# DESIGN OF LINEAR-PHASE FIR FILTERS WITH MINIMUM HAMMING DISTANCE

*Oscar Gustafsson and Lars Wanhammar*

Department of Electrical Engineering, Linköping University,
SE-581 83 Linköping, SWEDEN
E-mail: {oscarg, larsw}@isy.liu.se,
Tel: +46-13-28{4059, 1344}, Fax: +46-13-139282.

**ABSTRACT**

In this paper a new approach for the design of linear phase FIR filters with discrete coefficients is proposed. A mixed integer linear programming (MILP) problem is formulated that minimizes the total Hamming distance between the adjacent coefficients, i.e., the number of bit switches of the coefficients. The Hamming distance between the coefficients is a good measure of the power consumption, when the FIR filters is implemented on a programmable architecture. The method is applicable both for filters with a specified passband gain and filters where the normalized peak ripple magnitude is of interest. In both cases the globally minimal solution is found subject to the filter specification, filter order, and number of coefficient bits. A preprocessing method that removes a significant part of the variables is also proposed and it is shown by an example that this method speeds up the optimization process significantly. Both two's complement and signed magnitude representation of the coefficients are considered.

## 1. INTRODUCTION

Finding good fixed-point coefficients for FIR filters has received considerable attention during the last 20 years. The measure of good coefficients depends on the filter implementation, much of the work this far has been considering the number of signed-power-of-two (SPT) terms for the coefficients [1]–[3]. This is motivated by that a multiplication can be separated into a number of shift-and-add operations. Each shift-and-add operation corresponds to a SPT term which makes this an interesting variable either to minimize or to have as a constraint.

If the FIR filter is implemented on a programmable architecture with one general multiplier and an accumulator, as shown in Fig. 1, the Hamming distance between the coefficients has been shown to be a good measure of the power consumption in the multiplier [4]. The Hamming distance between two coefficients is the number of bits that are different. Other previous work is this area has been aimed at reordering the coefficients to minimize the Hamming distance [5]. In this work we focus on minimizing the Hamming distance between those coefficients that are naturally processed after each other, i.e., those coefficients that are adjacent in the impulse response of the filter. The motivation is that by reordering the coefficients, the data must also be reordered removing the advantage of possible correlation between the samples. Further, to effectively utilize features like hardware looping on general purpose DSPs, the multiplications should be performed in order.

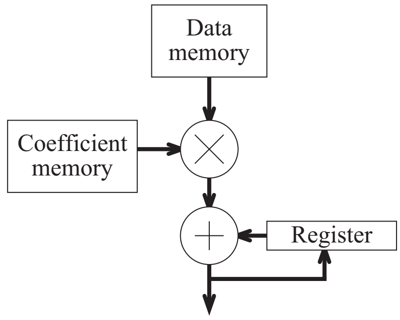


Figure 1: Programmable multiply-accumulate architecture suitable for FIR filtering.

In this paper we formulate an MILP problem that minimizes the Hamming distance between adjacent coefficients for a given filter specification subject to the filter order and coefficient wordlength. A preprocessing technique to reduce the number of variables that was previously proposed by the authors is shown useful for this case as well. The proposed method has been modelled using OPL and solved using the general purpose optimization package CPLEX [6]. An example shows the usefulness of the proposed technique.

## 2. 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

$$H(e^{j\omega T}) = H_R(\omega T)e^{j\Theta(\omega T)} \quad (1)$$

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

For an $N$th-order linear-phase FIR filter the zero-phase frequency response can be written as

$$H_R(\omega T) = \sum_{m=1}^{\left\lceil \frac{N+1}{2} \right\rceil} h_m c(m, \omega T) \quad (2)$$

where $c(m, \omega T)$ is a proper trigonometric function depending on if the filter order is odd or even and if the impulse response is symmetric or anti-symmetric.

Let the specifications of the filter be

$$-\delta(\omega T) \leq H_R(\omega T) - D(\omega T) \leq \delta(\omega T) \quad (3)$$

where $D(\omega T)$ is the desired magnitude and $\delta(\omega T)$ is the maximum allowed ripple for a given angle $\omega T$. Typically, $D(\omega T)$ is one in the passband and zero in the stopband of the filter.

Note that (3), for each angle $\omega T$, describes an inequality that is linear in the filter coefficients.

When fixed-point coefficient is considered it is often useful to allow a non-unity passband gain [7]. This can be done by introducing a scaling factor, $s$, as

$$-\delta(\omega T) \leq H_R(\omega T)/s - D(\omega T) \leq \delta(\omega T) \quad (4)$$

The scaling factor is always larger than 0, so (4) can be rewritten as

$$s[D(\omega T) - \delta(\omega T)] \leq H_R(\omega T) \leq s[D(\omega T) + \delta(\omega T)] \quad (5)$$

to obtain a linear inequality.

## 3. PROPOSED DESIGN METHOD

The objective is to form an MILP formulation of linear-phase FIR filter design that minimizes the Hamming distance between adjacent coefficients. This will be done by first introducing a linear expression for the coefficient bits, where the bits corresponds to binary variables, $x$. Then, a second set of binary variables, $y$, will be introduced that corresponds to if a bit position is switched between two coefficients. As we want to minimize the number of switches, the optimization problem can be stated as minimizing the sum of all $y$ variables.

### 3.1. Two's Complement Coefficients

Two's complement representation is commonly used in DSP applications. A coefficient, $h_m$, of wordlength $B$ expressed using two's complement can be written as

$$h_m = -x_{m,0} + \sum_{i=1}^{B-1} x_{m,i} 2^{-i} \quad (6)$$

where $x_{m,i} \in \{0, 1\}$. Thus, the coefficients of the FIR filter can be described using a linear expression with binary variables.

The Hamming distance between two coefficients is equal to the number of bits that are different. Introduce a binary variable, $y_{m,i} \in \{0, 1\}$, as

$$y_{m,i} = \begin{cases} 1, & x_{m,i} \neq x_{m+1,i} \\ 0, & x_{m,i} = x_{m+1,i} \end{cases}, m = 1, \ldots, M-1 \quad (7)$$

i.e., if bit $i$ changes value between $h_m$ and $h_{m+1}$, $y_{m,i}$ is one otherwise it is zero. This can be expressed using linear expressions as

$$\begin{matrix} y_{m,i} \geq x_{m,i} - x_{m+1,i} \\ y_{m,i} \geq x_{m+1,i} - x_{m,i} \end{matrix}, \begin{cases} m = 1, \ldots, M-1 \\ i = 0, \ldots, B-1 \end{cases} \quad (8)$$

Only one half of the impulse response needs to be considered due to the symmetry of the magnitude of the coefficients.

### 3.2. Signed Magnitude Representation

An alternative to two's complement is signed magnitude representation. A $B$ bit signed magnitude coefficient can be written as

$$h_m = (-1)^{x_{m,0}} \sum_{i=1}^{B-1} x_{m,i} 2^{-i} \quad (9)$$

Hence, the first bit corresponds to the sign, while the remaining bits corresponds to the magnitude. Signed magnitude representation has a low switching probability when the numbers are small and with varying sign. Multiplication is straight-forward using signed magnitude representation. The magnitudes can be multiplied using unsigned multiplication, while the sign of the result is determined with a simple xor-operation of the sign bits of the multiplier and multiplicand. However, addition and subtraction is more complex. It is possible to use two's complement representation of the data and signed magnitude representation of the coefficients. The multiplication is then performed in two's complement, with the coefficient always being positive. The sign bit of the coefficient controls if the result of the multiplication should be added or subtracted to the accumulator.

Equation (9) can not be used directly in a linear formulation, instead for each coefficient a binary sign variable $s_m$ is introduced along with $B-1$ binary variables, $x^+_{m,i}$, for positive coefficients and $B-1$ binary variables, $x^-_{m,i}$, for negative coefficients. The coefficient can now be written as

$$h_m = \sum_{i=1}^{B-1} (x^+_{m,i} - x^-_{m,i}) 2^{-i} \quad (10)$$

with the constraints

$$\sum_{i=1}^{B-1} x^+_{m,i} \leq (B-1)(1-s_m), \quad 1 \leq m \leq M \quad (11)$$

and

$$\sum_{i=1}^{B-1} x_{m,i}^{-} \le (B-1)s_m, \quad 1 \le m \le M \tag{12}$$

These constraints make sure that if the coefficient $h_m$ is positive, $s_m = 0$ and $x_{m,i}^{-} = 0$ for all $i$, and similarly, if the coefficient $h_m$ is negative, $s_m = 1$ and $x_{m,i}^{+} = 0$ for all $i$. The new variables are related to the formulation in (9) as

$$\begin{aligned} x_{m,0} &= s_m \\ x_{m,i} &= x_{m,i}^{+} + x_{m,i}^{-}, \quad i = 1, \ldots, B-1 \end{aligned} \tag{13}$$

The number of switched bits can now be expressed as in (8).

### 3.3. Problem Formulation

The resulting optimization problem can for two's complement representation be formulated as

$$\begin{aligned} &\text{minimize} && \sum_{m=1}^{M-1} \sum_{i=0}^{B-1} y_{m,i} \\ &\text{subject to} && s[D(\omega T) - \delta(\omega T)] \le \sum_{m=1}^{M} h_m c(m, \omega T) \\ &&& \sum_{m=1}^{M} h_m c(m, \omega T) \le s[D(\omega T) + \delta(\omega T)] \\ &&& y_{m,i} \ge x_{m,i} - x_{m+1,i} \\ &&& y_{m,i} \ge x_{m+1,i} - x_{m,i} \\ &&& h_m = -x_{m,0} + \sum_{i=1}^{B-1} x_{m,i} 2^{-i} \end{aligned} \tag{14}$$

For signed magnitude representation the last row of (14) can be exchanged for (10)–(13).

As the constraints are continuous in $\omega T$ it is necessary formulate the constraint for discrete values of $\omega T$. This is easily done by selecting a number of values in the passband and the stopband to obtain a grid of values which each leads to one constraint. As the specifications are checked only in these points it is necessary to check the resulting transfer function with a much finer grid to verify that a valid coefficient set is obtained. If not, additional points must be added to the constraints.

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 optimal solution with all variables equal to 0. In general it is possible to constrain $s$ to be larger than 0.1, say, without losing optimality in the solution. This is due to the fact that scaling $s$ with 2 is equal to shifting each coefficient one step.

## 4. REDUCING THE NUMBER OF VARIABLES

In the optimization problem in (14) the number of 0/1-variables is $(2M-1)B$ and $3MB - M - B$ for two's complement and signed magnitude representation, respectively. This number will be high for filters with stringent specifications and the execution time will therefore be long. Hence, it is of interest to find ways to decrease the number of variables before the actual optimization starts. We will apply a method similar to that proposed in [3].

To obtain the ranges for the coefficients it is possible to use linear programming. The problem can be stated as first a separate maximization of each variable $h_m$, subject to (3). These values are assigned to the variables $h_m^{(ub)}$. Then a minimization of each variable $h_m$ is performed and the result is assigned the variables $h_m^{(lb)}$. However, these values are for $s = 1$, so the possible range of $s$ must be found to obtain the true bounds for $h_m$.

For a both a two's complement and a signed-magnitude number with wordlength B, the maximal value for any positive coefficient is bounded approximately by one. A negative coefficient is bounded by minus one. To use the complete range of the coefficients the absolute value for the coefficient with the largest magnitude should be at least 1/2.

The minimal value for the scaling factor is obtained when the maximal absolute value for the upper or lower bound times the scaling factor is 1/2. Hence

$$s^{(lb)} = \frac{1}{2\max\{\max\{|h_m^{(ub)}|, |h_m^{(lb)}|\}\}} \tag{15}$$

The maximal value for the scaling factor is when the maximal value of the coefficients minimum absolute bounds are 1. This is for coefficients that do not change sign. The upper bound is then

$$s^{(ub)} = \frac{1}{\max\{\min\{|h_m^{(ub)}|, |h_m^{(lb)}|\}\}} \tag{16}$$

where the coefficients concerned are those that do not change sign. This reasoning is only valid when at least one coefficient have the same sign on its maximum and minimum value. However, this is the case for most filters unless the design margin is very large.

New bounds on the coefficients can now be derived as

$$\hat{h}_m^{(ub)} = \begin{cases} \min(1, h_m^{(ub)} s^{(lb)}), & h_m^{(ub)} \ge 0 \\ \max(-1, h_m^{(ub)} s^{(ub)}), & h_m^{(ub)} < 0 \end{cases} \tag{17}$$

and

$$\hat{h}_m^{(lb)} = \begin{cases} \min(1, h_m^{(lb)} s^{(ub)}), & h_m^{(lb)} \geq 0 \\ \max(-1, h_m^{(lb)} s^{(lb)}), & h_m^{(lb)} < 0 \end{cases} \quad (18)$$

If the passband gain is specified we have $s^{(ub)} = s^{(lb)} = s_s$, where $s_s$ is the specified passband gain. Finally, the following inequality is obtained

$$\hat{h}_m^{(lb)} \leq h_m \leq \hat{h}_m^{(ub)} \quad (19)$$

This inequality can either be added to the constraints of the optimization or used to preevaluate possible values of the variables. Here, (19) is utilized to determine which bits that do have a fixed value and remove these from the optimization problem. This can be done by searching all possible values for the filter coefficient and keep track of which values the bits assume. All variables that only assign one value can be fixed and thereby removed from the optimization problem.

As will be shown in the example below, the number of variables that can be removed depends on the available design margin and the representation used. The number of variables that can be removed are in general not dependent on the coefficient wordlength, as typically it is the most significant bits of a coefficient that are fixed. However, the relative savings will be smaller when the wordlength is increased.

## 5. EXAMPLE

To show the usefulness of the proposed idea we will consider one filter design example. The filter under consideration is a lowpass filter with passband edge at 0.3π and stopband edge at 0.5π. The allowed ripple is 0.001 in both the passband and the stopband. A filter order of 33 is used. Hence, there are 17 coefficients to be designed. Using the Remez exchange algorithm and rounding, 15 coefficient bits are required to satisfy the specification. For two's complement representation the sum of the Hamming distances between subsequent coefficients are 118. Using the proposed design method a solution with 13 bits can be found with a total Hamming distance of 79. The original optimization problem has 448 variables, while after the proposed preprocessing it has 292 variables. Thus, a significant speed-up was obtained for the solution. The problem was solved on an AMD Athlon 1800+ with 1.5 GB of RAM. The pre-processing took approximately 30 s, while the MILP stage took 55 minutes. For signed magnitude representation, the rounded solution has a total Hamming distance of 78. The original optimization problem has 652 variables, while after preprocessing 255 variables remain. The optimized solution has a total Hamming distance of 60. For this problem the MILP stage took approximately 150 minutes. The magnitude responses for the three designed filters are shown in Fig. 2.

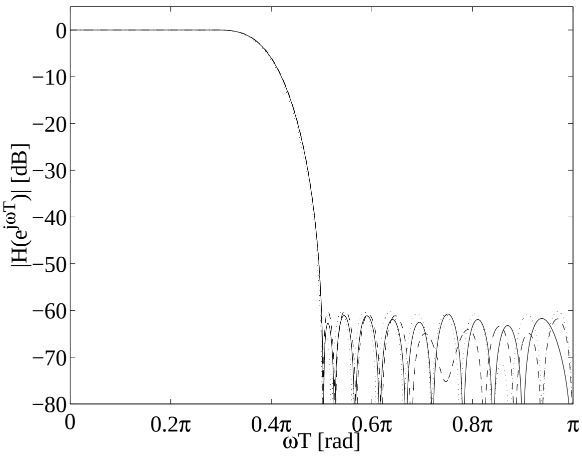


Figure 2: Magnitude responses for example filter designed using (solid) Remez and rounding, (dashed) proposed approach with two's complement representation, and (dotted) proposed approach with signed magnitude representation.

## 6. CONCLUSIONS

In this paper a mixed integer linear programming problem for design of linear-phase FIR filters with minimum Hamming distance between adjacent coefficients was formulated. It has been shown in earlier work that the Hamming distance is a good measure of the power consumption for a programmable FIR filter. A preprocessing technique to remove 0/1-variables, and, thus, decrease the solution time, was discussed. Both two's complement and signed magnitude representation were considered. A design example showed the usefulness of the proposed method.

## 7. REFERENCES


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

[2] H. Samuli, "An improved search algorithm for the design of multiplierless FIR filters with powers-of-two coefficients," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 1044–1047, July 1989.

[3] 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.

[4] M. Mehendale, S. D. Sherlekar, and G. Venkatesh, "Coefficient optimization for low power realization of FIR filters," in *Proc. IEEE Workshop VLSI Signal Processing*, 1995, pp. 352–361.

[5] P. K. Merakos et al, "A novel transformation for reduction of switching activity in FIR filters implementation," in *Proc. Int. Conf. Digital Signal Processing*, 1997, vol. 2, pp. 653–656.

[6] ILOG, http://www.ilog.com/, June 2002.

[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.