AN ADAPTIVE PROJECTION ALGORITHM FOR MULTIRATE FILTER BANK OPTIMIZATION

Dong-Yan Huang and Phillip A. Regalia
Département Signal & Image
Institut National des Télécommunications
9, rue Charles Fourier
F-91011 Evry cedex France

ABSTRACT

We develop a new algorithm for multirate filter bank optimization, which finds application in subband coding or wavelet signal analysis. Although some impressive off-line algorithms have recently been developed for this purpose, the computation demand of such algorithms often renders them prohibitive for real-time applications. In this vein, adaptive filtering solutions remain of interest. A simple gradient descent algorithm may be ill suited due to the nonquadratic nature of the cost function to be minimized, and accordingly non gradient algorithms may offer some attractive alternatives. The present paper describes a projection type algorithm, which aims to construct a lossless filter bank such that one of its impulse responses lies close to an extremal eigenvector of the input signal autocorrelation matrix. Though a formal convergence proof of the algorithm is not offered, simulations show that the algorithm converges to an acceptable vicinity of the global minimum point of the cost function.

1 INTRODUCTION

The two-channel lossless FIR filter bank of Figure 1 has been used in various applications of signal processing, including information coding and wavelet analysis [1]. In this context, a standard problem is to optimize the rotation angles $\{\theta_k\}_{k=0}^{M-1}$ such that the variance of the output $E[y_2^2(n)]$ is "small" or minimized with repect to some criterion.

The approach which we adopt here is based on an adaptive filtering formulation, of interest in real-time applications where off-line optimization methods may be costly, or ill-suited to the constraints of real-time signal processing.

The most direct approach towards this problem is to a gradient descent procedure applied to the cost function $E[y_2^2(n)]$ [2], [3]. As the cost function $E[y_2^2(n)]$ is nonquadratic in the rotation angles θ_k, local minima may exist and a gradient descent algorithm may become trapped in one. Although the global minimum will certainly give the "best" solution, local minima may provide suboptimal performance. This has motivated alternate adaptive filtering algorithms which are based on eigenvalue embedding rather than gradient descent procedures [4], as well as off-line approaches [5] which formulate the problem as one of convex optimization. We should note that the method of [5] requires that the solution be spectrally factored in order to obtain the "wavelet filter"; as out approach is based on adapting the lossless filter directly, such spectral factorization is obviated.

The purpose of this article is to propose an adaptive projection algorithm which aims to construct a lossless filter bank whose impulse response is close to an extremal eigenvector of the input signal autocorrelation matrix. A two-step algorithm is then developed, in which the first step uses well-known Rayleigh quotient type algorithms (e.g., [6], [7]) to obtain an extremal eigenvector of the input autocorrelation matrix. The lossless filter is then adapted to fit one of its impulse responses close to this eigenvector. Rather than use a gradient descent procedure for this purpose, the algorithm aims to project the "error" in the eigenvector fit onto the orthogonal complement subspace to the extended reachability space of the lossless filter. Although a formal convergence proof of this algorithm is not offered, numerous simulation examples indicate that the algorithm converges to a point acceptably close to the global minimum of the cost function, while at the same time avoiding potential local minima due to its non gradient nature.

2 PROBLEM STRUCTURE

We begin with the two-channel lossless FIR filter bank of Figure 1, using M rotation angles $\Theta = [\theta_0, \cdots, \theta_{M-1}]$. This system may be described as

$$\begin{bmatrix} \mathbf{x}(n+1) \\ \mathbf{y}(n) \end{bmatrix} = \underbrace{\begin{bmatrix} A(\Theta) & B(\Theta) \\ C(\Theta) & D(\Theta) \end{bmatrix}}_{Q(\Theta)} \begin{bmatrix} \mathbf{x}(n) \\ \mathbf{u}(n) \end{bmatrix} \quad (1)$$

where $\mathbf{x}(\cdot) = [x_1(\cdot), \cdots, x_M(\cdot)]^T$ is the state vector, $\mathbf{u}(\cdot) = [u_1(n), u_2(n)]^T$ is the input vector, and $\mathbf{y}(\cdot) = [y_1(\cdot), y_2(\cdot)]^T$ is ouput vector. The matrix $Q(\Theta)$ is the composite interconnection of the rotations of Figure 1, and so is orthogonal irrespective of the exact values of the rotation angles Θ. We assume that $\{u(\cdot)\}$ is a stationary second-order process, and that the input vector derives from this scalar process by way of subsampling, i.e.,

$$u_1(n) = u(2n), \quad u_2(n) = u(2n-1)$$

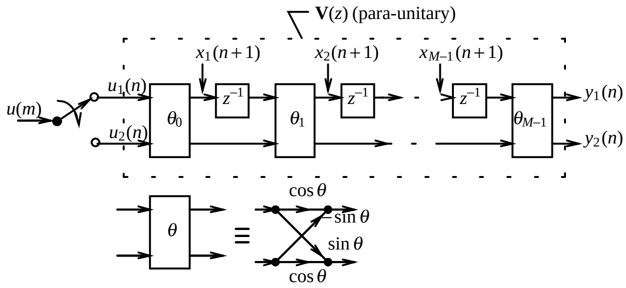

Figure 1: Two-channel lossless filter bank

The transfer matrix $V(z)$ mapping the input sequence to the output sequence, as in

$$\begin{bmatrix} y_1(z) \\ y_2(z) \end{bmatrix} = V(z) \begin{bmatrix} u_1(n) \\ u_2(n) \end{bmatrix} \tag{2}$$

is para-unitary, i.e.,

$$V(z)V^T(z^{-1}) = V^T(z^{-1})V(z) = I_2$$

for all choices of the rotation angles Θ.

If we partition the terms $C(\Theta)$ and $D(\Theta)$ from (1) into two row vectors as $C(\Theta) = \left[\begin{smallmatrix} c_1 \\ c_2 \end{smallmatrix}\right]$ and $D(\Theta) = \left[\begin{smallmatrix} d_1 \\ d_2 \end{smallmatrix}\right]$, then the filter output $y_2(n)$ becomes

$$y_2(n) = \underbrace{[\,d_2 \quad c_2B \quad c_2AB \quad \cdots \quad c_2A^{M-1}B\,]}_{\triangleq\, \mathbf{v}_2} \underbrace{\begin{bmatrix} \mathbf{u}(n) \\ \mathbf{u}(n-1) \\ \mathbf{u}(n-2) \\ \vdots \\ \mathbf{u}(n-M) \end{bmatrix}}_{\triangleq\, \mathcal{U}(n)}.$$

We consider the problem of adjusting the rotation angles according to the input signal so that

$$E\left[y_2^2(n)\right] = \mathbf{v}_2(\Theta) R_u \mathbf{v}_2^T(\Theta) \tag{3}$$

is rendered small. Here

$$R_u = E\left\{ \begin{bmatrix} \mathbf{u}(n) \\ \mathbf{u}(n-1) \\ \vdots \\ \mathbf{u}(n-M) \end{bmatrix} [\mathbf{u}^t(n) \quad \mathbf{u}^T(n-1) \quad \cdots \quad \mathbf{u}^T(n-M)] \right\}$$

is the input autocorrelation matrix whose (k,l) element (counting from zero) is $E[u(n-k)\,u(n-l)]$,

Because the realization of Figure 1 is lossless, the impulse response vector $\mathbf{v}_2$ has unit norm. As such, the cost function (3) appears related to a Rayleigh quotient of R_u. The vector $\mathbf{v}_2$ cannot be an arbitrary unit-norm vector, though. Indeed, if we introduce the shift matrix as

$$\mathcal{Z} \triangleq \begin{bmatrix} 0_2 & 0_2 & \cdots & 0_2 \\ \mathbf{I}_2 & 0_2 & \cdots & 0_2 \\ \vdots & \ddots & \ddots & \vdots \\ 0_2 & \cdots & \mathbf{I}_2 & 0_2 \end{bmatrix}, \qquad [2(M+1)\times 2(M+1)] \tag{4}$$

then losslessness implies that $\mathbf{v}_2$ is orthogonal to shifted versions of itself, i.e.,

$$\mathbf{v}_2 \mathcal{Z}^k \mathbf{v}_2^T = \begin{cases} 1, & k = 0; \\ 0, & k = 1, 2, 3, \ldots. \end{cases} \tag{5}$$

Now, given only that $\mathbf{v}_2$ has unit norm, the cost function (3) can be no smaller than $\lambda_{2M+2}(R_u)$, i.e., the smallest eigenvalue of R_u. This follows because for any unit-norm vector, say $\hat{\mathbf{v}}$, one has $\hat{\mathbf{v}} R_u \hat{\mathbf{v}}^T \geq \lambda_{2M+2}$, with equality holding if and only if $\hat{\mathbf{v}}$ is an eigenvector corresponding to λ_{2M+2}. Such an extremal eigenvector, though, will not in general verify the “shift-orthogonality” constraint as in (5), and so cannot be associated with the impulse response vector of a lossless system.

The Rayleigh quotient nature of the cost function nonetheless suggests adjusting the rotation angles Θ such that the resulting impulse response vector $\mathbf{v}_2$ is close to an extremal eigenvector of R_u in some sense. To this end, let

$$\hat{\mathbf{v}} = [\,\hat{v}_0 \quad \hat{v}_1 \quad \cdots \hat{v}_M\,]$$

with each term $\hat{v}_k$ of dimensions 1×2, be an adjustable unit norm vector, and set

$$\hat{y}_2(n) = \sum_{k=0}^{M} \hat{v}_k\, \mathbf{u}(n-k) = \hat{\mathbf{v}}\,\mathcal{U}(n).$$

Then

$$E[\hat{y}_2^2(n)] = \hat{\mathbf{v}} R_u \hat{\mathbf{v}}^T, \tag{6}$$

and this cost function will be minimized if and only if $\hat{\mathbf{v}}$ aligns with an extremal eigenvector of R_u (e.g., [7], [6])

$$R_u \hat{\mathbf{v}}^T = \lambda_{2M+2} \hat{\mathbf{v}}^T.$$

We can then consider adjusting the rotation angles Θ such that the filter output $y_2(n)$ is close in some sense to $\hat{y}_2(n)$, which suggests a cost function of the form

$$E\left[\left(\hat{y}_2(n) - y_2(n)\right)^2\right] = \left(\hat{\mathbf{v}} - \mathbf{v}_2\right) R_u \left(\hat{\mathbf{v}} - \mathbf{v}_2\right)^T. \tag{7}$$

Note that this cost function, as with $E[y_2^2(n)]$, is nonquadratic in the rotation angles Θ, so that a gradient descent procedure may prove inappropriate. For this reason, an projection-type algorithm is proposed in the next section.

3 ADAPTIVE ALGORITHM

Our adaptive algorithm simultaneously implements two procedures:

- Adapt a unit-norm vector $\hat{\mathbf{v}}(n)$ towards a minimum eigenvector of the input autocorrelation matrix;
- Adapt the impulse response $\mathbf{v}_2(\Theta)$ of a lossless FIR filter bank such that it best aligns with the eigenvector $\hat{\mathbf{v}}$ from the previous step.

For the first step, we can apply a classical update algorithm for an FIR vector of unit norm according to

$$\hat{\mathbf{v}}(n+1) = \frac{\hat{\mathbf{v}}(n) - \mu\,\mathcal{U}(n)\hat{y}_2(n)}{||\hat{\mathbf{v}}(n) - \mu\,\mathcal{U}(n)\hat{y}_2(n)||} \tag{8}$$

which is a gradient descent procedure applied to the Rayleigh quotient cost function from (6); other possibilities may be found in [7], [6]. It is fairly well known that this cost function has no local minima, so that global asymptotic convergence will apply for slow adaptation.

Our next objective is to find a $\mathbf{v}_2(\Theta)$ such that it lies near the vector $\mathbf{v}(n)$. For this purpose, we propose the following algorithm:

$$\Theta(n+1) = \Theta(n) + \mu\Gamma(\Theta)\begin{bmatrix}\mathbf{x}(n+1,\Theta)\\ y_1(n)\end{bmatrix}\Big(\hat{y}_2(n) - y_2(n)\Big) \tag{9}$$

with

$$\begin{cases}\gamma_M &= 1\\ \gamma_i &= \gamma_{i+1}\cos\theta_i \quad i = 0,\cdots,M-1\end{cases}$$

and $\Gamma(\Theta) = \mathrm{diag}\left[\gamma_1,\cdots,\gamma_M\right]$. All the necessary signals are directly available from the lattice filter. We should emphasize that this algorithm is *not* a gradient descent procedure applied to the cost function (7). It is closely related to, and in part inspired from, an algorithm proposed for rational subspace estimation in [8], except that the lossless filter is now FIR instead of the more general IIR structure used in [8], and moreover, the singularity of the input autocorrelation matrix—a fundamental assumption in [8]—is in general inapplicable in our context. The nongradient character implies that the convergence properties cannot be deduced by examining the minima of the cost function (7). We thus examine the convergence properties by an alternate analysis.

For this purpose, return to the partitioned form of $Q(\Theta)$ from (1). It is straightforward to check that

$$\begin{bmatrix}\mathbf{x}(n+1)\\ y_1(n)\\ y_2(n)\end{bmatrix} = \begin{bmatrix}B & AB & \cdots & A^{M-1}B & 0\\ d_1 & c_1B & \cdots & c_1A^{M-2}B & c_1A^{M-1}B\\ d_2 & c_2B & \cdots & c_2A^{M-2}B & c_2A^{M-1}B\end{bmatrix}\begin{bmatrix}\mathbf{u}(n)\\ \vdots\\ \mathbf{u}(n-M)\end{bmatrix}$$

$$= \begin{bmatrix}\mathcal{C}\\ \mathbf{v}_1\\ \mathbf{v}_2\end{bmatrix}\begin{bmatrix}\mathbf{u}(n)\\ \vdots\\ \mathbf{u}(n-M)\end{bmatrix}$$

in which $\mathcal{C}$ is the controllability matrix, while $\mathbf{v}_1$ and $\mathbf{v}_2$ are the impulse response vectors corresponding to the filter outputs $y_1(n)$ and $y_2(n)$, respectively. Owing to the orthogonality of the filter computations, we have the orthonormality property

$$\begin{bmatrix}\mathcal{C}\\ \mathbf{v}_1\\ \mathbf{v}_2\end{bmatrix}[\,\mathcal{C}^T \quad \mathbf{v}_1^T \quad \mathbf{v}_2^T\,] = \mathbf{I}_{M+2}$$

But since the row vectors comprising $\mathcal{C}$, as well as $\mathbf{v}_1$ and $\mathbf{v}_2$, are of length $2M+2$, they do not constitute a complete basis for the Euclidan space $\mathbb{R}^{2M+2}$. To complete this set of vectors, one may show that, since $Q(\Theta)$ in (1) is orthogonal,

$$\begin{bmatrix}\mathcal{C}\\ \mathbf{v}_1\\ \mathbf{v}_2\end{bmatrix}\mathcal{Z}^k\mathbf{v}_2^T = \mathbf{0}, \qquad k = 1,2,\ldots,M,$$

so that the (column) vectors $\mathcal{Z}\mathbf{v}_2^T$, $\mathcal{Z}^2\mathbf{v}_2^T$, $\ldots$, $\mathcal{Z}^M\mathbf{v}_2^T$ lie in the complementary space. From (5), one deduces that the vectors $\mathcal{Z}\mathbf{v}_2^T,\ldots,\mathcal{Z}^M\mathbf{v}_2^T$ are mutually orthogonal (though not orthonormal), such that the row vectors of the square matrix

$$\begin{bmatrix}\mathcal{C}\\ \mathbf{v}_1\\ \mathbf{v}_2\\ \mathbf{v}_2\mathcal{Z}^T\\ \vdots\\ \mathbf{v}_2(\mathcal{Z}^T)^M\end{bmatrix}$$

form a complete orthogonal basis for the Euclidean space $\mathbb{R}^{2M+2}$.

Now, the mean value of the correction term of the adaptation algorithm (9) appears as

$$E\left\{\begin{bmatrix}\mathbf{x}(n+1)\\ y_1(n)\end{bmatrix}\Big(\hat{y}_2(n) - y_2(n)\Big)\right\} = \begin{bmatrix}\mathcal{C}\\ \mathbf{v}_1\end{bmatrix}R_u(\hat{\mathbf{v}} - \mathbf{v}_2)^T.$$

After the first step has converged, we have $R_u\hat{\mathbf{v}}^T = \lambda_{2M+2}\hat{\mathbf{v}}^T$; upon recognizing that the stationary points of the algorithm correspond to the mean value of the correction term vanishing, we find that the stationary points are characterized by the equation

$$\mathbf{0} = \begin{bmatrix}\mathcal{C}\\ \mathbf{v}_1\end{bmatrix}\Big(\lambda_{2M+2}\hat{\mathbf{v}}^T - R_u\mathbf{v}_2^T\Big).$$

Since the space orthogonal to the row vectors of $\left[\begin{smallmatrix}\mathcal{C}\\ \mathbf{v}_1\end{smallmatrix}\right]$ is spanned by the vectors $\mathbf{v}_2$, $\mathbf{v}_2\mathcal{Z}^T$, $\ldots$, $\mathbf{v}_2^T(\mathcal{Z}^T)^M$, we find at any stationary point the relation

$$\lambda_{2M+1}\hat{\mathbf{v}} - R_u\mathbf{v}_2^T = [\,\mathbf{v}_2^T \quad \mathcal{Z}\mathbf{v}_2^T \quad \cdots \quad \mathcal{Z}^M\mathbf{v}_2^T\,]\begin{bmatrix}g_0\\ g_1\\ \vdots\\ g_M\end{bmatrix}$$

for some set of coefficients $g_0,\ldots,g_M$. Upon multiplying this expression from the left by $\mathbf{v}_2$, and recognizing that $E[y_2^2(n)] = \mathbf{v}_2R_u\mathbf{v}_2^T$, we obtain

$$\lambda_{2M+2}\mathbf{v}_2\hat{\mathbf{v}}^T - E[y_2^2(n)] = g_0.$$

where $|\mathbf{v}_2\hat{\mathbf{v}}^T| \leq 1$ since both $\hat{\mathbf{v}}$ and $\mathbf{v}_2$ have unit norm. This can be arranged as

$$E[y_2^2(n)] = \lambda_{2M+2}\,\mathbf{v}_2\hat{\mathbf{v}} - g_0.$$

Note that it is difficult to obtain a precise *a priori* bound on the coefficient g_0.

4 SIMULATION RESULTS

To test the behavior of the proposed algorithm, the sequence $\{u(\cdot)\}$ was generated from the output of a filter driven by white noise; the poles and zeros of the filter, as with [4], are given by

$$\text{poles}\begin{cases}-0.5+0.8307i\\ -0.5-0.8307i\\ 0.75+0.5809i\\ 0.75-0.5809i\\ -0.3\end{cases} \qquad \text{zeros}\begin{cases}0.9+0.4472i\\ 0.9-0.4472i\\ 1.05\\ -0.2\end{cases}$$

The resulting sequence $\{u(\cdot)\}$ is subsampled, and then applied to the lossless filter bank of Figure 1. $M = 4$ rotation angles were used in this simulation.

The rotation angles Θ corresponding to the global minimum of the cost function $E[y_2^2(n)]$ lie at the values

$$\begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \begin{bmatrix} 0.5704 \\ 0.4755 \\ -0.2115 \\ -0.1556 \end{bmatrix} \quad \text{giving} \quad E[y_2^2(n)] = 0.0219$$

The proposed adaptive algorithm was observed to converge to values whose time averages after convergence were

$$\begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \begin{bmatrix} 0.6509 \\ 0.4574 \\ -0.1885 \\ -0.1295 \end{bmatrix} \quad \text{giving} \quad E[y_2^2(n)] = 0.0228$$

The converged solution lies very close to the global minimum of the cost function $E[y_2^2(n)]$.

A second example involves generating $\{u(\cdot)\}$ from the output of an FIR filter whose transfer function is $1 + 1.5954z^{-1} + 0.9499z^{-2}$, and adapting $M = 3$ rotation angles. The global minimum of the cost function $E[y_2^2(n)]$ lies at

$$\begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} -0.7246 \\ -0.3374 \\ 0.2765 \end{bmatrix} \quad \text{giving} \quad E[y_2^2(n)] = 0.0401$$

The proposed algorithm was observed to converge to a solution at

$$\begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} -0.7873 \\ -0.3516 \\ 0.2498 \end{bmatrix} \quad \text{giving} \quad E[y_2^2(n)] = 0.0407$$

lying once again in an acceptable vicinity of the global minimum point.

Though no proof is given that the convergent point of the proposed algorithm will always lie near the global minimum of the cost function $E[y_2^2(n)]$, this phenomenon has nonetheless been observed in numerous simulation examples.

5 CONCLUSION

In this paper, we have proposed a projection-type adaptive algorithm which aims to adapt the impulse response vector of a lossless filter towards an extremal eigenvector of the input signal autocorrelation matrix. The algorithm features modest complexity (order M operations per time step), in contrast to a gradient descent algorithm whose complexity is typically order M^2 operations per time step to generate the gradient signals $\partial y_2(n)/\partial\theta_k$. Although global convergence to the minimum point of the cost function is not proved (in contrast to the off-line method of [5]), numerous simulations indicate that the algorithm converges to an acceptable vicinity of the global minimum point. A formal analysis of this phenomenon is complicated by the non gradient nature of the algorithm, and may be considered a topic for future research.

References

[1] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Prentice-Hall, Englewood Cliffs, NJ, 1992.

[2] P. Delsarte, B. Macq, and D. T. M. Slock, “Signal adapted multiresolution transform for image coding,” *IEEE Trans. Signal Processing*, 42(11):2955–2966, Nov. 1992.

[3] B. Macq and J. Y. Mertes, “Optimization of linear multiresolution transforms for scene adaptive coding,” *IEEE Trans. Signal Processing*, 41(12):3568–3572, Dec. 1993.

[4] P.A. Regalia and D. Y. Huang, “Attainable error bounds in multirate adaptive lossless FIR filters,” in: *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. V, pp. 1460–1463, Detroit, MI, May 1995.

[5] J.-C. Pesquet and P. L. Combettes, “Wavelet synthesis by alternating projections,” *IEEE Trans. Signal Processing* (to appear).

[6] P. A. Regalia, “An adaptive unit morm filter with application to signal analysis and Karhunen-Loeve transformations,” *IEEE Trans. Circuit and Systems*, 37(5):646–649, May 1990.

[7] V. U. Reddy, B. Eghardt, and T. Kailath, “Least-squares type algorithm for adaptive implementation of Pisarenko’s harmonic retrieval method,” *IEEE Trans. Acoustics, Speech, and Signal Processing*, 30:399–405, June 1982.

[8] I. Fijalkow, *Estimation de Sous-Espaces Rationnels*, doctoral thesis, Ecole Nationale Supérieure des Télécommunications, Paris, Oct. 1993.