

BACKWARD ADAPTIVE BIORTHOGONALIZATION

L. REBOLLO-NEIRA

Received 7 February 2003

A backward biorthogonalization approach is proposed, which modifies biorthogonal functions so as to generate orthogonal projections onto a reduced subspace. The technique is relevant to problems amenable to be represented by a general linear model. In particular, problems of data compression, noise reduction, and sparse representations may be tackled by the proposed approach.

2000 Mathematics Subject Classification: 46E99.

1. Introduction. We introduce a backward biorthogonalization technique relevant to the general linear model. Any data which may be described in terms of a linear combination of waveforms satisfies this model. For instance, the response of a physical system to a particular interaction varying as a function of a parameter, say t , is often represented by a measurable quantity $f(t)$, which is amenable to be expressed in the fashion

$$f(t) = \sum_n^N c_n^N \alpha_n(t), \quad (1.1)$$

where the model waveforms α_n , $n = 1, \dots, N$, are derivable by recourse to physical considerations. The determination of the coefficients c_n^N , $n = 1, \dots, N$, entails solving the inverse problem when the function $f(t)$ is measured. The superscript N indicates that, unless the waveforms α_n are orthogonal, the appropriate coefficients c_n^N depend on the order of the model, that is, the number N of waveforms α_n being considered in (1.1). If such waveforms are linearly independent, then there exists a set of reciprocal functions $\tilde{\alpha}_n^N$, $n = 1, \dots, N$, which is biorthogonal to the former, that is, $\langle \tilde{\alpha}_n^N | \alpha_m \rangle = \delta_{n,m}$ [5, 11]. Here the superscript N indicates that the biorthogonal functions $\tilde{\alpha}_n^N$ allow for constructing orthogonal projections onto the subspace V_N spanned by N waveforms α_n . Hence, the coefficients c_n^N of the linear expansion (1.1) approximating a function f at best (in a minimum distance sense) can be obtained by computing the inner products $c_n^N = \langle \tilde{\alpha}_n^N | f \rangle$ [5, 7].

Since the reciprocal set $\tilde{\alpha}_n^N$, $n = 1, \dots, N$ (and therefore the coefficients c_n^N), depends on the number N of total waveforms, they should be recalculated when this number is enlarged or reduced. This feature of nonorthogonal expansions is discussed in [5, 6, 7], where a recursive methodology is introduced for transforming the reciprocal set $\tilde{\alpha}_n^N$, $n = 1, \dots, N$, into a new one $\tilde{\alpha}_n^{N+1}$, $n = 1, \dots, N+1$. The latter is guaranteed to yield orthogonal projections onto the subspace V_{N+1} arising by the inclusion of a waveform α_{N+1} in V_N , that is, $V_{N+1} = V_N + \alpha_{N+1}$.

Here we wish to consider the converse situation: we suppose that the reciprocal waveforms $\tilde{\alpha}_n^N, n = 1, \dots, N$, are known and we want to modify them so as to obtain orthogonal projections onto a subspace V_{N/α_j} arising by eliminating an element, say the j th one, from V_N . Then, $V_{N/\alpha_j} = \text{span}\{\alpha_n; n = 1, \dots, j-1, j+1, \dots, N\}$. Our aim is to construct the corresponding reciprocal functions $\tilde{\alpha}_n^{N/j}, n = 1, \dots, j-1, j+1, \dots, N$, by modifying the previous $\tilde{\alpha}_n^N, n = 1, \dots, N$.

We suppose that in the summation of (1.1), we want to retain only some terms and approximate $f(t)$ by a linear combination of those elements. Thus, to obtain the best approximation of $f(t)$ by a linear combination of such a nature, we need to recalculate the coefficients corresponding to the waveforms we wish to retain. If we simply disregarded the coefficients of the unwanted terms, but did not recalculate the remaining ones, the approximation would not be optimal in a minimum distance sense. The approach we propose in this communication allows for the necessary modifications of coefficients so as to achieve the optimal approximation. The method is based on an iterative technique capable of adapting biorthogonal functions in order to generate orthogonal projections onto a reduced subspace.

The paper is organized as follows. Section 2 introduces the notation, discusses the motivation to the proposed approach, and summarizes a previously introduced forward biorthogonalization method [5, 7]. Section 3 discusses the proposed biorthogonalization technique to transform biorthogonal functions in order to build orthogonal projections onto a reduced subspace. The conclusions are drawn in Section 4.

2. Notation, background, and motivation to the approach. Adopting Dirac’s vector notation [2], we represent an element f of a Hilbert space \mathcal{H} as a vector $|f\rangle$ and its dual as $\langle f|$. Given a set of δ -normalized continuous orthogonal vectors $\{|t\rangle; -\infty < t < \infty; \langle t|t'\rangle = \delta(t-t')\}$, the unity operator in \mathcal{H} is expressed as follows:

$$\hat{I}_{\mathcal{H}} = \lim_{T \rightarrow \infty} \int_{-T}^T |t\rangle\langle t| dt. \tag{2.1}$$

Thus, for all $|f\rangle$ and $|g\rangle \in \mathcal{H}$, by inserting $\hat{I}_{\mathcal{H}}$ in $\langle f|g\rangle$, that is,

$$\langle f|\hat{I}_{\mathcal{H}}|g\rangle = \lim_{T \rightarrow \infty} \int_{-T}^T \langle f|t\rangle\langle t|g\rangle dt, \tag{2.2}$$

one is led to a representation of \mathcal{H} in terms of the space of square-integrable functions, with $\langle t|g\rangle = g(t)$ and $\langle g|t\rangle = \langle t|g\rangle^* = g^*(t)$, where $g^*(t)$ indicates the complex conjugate of $g(t)$.

Let vectors $|\alpha_n\rangle \in \mathcal{H}, n = 1, \dots, \infty$, be a Riesz basis for \mathcal{H} . Hence, all $|f\rangle \in \mathcal{H}$ can be expressed as the linear span

$$|f\rangle = \sum_{n=1}^{\infty} c_n |\alpha_n\rangle \tag{2.3}$$

and there exists a reciprocal basis $|\tilde{\alpha}_n\rangle, n = 1, \dots, \infty$, for \mathcal{H} , to which the former basis is biorthogonal, that is, $\langle \tilde{\alpha}_n|\alpha_m\rangle = \delta_{n,m}$ [11]. The reciprocal basis allows to compute

the coefficients c_n in (2.3) as the inner products

$$c_n = \langle \tilde{\alpha}_n | f \rangle = \lim_{T \rightarrow \infty} \int_{-T}^T \tilde{\alpha}_n^*(t) f(t) dt. \tag{2.4}$$

Thus,

$$|f\rangle = \sum_{n=1}^{\infty} |\alpha_n\rangle \langle \tilde{\alpha}_n | f \rangle \tag{2.5}$$

so that, by denoting

$$\hat{I} = \sum_{n=1}^{\infty} |\alpha_n\rangle \langle \tilde{\alpha}_n |, \tag{2.6}$$

(2.3) can be recast as $f = \hat{I}f$, which implies that, at least in the weak sense, \hat{I} is a representation of the identity operator in \mathcal{H} . Thus, we have the following generalization of the Plancherel-Parseval identity:

$$\| |f\rangle \|^2 = \langle f | \hat{I} | f \rangle = \sum_{n=1}^{\infty} \tilde{c}_n^* c_n, \tag{2.7}$$

with c_n as in (2.4) and $\tilde{c}_n^* = \langle f | \alpha_n \rangle$.

If the basis $|\alpha_n\rangle, n = 1, \dots, \infty$, is orthogonalized and we denote by $|\tilde{\psi}_n\rangle, n = 1, \dots, \infty$, the corresponding orthogonal vectors after normalization to unity, then the new basis is self-reciprocal, that is, it satisfies the orthonormality condition $\langle \tilde{\psi}_m | \tilde{\psi}_n \rangle = \delta_{m,n}$ and provides a representation for the identity operator as given by

$$\hat{I} = \sum_{n=1}^{\infty} |\tilde{\psi}_n\rangle \langle \tilde{\psi}_n |. \tag{2.8}$$

This representation of the identity operator can be seen as a particular case of (2.6) by considering the basis and its reciprocal identical to $|\tilde{\psi}_n\rangle, n = 1, \dots, \infty$. The equivalence between (2.6) and (2.8) holds only when both sums run to infinity. Because, on the one hand, if the sum in (2.8) is truncated up to N terms, we obtain an operator \hat{P} given by

$$\hat{P} = \sum_{n=1}^N |\tilde{\psi}_n\rangle \langle \tilde{\psi}_n |, \tag{2.9}$$

which is the orthogonal projector onto the subspace V_N spanned by N vectors $|\alpha_n\rangle, n = 1, \dots, N$. On the other hand, by truncating (2.6) up to N terms, one obtains an operator

$$\hat{Q} = \sum_{n=1}^N |\alpha_n\rangle \langle \tilde{\alpha}_n |, \tag{2.10}$$

which is idempotent, and hence a projector, but as it fails to be selfadjoint, it is not an orthogonal projector operator. As a consequence, the approximation of $|f\rangle$ that we

obtain by truncating the expansion (1.1) up to N terms is *not* the best approximation of $|f\rangle$ that can be obtained by a linear superposition of N vectors $|\alpha_n\rangle$. If one wishes to construct orthogonal projections by means of biorthogonal families, then biorthogonal vectors $|\tilde{\alpha}_n^N\rangle$, $n = 1, \dots, N$, specially devised for such a purpose must be constructed. The superscript N indicates that if the subspace V_N is enlarged (or reduced), each function should be recalculated.

2.1. Forward adaptive biorthogonalization. Let $|\alpha_n\rangle$ be a set of linearly independent vectors and let vectors $|\psi_n\rangle$ be obtained by orthogonalizing the formers in such a way that $|\psi_n\rangle = |\alpha_n\rangle - \hat{P}_{V_{n-1}}|\alpha_n\rangle$, where $\hat{P}_{V_{n-1}}$ is the orthogonal projector operator onto the subspace V_{n-1} spanned by $|\alpha_l\rangle$, $l = 1, \dots, n-1$. Then, it is proved in [5, 7] that vectors $|\tilde{\alpha}_n^{k+1}\rangle$ arising from $|\psi_1\rangle = |\alpha_1\rangle$ through the recursive equation

$$\begin{aligned}
 |\tilde{\alpha}_n^{k+1}\rangle &= |\tilde{\alpha}_n^k\rangle - \frac{|\psi_{k+1}\rangle}{\| |\psi_{k+1}\rangle \|^2} \langle \alpha_{k+1} | \tilde{\alpha}_n^k \rangle, \quad n = 1, \dots, k, \\
 |\tilde{\alpha}_{k+1}^{k+1}\rangle &= \frac{|\psi_{k+1}\rangle}{\| |\psi_{k+1}\rangle \|^2}
 \end{aligned}
 \tag{2.11}$$

are biorthogonal to vectors $|\alpha_n\rangle$, $n = 1, \dots, k+1$, and provide a representation of the orthogonal projection operator onto V_{k+1} , that is,

$$\hat{P}_{V_{k+1}} = \sum_{n=1}^{k+1} |\alpha_n\rangle \langle \tilde{\alpha}_n^{k+1} | = \hat{P}_{V_{k+1}}^\dagger = \sum_{n=1}^{k+1} |\tilde{\alpha}_n^{k+1}\rangle \langle \alpha_n|.
 \tag{2.12}$$

As discussed in [6], in order to reduce numerical errors, the vectors $|\psi_k\rangle$ are conveniently computed by modified Gram-Schmidt procedure or modified Gram-Schmidt procedure with pivoting [6, 9]. In some cases, it may be necessary to include a re-orthogonalization step [10].

Since the unique vector in V_{k+1} minimizing the distance to an arbitrary vector $|f\rangle \in \mathcal{H}$ is obtained by the operation $\hat{P}_{V_{k+1}}|f\rangle$ [5, 7], it follows from (2.12) that the coefficients c_n^{k+1} of the linear expansion

$$\sum_{n=1}^{k+1} c_n^{k+1} |\alpha_n\rangle,
 \tag{2.13}$$

which approximates an arbitrary $|f\rangle \in \mathcal{H}$ at best in a minimum distance sense, can be recursively obtained as

$$\begin{aligned}
 c_n^{k+1} &= c_n^k - \langle \tilde{\alpha}_n^k | \alpha_{k+1} \rangle \frac{\langle \psi_{k+1} | f \rangle}{\| |\psi_{k+1}\rangle \|^2}, \quad n = 1, \dots, k, \\
 c_{k+1}^{k+1} &= \frac{\langle \psi_{k+1} | f \rangle}{\| |\psi_{k+1}\rangle \|^2},
 \end{aligned}
 \tag{2.14}$$

with $c_1^1 = \langle \alpha_{l_1} | f \rangle / \| |\alpha_{l_1}\rangle \|^2$.

This technique, yielding forward approximations, has been shown to be of assistance in sparse signal representation by waveforms selection [7] as well as data subset selection [8]. Nevertheless, in those and other application areas, the need for a technique

yielding approximations in the opposite direction is clear. Hence, the motivation to the approach of the next section.

3. Backward adaptive biorthogonalization. Let V_{N/α_j} denote the subspace which is left by removing the vector $|\alpha_j\rangle$ from V_N , that is,

$$V_{N/\alpha_j} = \text{span}\{|\alpha_1\rangle, \dots, |\alpha_{j-1}\rangle, |\alpha_{j+1}\rangle, \dots, |\alpha_N\rangle\}, \quad (3.1)$$

and let $|\tilde{\alpha}_n^{N/j}\rangle, n = 1, \dots, j-1, j+1, \dots, N$, be the corresponding reciprocal family which allows expressing the orthogonal projector operator onto V_{N/α_j} as follows:

$$\hat{P}_{V_{N/\alpha_j}} = \sum_{\substack{n=1 \\ n \neq j}}^N |\alpha_n\rangle \langle \tilde{\alpha}_n^{N/j}| = \sum_{\substack{n=1 \\ n \neq j}}^N |\tilde{\alpha}_n^{N/j}\rangle \langle \alpha_n|. \quad (3.2)$$

Assuming that the biorthogonal vectors $|\tilde{\alpha}_n^N\rangle, n = 1, \dots, N$, yielding a representation of \hat{P}_{V_N} as given by

$$\hat{P}_{V_N} = \sum_{n=1}^N |\alpha_n\rangle \langle \tilde{\alpha}_n^N| = \sum_{n=1}^N |\tilde{\alpha}_n^N\rangle \langle \alpha_n|, \quad (3.3)$$

are known, our goal is to modify such vectors so as to obtain the corresponding set $|\tilde{\alpha}_n^{N/j}\rangle, n = 1, \dots, j-1, j+1, \dots, N$, yielding $\hat{P}_{V_{N/\alpha_j}}$ as in (3.2).

We start by writing

$$\hat{P}_{V_N} = \hat{P}_{V_{N/\alpha_j}} + \hat{P}_{V_{N/\alpha_j}^\perp}, \quad (3.4)$$

where $\hat{P}_{V_{N/\alpha_j}^\perp}$ is the orthogonal projector onto V_{N/α_j}^\perp , the orthogonal complement of V_{N/α_j} in V_N . Thus, V_{N/α_j}^\perp contains only one linear independent vector, arising by subtracting from $|\alpha_j\rangle$ its component in V_{N/α_j} , that is,

$$\hat{P}_{V_{N/\alpha_j}^\perp} = |\tilde{\psi}_j^f\rangle \langle \tilde{\psi}_j^f|, \quad (3.5)$$

where

$$|\tilde{\psi}_j^f\rangle = |\alpha_j\rangle - \hat{P}_{V_{N/\alpha_j}}|\alpha_j\rangle \quad (3.6)$$

and $|\tilde{\psi}_j^f\rangle = |\psi_j^f\rangle / \|\psi_j^f\|$.

(Note that we use the notation $|\psi_j^f\rangle$ to differentiate this new vector from the previous $|\psi_j\rangle$ introduced in Section 2.1.)

Using now (3.2), (3.3), and (3.5), we express (3.4) as follows:

$$\sum_{n=1}^N |\alpha_n\rangle \langle \tilde{\alpha}_n^N| = \sum_{\substack{n=1 \\ n \neq j}}^N |\alpha_n\rangle \langle \tilde{\alpha}_n^{N/j}| + |\tilde{\psi}_j^f\rangle \langle \tilde{\psi}_j^f|. \quad (3.7)$$

Taking the inner product of both sides of (3.7) with $\langle \tilde{\psi}_j^f |$, and using the fact that $\langle \tilde{\psi}_j^f | \alpha_n \rangle = 0$ for $n \neq j$, we obtain

$$\langle \tilde{\psi}_j^f | \alpha_j \rangle \langle \tilde{\alpha}_j^N | = \langle \tilde{\psi}_j^f |. \tag{3.8}$$

Moreover, since $\langle \tilde{\psi}_j^f | \alpha_j \rangle = \langle \alpha_j | \alpha_j \rangle - \langle \alpha_j | \hat{P}_{V_{N/\alpha_j}} | \alpha_j \rangle = \langle \tilde{\psi}_j^f | \psi_j^f \rangle$, it follows from (3.8) that $\| \langle \tilde{\psi}_j^f | \| = \| \langle \tilde{\alpha}_j^N | \|^{-1}$. Hence, the vector $\langle \tilde{\psi}_j^f |$ turns out to be

$$\langle \tilde{\psi}_j^f | = \frac{\langle \tilde{\alpha}_j^N |}{\| \langle \tilde{\alpha}_j^N | \|}. \tag{3.9}$$

Taking now the inner product of both sides of (3.7) with every $\langle \tilde{\alpha}_n^N |$, $n = 1, \dots, j - 1, j + 1, \dots, N$, we obtain the equation we wanted to find:

$$\langle \tilde{\alpha}_n^{N/j} | = \langle \tilde{\alpha}_n^N | - \langle \tilde{\alpha}_n^N | \tilde{\psi}_j^f \rangle \langle \tilde{\psi}_j^f |, \quad n = 1, \dots, j - 1, j + 1, \dots, N. \tag{3.10}$$

The following theorem demonstrates that the modification of vectors $|\tilde{\alpha}_n^N\rangle$ as prescribed in (3.10) provides us with biorthogonal vectors $|\tilde{\alpha}_n^{N/j}\rangle$, $n = 1, \dots, j - 1, j + 1, \dots, N$, rendering orthogonal projections.

THEOREM 3.1. *Given a set of vectors $|\tilde{\alpha}_n^N\rangle$, $n = 1, \dots, N$, biorthogonal to vectors $|\alpha_n\rangle$, $n = 1, \dots, N$, and yielding a representation of \hat{P}_{V_N} as given in (3.3), a new set of biorthogonal vectors $|\tilde{\alpha}_n^{N/j}\rangle$, $n = 1, \dots, j - 1, j + 1, \dots, N$, yielding a representation for $\hat{P}_{V_{N/\alpha_j}}$, as given in (3.2), can be obtained from the following equations:*

$$|\tilde{\alpha}_n^{N/j}\rangle = |\tilde{\alpha}_n^N\rangle - \frac{|\tilde{\alpha}_j^N\rangle \langle \tilde{\alpha}_j^N | \tilde{\alpha}_n^N \rangle}{\| |\tilde{\alpha}_j^N\rangle \|^2}, \quad n = 1, \dots, j - 1, j + 1, \dots, N. \tag{3.11}$$

PROOF. We first use (3.11) to write

$$\hat{P}_{V_{N/\alpha_j}} = \sum_{\substack{n=1 \\ n \neq j}}^N |\alpha_n\rangle \langle \tilde{\alpha}_n^{N/j} | = \sum_{\substack{n=1 \\ n \neq j}}^N |\alpha_n\rangle \langle \tilde{\alpha}_n^N | - \sum_{\substack{n=1 \\ n \neq j}}^N |\alpha_n\rangle \frac{\langle \tilde{\alpha}_n^N | \tilde{\alpha}_j^N \rangle \langle \tilde{\alpha}_j^N |}{\| |\tilde{\alpha}_j^N\rangle \|^2}. \tag{3.12}$$

To prove that (3.12) is the orthogonal projector onto V_{N/α_j} , we show that (a) $\hat{P}_{V_{N/\alpha_j}} |g\rangle = |g\rangle$ for all $|g\rangle \in V_{N/\alpha_j}$ and (b) $\hat{P}_{V_{N/\alpha_j}} |g^\perp\rangle = 0$ for all $|g^\perp\rangle$ in the orthogonal complement of V_{N/α_j} in \mathcal{H} .

Indeed, every $|g\rangle \in V_{N/\alpha_j}$ can be expressed as a linear combination $|g\rangle = \sum_{\substack{n=1 \\ n \neq j}}^N a_n |\alpha_n\rangle$ for some coefficients a_n , $n = 1, \dots, j - 1, j + 1, \dots, N$, and since, by hypothesis, $\langle \tilde{\alpha}_n^N | \alpha_l \rangle = \delta_{n,l}$, from (3.12), we have

$$\hat{P}_{V_{N/\alpha_j}} |g\rangle = \sum_{\substack{n=1 \\ n \neq j}}^N \sum_{\substack{l=1 \\ l \neq j}}^N |\alpha_n\rangle a_l \langle \tilde{\alpha}_n^N | \alpha_l \rangle = \sum_{\substack{n=1 \\ n \neq j}}^N a_n |\alpha_n\rangle = |g\rangle, \tag{3.13}$$

which proves (a).

To prove (b), we write $\sum_{n \neq j}^N |\alpha_n\rangle \langle \tilde{\alpha}_n^N| = \sum_{n=1}^N |\alpha_n\rangle \langle \tilde{\alpha}_n^N| - |\alpha_j\rangle \langle \tilde{\alpha}_j^N| = \hat{P}_{V_N} - |\alpha_j\rangle \langle \tilde{\alpha}_j^N|$ and recast (3.12) as follows:

$$\begin{aligned} \hat{P}_{V_N/\alpha_j} &= \hat{P}_{V_N} - |\alpha_j\rangle \langle \tilde{\alpha}_j^N| - \frac{\hat{P}_{V_N} |\tilde{\alpha}_j^N\rangle \langle \tilde{\alpha}_j^N|}{\| |\tilde{\alpha}_j^N\rangle \|^2} + \frac{|\alpha_j\rangle \langle \tilde{\alpha}_j^N| \tilde{\alpha}_j^N \langle \tilde{\alpha}_j^N|}{\| |\tilde{\alpha}_j^N\rangle \|^2} \\ &= \hat{P}_{V_N} - \frac{\hat{P}_{V_N} |\tilde{\alpha}_j^N\rangle \langle \tilde{\alpha}_j^N|}{\| |\tilde{\alpha}_j^N\rangle \|^2}. \end{aligned} \tag{3.14}$$

Now, since $\langle \alpha_n | g^\perp \rangle = \delta_{n,j} \langle \alpha_j | g^\perp \rangle$, $n = 1, \dots, N$, it follows that $\hat{P}_{V_N} |g^\perp\rangle = |\tilde{\alpha}_j^N\rangle \langle \alpha_j | g^\perp \rangle$ and, since $|\tilde{\alpha}_j^N\rangle \in V_N$, it follows that $\langle \tilde{\alpha}_j^N | g^\perp \rangle = \langle \tilde{\alpha}_j^N | \hat{P}_{V_N} |g^\perp \rangle = \langle \tilde{\alpha}_j^N | \tilde{\alpha}_j^N \rangle \langle \alpha_j | g^\perp \rangle$. Hence,

$$\begin{aligned} \hat{P}_{V_N/\alpha_j} |g^\perp\rangle &= \hat{P}_{V_N} |g^\perp\rangle - \frac{|\tilde{\alpha}_j^N\rangle \langle \tilde{\alpha}_j^N | \hat{P}_{V_N} |g^\perp\rangle}{\| |\tilde{\alpha}_j^N\rangle \|^2} \\ &= |\tilde{\alpha}_j^N\rangle \langle \alpha_j | g^\perp \rangle - |\tilde{\alpha}_j^N\rangle \langle \alpha_j | g^\perp \rangle = 0. \end{aligned} \tag{3.15}$$

The biorthogonality property of vectors $|\tilde{\alpha}_n^{N/j}\rangle$, $n = 1, \dots, j-1, j+1, \dots, N$, is an immediate consequence of the biorthogonality property of vectors $|\tilde{\alpha}_n^N\rangle$, as readily follows by taking the inner product of both sides of (3.11) with each vector $\langle \alpha_n |$, $n = 1, \dots, j-1, j+1, \dots, N$.

Since $\hat{P}_{V_N/\alpha_j} = \sum_{n \neq j}^N |\alpha_n\rangle \langle \tilde{\alpha}_n^N|$ has been proved to be a projector, it is selfadjoint. Hence, (3.2) holds. □

COROLLARY 3.2. *Let $|f_N\rangle$ be the orthogonal projection of an arbitrary $|f\rangle \in \mathcal{H}$ onto V_N , that is,*

$$|f_N\rangle = \hat{P}_{V_N} |f\rangle = \sum_{n=1}^N c_n^N |\alpha_n\rangle, \tag{3.16}$$

with $c_n^N = \langle \tilde{\alpha}_n^N | f \rangle$, $n = 1, \dots, N$, assumed to be known. Hence, the coefficients $c_n^{N/j}$ of the orthogonal projection of $|f\rangle$ onto V_N/α_j are obtained from the known coefficients c_n^N as follows:

$$c_n^{N/j} = c_n^N - \frac{\langle \tilde{\alpha}_n^N | \tilde{\alpha}_j^N \rangle c_j^N}{\| |\tilde{\alpha}_j^N\rangle \|^2}. \tag{3.17}$$

The proof trivially stems from (3.12) since $\hat{P}_{V_N/\alpha_j} |f\rangle = \sum_{n \neq j}^N c_n^{N/j} |\alpha_n\rangle$ implies $c_n^{N/j} = \langle \tilde{\alpha}_n^{N/j} | f \rangle$.

COROLLARY 3.3. *For $|f\rangle \in \mathcal{H}$, let $|f_N\rangle$ be as above and $|f_{N/j}\rangle = \hat{P}_{V_N/\alpha_j} |f\rangle$. Then, the following relation between $\| |f_N\rangle \|$ and $\| |f_{N/j}\rangle \|$ holds:*

$$\| |f_{N/j}\rangle \|^2 = \| |f_N\rangle \|^2 - \frac{|c_j^N|^2}{\| |\tilde{\alpha}_j^N\rangle \|^2}. \tag{3.18}$$

PROOF. Using (3.14) and the fact that orthogonal projectors are selfadjoint and idempotent, it follows that

$$\begin{aligned} \| |f_{N/j}\rangle \|^2 &= \langle f | \hat{P}_{V_N/\alpha_j} | f \rangle = \langle f | \hat{P}_{V_N} | f \rangle - \frac{\langle f | \tilde{\alpha}_j^N \rangle \langle \tilde{\alpha}_j^N | f \rangle}{\| | \tilde{\alpha}_j^N \rangle \|^2} \\ &= \| |f_N\rangle \|^2 - \frac{|c_j^N|^2}{\| | \tilde{\alpha}_j^N \rangle \|^2}. \end{aligned} \tag{3.19}$$

□

So far, we have discussed how to modify the coefficients of a linear expansion when one of its components is removed. Nevertheless, we have given no specification on how to choose such an element. We are now in a position to address this point since Corollary 3.3 suggests how the selection could be made optimal. The following proposition is in order.

PROPOSITION 3.4. *Let*

$$|f_N\rangle = \hat{P}_{V_N} |f\rangle = \sum_{n=1}^N c_n^N | \alpha_n \rangle \tag{3.20}$$

be given by the coefficients c_n^N , $n = 1, \dots, N$, and let

$$|f_{N/j}\rangle = \hat{P}_{V_N/\alpha_j} |f\rangle = \sum_{\substack{n=1 \\ n \neq j}}^N c_n^{N/j} | \alpha_n \rangle \tag{3.21}$$

be obtained by eliminating the coefficient c_j^N from (3.20) and modifying the remaining coefficients as prescribed in (3.17). The coefficient c_j^N to be removed as minimizing the norm of the residual error $|\Delta\rangle = |f_N\rangle - |f_{N/j}\rangle$ is the one yielding a minimum value of

$$\frac{|c_j^N|^2}{\| | \tilde{\alpha}_j^N \rangle \|^2}. \tag{3.22}$$

PROOF. Since $\hat{P}_{V_N} \hat{P}_{V_N/\alpha_j} = \hat{P}_{V_N/\alpha_j} \hat{P}_{V_N} = \hat{P}_{V_N/\alpha_j}$, we have

$$\| |f_N\rangle - |f_{N/j}\rangle \|^2 = \langle f | \hat{P}_{V_N} | f \rangle - \langle f | \hat{P}_{V_N/\alpha_j} | f \rangle = \| |f_N\rangle \|^2 - \| |f_{N/j}\rangle \|^2. \tag{3.23}$$

Hence, making use of (3.18), we further have

$$\| |f_N\rangle - |f_{N/j}\rangle \|^2 = \frac{|c_j^N|^2}{\| | \tilde{\alpha}_j^N \rangle \|^2}, \tag{3.24}$$

from which we gather that $\| |f_N\rangle - |f_{N/j}\rangle \|^2$ is minimum if $|c_j^N|^2 / \| | \tilde{\alpha}_j^N \rangle \|^2$ is minimum. □

Proposition 3.4 is relevant to backward approximation of a signal, a common procedure in compression and noise reduction techniques. The goal in this case is to diminish

the number of coefficients so as to have a more economical representation and/or reduce spurious information (noise). Successive applications of criterion (3.22) leads to an algorithm for recursive coarser approximations. Indeed, we assume that at the first iteration, we eliminate the j th term yielding a minimum of (3.22). We then construct the new reciprocal vectors as prescribed in (3.11) and the corresponding new coefficients as prescribed in (3.17). We are thus in a position to repeat the process and obtain a coarser approximation of the previous one. If we denote by $|f^{(k)}\rangle$ the approximation arising at the k -step, a common stopping criterion would imply ceasing the iteration process when the following situation is reached:

$$\| |f\rangle - |f^{(k)}\rangle \|^2 > \delta, \tag{3.25}$$

where δ is estimated according to the desired precision. If the aim is to denoise a signal, the value of δ may be set as the variance of the noise, when available. It is appropriate to remark, however, that in the context of some applications, the selection criterion (3.22) may not be the adequate one. Instead, other criteria based on statistical properties may be required [1, 3, 4]. In any case, regardless of the criterion for selecting the coefficient c_j^N to be deleted, if one wishes the remaining ones to yield the optimal approximation in a minimum distance sense, such coefficients should be modified as indicated in (3.17). We illustrate next, by a simple example, the gain that results in following this prescription.

We consider $N = 13$ elements $|\alpha_n\rangle$, $n = 1, \dots, 13$, whose functional representations are given by the following shifted Mexican hat wavelet:

$$\begin{aligned} \alpha_{2n}(t) &= \langle t | \alpha_{2n} \rangle = A_{2n} e^{-(t+2n)^2} (1 - (t + 2n)^2), \\ \alpha_{2n+1}(t) &= \langle t | \alpha_{2n+1} \rangle = A_{2n+1} e^{-(t-2n-1)^2} (1 - (t - 2n - 1)^2), \quad n = 0, \dots, 6, \end{aligned} \tag{3.26}$$

where each A_n is a constant which normalizes the corresponding function $|\alpha_n\rangle$ to unity in the interval $[-4, 4]$. We construct the biorthogonal functions $\tilde{\alpha}_n^{13}(t)$ (assumed to be known) by applying the forward biorthogonalization technique of Section 2.1, but they could be constructed by any other available method.

The signal $f(t)$ is considered to be the function plotted by the continuous line of Figure 3.1. Such a signal is also expressible by a linear combination of the waveforms given in (3.26). A high-quality fitting results by using the corresponding 13 coefficients c_n^{13} , each of which is calculated as $c_n^{13} = \langle \tilde{\alpha}_n^{13} | f \rangle$. The actual numbers are as follows: 2.8273, 2.4954, 2.4954, 1.9988, 1.9988, 1.4989, 1.4989, 0.8630, 0.8630, 0.2957, 0.2957, 0.0648, 0.0648. We now disregard the last two coefficients, the ones of smallest magnitude, and use the remaining ones without modification. Although the neglected coefficients are quite small in comparison with some of the others, the approximation that results, represented by the dotted line in Figure 3.1, does not correctly fit the distribution tails. Nevertheless, if we disregard the same coefficients but modify the others by applying (3.17) twice, the resulting approximation happens to coincide with the continuous line of Figure 3.1. To magnify the effect we wish to show, we now disregard two more coefficients, those of value 0.2957. The approximation that results from a simple truncation is shown by the darker dotted line in Figure 3.2. The lighter dotted line plots

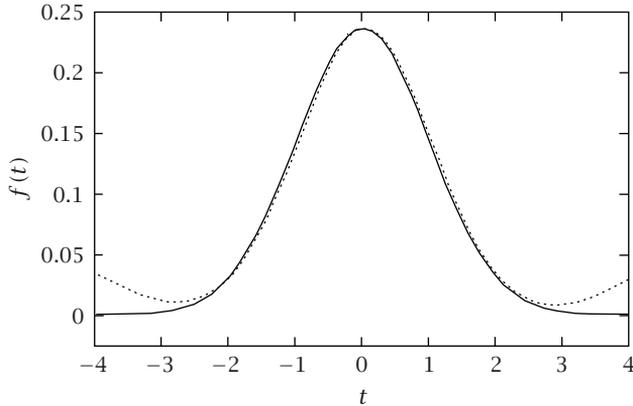


FIGURE 3.1. The continuous line represents a signal $f(t)$, which is also expressible as a linear combination of the 13 waveforms given in (3.26). The dotted line is the approximation arising by disregarding two coefficients in such linear expansion. Our approach coincides with the continuous line.

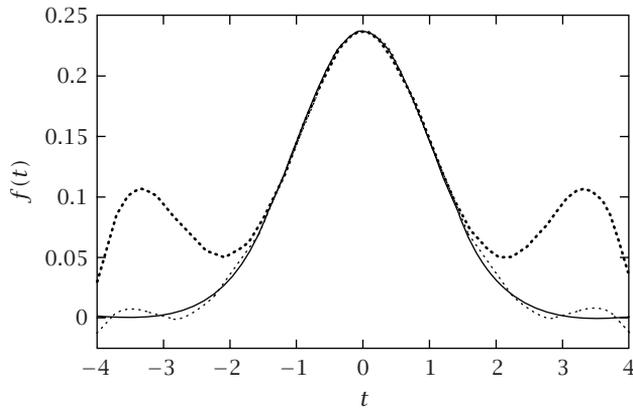


FIGURE 3.2. Here the darker dotted line is obtained by disregarding 4 coefficients. Our approach is represented by the lighter dotted line.

our approximation. This very simple example clearly illustrates the significance of the proposed modification of coefficients.

4. Conclusions. A recursive approach for adapting biorthogonal functions so as to obtain orthogonal projections onto a reduced subspace has been proposed. The required modifications are simple and easy to implement. The modified functions are used to adapt coefficients of a lower-order linear model in order to obtain an optimal approximation in a minimum distance sense.

A criterion for disregarding coefficients has been discussed. Such a criterion leads to an iterative procedure for successive backward approximations which yields, at each

iteration, a minimal residual norm. It should be stressed that, regardless of the criterion used for neglecting coefficients, the proposed approach may be applied to guarantee optimality (in a minimum distance sense) of the remaining approximation. We believe, thereby, that this technique is potentially applicable to a broad range of problems including data compression, noise reduction, and sparse representation.

MATLAB codes for implementation of the proposed approach can be found at <http://www.ncrg.aston.ac.uk/Projects/BiOrthog>

ACKNOWLEDGMENTS. Support from the Engineering and Physical Sciences Research Council (EPSRC) (GR/R86355/01) is acknowledged. I wish to thank Dr. S. Jain for corrections to the manuscript.

REFERENCES

- [1] A. Antoniadis and G. Oppenheim (eds.), *Wavelets and Statistics*, Lecture Notes in Statistics, vol. 103, Springer-Verlag, New York, 1995.
- [2] P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford University Press, London, 1958.
- [3] M. Jansen, *Noise Reduction by Wavelet Thresholding*, Lecture Notes in Statistics, vol. 161, Springer-Verlag, New York, 2001.
- [4] I. M. Johnstone and B. W. Silverman, *Wavelet threshold estimators for data with correlated noise*, J. Roy. Statist. Soc. Ser. B **59** (1997), no. 2, 319–351.
- [5] L. Rebollo-Neira, *Recursive biorthogonalization approach and orthogonal projectors*, On Non-Orthogonal Signal Representation, Progress in Mathematics Physics, Nova Science, New York, in press.
- [6] ———, *Dictionary redundancy elimination*, IEE Proceedings-Vision, Image and Signal Processing **151** (2004), no. 1, 31–34.
- [7] L. Rebollo-Neira and D. Lowe, *Optimized orthogonal matching pursuit approach*, IEEE Signal Processing Letters **9** (2002), no. 4, 137–140.
- [8] L. Rebollo-Neira and A. Plastino, *Recursive approach for constructing the $q = 1/2$ maximum entropy distribution from redundant data*, Phys. Rev. E **66** (2002), no. 3, 032102.
- [9] J. R. Rice, *Experiments on Gram-Schmidt orthogonalization*, Math. Comp. **20** (1966), 325–328.
- [10] D. S. Watkins, *Fundamentals of Matrix Computations*, Pure and Applied Mathematics, Wiley-Interscience, New York, 2002.
- [11] R. M. Young, *An Introduction to Nonharmonic Fourier Series*, Pure and Applied Mathematics, vol. 93, Academic Press, New York, 1980.

L. Rebollo-Neira: Neural Computing Research Group, Aston University, Birmingham B4 7ET, UK

E-mail address: rebol101@aston.ac.uk