# DESIGN OF REDUCED COMPLEXITY LINEAR-PHASE POLYPHASE FIR FILTERS USING MIXED INTEGER LINEAR PROGRAMMING

*Oscar Gustafsson and Lars Wanhammar*

Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, SWEDEN
Tel: +46-13-28 {4059, 1344}, Fax: +46-13-13 92 82, E-mail: {oscarg, larsw}@isy.liu.se

## ABSTRACT

In this work a mixed integer linear programming (MILP) formulation for the design of a class of linear-phase FIR filters are presented. The formulation can be solved using general purpose MILP solvers to obtain filter implementation with a minimum number of signed-power-of-two (SPT) terms given a filter specification. The filter structures considered are based on reduced complexity polyphase decomposition. It is shown that the total number of SPT terms per sample can be reduced using this filter architecture. However, the savings are not as large as other work propose, when optimal design techniques are used.

## 1. INTRODUCTION

The problem of finding fixed-point multiplier coefficients for digital filters has been extensively studied during the years [1]. A large part of the work has considered finding filter coefficients with few signed-power-of-two (SPT) terms. This is motivated by that the number of additions required to realize a multiplication is the number of SPT terms minus one.

On the other hand, work has been done that consider the number of multiplications in an FIR filter. By introducing new filter structures, the number of arithmetic operations has been reduced [2].

In this work we present an approach to combine some of these two efforts. We formulate a mixed integer linear programming (MILP) problem for a class of reduced complexity polyphase FIR filters. These filter architectures are sometimes referred to as fast FIR algorithms (FFA) or parallel FIR filters [3]–[5]. Solving the MILP problem leads to a minimum number of SPT terms given a filter specification.

## 2. REDUCED COMPLEXITY POLYPHASE FIR FILTERS

The output of an FIR filter of order $N$ can be written as

$$y(n) = \sum_{m=0}^{N} h_m x(n-m) \tag{1}$$

Consider the same filter using polyphase decomposition with a factor of two. This can now be written as

$$\begin{aligned} Y_0 + z^{-1}Y_1 &= (H_0 + z^{-1}H_1)(X_0 + z^{-1}X_1) \\ &= H_0X_0 + z^{-1}(H_0X_1 + H_1X_0) + z^{-2}H_1X_1 \end{aligned} \tag{2}$$

where $H_0$, $H_1$, $X_0$, $X_1$, $Y_0$, and $Y_1$ are functions of $z^2$. Identification gives

$$Y_0 = H_0X_0 + z^{-2}H_1X_1 \tag{3}$$

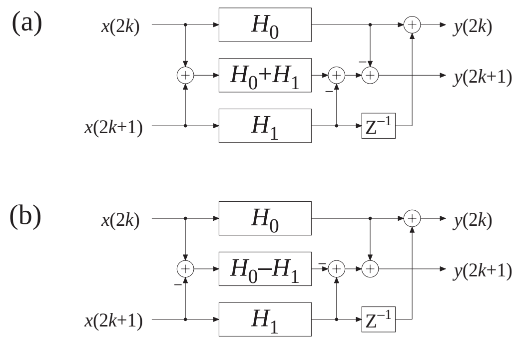


**Figure 1.** Reduced complexity polyphase FIR filter architectures considered in this work.

and

$$Y_1 = H_0X_1 + H_1X_0 \tag{4}$$

Equation (4) can now be rewritten as

$$Y_1 = (H_0 + H_1)(X_0 + X_1) - H_0X_0 - H_1X_1 \tag{5}$$

or

$$Y_1 = -(H_0 - H_1)(X_0 - X_1) + H_0X_0 + H_1X_1 \tag{6}$$

resulting in the filter architectures in Fig. 1 (a) and Fig. 1 (b), respectively [3]–[5].

This leads to that the number of multiplications has decreased from $N + 1$ per sample to $3N/4 + 1$ for even $N$ and $3(N + 1)/4$ for odd $N$. The number of additions per sample has decreased from $N$ to $3N/4$ for even $N$ and $3N/4 + 1/2$ for odd $N$.

However, for linear-phase FIR filters the coefficient symmetry can be utilized to decrease the number of multiplications per sample to $N/2 + 1$ for even $N$ and $(N + 1)/2$ for odd $N$ using a direct realization.

If the considered reduced complexity polyphase FIR filter architecture is used all subfilters will not have symmetric coefficients. The number of multiplications per sample is now $N/2 + 1$ for even $N$, while for odd $N$ it is $5N/8 + 1/2$. Hence, no reduction in the number of multiplications per sample is obtained using the polyphase architecture for linear-phase FIR filters. For odd order filters the number of multiplications per sample is even increasing. The number of additions is not changed, so the number of additions per sample is decreased.

However, the multiplication coefficients in the filters $H_0 + H_1$ and $H_0 - H_1$ are likely to have a shorter wordlength on average, so the arithmetic complexity can still be decreased using the considered reduced complexity polyphase architecture for linear-phase FIR filters.

Whether the $H_0 + H_1$ or $H_0 - H_1$ case should be used depends on how fast values of the impulse response of the filter changes. For slow changes (narrow-band filters) the $H_0 - H_1$ case is expected to require fewer SPT terms, while for fast changes (wide-band filters) the $H_0 + H_1$ case should work better [6].

In this work we will restrict ourselves to even order linear-phase filters as the number of multiplications per sample is smaller. However, a similar approach can be derived for odd order filters. Furthermore, it is possible to formulate problems for the considered reduced complexity polyphase filters decomposed in more than two polyphase stages.

## 3. LINEAR-PHASE FIR FILTERS

The frequency response of a linear-phase FIR filter can be separated into a real-valued function $H_R(\omega T)$ and a real-valued phase function $\Theta(\omega T)$ as

$$H(e^{j\omega T}) = H_R(\omega T)e^{j\Theta(\omega T)} \tag{7}$$

where $H_R(\omega T)$ is the *zero-phase frequency response*. We have $|H(e^{j\omega T})| = |H_R(\omega T)|$, as $|e^{jx}| = 1$ for real-valued $x$.

For an even order linear-phase FIR filter with filter order $N$ and symmetric impulse response, the zero-phase frequency response can be written as

$$H_R(\omega T) = h_{N/2} + 2\sum_{m=1}^{N/2} h_{N/2-i}\cos(\omega Tm) \tag{8}$$

Denote the desired zero-phase frequency response by $D(\omega T)$. $D(\omega T)$ is normally 1 in the passband and 0 in the stopband. If the maximal allowed deviation is denoted by $\delta(\omega T)$ the constraints on the filter design can be written as

$$|H_R(\omega T) - D(\omega T)| \le \delta(\omega T) \tag{9}$$

or, equivalently,

$$-\delta(\omega T) \le H_R(\omega T) - D(\omega T) \le \delta(\omega T) \tag{10}$$

For a given angle, $\omega T$, (10) is an inequality that is linear in the filter coefficients, $h_n$. Thus, by expressing (10) for a grid of different angles, $\omega T$, a set of inequalities are obtained that can be used for the filter design. However, even if a solution is found that is valid for all angles on the grid it can still violate the original specification for an angle that is not on the grid. Hence, the obtained solution must be checked on a much denser grid.

Note that there are no restrictions on the functions $D(\omega T)$ and $\delta(\omega T)$. It is therefore possible to design any type of linear-phase FIR filter as long as the specifications can be expressed using the zero-phase frequency function.

## 4. PROPOSED DESIGN METHOD

### 4.1. Signed-Powers-of-Two Coefficients

A fixed-point coefficient of wordlength $B$ can be represented as a sum of SPT terms in the general form

$$h_m = \sum_{i=1}^{B} s_i 2^{-i} \tag{11}$$

where $s_i \in \{-1, 0, 1\}$. Here we assume $-1 < h_m < 1$.

A minimum representation refers to the representation with the minimum required number of SPT terms. One minimum representation is the canonic signed digit (CSD) representation. Here, no two SPT terms can be adjacent.

It is possible to transform the SPT representation in (11) to 0/1-variables as

$$h_m = \sum_{i=1}^{B} (x^+_{m,i} - x^-_{m,i})2^{-i} \tag{12}$$

where $x^+_{m,i}, x^-_{m,i} \in \{0, 1\}$. This is advantageous as the formulation of the optimization goal function will be linear, otherwise it would have to be formulated as a non-linear programming problem. Furthermore, it will make it possible to have linear constraints on the number of SPT terms per coefficient.

To obtain a CSD solution the following constraints can be added

$$x^+_{m,i} + x^-_{m,i} + x^+_{m,i+1} + x^-_{m,i+1} \le 1 \tag{13}$$

for $m = 0, 1, ..., N/2$ and $i = 1, 2, ..., B - 1$. Adding these constraints yield a significant decrease of the solution time [8].

### 4.2. Normalized Peak Ripple Magnitude

The previous discussion assumed that the passband gain of the filter is equal to one. However, when the filter coefficients are expressed using a fixed-point format it is more convenient to allow an arbitrary passband gain. The stopband attenuation is then related to the passband gain. This introduces the opportunity to find “better” fixed-point coefficients, as it allows more flexibility [7]. The *normalized peak ripple magnitude* (NPRM) is commonly used for this type of problems. The passband gain is then computed as

$$g = \frac{\max\{|H_R(\omega T)|\} + \min\{|H_R(\omega T)|\}}{2} \tag{14}$$

where $\omega T$ is in the passband of the filter.

If (14) is introduced in (10) the inequalities will be non-linear, and, thus, it is not always possible to find an optimal solution. Instead, we introduce a continuous scaling variable, $s$, and rewrite (10) as

$$-\delta(\omega T) \le \frac{H_R(\omega T)}{s} - D(\omega T) \le \delta(\omega T) \tag{15}$$

The passband gain is always larger than zero, hence (15) can be rewritten as

$$s[D(\omega T) - \delta(\omega T)] \le H_R(\omega T) \le s[D(\omega T) + \delta(\omega T)] \tag{16}$$

to obtain a linear inequality with the passband gain as a variable.

The difference between the introduced form and the original NPRM is that in (16) the maximum value of the passband ripple can be different between the positive and negative ripple.

It could be argued that (14) must hold, but if a filter design with a specified passband gain is considered, it is

$$
\begin{aligned}
&\text{minimize} \quad \sum_{m=1}^{N/2} \sum_{i=1}^{B} x^{+}_{m,i} + x^{-}_{m,i} + y^{+}_{m,i} + y^{-}_{m,i} \\
&\text{subject to} \quad h_{N/2} + 2\sum_{m=1}^{N/2} h_{N/2-i}\cos(\omega T m) \le s[D(\omega T) - \delta(\omega T)] \\
&\quad h_{N/2} + 2\sum_{m=1}^{N/2} h_{N/2-i}\cos(\omega T m) \ge -s[D(\omega T) - \delta(\omega T)] \\
&\quad h_m = \sum_{i=1}^{B} (x^{+}_{m,i} - x^{-}_{m,i}) 2^{-i} \qquad m = 0, 1, \ldots, N/2 \\
&\quad g_m = h_m + h_{m-1} \qquad m = 1, 2, \ldots, N/2 \\
&\quad g_0 = h_0 \\
&\quad g_m = \sum_{i=1}^{B} (y^{+}_{m,i} - y^{-}_{m,i}) 2^{-i} \qquad m = 0, 1, \ldots, N/2 \\
&\quad x^{+}_{m,i} + x^{-}_{m,i} + x^{+}_{m,i+1} + x^{-}_{m,i+1} \le 1 \qquad \begin{cases} m = 0, 1, \ldots, N/2 \\ i = 0, 1, \ldots, B-1 \end{cases} \\
&\quad y^{+}_{m,i} + y^{-}_{m,i} + y^{+}_{m,i+1} + y^{-}_{m,i+1} \le 1 \qquad \begin{cases} m = 0, 1, \ldots, N/2 \\ i = 0, 1, \ldots, B-1 \end{cases}
\end{aligned} \tag{17}
$$

rarely the case that symmetric passband ripple is required. Hence, the proposed form not only provides linear inequalities, it could also be argued that it is a more natural way of expressing the passband gain. Indeed, (16) is just a scaled version of (10).

### 4.3. Problem Formulation

For the problem formulation we use the variables $h_m$, $x^{+}_{m,i}$, and $x^{-}_{m,i}$ to represent the filter coefficients and the SPT terms for the filter coefficients, respectively. For the filter $H_0 + H_1$ (or $H_0 - H_1$) we use $g_m$, $y^{+}_{m,i}$, and $y^{-}_{m,i}$ in a similar way. The optimization problem can then be formulated as in (17). Note that the $g_m$ variables are not defined in the order of the impulse response of that filter, but are utilizing the coefficient symmetries to obtain a simpler formulation. If the $H_0 - H_1$ is required the equation $g_m = h_m + h_{m-1}$ in (17) is replaced with $g_m = h_m - h_{m-1}$. These $g_m$ variables are not defined in the order of the impulse response either and half of them will not have the same sign. However, the minimum number of SPT terms are the same for $c$ and $-c$ for any fixed-point coefficient.

The solution to this optimization problem yields a coefficient set that is minimal in the number of SPT terms when the NPRM is considered. This problem can be solved using standard MILP solvers using e. g. branch-and-bound.

If a prescribed passband gain is required the value of $s$ can be fixed before the optimization starts. Note also that $s$ must be larger than 0 as this otherwise would yield an incorrect optimal solution with all variables equal to 0.

### 4.4. Reducing the Number of Variables

It is possible to reduce the number of variables before the actual optimization starts [8]. This due to the fact that the filter coefficient values are usually small at the start and end of the impulse response (for, e.g., a lowpass filter).

Upper and lower bounds for all filter coefficients (including $H_0 + H_1$ or $H_0 - H_1$) can be determined for a fixed passband gain using linear programming. From these bounds the maximum and minimum value of the passband gain scaling coefficient, $s$, can be derived utilizing the complete dynamic range of the representation. Now, it is possible to derive the upper and lower bounds on the scaled values of the filter coefficients. When these bounds have been calculated it can be determined which SPT terms that can be fixed to a certain value, usually zero. This method is discussed in greater detail in [8]. It is worth noting that for the reduced complexity polyphase FIR filters one of the filter coefficients in the $H_0 + H_1$ or $H_0 - H_1$ filter can be the one determining the scaling bounds.

To decrease the solution time it is also useful to have an upper bound for the objective function, which can be obtained from, e.g., a heuristic approach [9].

## 5. EXAMPLES

In this section a number of filters is designed to show the properties of the design method and, to some extent, the filter architecture. For all filters, solutions with minimum number of SPT terms are obtained for both the $H_0 + H_1$ and $H_0 - H_1$ cases, as well as a direct realization. The numbers given are for all multiplications in the filter. There are design examples given in [6], but these are hard to reproduce since a filter specification is not given. However, as we obtain optimal results our approach will give as good or better solutions. All problems are solved using the general purpose MILP solver ILOG CPLEX [10].

### 5.1. Example 1

Consider a linear-phase lowpass FIR filter with passband edge at 0.1π rad, stopband edge at 0.35π rad, passband ripple of 0.01, and stopband ripple 0.01. The minimum filter order is 15. To obtain a design margin we select a filter order of 17, i. e., nine filter taps. We use eight coefficient bits, except for the $H_0 + H_1$ case where nine coefficient bits are allowed, as the sum will require more bits. As this is a rather narrow-band filter it is expected that the $H_0 - H_1$ case will be most efficient [6].

The minimum number of SPT terms for the three different filters are presented in Table 1. Here, it is obvious that there is a small decrease in the number of SPT terms per sample for the $H_0 - H_1$ case. However, the direct realization is almost as good when the number of SPT terms are considered. The magnitude responses for the designed filters are shown in Fig. 2.

**Table 1.** Number of SPT terms for the filters in Example 1.

| **SPT Terms** | **Filter type** | | |
|---|---|---|---|
| | $H_0 - H_1$ | $H_0 + H_1$ | **Direct** |
| Total | 31 | 35 | 16 |
| Per sample | 15.5 | 17.5 | 16 |

### 5.2. Example 2

Consider a linear-phase lowpass FIR filter with passband edge at 0.65π rad, stopband edge at 0.9π rad, passband ripple of 0.01, and stopband ripple 0.01. The minimum filter order is 16. Again, to obtain a design margin we select a filter order of 17, i.e., nine filter taps. As for the filter in Example 2 eight coefficient bits are used, except for the $H_0 + H_1$ case where nine coefficient bits are allowed, as the sum will require more bits. This is a wideband filter so the $H_0 + H_1$ case is expected to be the most efficient [6].

The minimum number of SPT terms for the three different filters are presented in Table 2. Here, it is a slightly larger decrease in the number of SPT terms per sample for the $H_0 + H_1$ case compared with the direct realization. The $H_0 - H_1$ case and the direct realization yields exactly the same filter. The magnitude responses of the filters are shown in Fig. 3.

**Table 2.** Number of SPT terms for the filters in Example 2.

| **SPT Terms** | **Filter type** | | |
|---|---|---|---|
| | $H_0 - H_1$ | $H_0 + H_1$ | **Direct** |
| Total | 37 | 33 | 18 |
| Per sample | 18.5 | 16.5 | 18 |

## 6. CONCLUSIONS

In this work an approach to the design of linear-phase FIR filters implemented using a reduced complexity polyphase architecture is proposed. It is shown that a mixed integer linear programming (MILP) problem can be formulated and solved. This yields optimal results, minimizing the number of signed-power-of-two (SPT) terms required for the filter coefficients. Examples showed that it is possible to reduce the number of SPT terms required per sample compared with a direct realization, although the savings were marginal in some cases. This, to some extent, contradicts the results in [6], where large savings were obtained.

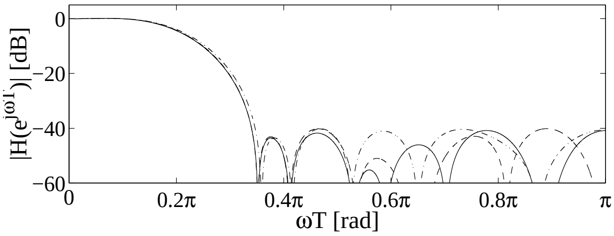


**Figure 2.** Magnitude responses of the filters in Example 1. Direct (dashed), $H_0 - H_1$ (solid), and $H_0 + H_1$ (dash-dotted).

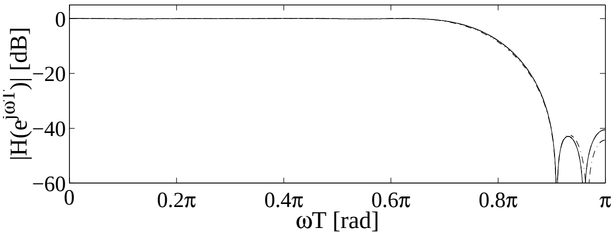


**Figure 3.** Magnitude responses of the filters in Example 2. Direct and $H_0 - H_1$ (solid) and $H_0 + H_1$ (dash-dotted).

Hence, we can conclude that the discussed filter structure may reduce the complexity for linear-phase FIR filters, but the reduction is not as large as previous work has proposed when optimal designs are compared.

## 7. REFERENCES


[1] T. Saramäki and J. Yli-Kaakinen, “Design of digital filters and filter banks by optimization: applications,” *Proc. X European Signal Processing Conf.*, Tampere, Finland, Sept. 4–8, 2000.

[2] L. Wanhammar and H. Johansson, *Digital Filters*, Linköping University, 2003.

[3] Z. J. Mou and P. Duhamel, “Short-length FIR filters and their use in fast nonrecursive filtering,” *IEEE Trans. Signal Processing*, vol. 39, pp. 1322–1332, June 1991.

[4] D. A. Parker and K. K. Parhi, “Low-area/power parallel FIR digital filter implementation,” *J. VLSI Signal Processing*, vol. 17, pp. 75–92, 1997.

[5] K. K. Parhi, *VLSI Digital Signal Processing Systems: Design and Implementation*, Wiley, 1999.

[6] J.-G. Chung et al., “Efficient parallel FIR filter implementations using frequency spectrum characteristics”, *Proc. IEEE Int. Symp. Circuits Syst.*, Monterey, May 31–June 3, 1998, vol. V, pp. 354–358.

[7] Y. C. Lim, “Design of discrete-coefficient-value linear phase FIR filters with optimum normalized peak ripple magnitude,” *IEEE Trans. Circuits Syst.*, vol. 37, pp. 1480–1486, Dec. 1990.

[8] O. Gustafsson, H. Johansson, and L. Wanhammar, “An MILP approach for the design of linear-phase FIR filters with minimum number of signed-power-of-two terms,” in *Proc. European Conf. Circuit Theory Design*, Espoo, Finland, Aug. 28–31, 2001.

[9] O. Gustafsson, *Contributions to Low-Complexity Digital Filters*, Linköping Studies in Science and Technology, Diss., No. 837, Linköping University, Sweden, 2003.

[10] ILOG CPLEX, http://www.ilog.com/products/cplex/, Oct. 2003.