A Matrix-Oriented Approach for Analysis and Optimisation of Block Digital Filters

Gilles Burel

in "Advances in Simulation, Systems Theory, and Systems Engineering", N. Mastorakis, V. Kluev, D. Koruga Editors, Published by WSEAS Press, 2002, pp. 245-249, ISBN 960-8052-70-X

A Matrix-Oriented Approach for Analysis and Optimisation of Block Digital Filters

Gilles BUREL
LEST, Université de Bretagne Occidentale
BP 809, 29285 Brest cedex, France
Gilles.Burel@univ-brest.fr   http://www.univ-brest.fr/lest/tst/

Abstract : — Block Digital Filtering (BDF) is a well known method for fast digital filtering. Indeed, it decomposes the input signal into blocks and takes profit of complexity reduction provided by the Fast Fourier Transform algorithm to considerably reduce the computational cost. In this paper, we propose a simple and efficient matrix-oriented approach to compute and visualize the frequency response of a BDF, as well as to optimise the BDF coefficients. This approach can be considered as an interesting alternative to the traditional overlap-save method, because it provides better frequency responses.

Key-Words : — Block Digital Filters, Overlap-Save, Optimal filter, Aliasing, Time-varying response.

1 Introduction

In many signal processing applications, fast digital filtering of a signal is required [1][2]. The usual solution to reduce computational complexity of digital filters is to take profit of the complexity reduction provided by the FFT (Fast Fourier Transform) algorithm, as illustrated on figure 1.

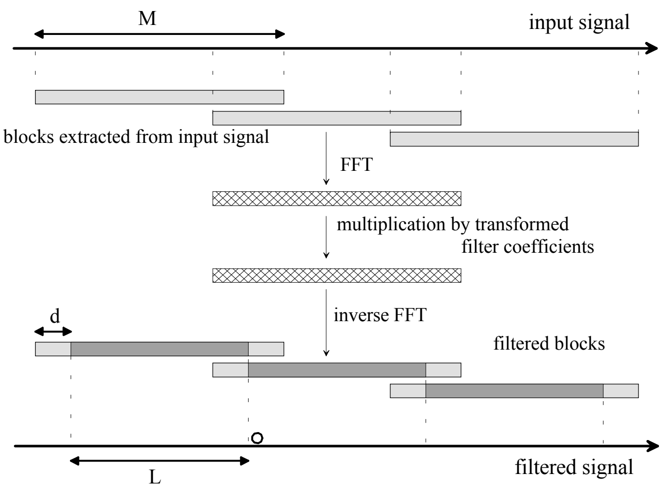

Figure 1: Principle of a Block Digital Filter

An FFT-based Block Digital Filter (BDF) works as follows [6]. The signal is divided into overlapping blocks of M samples (the amount of overlapping is $M - L$). Then, for each block, the following computations are performed:

- The block is transformed through an M-points FFT;
- The transformed block is multiplied by the M-points FFT of the filter impulse response;
- An inverse M-points FFT is performed to obtain the filtered block;
- Only the L central points of the transformed block are kept.

The concatenation of the L points kept in each filtered block is the filtered signal. Under certain conditions, the obtained signal is equal to the signal that would be obtained by direct filtering. The interest of the approach is the large complexity reduction provided by the FFT algorithm.

However, the obtained signal is correct only if the filter impulse response is restricted to the interval $[-d, +d]$, where $d = (M - L)/2$. To illustrate this point, let us consider the sample of the filtered signal which is marked by a small circle on figure 1. If the filter extension is larger than d, this sample depends partially on samples of the input signal which are outside the input block. Hence, it is clear that the block digital filter cannot compute its correct value. In other words, the problem is to ensure equality between the linear convolution (provided by direct filtering) and the circular convolution (provided by the block digital filter). If this is not realised, it can easily be proved that the BDF impulse response is not time-invariant, hence aliasing is present [3][5].

This paper was initially presented at WSEAS Int. Conf. on Signal, Speech and Image Processing (ICOSSIP 2002), Skiathos Island, Greece, September 25-28, 2002, then selected for publication in the book referenced above.

The method known as Overlap-Save ([4] p. 558) is based on the principle shown on figure 1. The filter impulse response is restricted to the interval $[-d, +d]$ in order to ensure that the obtained filtered signal is exactly the same as the signal that would be obtained by direct filtering.

In this paper, we use a different approach, in which difference between linear convolution and circular convolution (i.e. aliasing) is tolerated. The interest may not appear at a first glance, and a classical objection could be: why use an imperfect method when an aliasing-free method (overlap-save) is known? In fact, this would be forgetting that a good filter is a filter which frequency response is as close as possible to the desired frequency response, and not necessarily a filter that is perfect in the sense of equivalence between linear and circular convolutions.

The interest of tolerating differences between linear convolution and circular convolution is that the filter impulse response is not restricted anymore to the interval $[-d, +d]$. Hence, it is possible to better match the desired frequency response. In this paper, we propose a fast method to compute the optimal coefficients of such a filter using a mean square criterion. Since the filter is optimal, it is always better than overlap-save, with the same computational cost.

Simple matrix-oriented tools are proposed in this paper to evaluate and design block digital filters. The article is organised as follows. In Section 2, a matrix-oriented mathematical model of a BDF is provided. Then, in Section 3, the frequency response of a BDF is established, and a method for fast computation of the optimal BDF parameters is proposed in Section 4. Finally, experimental results are provided in Section 5 to illustrate the approach, and a conclusion is drawn in Section 6.

2 Mathematical model

Let us define some notations:

- e: an M-dimensional vector containing the samples of an input block;
- f: an L-dimensional vector containing the samples of an output block;
- W_M: the $M \times M$ unitary matrix corresponding to the normalised (i.e. norm preserving) M-points DFT (Discrete Fourier Transform) matrix;
- G: a diagonal $M \times M$ matrix, the diagonal of which contains the filter coefficients;
- S: the $L \times M$ selection matrix.

The selection matrix is defined as:

$$S = [0_{L\times b} \; I_{L\times L} \; 0_{L\times b}] \tag{1}$$

That is b null columns, an $L \times L$ identity matrix, and b null columns again. From figure 1 it is clear that we have:

$$f = Ae \tag{2}$$

where A is the $L \times M$ matrix below (* stands for the conjugate transpose):

$$A = SW_M^* G W_M \tag{3}$$

Let us note $K = bL$, where b is an integer, the number of points we want to consider on the frequency response. Due to the well known properties of the DFT, this allows considering periodical signals x of period K (indeed, with a K-points frequency resolution, considering signals of period K or non-periodical signals provide the same result). The filtered signal is:

$$y = Hx \tag{4}$$

where x and y are K-dimensional vectors and H is a $K \times K$ matrix containing b copies of matrix A, as shown on figure 2.

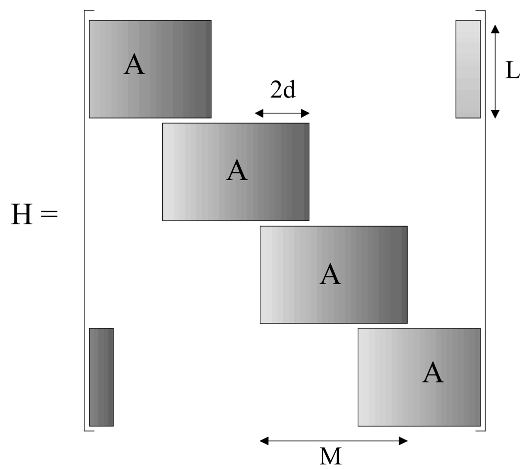

Figure 2: Structure of matrix H (for $b = 4$)

3 Frequency response of a BDF

Let us note $\overline{y}$ and $\overline{x}$ the DFTs of y and x. The frequency response of the BDF will be characterised by a $K \times K$ matrix $\overline{\overline{H}}$ such that:

$$\overline{y} = \overline{\overline{H}}\,\overline{x} \tag{5}$$

From equation 4 we have:

$$\begin{aligned} W_K\, y &= W_K\, H\, x && (6)\\ &= \left(W_K\, H\, W_K^*\right)\,(W_K\, x) && (7)\end{aligned}$$

Hence:

$$\overline{\overline{H}} = W_K\, H\, W_K^* \tag{8}$$

4 Computation of the optimal filter

Let us note h_d the K-dimensional vector containing the desired frequency response, and $\overline{\overline{H}}_d$ a diagonal matrix, the diagonal of which is h_d. The quadratic error e_Q is a good measure of the quality of the obtained frequency response:

$$e_Q = \left\| \overline{\overline{H}} - \overline{\overline{H}}_d \right\|^2 \tag{9}$$

where $\|.\|^2$ stands for the Frobenius norm. Using equation 8 and the norm preserving properties of unitary matrices, we can write:

$$\begin{aligned} e_Q &= \left\| W_K H W_K^* - \overline{\overline{H}}_d \right\|^2 && (10)\\ &= \left\| H - W_K^* \overline{\overline{H}}_d W_K \right\|^2 && (11)\\ &= \|H - H_d\|^2 && (12)\end{aligned}$$

From the structure of H (see figure 2), it is clear that minimizing e_Q is equivalent to minimizing $\|A - A_d\|^2$ (A_d is extracted from H_d, according to the structure shown on the figure). Then, we can write:

$$\begin{aligned} \|A - A_d\|^2 &= \left\| S W_M^* G W_M - A_d \right\|^2 && (13)\\ &= \left\| W_M S^* S W_M^* G - W_M S^* A_d W_M^* \right\|^2 && (14)\\ &= \|PG - Q\|^2 && (15)\\ &= \sum_{n=1}^{M} \|g_{nn} p_n - q_n\|^2 && (16)\end{aligned}$$

where:

$$P = W_M S^* S W_M^* \tag{17}$$

and

$$Q = W_M S^* A_d W_M^* \tag{18}$$

Vectors p_n and q_n stand, respectively, for column n of P and Q. Using the pseudo-inverse, we obtain the optimal coefficients of matrix G:

$$g_{nn} = \|p_n\|^{-2}\ (p_n)^*\ q_n \tag{19}$$

Since $P^* P = P$ and $P^* Q = Q$, this equation can be simplified:

$$g_{nn} = \frac{q_{nn}}{p_{nn}} \tag{20}$$

This matrix-oriented method to compute the optimal filter is simple and can be easily programmed. Using matrix-based languages, such as Matlab or Octave, it requires only a few lines of program. Furthermore, the method is very fast because it requires no iterative algorithm and uses FFT. Let us summarize it:

- Choose the desired filter frequency response h_d (K-dimensional vector);
- Build a $K \times K$ diagonal matrix $\overline{\overline{H}}_d$, the diagonal of which is h_d;
- Compute $H_d = W_K^* \overline{\overline{H}}_d W_K$ (K-points FFT on the columns of $\overline{\overline{H}}_d$ followed by a K-points inverse FFT on the rows of the obtained matrix);
- From H_d extract A_d (from any location shown on figure 2);
- Compute $P = W_M S^* S W_M^*$ and $Q = W_M S^* A_d W_M^*$ (please note that P could be precomputed because it does not depend on $\overline{\overline{H}}_d$). For fast computation, remind that left multiplication by W_M is equivalent to an M-points FFT on the rows, and right-multiplication by W_M^* is equivalent to an M-points inverse FFT on the columns;
- Compute matrix G. It is a diagonal matrix, the diagonal elements of which are $g_{nn} = q_{nn}/p_{nn}$.

5 Experimental results

In order to illustrate the approach and to make visualisations easier, we use a relatively small block size: $M = 32$. For each block, $L = 24$ samples are kept. The frequency response is analysed with $K = 72$ points (that is, here, $K = 3L$).

The signal sampling frequency is $F_e = 100\,kHz$. The desired frequency response is a raised cosine with centre frequency $25\,kHz$, roll-off 0.2, and bandwidth $20\,kHz$.

Using the approach described above, we can compute the optimal matrix G. Figure 3 shows the values obtained on the diagonal of G. For comparison, the values obtained by overlap-save are also shown.

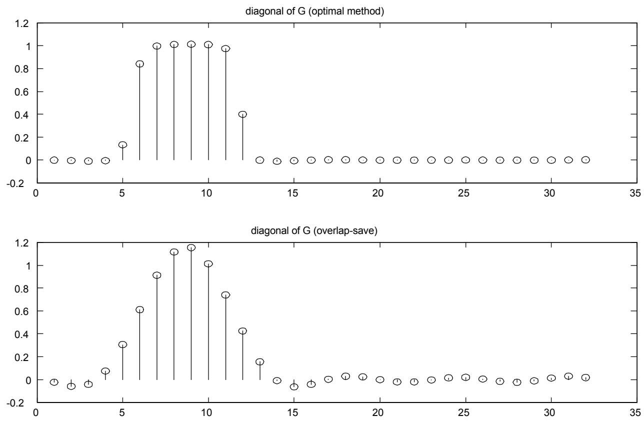

Figure 3: Diagonal of matrix G (top: optimal method, bottom: overlap-save).

From matrix G, we can compute matrix A (see Eq. 3), which is shown on figure 4. The gray level is an increasing function of the modulus of the corresponding matrix element.

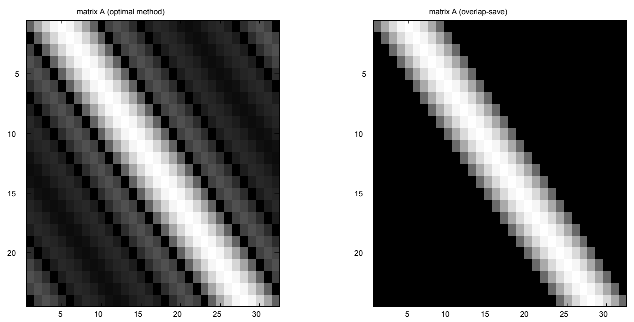

Figure 4: Matrix A (left: optimal method, right: overlap-save)

Matrix A determines matrix H, as was explained on figure 2. Here, we have $b = K/L = 3$ and the obtained matrices H are shown on figure 5. We remind that matrix H determines the link between the input signal x and the output signal y (see Eq. 4). We can see that overlap-save provides a matrix H with all rows equal (up to a circular permutation). As a consequence, the impulse response of the obtained BDF is time-invariant. This is not the case with matrix H provided by the optimal method. As a consequence, the BDF impulse response is time-varying. This may appear as a drawback of the optimal method. However, as we will see below, the global performances are better with this method.

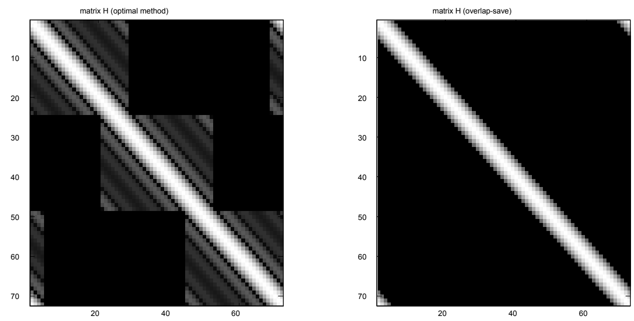

Figure 5: Matrix H (left: optimal method, right: overlap-save).

From matrix H, we can compute matrix $\overline{\overline{H}}$ using Eq. 8. Figure 6 shows matrix $\overline{\overline{H}}$, for both methods. We remind that this matrix determines the link between the input and output spectra (see Eq. 5). As expected, overlap-save provides a diagonal matrix (hence, there is no aliasing), while the optimal method provides a non-diagonal one (hence, aliasing is present).

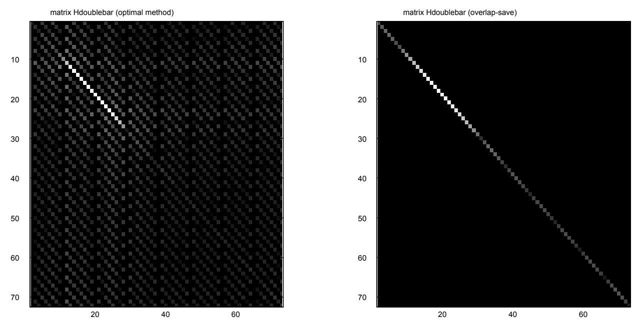

Figure 6: Bidimensional frequency response (matrix $\overline{\overline{H}}$). Left: optimal method. Right: overlap-save.

The table below shows the diagonal and off-diagonal quadratic error (between the obtained and desired frequency response), and the global quadratic error.

method	optimal	overlap-save
diagonal error	0.38	0.88
off-diagonal error	0.14	0
global error	0.52	0.88

We can see that tolerating a small amount of off-diagonal error allows the optimal method to provide a large improvement on time-invariant (i.e. diagonal) error.

Figure 7 shows the obtained and desired time-invariant frequency responses (i.e. the diagonals of matrices $\overline{\overline{H}}$ and $\overline{\overline{H}}_d$) for the optimal method and the overlap-save. It is clear that the optimal method provides the response that best matches the desired one.

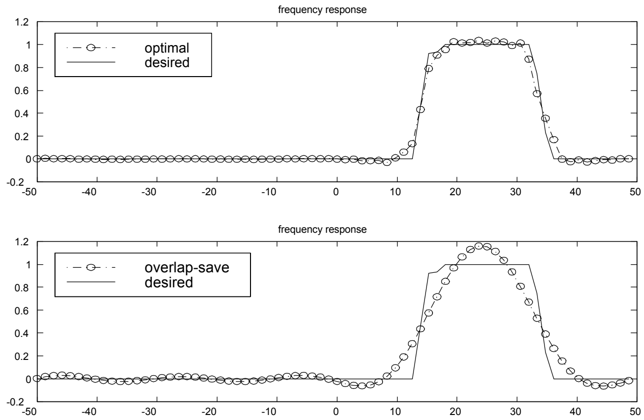

Figure 7: Obtained and desired time-invariant frequency response (top: optimal method, bottom: overlap-save).

6 Conclusion

In this article, we have described a simple and efficient matrix-oriented method to analyse and optimise block digital filters. We have also shown that tolerating a small amount of aliasing yields a large improvement of obtained frequencies responses. Hence, except for applications for which there is a good reason to refuse to tolerate a small amount of aliasing, this simple approach is an interesting alternative to overlap-save because it provides improvement of obtained frequency response, without increasing complexity.

References

[1] I.S. Lin, S.K. Mitra, “Overlapped Block Digital Filtering”, IEEE Trans. on Circuits and Systems - II: Analog and Digital Signal Processing, Vol. 43, No. 8, August 1996

[2] C.M. Loeffler, C.S. Burrus, “Optimal Design of Periodically Time-Varying and Multirate Digital Filters”, IEEE Trans. on Acoustic, Speech, and Signal Processing, Vol. 32, No. 5, October 1984

[3] H.S. Malvar, “Signal Processing with Lapped Transforms”, Artech House, 1992, ISBN 089006-467-9

[4] Oppenheim and Shaffer, “Discrete-time signal processing”, Prentice-Hall, 1989

[5] R.G. Shenoy, D. Burnside, T.W. Parks, “Linear Periodic Systems and Multirate Filter Design”, IEEE Trans. on Signal Processing, Vol. 42, No. 9, September 1994

[6] P.P. Vaidyanathan, “Multirate Systems and Filters”, Prentice Hall Signal Processing series (1993).