rates. The same sort of computation can be done for the PBFDRAP algorithm to check what happens if the number of iteration steps R goes to infinity.

5.1 Error output and weight updating

In general one can state that at the end of iteration step R the error output $\mathbf{e}^{(n,R)}$ of the PBFDRAP algorithm can be written as

$$\mathbf{e}^{(n,R)} = \left(\mathbf{I}_M - \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G} \boldsymbol{\Delta} \mathbf{X}_p^{(n)*} \mathbf{F} \right)^{R-1} \mathbf{e}^{(n,1)} \tag{20}$$

and that the adaptive filter weights are given by

$$\mathbf{w}_p^{(n,R+1)} \overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G} \boldsymbol{\Delta} \mathbf{X}_p^{(n)*} \mathbf{F} \sum_{r=1}^{R} \mathbf{e}^{(n,r)}, \quad p = 0 : \frac{N}{P} - 1. \tag{21}$$

Proof :
Taking in account Eq. 17, Eq. 20 and 21 are trivially fulfilled for $R = 1$. Now assuming that Eq. 20 and 21 hold for iteration step $R - 1$, i.e.

$$\mathbf{e}^{(n,R-1)} = \left(\mathbf{I}_M - \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G} \boldsymbol{\Delta} \mathbf{X}_p^{(n)*} \mathbf{F} \right)^{R-2} \mathbf{e}^{(n,1)} \tag{22}$$

and

$$\mathbf{w}_p^{(n,R)} \overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G} \boldsymbol{\Delta} \mathbf{X}_p^{(n)*} \mathbf{F} \sum_{r=1}^{R-1} \mathbf{e}^{(n,r)}, \tag{23}$$

it needs to be proven that Eq. 20 and 21 also hold for iteration step R. After the $(R-1)$–th iteration step (see Eq. 17)

$$\mathbf{w}_p^{(n,R)} \stackrel{\forall p}{=} \mathbf{w}_p^{(n,R-1)} + \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)^*}\mathbf{F}\mathbf{e}^{(n,R-1)}, \tag{24}$$

such that after iteration step R (using Eq. 14, 16, 22 and 24)

$$\begin{aligned}
\mathbf{e}^{(n,R)} &= \mathbf{d}^{(n)} - \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{w}_p^{(n,R)} && (25)\\
&= \mathbf{e}^{(n,R-1)} - \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)^*}\mathbf{F}\mathbf{e}^{(n,R-1)} && (26)\\
&= \left(\mathbf{I}_M - \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)^*}\mathbf{F} \right) \mathbf{e}^{(n,R-1)} && (27)\\
&= \left(\mathbf{I}_M - \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)^*}\mathbf{F} \right)^{R-1} \mathbf{e}^{(n,1)} && (28)
\end{aligned}$$

and by Eq. 17 and 23,

$$\begin{aligned}
\mathbf{w}_p^{(n,R+1)} &\stackrel{\forall p}{=} \mathbf{w}_p^{(n,R)} + \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)^*}\mathbf{F}\mathbf{e}^{(n,R)} && (29)\\
&\stackrel{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)^*}\mathbf{F} \sum_{r=1}^{R} \mathbf{e}^{(n,r)}, && (30)
\end{aligned}$$

which completes the proof. Hence, Eq. 20 and 21 are fulfilled for all $R > 0$.

5.2 Computation of $\lim_{R\to\infty} \mathbf{w}_p^{(n,R)}$

In order to compute $\lim_{R\to\infty} \mathbf{w}_p^{(n,R)}$ for the PBFDRAP,

$$\mathbf{N}_n = \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)^*}\mathbf{F} \tag{31}$$

and

$$\mathbf{P}_n = \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)^*}\mathbf{F} = \begin{bmatrix} \boldsymbol{\mathcal{A}} & \boldsymbol{\mathcal{B}} \\ \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{D}} \end{bmatrix} \tag{32}$$

are defined with $\boldsymbol{\mathcal{D}}$ an $L \times L$ matrix. It is further assumed that the eigenvalue decomposition of $\mathbf{N}_n$ is given by

$$\underbrace{\begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{D}} \end{bmatrix}}_{\mathbf{N}_n} \underbrace{\begin{bmatrix} \mathbf{V}_\alpha & \mathbf{V}_\beta \\ \mathbf{V}_\gamma & \mathbf{V}_\delta \end{bmatrix}}_{\mathbf{V}} = \underbrace{\begin{bmatrix} \mathbf{V}_\alpha & \mathbf{V}_\beta \\ \mathbf{V}_\gamma & \mathbf{V}_\delta \end{bmatrix}}_{\mathbf{V}} \underbrace{\begin{bmatrix} \boldsymbol{\Gamma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{P-1+s} \end{bmatrix}}_{\boldsymbol{\Gamma}}, \tag{33}$$

in which $\mathbf{\Gamma}_1$ is a full rank diagonal $L \times L$ matrix, i.e. the rank of $\mathbf{N}_n$ is L. In section 5.3 and 5.4 it will be pointed out under which conditions this assumption is true.
The structure of $\mathbf{N}_n$ and its rank deficiency force $\mathbf{V}_\alpha$ to be $\mathbf{0}$. $\mathbf{V}_\beta$ and $\mathbf{V}_\gamma$ are square and invertible matrices. Hence, the eigenvector matrix

$$\mathbf{V} = \begin{bmatrix} \mathbf{0} & \mathbf{V}_\beta \\ \mathbf{V}_\gamma & \mathbf{V}_\delta \end{bmatrix} \tag{34}$$

has a specific structure. Furthermore, it is easily verified using Eq. 33 that the eigenvalue decomposition of $\boldsymbol{\mathcal{D}}$ is given by

$$\boldsymbol{\mathcal{D}} = \mathbf{V}_\gamma \mathbf{\Gamma}_1 \mathbf{V}_\gamma^{-1} \tag{35}$$

and that

$$\mathbf{V}^{-1} = \begin{bmatrix} -\mathbf{V}_\gamma^{-1}\mathbf{V}_\delta\mathbf{V}_\beta^{-1} & \mathbf{V}_\gamma^{-1} \\ \mathbf{V}_\beta^{-1} & \mathbf{0} \end{bmatrix}. \tag{36}$$

From Eq. 20 and 33,

$$\sum_{r=1}^{R} \mathbf{e}^{(n,r)} = \sum_{r=1}^{R} \left(\mathbf{I}_M - \mathbf{N}_n\right)^{r-1} \mathbf{e}^{(n,1)} \tag{37}$$

$$= \mathbf{V} \sum_{r=1}^{R} \left(\mathbf{I}_M - \mathbf{\Gamma}\right)^{r-1} \mathbf{V}^{-1} \mathbf{e}^{(n,1)} \tag{38}$$

$$= \mathbf{V} \sum_{r=1}^{R} \begin{bmatrix} \mathbf{I}_L - \mathbf{\Gamma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{P-1+s} \end{bmatrix}^{r-1} \mathbf{V}^{-1} \mathbf{e}^{(n,1)} \tag{39}$$

can be obtained. As $\mathbf{\Gamma}_1$ is assumed to be a full rank diagonal matrix

$$\sum_{r=1}^{R} \mathbf{e}^{(n,r)} = \mathbf{V} \begin{bmatrix} \mathbf{\Gamma}_1^{-1}\left(\mathbf{I}_L - \left(\mathbf{I}_L - \mathbf{\Gamma}_1\right)^R\right) & \mathbf{0} \\ \mathbf{0} & R\mathbf{I}_{P-1+s} \end{bmatrix} \mathbf{V}^{-1} \mathbf{e}^{(n,1)}. \tag{40}$$

By inserting Eq. 34 and 36 in Eq. 40 and defining $\mathbf{\Xi}^{(R)} = \mathbf{\Gamma}_1^{-1}\left(\mathbf{I}_L - \left(\mathbf{I}_L - \mathbf{\Gamma}_1\right)^R\right)$

$$\sum_{r=1}^{R} \mathbf{e}^{(n,r)} = \begin{bmatrix} R\mathbf{I}_{P-1+s} & \mathbf{0} \\ \left(-\mathbf{V}_\gamma \mathbf{\Xi}^{(R)} \mathbf{V}_\gamma^{-1} + R\mathbf{I}_L\right)\mathbf{V}_\delta\mathbf{V}_\beta^{-1} & \mathbf{V}_\gamma\mathbf{\Xi}^{(R)}\mathbf{V}_\gamma^{-1} \end{bmatrix} \mathbf{e}^{(n,1)} \tag{41}$$

is found.
From Eq. 14, 15 and 16 it can be verified that $\mathbf{e}^{(n,1)}$ can be written as

$$\mathbf{e}^{(n,1)} = \begin{bmatrix} \mathbf{0}_{P-1+s\times 1} \\ \mathbf{e}_n \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} \mathbf{e}_n. \tag{42}$$

Hence,

$$\sum_{r=1}^{R} \mathbf{e}^{(n,r)} = \begin{bmatrix} \boldsymbol{\mathcal{K}} & \mathbf{0} \\ \boldsymbol{\mathcal{L}} & \mathbf{V}_\gamma \mathbf{\Xi}^{(R)} \mathbf{V}_\gamma^{-1} \end{bmatrix} \mathbf{e}^{(n,1)}, \tag{43}$$

in which $\boldsymbol{\mathcal{K}}$ and $\boldsymbol{\mathcal{L}}$ can be any matrix having the correct dimensions.
If γ_i, the diagonal elements of $\boldsymbol{\Gamma}_1$, are such that $0 \leqslant |1-\gamma_i| < 1$,

$$\lim_{R\to\infty} \boldsymbol{\Xi}^{(R)} = \boldsymbol{\Gamma}_1^{-1} \tag{44}$$

and from Eq. 35

$$\begin{aligned} \lim_{R\to\infty} \sum_{r=1}^{R} \mathbf{e}^{(n,r)} &= \begin{bmatrix} \boldsymbol{\mathcal{K}} & \mathbf{0} \\ \boldsymbol{\mathcal{L}} & \mathbf{V}_\gamma \boldsymbol{\Gamma}_1^{-1} \mathbf{V}_\gamma^{-1} \end{bmatrix} \mathbf{e}^{(n,1)} && (45) \\ &= \begin{bmatrix} \boldsymbol{\mathcal{K}} & \mathbf{0} \\ \boldsymbol{\mathcal{L}} & \boldsymbol{\mathcal{D}}^{-1} \end{bmatrix} \mathbf{e}^{(n,1)}. && (46) \end{aligned}$$

Notice that the eigenvalues γ_i depend on the stepsize matrix $\boldsymbol{\Delta}$.
From Eq. 21 and 46,

$$\begin{aligned} \lim_{R\to\infty} \mathbf{w}_p^{(n,R)} &\overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)*} \mathbf{F} \begin{bmatrix} \boldsymbol{\mathcal{K}} & \mathbf{0} \\ \boldsymbol{\mathcal{L}} & \boldsymbol{\mathcal{D}}^{-1} \end{bmatrix} \mathbf{e}^{(n,1)} && (47) \\ &\overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)*} \left(\mathbf{F} \begin{bmatrix} \boldsymbol{\mathcal{K}}^{-1} & \mathbf{0} \\ -\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{K}}^{-1} & \boldsymbol{\mathcal{D}} \end{bmatrix} \mathbf{F}^{-1} \right)^{-1} \mathbf{F}\mathbf{e}^{(n,1)} && (48) \end{aligned}$$

is found. Choose for instance $\boldsymbol{\mathcal{L}} = -\boldsymbol{\mathcal{D}}^{-1}\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{K}}$. Then,

$$\begin{aligned} \lim_{R\to\infty} \mathbf{w}_p^{(n,R)} &\overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)*} \left(\mathbf{F} \begin{bmatrix} \boldsymbol{\mathcal{K}}^{-1} & \mathbf{0} \\ \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{D}} \end{bmatrix} \mathbf{F}^{-1} \right)^{-1} \mathbf{F}\mathbf{e}^{(n,1)} && (49) \\ &\overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)*} \left(\mathbf{F} \begin{bmatrix} \boldsymbol{\mathcal{G}} & \boldsymbol{\mathcal{H}} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{A}} & \boldsymbol{\mathcal{B}} \\ \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{D}} \end{bmatrix} \mathbf{F}^{-1} \right)^{-1} \mathbf{F}\mathbf{e}^{(n,1)} && (50) \end{aligned}$$

with $\boldsymbol{\mathcal{G}} = \left((\boldsymbol{\mathcal{A}} - \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{D}}^{-1}\boldsymbol{\mathcal{C}})\boldsymbol{\mathcal{K}}\right)^{-1}$ and $\boldsymbol{\mathcal{H}} = -\left((\boldsymbol{\mathcal{A}} - \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{D}}^{-1}\boldsymbol{\mathcal{C}})\boldsymbol{\mathcal{K}}\right)^{-1} \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{D}}^{-1}$, provided $(\boldsymbol{\mathcal{A}} - \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{D}}^{-1}\boldsymbol{\mathcal{C}})\boldsymbol{\mathcal{K}}$ is invertible. Hence from Eq. 32,

$$\lim_{R\to\infty} \mathbf{w}_p^{(n,R)} \overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)*} \left(\boldsymbol{\mathcal{P}} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)*} \right)^{-1} \mathbf{F}\mathbf{e}^{(n,1)}, \tag{51}$$

with $\boldsymbol{\mathcal{P}}$ a projection matrix

$$\boldsymbol{\mathcal{P}} = \mathbf{F} \begin{bmatrix} \boldsymbol{\mathcal{G}} & \boldsymbol{\mathcal{H}} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1}. \tag{52}$$

Both $\boldsymbol{\Delta}$ and $\mathbf{X}_p^{(n)}$ are diagonal matrices. Therefore Eq. 51 reduces to

$$\lim_{R\to\infty} \mathbf{w}_p^{(n,R)} \overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G}\mathbf{X}_p^{(n)*} \left(\boldsymbol{\mathcal{P}} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G}\mathbf{X}_p^{(n)*} \right)^{-1} \mathbf{F}\mathbf{e}^{(n,1)}. \tag{53}$$

$\boldsymbol{\Delta}$ has disappeared from the weight updating equation. However in general, $\boldsymbol{\mathcal{P}}$ is still dependent on $\boldsymbol{\Delta}$.
Observe that equation 53 can also be written as

$$\lim_{R\to\infty} \mathbf{w}_p^{(n,R)} \stackrel{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G}\mathbf{X}_p^{(n)*} \left(\sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G}\mathbf{X}_p^{(n)*} \right)^{-1} \boldsymbol{\mathcal{P}}^{-1}\mathbf{F}\mathbf{e}^{(n,1)}, \tag{54}$$

with

$$\boldsymbol{\mathcal{P}}^{-1} = \mathbf{F} \begin{bmatrix} \boldsymbol{\mathcal{G}}^{-1} & -\boldsymbol{\mathcal{G}}^{-1}\boldsymbol{\mathcal{H}} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1}. \tag{55}$$

Choosing $\boldsymbol{\mathcal{K}} = \mathbf{0}$ and using Eq. 42,

$$\lim_{R\to\infty} \mathbf{w}_p^{(n,R)} \stackrel{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G}\mathbf{X}_p^{(n)*} \left(\sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G}\mathbf{X}_p^{(n)*} \right)^{-1} \mathbf{F} \begin{bmatrix} \mathbf{0} & \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{D}}^{-1} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{e}^{(n,1)} \tag{56}$$

$$\stackrel{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G}\mathbf{X}_p^{(n)*} \left(\sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G}\mathbf{X}_p^{(n)*} \right)^{-1} \mathbf{F} \begin{bmatrix} \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{D}}^{-1}\mathbf{e}_n \\ \mathbf{e}_n \end{bmatrix}. \tag{57}$$

Both Eq. 53 and 57 give a closed expression for $\lim_{R\to\infty} \mathbf{w}_p^{(n,R)}$ under the assumption that $\boldsymbol{\Gamma}_1$ is a full rank diagonal $L \times L$ matrix, i.e.

$$\gamma_i \neq 0, \qquad i = 0 : L-1, \tag{58}$$

that

$$\sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{G}\mathbf{X}_p^{(n)*} \tag{59}$$

is invertible and that all eigenvalues γ_i of $\boldsymbol{\Gamma}_1$ are such that (see Eq. 44)

$$0 \leqslant |1-\gamma_i| < 1, \qquad i = 0 : L-1. \tag{60}$$

From Eq. 53 and 57 it appears that under these conditions the PBFDRAP approaches a normalised version of the PBFDAF that is based on projected subband energies with maximum stepsize $\mu = 1$, if the number of iteration steps R goes to infinity. In section 5.3 and 5.4 it is verified whether and when these conditions are met.

5.3 Unconstrained PBFDRAP : $\lim_{R\to\infty} \mathbf{w}_p^{(n,R)}$

For the unconstrained PBFDRAP, $\mathbf{G} = \mathbf{I}_M$ and hence equation 32 becomes

$$\mathbf{P}_n = \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \boldsymbol{\Delta} \mathbf{X}_p^{(n)*} \mathbf{F} = \begin{bmatrix} \boldsymbol{\mathcal{A}} & \boldsymbol{\mathcal{B}} \\ \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{D}} \end{bmatrix}. \tag{61}$$

$\boldsymbol{\Delta} = \mathrm{diag}(\boldsymbol{\mu}_s^{(n)})$ contains the subband dependent stepsizes $\boldsymbol{\mu}_s^{(n)}(m)$ and is therefore real–valued. If the PBFDRAP is adapted, all $\boldsymbol{\mu}_s^{(n)}(m) > 0$. Further, $\mathbf{P}_n$ is a real–valued matrix.

Proof :

If $\mathbf{P}_n$ is multiplied with a random real vector $\mathbf{x}$,

$$\mathbf{P}_n\mathbf{x} = \mathbf{F}^{-1} \underbrace{\sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \boldsymbol{\Delta} \mathbf{X}_p^{(n)\,*}}_{\boldsymbol{\Upsilon}} \underbrace{\mathbf{F}\mathbf{x}}_{\boldsymbol{\chi}} \tag{62}$$

is obtained. As $\mathbf{x}$ is real, its frequency–domain representation $\boldsymbol{\chi}$ has the following symmetry property[2] :

$$\begin{cases} \boldsymbol{\chi}(0) = \boldsymbol{\chi}^*(0) \\ \boldsymbol{\chi}(m) = \boldsymbol{\chi}^*(M-m), \quad \text{for } 0 < m < M. \end{cases} \tag{63}$$

Both $\boldsymbol{\Delta}$ and $\mathbf{X}_p^{(n)}$ are diagonal matrices and as x, d and e (see figure 1) are assumed to be real, the m–th and $(M-m)$–th diagonal elements of $\boldsymbol{\Delta}$ and $\mathbf{X}_p^{(n)}$ are complex conjugated numbers. Therefore, $\boldsymbol{\Upsilon}$ is a real diagonal matrix and $\boldsymbol{\Upsilon}(m) = \boldsymbol{\Upsilon}(M-m)$.[3] If $\boldsymbol{\xi} = \boldsymbol{\Upsilon}\boldsymbol{\chi}$ is defined,

$$\begin{aligned} \boldsymbol{\xi}(m) &= \boldsymbol{\Upsilon}(m)\boldsymbol{\chi}(m) && (64) \\ &= \boldsymbol{\Upsilon}(M-m)\boldsymbol{\chi}^*(M-m) && (65) \\ &= \boldsymbol{\xi}^*(M-m) && (66) \end{aligned}$$

if $0 < m < M$ and $\boldsymbol{\xi}(0) = \boldsymbol{\xi}^*(0)$. Hence, $\mathbf{F}^{-1}\boldsymbol{\xi} = \mathbf{P}_n\mathbf{x}$ will be real–valued. As this holds for any random vector $\mathbf{x} \in \mathbb{R}^M$, the imaginary part of $\mathbf{P}_n$ must be zero.

Taking in account that $\mathbf{F}^H = M\mathbf{F}^{-1}$, it can be easily verified that $\mathbf{P}_n = \mathbf{P}_n^H = \mathbf{P}_n^T$, hence $\mathbf{P}_n$ is symmetric. As $\mathbf{F}\mathbf{P}_n\mathbf{F}^{-1}$ is diagonal, $\mathbf{P}_n$ is a right–circulant matrix and therefore also Toeplitz. The diagonal elements of $\mathbf{F}\mathbf{P}_n\mathbf{F}^{-1}$ correspond to the eigenvalues of the symmetric matrix $\mathbf{P}_n$ and they are all real and positive. If the data $\mathbf{X}_p^{(n)}$ are randomly chosen and all stepsizes are different from zero, the diagonal elements of $\mathbf{F}\mathbf{P}_n\mathbf{F}^{-1}$ are strictly positive. Hence, $\mathbf{P}_n$ is a full rank positive definite matrix, i.e.

$$\forall \mathbf{x} \in \mathbb{R}^M \setminus \{\mathbf{0}\}, \qquad \mathbf{x}^T\mathbf{P}_n\mathbf{x} > 0. \tag{67}$$

As $\mathbf{P}_n$ has full rank, the rank of $\mathbf{N}_n$ (Eq. 31) will be L. As a consequence, $\boldsymbol{\Gamma}_1$ will be a full rank diagonal matrix (first condition, Eq. 58). Also the second condition (Eq. 59) is fulfilled as it is assumed that the data $\mathbf{X}_p^{(n)}$ are randomly chosen.

Now, by inserting Eq. 61 in Eq. 67 and making a special choice for $\mathbf{x}$:

$$\forall \mathbf{x}_2 \in \mathbb{R}^L \setminus \{\mathbf{0}\}, \qquad \begin{bmatrix} \mathbf{0} & \mathbf{x}_2^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{A}} & \boldsymbol{\mathcal{B}} \\ \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{D}} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{x}_2 \end{bmatrix} > 0 \tag{68}$$

[2] $\boldsymbol{\chi}(m)$ is the m–th element of vector $\boldsymbol{\chi}$.

[3] $\boldsymbol{\Upsilon}(m)$ is the m–th diagonal element of matrix $\boldsymbol{\Upsilon}$.

is obtained which is equivalent to

$$\forall \mathbf{x}_2 \in \mathbb{R}^L \setminus \{\mathbf{0}\}, \qquad \mathbf{x}_2^T \boldsymbol{\mathcal{D}} \mathbf{x}_2 > 0. \tag{69}$$

From Eq. 69 it can be concluded that also $\boldsymbol{\mathcal{D}}$ is positive definite. Furthermore, as $\mathbf{P}_n$ is a real symmetric matrix, also $\boldsymbol{\mathcal{D}}$ is symmetric and real–valued. Hence, the eigenvalues of $\boldsymbol{\mathcal{D}}$, i.e. the diagonal elements of $\boldsymbol{\Gamma}_1$ (see Eq. 35), are real and strictly positive. If the stepsizes $\boldsymbol{\mu}_s^{(n)}(m)$ and hence $||\boldsymbol{\Delta}||$ are sufficiently small all eigenvalues of $\boldsymbol{\Gamma}_1$ are between 0 and 2. Then also the third condition (Eq. 60) is fulfilled and $\mathbf{w}_p^{(n,R)}$ converges to equation 53 if R goes to infinity.

5.3.1 Unnormalised and globally normalised unconstrained PBFDRAP

For the unnormalised PBFDRAP, $\boldsymbol{\Delta}$ is defined as $\mu \mathbf{I}_M$, with μ a constant. More generally, $\boldsymbol{\Delta} = \mu_n \mathbf{I}_M$, with μ_n a block dependent stepsize. For instance, if[4]

$$\mu_n = \frac{\mu}{\frac{1}{M}\sum_{m=0}^{M-1}\sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)}(m)\mathbf{X}_p^{(n)*}(m)} = \frac{\mu}{\sum_{m=0}^{M-1}\sum_{p=0}^{\frac{N}{P}-1} x[(n+1)L - pP - m]^2}, \tag{70}$$

the algorithm performs a global normalisation, leading to better convergence if signals with a high dynamic range (such as speech) are involved.
As $\boldsymbol{\Delta} = \mu_n \mathbf{I}_M$, the elements of $\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}, \boldsymbol{\mathcal{C}}, \boldsymbol{\mathcal{D}}$ and $\mathbf{N}_n$ linearly depend on μ_n. $\boldsymbol{\mathcal{P}}$ will be independent of μ_n by choosing $\boldsymbol{\mathcal{K}}$ inversely proportional to μ_n. Hence, Eq. 53

$$\lim_{R\to\infty} \mathbf{w}_p^{(n,R)} \stackrel{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{X}_p^{(n)*} \left(\boldsymbol{\mathcal{P}} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{X}_p^{(n)*} \right)^{-1} \mathbf{F}\mathbf{e}^{(n,1)} \tag{71}$$

becomes independent of μ_n with an appropriate choice for $\boldsymbol{\mathcal{K}}$. From Eq. 71 it can be concluded that the unnormalised and globally normalised unconstrained PBFDRAP approach a normalised version of the unconstrained PBFDAF that is based on projected subband energies with maximum stepsize $\mu_n = 1$, if the number of iteration steps R goes to infinity. This result is similar to the conclusions formulated in [2] and [6], which suggests that in general iterated unnormalised algorithms converge to their normalised counterparts.

5.3.2 Subband–normalised unconstrained PBFDRAP

The PBFDRAP may be considered as a subband adaptive system [3]. The convergence properties can be optimised by applying a different stepsize to each of the subband adaptive filters. For the subband–normalised PBFDRAP, matrix $\boldsymbol{\Delta}$ is typically chosen as follows :

$$\boldsymbol{\Delta} = \mu \left(\sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{X}_p^{(n)*} \right)^{-1}. \tag{72}$$

[4] $\mathbf{X}_p^{(n)}(m)$ is the m–th diagonal element of $\mathbf{X}_p^{(n)}$.

In this way in each subband the adaptation stepsize μ is divided by the subband energy. Also in this case the elements of $\mathbf{N}_n$ linearly depend on μ, and by taking $\boldsymbol{\mathcal{K}}$ inversely proportional to μ, projection matrix $\boldsymbol{\mathcal{P}}$ will be independent of μ.
By inserting Eq. 72 in Eq. 20,

$$\mathbf{e}^{(n,r)} = \begin{bmatrix} \mathbf{I}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & (1-\mu)^{r-1}\mathbf{I}_L \end{bmatrix} \mathbf{e}^{(n,1)} \tag{73}$$

is obtained. Therefore, if $\mu \neq 0$

$$\sum_{r=1}^{R} \mathbf{e}^{(n,r)} = \begin{bmatrix} R\mathbf{I}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \frac{1-(1-\mu)^R}{\mu}\mathbf{I}_L \end{bmatrix} \mathbf{e}^{(n,1)}. \tag{74}$$

If $0 < \mu < 2$ equation 74 converges to

$$\lim_{R\to\infty} \sum_{r=1}^{R} \mathbf{e}^{(n,r)} = \begin{bmatrix} \times & \mathbf{0} \\ \mathbf{0} & \frac{1}{\mu}\mathbf{I}_L \end{bmatrix} \mathbf{e}^{(n,1)}, \tag{75}$$

in which $\times$ can be any matrix having the correct size, due to the special structure of $\mathbf{e}^{(n,1)}$ (see also Eq. 42). If $\times$ is chosen equal to $\frac{1}{\mu}\mathbf{I}_{P-1+s}$,

$$\lim_{R\to\infty} \sum_{r=1}^{R} \mathbf{e}^{(n,r)} = \frac{1}{\mu}\mathbf{e}^{(n,1)} \tag{76}$$

which implies that $\boldsymbol{\mathcal{K}} = \frac{1}{\mu}\mathbf{I}_{P-1+s}$ and $\boldsymbol{\mathcal{P}} = \mathbf{I}_M$, so that finally by inserting Eq. 72 and 76 in Eq. 21

$$\lim_{R\to\infty} \mathbf{w}_p^{(n,R)} \overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \left(\sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{X}_p^{(n)^*} \right)^{-1} \mathbf{X}_p^{(n)^*} \mathbf{F} \mathbf{e}^{(n,1)} \tag{77}$$

is obtained.
It appears that the subband–normalised unconstrained PBFDRAP converges to the subband–normalised unconstrained PBFDAF with $\mu = 1$: increasing the number of iteration steps R has the same effect as applying a larger stepsize μ.

5.4 Constrained PBFDRAP : $\lim_{R\to\infty} \mathbf{w}_p^{(n,R)}$

If it is assumed that the data (x, d) are randomly chosen and that all stepsizes are different from zero, the rank of $\mathbf{X}_p^{(n)}\mathbf{G}\boldsymbol{\Delta}\mathbf{X}_p^{(n)^*}$ is P, due to the structure of $\mathbf{G}$ (see Eq. 19). The rank of $\mathbf{P}_n$ (Eq. 32) on the other hand will be $\min(M, P\frac{N}{P}) = \min(M, N)$ under the same assumptions. The rank of $\mathbf{N}_n$ is given by $\min(\min(M, N), L)$.
It can be verified that

$$\mathbf{P}_n = \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1} \boldsymbol{\Delta} \mathbf{X}_p^{(n)^*} \mathbf{F} = \begin{bmatrix} \boldsymbol{\mathcal{A}} & \boldsymbol{\mathcal{B}} \\ \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{D}} \end{bmatrix} \tag{78}$$

is real–valued :
Proof :
If $\mathbf{P}_n$ is multiplied with a random real vector $\mathbf{x}$,

$$\mathbf{P}_n\mathbf{x} = \sum_{p=0}^{\frac{N}{P}-1} \mathbf{F}^{-1}\mathbf{X}_p^{(n)}\mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1} \underbrace{\mathbf{\Delta}\mathbf{X}_p^{(n)*}}_{\boldsymbol{\Upsilon}} \underbrace{\mathbf{F}\mathbf{x}}_{\boldsymbol{\chi}} \tag{79}$$

is obtained. Again $\boldsymbol{\chi}(m) = \boldsymbol{\chi}^*(M-m)$ and $\boldsymbol{\Upsilon}(m) = \boldsymbol{\Upsilon}^*(M-m)$ for $0 < m < M$ and $\boldsymbol{\chi}(0) = \boldsymbol{\chi}^*(0)$ and $\boldsymbol{\Upsilon}(0) = \boldsymbol{\Upsilon}^*(0)$ as all signals entering the adaptive filter are assumed to be real. If $\boldsymbol{\xi} = \boldsymbol{\Upsilon}\boldsymbol{\chi}$ is defined, $\boldsymbol{\xi}(m) = \boldsymbol{\Upsilon}(m)\boldsymbol{\chi}(m) = \boldsymbol{\xi}^*(M-m)$, for $0 < m < M$ and $\boldsymbol{\xi}(0) = \boldsymbol{\xi}^*(0)$. Hence, $\mathbf{F}^{-1}\boldsymbol{\xi}$ and also

$$\boldsymbol{\eta} = \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1}\boldsymbol{\xi} \tag{80}$$

will be real–valued. Using the same arguments, it is found that also $\mathbf{F}^{-1}\mathbf{X}_p^{(n)}\mathbf{F}\boldsymbol{\eta}$ and hence $\mathbf{P}_n\mathbf{x}$ are real. As this holds for any random vector $\mathbf{x} \in \mathbb{R}^M$, the imaginary part of $\mathbf{P}_n$ must be zero.

5.4.1 Unnormalised and globally normalised constrained PBFDRAP

For the unnormalised or globally normalised constrained PBFDRAP for which $\mathbf{\Delta} = \mu_n\mathbf{I}_M$, it can be verified that

$$\mathbf{P}_n = \mu_n\mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)}\mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1}\mathbf{X}_p^{(n)*}\mathbf{F} = \begin{bmatrix} \boldsymbol{\mathcal{A}} & \boldsymbol{\mathcal{B}} \\ \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{D}} \end{bmatrix} \tag{81}$$

is symmetric and positive semi–definite as $\forall \mathbf{x} \in \mathbb{R}^M$,

$$\mathbf{x}^T\mathbf{P}_n\mathbf{x} = \mu_n \sum_{p=0}^{\frac{N}{P}-1} \mathbf{x}^H\mathbf{F}^{-1}\mathbf{X}_p^{(n)}\mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1}\mathbf{X}_p^{(n)*}\mathbf{F}\mathbf{x} \tag{82}$$

$$\mathbf{x}^T\mathbf{P}_n\mathbf{x} = \mu_n \sum_{p=0}^{\frac{N}{P}-1} \left(\begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{F}^{-1}\mathbf{X}_p^{(n)*}\mathbf{F}\mathbf{x} \right)^H \left(\begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{F}^{-1}\mathbf{X}_p^{(n)*}\mathbf{F}\mathbf{x} \right) \tag{83}$$

$$\mathbf{x}^T\mathbf{P}_n\mathbf{x} \geqslant 0. \tag{84}$$

As $\mathbf{P}_n$ is symmetric and positive semi–definite all eigenvalues are real and positive. Now, for a special choice of $\mathbf{x}$,

$$\forall \mathbf{x}_2 \in \mathbb{R}^L, \qquad \begin{bmatrix} \mathbf{0} & \mathbf{x}_2^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{A}} & \boldsymbol{\mathcal{B}} \\ \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{D}} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{x}_2 \end{bmatrix} \geqslant 0, \tag{85}$$

which is equivalent to

$$\forall \mathbf{x}_2 \in \mathbb{R}^L, \qquad \mathbf{x}_2^T\boldsymbol{\mathcal{D}}\mathbf{x}_2 \geqslant 0. \tag{86}$$

From Eq. 86 it can be concluded that $\boldsymbol{\mathcal{D}}$ is positive semi–definite. Furthermore, $\boldsymbol{\mathcal{D}}$ is symmetric and real–valued, just as $\mathbf{P}_n$. Hence, the eigenvalues of $\boldsymbol{\mathcal{D}}$, i.e. the diagonal elements of $\boldsymbol{\Gamma}_1$ are real and positive.

N* ⩾ *L

For the type of applications the PBFDRAP is typically used for, e.g. for the identification of high–order FIR systems, $N \geqslant L$ is a realistic assumption. Hence, the rank of $\mathbf{N}_n$ is $\min(\min(M,N),L) = L$. $\boldsymbol{\Gamma}_1$ will be a full rank diagonal matrix and $\boldsymbol{\mathcal{D}}$ is strictly positive definite. If the stepsizes μ_n and hence $||\boldsymbol{\Delta}||$ are sufficiently small all eigenvalues of $\boldsymbol{\Gamma}_1$ are between 0 and 2 and $\boldsymbol{\Xi}^{(R)}$ converges to $\boldsymbol{\Gamma}_1^{-1}$ if R goes to infinity. As conditions 58 and 60 are fulfilled for the unnormalised and globally normalised constrained PBFDRAP with $N \geqslant L$, equations 33–49 are applicable and the algorithm converges to equation 47 if R goes to infinity.

N* ⩾ *M

If furthermore $N \geqslant M \geqslant L$, which is a more restrictive, but still realistic assumption if the PBFDRAP is used for the identification of high–order FIR systems, $\mathbf{P}_n$ has rank M, hence Eq. 59 is invertible and also equations 50–57 are valid. All conditions (Eq. 58, 59 and 60) are fulfilled, hence the algorithm converges to equation 53 if R goes to infinity.

N* < *L

If $N < L$, the algorithm still converges, but the formulas derived in Eq. 34–36 and 40–57 aren't applicable anymore. Remark that if $N < L$ the rank of $\mathbf{N}_n$ is $\min(\min(M,N),L) = N$. $\boldsymbol{\Gamma}_1$ is not of full rank and can be expressed as

$$\boldsymbol{\Gamma}_1 = \begin{bmatrix} \boldsymbol{\Gamma}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-N} \end{bmatrix}, \tag{87}$$

with $\boldsymbol{\Gamma}_{11}$ a full rank $N \times N$ diagonal matrix. Based on Eq. 14–16, taking in account that $d = w \star x$ and partitioning both the unknown system w as its model $\hat{w}^{(n)}$ according to Eq. 7, $\mathbf{e}^{(n,1)}$ can be written as

$$\mathbf{e}^{(n,1)} = \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \tilde{\mathbf{w}}_p^{(n,1)}, \tag{88}$$

with $\tilde{\mathbf{w}}_p^{(n,1)}$ $M \times 1$ vectors of the form

$$\tilde{\mathbf{w}}_p^{(n,1)} \overset{\forall p}{=} \mathbf{F} \begin{bmatrix} \tilde{\mathbf{w}}^{(n,p)} \\ \mathbf{0} \end{bmatrix} \begin{matrix} \updownarrow P \\ \updownarrow L-1+s \end{matrix} \tag{89}$$

and the j–th element of $\tilde{\mathbf{w}}^{(n,p)}$ such that

$$\tilde{\mathbf{w}}^{(n,p)}(j) = w[j+pP] - \hat{w}^{(n)}[j+pP], \qquad p = 0 : \frac{N}{P} - 1. \tag{90}$$

Hence, equation 88 can also be written as

$$\mathbf{e}^{(n,1)} = \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{w}}^{(n,p)} \\ \mathbf{0} \end{bmatrix}. \tag{91}$$

It will be shown that a real $M \times 1$ vector $\mathbf{v}_n$ can be found, such that

$$\tilde{\mathbf{w}}^{(n,p)} \stackrel{\forall p}{=} \left[\begin{array}{cc} \mathbf{I}_P & \mathbf{0} \end{array} \right] \mathbf{F}^{-1} \mathbf{X}_p^{(n)^*} \mathbf{F} \mathbf{v}_n, \tag{92}$$

and Eq. 91 becomes (see also Eq. 31)

$$\mathbf{e}^{(n,1)} = \left[\begin{array}{cc} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{array} \right] \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{F} \left[\begin{array}{cc} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{array} \right] \mathbf{F}^{-1} \mathbf{X}_p^{(n)^*} \mathbf{F} \mathbf{v}_n \tag{93}$$

$$= \frac{1}{\mu_n} \mathbf{N}_n \mathbf{v}_n. \tag{94}$$

Proof :
Taking in account Eq. 6, equation 92 can be rewritten as

$$\underbrace{\left[\begin{array}{c} \tilde{\mathbf{w}}^{(n,0)} \\ \vdots \\ \tilde{\mathbf{w}}^{(n,\frac{N}{P}-1)} \end{array} \right]}_{\tilde{\mathbf{w}}_n} = \underbrace{\left[\begin{array}{ccccc} \mathbf{I}_P & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_P & \mathbf{0} \end{array} \right]}_{\mathbf{L}} \underbrace{\left[\begin{array}{c} \mathbf{F}^{-1} \mathbf{X}_0^{(n)^*} \mathbf{F} \\ \vdots \\ \mathbf{F}^{-1} \mathbf{X}_{\frac{N}{P}-1}^{(n)^*} \mathbf{F} \end{array} \right]}_{\mathbf{R}_n} \mathbf{v}_n. \tag{95}$$

For randomly chosen data the $\frac{N}{P}M \times M$ matrix $\mathbf{R}_n$ has full rank M as N is always equal to or greater than P for the PBFDRAP, whereas the rank of the $N \times \frac{N}{P}M$ matrix $\mathbf{L}$ is $P\frac{N}{P} = N$. Hence, the $N \times M$ matrix $\mathbf{L}\mathbf{R}_n$ has full rank $\min(N, M) = N$. For every $\tilde{\mathbf{w}}_n$ an $M \times 1$ vector $\mathbf{v}_n$ can be found as the product of the pseudo–inverse of $\mathbf{L}\mathbf{R}_n$ and vector $\tilde{\mathbf{w}}_n$:

$$\mathbf{v}_n = (\mathbf{L}\mathbf{R}_n)^\dagger \, \tilde{\mathbf{w}}_n, \tag{96}$$

which completes the proof.

Now, by inserting Eq. 33, 87 and 94 in equation 38,

$$\sum_{r=1}^{R} \mathbf{e}^{(n,r)} = \frac{1}{\mu_n} \mathbf{V} \sum_{r=1}^{R} \left(\mathbf{I}_M - \mathbf{\Gamma} \right)^{r-1} \mathbf{V}^{-1} \mathbf{N}_n \mathbf{v}_n \tag{97}$$

$$= \frac{1}{\mu_n} \mathbf{V} \sum_{r=1}^{R} \left[\begin{array}{ccc} \mathbf{I}_N - \mathbf{\Gamma}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{L-N} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{P-1+s} \end{array} \right]^{r-1} \left[\begin{array}{ccc} \mathbf{\Gamma}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-N} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}_{P-1+s} \end{array} \right] \mathbf{V}^{-1} \mathbf{v}_n \tag{98}$$

$$= \frac{1}{\mu_n} \mathbf{V} \left[\begin{array}{ccc} \mathbf{\Gamma}_{11}^{-1} \left(\mathbf{I}_N - \left(\mathbf{I}_N - \mathbf{\Gamma}_{11} \right)^R \right) \mathbf{\Gamma}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-N} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}_{P-1+s} \end{array} \right] \mathbf{V}^{-1} \mathbf{v}_n \tag{99}$$

is obtained. As $\boldsymbol{\mathcal{D}}$ is positive semi–definite and symmetric (Eq. 86), all elements of $\mathbf{\Gamma}_{11}$ are real and strictly positive. Hence, if μ_n is sufficiently small

$$\lim_{R \to \infty} \sum_{r=1}^{R} \mathbf{e}^{(n,r)} = \frac{1}{\mu_n} \mathbf{V} \left[\begin{array}{ccc} \mathbf{I}_N & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-N} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}_{P-1+s} \end{array} \right] \mathbf{V}^{-1} \mathbf{v}_n \tag{100}$$

and the filter weights (Eq. 21) converge to

$$\lim_{R\to\infty} \mathbf{w}_p^{(n,R)} \overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mathbf{G}\mathbf{X}_p^{(n)^*}\mathbf{F}\mathbf{V}\begin{bmatrix} \mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{M-N} \end{bmatrix}\mathbf{V}^{-1}\mathbf{v}_n. \tag{101}$$

Relation with the Affine Projection Algorithm

It can be proven that the weight updating of the unnormalised and globally normalised constrained PBFDRAP for which $N \geqslant L$ (i.e. Eq. 47) approaches the weight updating of the Partial Rank Algorithm (PRA) with maximum stepsize, if R goes to infinity. The PRA is a block version of the Affine Projection Algorithm (APA) [5] and can be interpreted as a block–normalised version of the block–LMS algorithm [2]. For each block of L samples the PRA performs the following steps :

$$\mathbf{X}_n = \begin{bmatrix} x[nL+1] & \dots & x[(n+1)L] \\ \vdots & \ddots & \vdots \\ x[nL-N+2] & \dots & x[(n+1)L-N+1] \end{bmatrix} \tag{102}$$

$$\mathbf{e}_n = \mathbf{d}_n - \mathbf{X}_n^T\hat{\mathbf{w}}_n \tag{103}$$

$$\hat{\mathbf{w}}_{n+1} = \hat{\mathbf{w}}_n + \mu_n\mathbf{X}_n\left(\mathbf{X}_n^T\mathbf{X}_n\right)^{-1}\mathbf{e}_n. \tag{104}$$

$\mathbf{X}_n$ is a (real) $N \times L$ Toeplitz matrix containing the x–samples, $\mathbf{d}_n$ is defined in Eq. 10. The $N \times 1$ vector $\hat{\mathbf{w}}_n$ contains the adaptive filter weights and models the unknown system. Also here, μ_n is a block dependent stepsize. If $L = 1$ the PRA reduces to the NLMS algorithm. Hence, the PRA is scalable and lies in between the NLMS algorithm ($L = 1$) and a block adaptive least squares estimate of the unknown system ($L = N$). In [2] it was proven that iterated block–LMS approaches PRA with stepsize $\mu_n = 1$, if the number of iteration steps goes to infinity. As the unnormalised constrained PBFDAF is an exact representation of block–LMS in the frequency domain, the unnormalised and also the globally normalised constrained PBFDRAP will approach the PRA, if R goes to infinity. This will be proven now, based on the equations derived in section 5.2.

Proof :

The updating equation of the constrained PBFDRAP ($N \geqslant L$) for $R \to \infty$ is given by equation 47. Equation 53 is just a special case hereof if furthermore $N \geqslant M$. If $\mathbf{\Delta} = \mu_n\mathbf{I}_M$, equation 47 can be written as

$$\lim_{R\to\infty} \mathbf{w}_p^{(n,R)} \overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mu_n\mathbf{F}\begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix}\mathbf{F}^{-1}\mathbf{X}_p^{(n)^*}\mathbf{F}\begin{bmatrix} \boldsymbol{\mathcal{K}} & \mathbf{0} \\ \boldsymbol{\mathcal{L}} & \boldsymbol{\mathcal{D}}^{-1} \end{bmatrix}\mathbf{e}^{(n,1)}. \tag{105}$$

Inserting Eq. 7 and 42,

$$\lim_{R\to\infty} \mathbf{w}_p^{(n,R)} \overset{\forall p}{=} \mathbf{F}\left(\begin{bmatrix} \mathbf{w}^{(n,p)} \\ \mathbf{0} \end{bmatrix} + \mu_n\begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix}\mathbf{F}^{-1}\mathbf{X}_p^{(n)^*}\mathbf{F}\begin{bmatrix} \mathbf{0} \\ \boldsymbol{\mathcal{D}}^{-1}\mathbf{e}_n \end{bmatrix}\right). \tag{106}$$

As $\mathbf{X}_p^{(n)}$ is a diagonal matrix, defined by Eq. 8, $\bar{\mathbf{X}}_{n,p} = \mathbf{F}^{-1}\mathbf{X}_p^{(n)^*}\mathbf{F}$ is a right–circulant $M \times M$ matrix. Hence, the discrete Fourier transform of the first column of $\bar{\mathbf{X}}_{n,p}$ are

the diagonal elements of $\mathbf{X}_p^{(n)^*}$. Taking in account Eq. 8,

$$\bar{\mathbf{X}}_{n,p} = \mathbf{F}^{-1}\mathbf{X}_p^{(n)^*}\mathbf{F} = \begin{bmatrix} x[(n+1)L-pP-M+1] & \dots & x[(n+1)L-pP] \\ x[(n+1)L-pP] & \dots & x[(n+1)L-pP-1] \\ \vdots & \ddots & \vdots \\ x[(n+1)L-pP-M+2] & \dots & x[(n+1)L-pP-M+1] \end{bmatrix}. \tag{107}$$

Plugging equation 107 into Eq. 106 and defining

$$\lim_{R\to\infty} \mathbf{w}_p^{(n,R)} \stackrel{\forall p}{=} \mathbf{F} \begin{bmatrix} \mathbf{w}^{(n+1,p)} \\ \mathbf{0} \end{bmatrix}, \tag{108}$$

reduces equation 106 to

$$\mathbf{w}^{(n+1,p)} \stackrel{\forall p}{=} \mathbf{w}^{(n,p)} + \mu_n \mathbf{X}_{n,p}\boldsymbol{\mathcal{D}}^{-1}\mathbf{e}_n \tag{109}$$

with $\mathbf{X}_{n,p}$ the following $P \times L$ Toeplitz matrix :

$$\mathbf{X}_{n,p} = \begin{bmatrix} x[nL-pP+1] & \dots & x[(n+1)L-pP] \\ \vdots & \ddots & \vdots \\ x[nL-pP-P+2] & \dots & x[(n+1)L-pP-P+1] \end{bmatrix}. \tag{110}$$

The rows of $\mathbf{X}_{n,p}$ are delayed versions of each other. This is true only if the time index of the left bottom element of $\mathbf{X}_{n,p}$ is larger than or equal to the time index of the left top element of $\bar{\mathbf{X}}_{n,p}$. This implies that

$$\begin{aligned} (n+1)L-pP-M+1 &\leqslant nL-pP-P+2 && (111) \\ &\Leftrightarrow \\ M &\geqslant P+L-1, && (112) \end{aligned}$$

which is the well–known condition that must be true for any PBFDAF algorithm (see section 3). From Eq. 109,

$$\begin{aligned} \begin{bmatrix} \mathbf{w}^{(n+1,0)} \\ \vdots \\ \mathbf{w}^{(n+1,\frac{N}{P}-1)} \end{bmatrix} &= \begin{bmatrix} \mathbf{w}^{(n,0)} \\ \vdots \\ \mathbf{w}^{(n,\frac{N}{P}-1)} \end{bmatrix} + \mu_n \begin{bmatrix} \mathbf{X}_{n,0} \\ \vdots \\ \mathbf{X}_{n,\frac{N}{P}-1} \end{bmatrix} \boldsymbol{\mathcal{D}}^{-1}\mathbf{e}_n && (113) \\ \mathbf{w}_{n+1} &= \mathbf{w}_n + \mu_n \mathbf{X}_n \boldsymbol{\mathcal{D}}^{-1}\mathbf{e}_n, && (114) \end{aligned}$$

in which $\mathbf{X}_n$ is as defined in Eq. 102.
$\boldsymbol{\mathcal{D}}$ is defined by equation 32 and can be written as

$$\boldsymbol{\mathcal{D}} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{P}_n \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix}. \tag{115}$$

This equation can be further simplified using Eq. 20, 102, 107 and 110 :

$$\begin{aligned}
\boldsymbol{\mathcal{D}} &= \mu_n \begin{bmatrix} \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{F}^{-1} \mathbf{X}_p^{(n)*} \mathbf{F} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} && (116) \\
&= \mu_n \sum_{p=0}^{\frac{N}{P}-1} \begin{bmatrix} \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \mathbf{X}_p^{(n)} \mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{F}^{-1} \mathbf{X}_p^{(n)*} \mathbf{F} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} && (117) \\
&= \mu_n \sum_{p=0}^{\frac{N}{P}-1} \left(\begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \bar{\mathbf{X}}_{n,p} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} \right)^H \left(\begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \bar{\mathbf{X}}_{n,p} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} \right) && (118) \\
&= \mu_n \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_{n,p}^T \mathbf{X}_{n,p} && (119) \\
&= \mu_n \mathbf{X}_n^T \mathbf{X}_n. && (120)
\end{aligned}$$

As $N \geqslant L$ and the data are randomly chosen, $\mathbf{X}_n^T \mathbf{X}_n$ is invertible and equation 114 becomes

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mathbf{X}_n \left(\mathbf{X}_n^T \mathbf{X}_n \right)^{-1} \mathbf{e}_n. \tag{121}$$

From Eq. 7, 14,15, 16, 42 and 107

$$\begin{aligned}
\mathbf{e}_n &= \mathbf{d}_n - \begin{bmatrix} \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{w}_p^{(n,1)} && (122) \\
&= \mathbf{d}_n - \begin{bmatrix} \mathbf{0} & \mathbf{I}_L \end{bmatrix} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{F}^{-1} \mathbf{X}_p^{(n)} \mathbf{F} \begin{bmatrix} \mathbf{w}^{(n,p)} \\ \mathbf{0} \end{bmatrix} && (123) \\
&= \mathbf{d}_n - \sum_{p=0}^{\frac{N}{P}-1} \begin{bmatrix} \mathbf{0} & \mathbf{I}_L \end{bmatrix} \bar{\mathbf{X}}_{n,p}^T \begin{bmatrix} \mathbf{w}^{(n,p)} \\ \mathbf{0} \end{bmatrix} && (124) \\
&= \mathbf{d}_n - \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_{n,p}^T \mathbf{w}^{(n,p)} && (125)
\end{aligned}$$

and finally using Eq. 102 and 114

$$\mathbf{e}_n = \mathbf{d}_n - \mathbf{X}_n^T \mathbf{w}_n \tag{126}$$

is obtained. By comparing Eq. 103 with 126 and Eq. 104 with 121, it is easily verified that if $\hat{\mathbf{w}}_0 = \mathbf{w}_0$, also $\hat{\mathbf{w}}_n = \mathbf{w}_n, \forall n$ and $\mu_n = 1$. Hence, the unnormalised and globally normalised constrained PBFDRAP converge to the PRA if $N \geqslant L$ and the number of iteration steps goes to infinity.
The same sort of computation can be done to check what happens with the unnormalised constrained PBFDRAP for finite R. Plugging in Eq. 20, which holds for all

$R > 0$, into Eq. 21 leads to

$$\mathbf{w}_p^{(n,R+1)} \overset{\forall p}{=} \mathbf{w}_p^{(n,1)} + \mu_n \mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1} \mathbf{X}_p^{(n)^*} \mathbf{F} \sum_{r=1}^{R} \mathbf{B}_n^{r-1} \mathbf{e}^{(n,1)} \tag{127}$$

in which

$$\mathbf{B}_n = \mathbf{I}_M - \mu_n \begin{bmatrix} \mathbf{0}_{P-1+s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1} \mathbf{X}_p^{(n)^*} \mathbf{F}. \tag{128}$$

Inserting Eq. 7 and 42 into Eq. 127 leads to

$$\mathbf{w}_p^{(n,R+1)} \overset{\forall p}{=} \mathbf{F} \left(\begin{bmatrix} \mathbf{w}^{(n,p)} \\ \mathbf{0} \end{bmatrix} + \mu_n \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{L-1+s} \end{bmatrix} \mathbf{F}^{-1} \mathbf{X}_p^{(n)^*} \mathbf{F} \sum_{r=1}^{R} \mathbf{B}_n^{r-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} \mathbf{e}_n \right). \tag{129}$$

Defining

$$\mathbf{w}_p^{(n,R+1)} \overset{\forall p}{=} \mathbf{F} \begin{bmatrix} \mathbf{w}^{(n+1,p)} \\ \mathbf{0} \end{bmatrix} \tag{130}$$

and plugging in equation 107 into Eq. 129, reduces equation 129 to

$$\mathbf{w}^{(n+1,p)} \overset{\forall p}{=} \mathbf{w}^{(n,p)} + \mu_n \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \end{bmatrix} \bar{\mathbf{X}}_{n,p} \sum_{r=1}^{R} \mathbf{B}_n^{r-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} \mathbf{e}_n. \tag{131}$$

Now, from Eq. 128

$$\mathbf{B}_n \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} = \left(\mathbf{I}_M - \mu_n \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{X}_p^{(n)} \mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{F}^{-1} \mathbf{X}_p^{(n)^*} \mathbf{F} \right) \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} \tag{132}$$

$$= \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} - \mu_n \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \sum_{p=0}^{\frac{N}{P}-1} \mathbf{F}^{-1} \mathbf{X}_p^{(n)} \mathbf{F} \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{F}^{-1} \mathbf{X}_p^{(n)^*} \mathbf{F} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} \tag{133}$$

$$= \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L - \boldsymbol{\mathcal{D}} \end{bmatrix} \tag{134}$$

with $\boldsymbol{\mathcal{D}}$ defined in Eq. 116. From Eq. 134 it appears that

$$\mathbf{B}_n^r \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} = \mathbf{B}_n^{r-1} \mathbf{B}_n \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} \tag{135}$$

$$= \mathbf{B}_n^{r-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L - \boldsymbol{\mathcal{D}} \end{bmatrix} \tag{136}$$

$$= \mathbf{B}_n^{r-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} (\mathbf{I}_L - \boldsymbol{\mathcal{D}}) \tag{137}$$

$$= \mathbf{B}_n^{r-2} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L - \boldsymbol{\mathcal{D}} \end{bmatrix} (\mathbf{I}_L - \boldsymbol{\mathcal{D}}). \tag{138}$$

By recursion

$$\mathbf{B}_n^r \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} (\mathbf{I}_L - \boldsymbol{\mathcal{D}})^r \tag{139}$$

is found. Using Eq. 120 and 139, Eq. 131 becomes

$$\mathbf{w}^{(n+1,p)} \overset{\forall p}{=} \mathbf{w}^{(n,p)} + \mu_n \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \end{bmatrix} \bar{\mathbf{X}}_{n,p} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_L \end{bmatrix} \sum_{r=1}^{R} \left(\mathbf{I}_L - \mu_n \mathbf{X}_n^T \mathbf{X}_n \right)^{r-1} \mathbf{e}_n \tag{140}$$

$$\overset{\forall p}{=} \mathbf{w}^{(n,p)} + \mu_n \mathbf{X}_{n,p} \sum_{r=1}^{R} \left(\mathbf{I}_L - \mu_n \mathbf{X}_n^T \mathbf{X}_n \right)^{r-1} \mathbf{e}_n \tag{141}$$

in which $\mathbf{X}_{n,p}$ is defined in Eq. 110. If $N \geqslant L$, Eq. 141 becomes

$$\begin{bmatrix} \mathbf{w}^{(n+1,0)} \\ \vdots \\ \mathbf{w}^{(n+1,\frac{N}{P}-1)} \end{bmatrix} = \begin{bmatrix} \mathbf{w}^{(n,0)} \\ \vdots \\ \mathbf{w}^{(n,\frac{N}{P}-1)} \end{bmatrix} + \mu_n \begin{bmatrix} \mathbf{X}_{n,0} \\ \vdots \\ \mathbf{X}_{n,\frac{N}{P}-1} \end{bmatrix} \sum_{r=1}^{R} \left(\mathbf{I}_L - \mu_n \mathbf{X}_n^T \mathbf{X}_n \right)^{r-1} \mathbf{e}_n \tag{142}$$

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu_n \mathbf{X}_n \sum_{r=1}^{R} \left(\mathbf{I}_L - \mu_n \mathbf{X}_n^T \mathbf{X}_n \right)^{r-1} \mathbf{e}_n \tag{143}$$

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu_n \mathbf{X}_n \left(\mu_n \mathbf{X}_n^T \mathbf{X}_n \right)^{-1} \left(\mathbf{I}_L - \left(\mathbf{I}_L - \mu_n \mathbf{X}_n^T \mathbf{X}_n \right)^R \right) \mathbf{e}_n \tag{144}$$

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mathbf{X}_n \left(\mathbf{X}_n^T \mathbf{X}_n \right)^{-1} \left(\mathbf{I}_L - \left(\mathbf{I}_L - \mu_n \mathbf{X}_n^T \mathbf{X}_n \right)^R \right) \mathbf{e}_n. \tag{145}$$

If the stepsize μ_n is small enough, all the eigenvalues of $\mathbf{I}_L - \mu_n \mathbf{X}_n^T \mathbf{X}_n$ lie within the unit circle : $\mathbf{X}_n^T \mathbf{X}_n$ is symmetric positive definite for randomly chosen data and $N \geqslant L$. Hence, if R goes to infinity,

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mathbf{X}_n \left(\mathbf{X}_n^T \mathbf{X}_n \right)^{-1} \mathbf{e}_n \tag{146}$$

is obtained, which confirms the conclusion made based on Eq. 121. So, it appears that the unnormalised and globally normalised constrained PBFDRAP, for which $\boldsymbol{\Delta} = \mu_n \mathbf{I}_M$, approach the PRA by series expansion of the matrix inverse in Eq. 104.

5.4.2 Subband–normalised constrained PBFDRAP

Unlike in section 5.4.1, matrix $\boldsymbol{\mathcal{D}}$, defined by Eq. 32, is typically not symmetric here. For the subband–normalised constrained PBFDRAP for which $\boldsymbol{\Delta} = \mathrm{diag}(\boldsymbol{\mu}_s^{(n)})$, being a diagonal matrix with subband dependent stepsizes $\boldsymbol{\mu}_s^{(n)}(m)$, the eigenvalues of $\boldsymbol{\mathcal{D}}$, i.e. the diagonal elements of $\boldsymbol{\Gamma}_1$, are not always real and positive. Hence, for $N \geqslant L$,

$$\lim_{R \to \infty} \boldsymbol{\Xi}^{(R)} = \lim_{R \to \infty} \boldsymbol{\Gamma}_1^{-1} \left(\mathbf{I}_L - (\mathbf{I}_L - \boldsymbol{\Gamma}_1)^R \right) \tag{147}$$

may not be bounded and the algorithm diverges for $R \to \infty$. However, this does not automatically imply divergence of the algorithm for finite R, but indicates that depending on the data set used iterating this algorithm is expected to invoke stability problems.

6 Experiments

Some experiments were done to check the validity of Eq. 53, 57 and 77. A white noise signal x was fed through a randomly generated FIR filter w of order 15. Filter w was adaptively estimated using the PBFDRAP with different settings. In all simulations $N = 16$, $M = 16$ and $P = L = 8$. The results for the unnormalised unconstrained PBFDRAP are shown in figure 2. The figure corresponding to the unnormalised

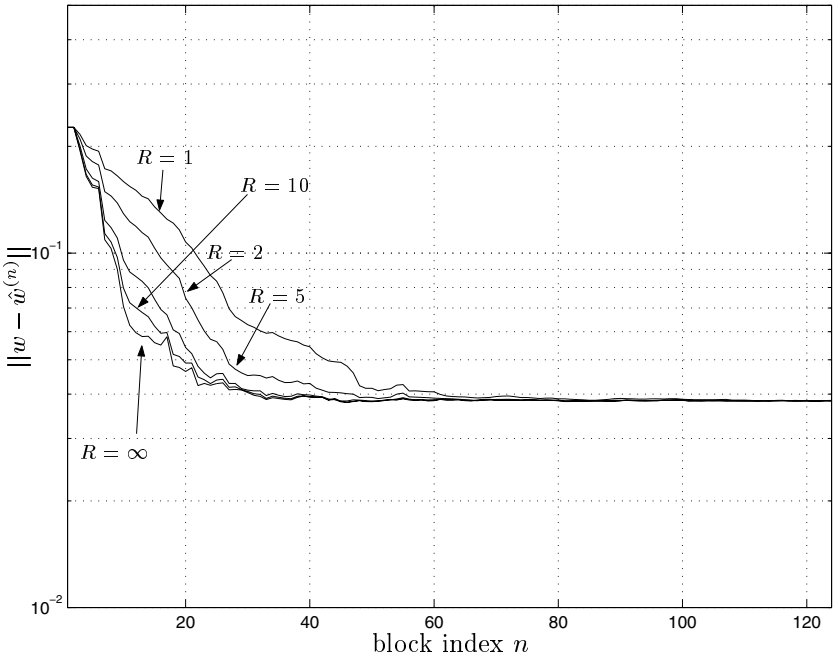


Figure 2: Unnormalised unconstrained PBFDRAP : the norm between w and $\hat{w}$, the equivalent time–domain representation of the adaptive filter coefficients, is plotted as a function of time. It appears that if R is increased the bottom curve ($R = \infty$) is approached, corresponding to a normalised version of the unconstrained PBFDAF that is based on projected subband energies with maximum stepsize $\mu = 1$ (Eq. 57).

constrained PBFDRAP is figure 3. For the subband–normalised PBFDRAP also 2 figures are shown : in figure 4 the convergence curves of the unconstrained algorithm are shown, in figure 5 it can be observed that the subband–normalised constrained PBFDRAP becomes unstable for large values of R, such that in practice Eq. 57 is not valid if R goes to infinity.

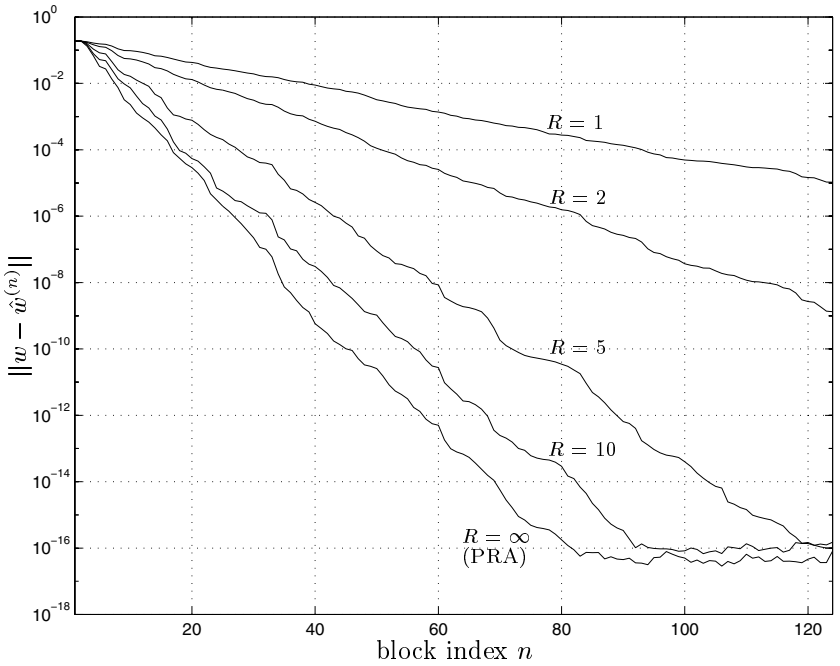


Figure 3: Unnormalised constrained PBFDRAP : the norm between w and $\hat{w}$, the equivalent time–domain representation of the adaptive filter coefficients, is plotted as a function of time. It appears that if R is increased the bottom curve ($R = \infty$) is approached, corresponding to a block version of the APA–algorithm with maximum stepsize $\mu = 1$.

7 Conclusions

In this paper the PBFDRAP algorithm was considered, an adaptive filtering technique combining the partitioned block frequency–domain adaptive filter and row action projection. It was shown that the PBFDRAP algorithm approaches a normalised version of the PBFDAF that is based on projected subband energies with maximum stepsize $\mu = 1$, if the number of iteration steps goes to infinity. For certain parameter settings the infinitely iterated PBFDRAP reduces to well–known adaptive filter algorithms such as the PRA.

8 Acknowledgements

This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of the Belgian State Interuniversity Poles of Attraction Programme – IUAP P4–02 (1997–2001) : Modeling, Identification, Simulation and Control of Complex Systems, Concerted Research Action GOA–MEFISTO–666 of the Flemish

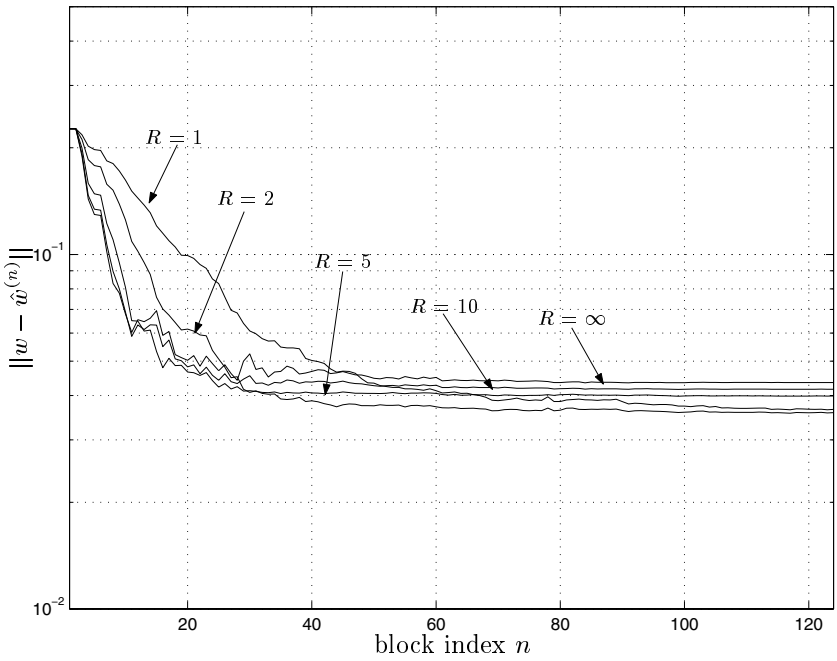


Figure 4: Subband–normalised unconstrained PBFDRAP : the norm between w and $\hat{w}$, the equivalent time–domain representation of the adaptive filter coefficients, is plotted as a function of time. It appears that if R is increased the bottom curve ($R = \infty$) is approached, corresponding to the subband–normalised unconstrained PBFDRAP with stepsize $\mu = 1$ (Eq. 77).

Government, Research Project F.W.O. nr. G.0295.97 ('Design and implementation of adaptive digital signal processing algorithms for broadband applications') and IWT project AUT/970517: MUSETTE ('Multimicrophone Signal Enhancement Techniques for handsfree telephony and voice controlled systems'). The scientific responsibility is assumed by its authors.

References

[1] G. Egelmeers, P. Sommen, and J. de Boer. Realization of an Acoustic Echo Canceller on a Single DSP. In *Proceedings of the European Signal Processing Conference (EUSIPCO–96)*, pages 33–36, Trieste, Italy, September 1996.

[2] K. Eneman, S. Doclo, and M. Moonen. A Comparison between Iterative block–LMS and Partial Rank Algorithm. Technical Report 99–52, ESAT–SISTA, Katholieke Universiteit Leuven, Belgium, March 1999.

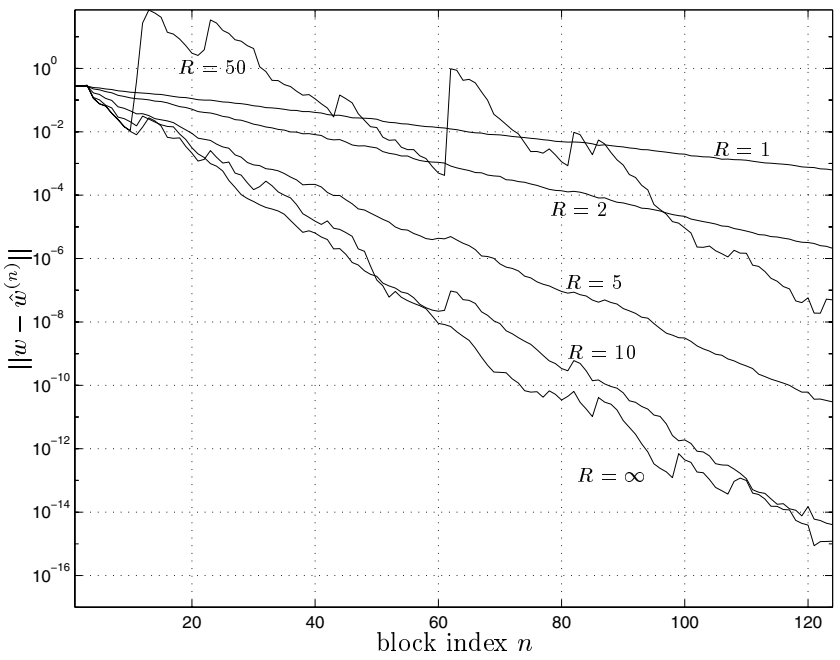


Figure 5: Subband–normalised constrained PBFDRAP : the norm between w and $\hat{w}$, the equivalent time–domain representation of the adaptive filter coefficients, is plotted as a function of time. It appears that if R is increased the subband–normalised constrained PBFDRAP becomes unstable.

[3] K. Eneman and M. Moonen. Hybrid Subband/Frequency–Domain Adaptive Systems. *Signal Processing*, 81(1):117–136, January 2001.

[4] K. Eneman and M. Moonen. Iterated Partitioned Block Frequency–Domain Adaptive Filtering for Acoustic Echo Cancellation. Technical Report 00–83 (ftp.esat.kuleuven.ac.be/pub/sista/eneman/reports/00-83.ps.gz), ESAT–SISTA, Katholieke Universiteit Leuven, Belgium, December 2000.

[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.

[6] D. Morgan and S. Kratzer. On a Class of Computationally Efficient, Rapidly Converging, Generalized NLMS Algorithms. *IEEE Signal Processing Letters*, 3(8):245–247, August 1996.

[7] J. Páez Borrallo and M. García Otero. On the implementation of a partitioned block frequency domain adaptive filter (PBFDAF) for long acoustic echo cancellation. *Signal Processing*, 27:301–315, June 1992.

[8] J. Shynk. Frequency–Domain and Multirate Adaptive Filtering. *IEEE Signal Processing Magazine*, 9(1):15–37, January 1992.

[9] J.–S. Soo and K. Pang. Multidelay Block Frequency Domain Adaptive Filter. *IEEE Trans. Acoust., Speech and Signal Processing*, 38(2):373–376, February 1990.

[10] B. Widrow and S. Stearns. *Adaptive Signal Processing*. Prentice Hall, Englewood Cliffs, New Jersey, USA, 1985.