<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Design of linear phase FIR filters in subexpression space using mixed integer linear programming( Published version )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Yu, Ya Jun; Lim, Yong Ching</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>2007</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/4659">http://hdl.handle.net/10220/4659</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© 2007 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE. This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder. <a href="http://www.ieee.org/portal/site">http://www.ieee.org/portal/site</a>.</td>
</tr>
</tbody>
</table>
Design of Linear Phase FIR Filters in Subexpression Space Using Mixed Integer Linear Programming
Ya Jun Yu, Member, IEEE, and Yong Ching Lim, Fellow, IEEE

Abstract—In this paper, a novel optimization technique is proposed to optimize filter coefficients of linear phase finite-impulse response (FIR) filter to share common subexpressions within and among coefficients. Existing approaches of common subexpression elimination optimize digital filters in two stages: first, an FIR filter is designed in a discrete space such as finite wordlength space or signed power-of-two (SPT) space to meet a given specification; in the second stage, an optimization algorithm is applied on the discrete coefficients to find and eliminate the common subexpressions. Such a two-stage optimization technique suffers from the problem that the search space in the second stage is limited by the finite wordlength or SPT coefficients obtained in the first stage optimization. The new proposed algorithm overcomes this problem by optimizing the filter coefficients directly in subexpression space for a given specification. Numerical examples of benchmark filters show that the required number of adders obtained using the proposed algorithm is much less than those obtained using two-stage optimization approaches.

Index Terms—Common subexpression sharing, finite-impulse response (FIR) filters, mixed integer linear programming (MILP), subexpression space.

I. INTRODUCTION

FINITE-IMPULSE RESPONSE (FIR) filters have many attractive virtues such as exact linear phase property, guaranteed stability, free of limit cycle oscillations, and low coefficient sensitivity. However, the complexity of an FIR filter in terms of the implementation cost in VLSI is generally higher than that of a corresponding infinite-impulse response (IIR) filter meeting the same magnitude response specifications. Therefore, during the past decades, continuous effort has been made to design low complexity FIR filters. One of the most successful strategies is to optimize the filter coefficients in signed power-of-two (SPT) [1]–[3] space, where each coefficient is represented as a sum of a limited number of SPT terms. Thus, coefficient multiplications can be replaced by shifters and adders, so that the implementation of the filter is essentially multiplierless.

The number of adders to implement the FIR filter can be further reduced by extracting common subexpressions from the coefficients [4], [5]. The adders used to realize the common subexpressions can be shared within a coefficient as well as among coefficients. For example, when two coefficients with binary numbers 100101 and 10101 are implemented independently, two adders are required for each coefficient, as shown in Fig. 1(a). Thus, in total four adders are used for the coefficients. If the common subexpression of 101 is extracted and implemented first, only an additional adder is required for each coefficient, ending up with three adders in total, as shown in Fig. 1(b). Even in such a simple example, an adder is saved.

In Fig. 1, transposed direct form implementation is used to share the common subexpression. Since direct form is the dual of the transposed direct form, the subexpression sharing can always be implemented using the direct form in the same complexity as that using the transposed direct form. In transposed direct form implementation, the adders can easily be classified into two categories: 1) those implementing the coefficients; and 2) those inserted between consecutive delays. Since the number of adders between consecutive delays remains invariant for different coefficient implementation scheme, the number of adders used to implement the coefficients is commonly taken as the coefficient implementation complexity. This is also the coefficient implementation complexity measure adopted in this paper.

Various optimization techniques [5]–[16] have been proposed to identify and extract as many common subexpressions from coefficients as possible to reduce the number of adders. Traditionally, the filters are designed in two stages. First, the FIR filter is designed in a discrete space such as finite wordlength space or SPT space to meet a given specification. In the second stage, an optimization technique is applied on the discrete coefficients to find and eliminate the common subexpressions. Many of the optimization techniques in the second stage are based on a heuristic search or combined exhaustive search.

There is an obvious disadvantage in such a two-stage optimization technique. The filter designed in the first stage is an optimum or suboptimum solution in the given discrete space, but is not in the discrete space constructed by subexpressions. For example, assume that the optimum value of a coefficient is 101, and the discrete value obtained in the first stage optimization is 100 if only 1 SPT term is allowed per coefficient; it is impossible to obtain the optimum solution of 101 in the second stage search, since 101 is not a permissible value in the 1st stage optimization. A recent publication [13] alleviates this disadvantage by perturbing the first stage (sub)optimum solution during the second stage search. Such a perturbation is equivalent to a local search.

Two techniques [17], [18] were proposed to optimize the filter coefficient in subexpression space directly for given filter specifications. In [17], the dynamic range of each coefficient is determined; exhaustive search is then used to check all possible
combinations to find the best solution. In [18], a 0/1 integer programming is used to minimize the number of adders subject to a given filter specification. Very good results were shown in both papers. However, the computation time of these two approaches are huge. They are only suitable to design short filters with short wordlength. For example, in [18], it is stated that the reasonable solving time to optimize a filter with 24 taps and 10 coefficient bits is within one day. Since the computation time of these optimization techniques increase at least exponentially with the filter length and bit width, it is impractical to design longer filters with wider bit width.

In this paper, a technique directly optimizing the filter coefficients in the discrete space constructed by a given set of subexpressions is proposed. Mixed integer linear programming (MILP) is the technique employed to optimize the filter coefficients to meet the given specifications. In this technique, the frequency response ripple is minimized subject to a given number of adders. The obtained results may not be the optimum in terms of the number of adders, but the saving in the number of adders achieved using the proposed technique is significant compared with those obtained using other techniques that can be used to design filters with respectable length and bit width. The computation time to optimize a filter with order 60 and coefficient bit width 12 is typically within a few minutes to a few hours depending on specific examples.

The remaining of the paper is organized as follows. Section II introduces the space constructed by a given set of subexpressions. An algorithm to find the best approximation of a real number in the subexpression space is proposed. Section III introduces the branch and bound MILP to optimize the filter coefficients in subexpression spaces. Numerical examples are shown in Section IV to illustrate the advantage of the proposed technique. Section V proposes a multi-step optimization approach to circumvent the difficulty of long filter design using MILP. Section VI shows the flexibility of MILP to design filters in subexpression space with other constraints such as limiting the maximum adder step of coefficients.

II. CONSTRUCTION OF DISCRETE SUBEXPRESSION SPACE

Since a finite word-length coefficient values can be converted into an integer value by multiplying a suitable integer power-of-two, it is sufficient to consider only coefficients with integer values.

A. Subexpression Space

It is well known that an integer $n$ can be represented in SPT number space as [3], [19]

$$n = \sum_{i=0}^{K-1} y(i)2^{q(i)}, \quad y(i) \in \{-1, 1\} \tag{1}$$

where $K$ is the number of SPT terms, and $q(i)$ is a non-negative integer.

Similarly, a discrete subexpressions space can be constructed by defining its element as

$$n = \sum_{i=0}^{K-1} y(i)2^{q(i)}, \quad y(i) \in S \tag{2}$$

where $S$ is a set of permissible subexpression bases. The set of subexpression bases is also called basis set in the remaining of this paper for expository convenience. In (2), $y(i)2^{q(i)}$ is a shifted version of a subexpression basis, named a subexpression term, and $K$ is redefined as the number of subexpression terms.

Usually, a subexpression basis refers to a trinary string with more than 1 nonzero digit, such as 1011 or 1001, where 1 represents $-1$. In this paper, a single digit 1, 0, or 1 is also broadly referred to as a subexpression basis. Therefore, a typical basis set may be $S_1 = \{0, 1, \bar{1}, 101, 10\bar{1}, 1\bar{0}1, \bar{1}01\}$. Thus, we say that a subexpression space is constructed based on a basis set according to (2). We stipulate that if a digit string is in a basis set, its negative value must also be in the same set. This stipulation is reasonable because inclusion of the negative value into the set does not incur the use of additional adders. It is noted further that element 0 in $S_1$ has no contributions in the construction of subexpression space, but it is included in every subexpression basis set for the convenience of the description of algorithm, of which we will see in the following derivation.

The elements of a basis set can alternatively be represented by decimal numbers. For example, $S_1$ can be written as $\{0, \pm1, \pm3, \pm5\}$. In this paper, these two representations are used interchangeably for expository convenience.

B. Some Definitions About Basis Set

The basis set is said to be contiguous if the basis set contains contiguous odd numbers starting from $\pm1$, i.e., any odd
number whose absolute value is smaller than the maximum element is an element of the set. Thus, \( S_1 \) is contiguous, but \( S_2 = \{0, \pm 1, \pm 5, \pm 7\} \) is not contiguous.

The number of adders required to construct a basis is called the order of the basis. For example, \( 10\overline{1} \) is a first-order basis, whereas \( 10101 \) is a second-order basis.

The order of the basis set is defined as the number of adders required to generate the values of all elements in the set. It is obvious that the order of \( S_1 \) is 2 since 1, 0 and 2 are/are increasing.

Let \( \overline{1} \). Go to Step 2.

Step 2: \( \text{Find } p \in \text{subexpression basis set.} \) Let \( p = 2 \) since 1, 0 and 2 are/are the negatives of 101 and \( \overline{1} \), and when the basis set is contiguously contiguous, no additional adder is required.

It is apparent by inspection that the order of a contiguous basis set is equal to \((P - 1)/2\), where \( P \) is the maximum elements of the set.

From the above definitions, it is interesting to note that SPT space is a special subexpression spaces whose basis set has order 0.

C. Finding the Best Approximation of a Number in Subexpression Space

Having the subexpression space constructed based on a basis set, we wish to find the best approximation of a real number \( x \) using a sum of not more than \( K \) subexpression terms \( y(i)2^{q(i)} \), where \( y(i) \) is an element of the set of permissible subexpression bases and \( q(i) \) is a non-negative integer. Let \([x]_K \) be the \( K \)-term subexpression number that is a best approximation to \( x \). Since the number of elements in the basis set is limited, an exhaustive search can be employed to find the combination of \( K \) pairs of \( y(i) \) and \( q(i) \) which minimize the error between \( x \) and the approximation.

When the order of the basis set and/or \( K \) are/is increasing, exhaustive search becomes computational impractical. A greedy algorithm for computing \([x]_K \) is proposed. In this algorithm, initially, the approximation to \( x \), denoted as \([x]_1 \), is set to 0. Subexpression terms are assigned to \([x]_k \) one at a time. Each time, the subexpression term which can minimize the difference between \( x \) and \([x]_k \) is assigned to \([x]_k \). The algorithm runs as follows.

Step 1) Initialize \( m = 1 \) and \( z_0 = x \).

Step 2) Find \( y(m)2^q(m) \) which minimizes \(|z_{m-1} - y(m)2^q(m)| \).

Step 3) If either \( y(m) = 0 \) or \( m = K \), go to Step 6; otherwise go to Step 4.

Step 4) Update \( z_m = z_{m-1} - y(m)2^q(m) \).

Step 5) Increment \( m \). Go to Step 2.

Step 6) \([x]_K = \sum_{i=1}^{K} y(i)2^{q(i)}. \) Stop.

In step 2, \( y(m) \) is selected from the subexpression basis set. \( y(m) = 0 \) is also a possible choice. When \( y(m) = 0 \) minimizes \(|z_{m-1} - y(m)2^q(m)| \), it implies that any further subexpression terms from the given basis set will not reduce the approximation error. For this reason, 0 is included in every basis set, as stated earlier. In the following, 0 is omitted from the basis set for expository convenience.

From our experience derived from a large number of examples when \( K = 2 \) and \( K = 3 \), and when the basis set is contiguous, the approximation obtained by using this greedy algorithm is the best approximation. Unfortunately, no mathematical proof is available so far to prove the optimality of the algorithm.

However, it is possible to find counter examples showing that the approximation obtained using the greedy algorithm is not the best approximation when the basis set is not contiguous. An example of which is to round the integer 23 to a subexpression space using two subexpression terms from the noncontiguous basis set of \( S_1 = \{\pm 1, \pm 5, \pm 7, \pm 9\} \); the approximation obtained using the greedy algorithm is \( 5 \times 2^2 + 1 \times 2^2 = 24 \), but the best approximation is \( 1 \times 2^4 + 7 \times 2^0 = 23 \).

III. OPTIMIZING FILTER COEFFICIENTS USING BRANCH-AND-BOUND MILP

The zero-phase frequency response of a linear-phase FIR filter of length \( N \) can be expressed as

\[
H(\omega) = \sum_{n=0}^{N-1} h(n) \text{Trig}(\omega, n)
\]

where, \( h(n)'s \) are the filter coefficients and \( \text{Trig}(\omega, n) \) is an appropriate trigonometric function depending on the parity of \( N \) and the symmetry of the impulse response [20].

For a given set of filter specifications, such as passband edge \( \omega_p \), stopband edge \( \omega_s \), passband and stopband ripple ratio \( \delta_p/\delta_s \), and filter length \( N \), the optimum continuous coefficients minimizing the magnitude ripple can be found by formulating the problem as a linear programming problem as follows:

\[
\text{minimize: } \delta \\
\text{subject to: } 1 - \delta \leq H(\omega) \leq 1 + \delta, \quad \text{for } \omega \in [0, \omega_p] \\
\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad - \frac{\delta_p}{\delta_s} \leq H(\omega) \leq \frac{\delta_s}{\delta_p}, \quad \text{for } \omega \in [\omega_s, \pi]
\]

To find the coefficient values in a discrete space for the same set of filter specifications, the linear programming problem is replaced by a MILP problem. Branch and bound algorithm is one of the most efficient techniques to solve MILP problem.

First we review the depth-first search branch and bound algorithm [21]–[23] which has successfully optimized filter coefficients in integer and SPT spaces.

In the branch and bound (B&B) algorithm, after the continuous optimum solution is obtained, a coefficient \( x_k \) is selected for branching. Suppose that \( x_2 \) is selected and that the integer space is the desired discrete coefficient space. Suppose also that the continuous optimum value of \( x_2 \) is 3.4. Two subproblem \( P_1 \) and \( P_2 \) are created by imposing the bounds \( x_2 \leq 3 \) and \( x_2 \geq 4 \), respectively (see Fig. 2). Problem \( P_2 \) is stored and \( P_1 \) is solved. Another coefficient (say \( x_1 \)) is then selected for partitioning into two subproblems \( P_3 \) and \( P_4 \) by imposing bounds on the selected coefficient (say \( x_1 \leq [x_1] \) and \( x_1 \geq [x_1] \), respectively, where \([x] \) is the permitted discrete value immediately larger than or equal to \( x \) and \([x] \) is the permitted discrete value of \( x \) immediately less than or equal to \( x \)). \( P_2 \) is stored and \( P_3 \) is solved. \( P_3 \) is then further partitioned into \( P_4 \) and \( P_6 \). In a similar way, \( P_6 \) is stored and \( P_5 \) is solved. Suppose that \( P_5 \) yields a discrete solution. \( P_5 \) is then fathomed and \( P_6 \) is solved. In Fig. 2, a line underneath a node indicates that no further exploration from that node can be profitable. Such a node is said to be fathomed. If \( P_6 \) yields a discrete solution, \( P_6 \) is fathomed and
the algorithm backtracks to $P_3$ and switches to solve $P_4$. The branching, backtracking and searching process continues until all the nodes are fathomed. The algorithm searches the tree in a depth-first manner and earns its name “depth-first” search.

When the discrete space has a highly nonuniform grid as in the case of SPT space or the subexpression space, it is more appropriate to minimize the normalized peak ripple magnitude [22]. In order to optimize the normalized peak ripple, the passband gain must be left as a continuous variable. Technique for determining the optimal passband gain has been reported in [22].

In applying the depth-first search B&B algorithm to optimize the coefficient in subexpression space, it is necessary to:

1) determine $\lfloor x \rfloor$ and $x$ for any given coefficient value $x$ subject to a given basis set and a given number of subexpression terms used;
2) determine the basis set of the subexpression space; and
3) determine the number of subexpression terms used for each coefficient.

The basis set and the number of subexpression terms used for each coefficient determine the number of adders required for the implementation of the filter.

In the following, all the above are addressed.

### A. Determination of $\lfloor x \rfloor$ and $x$

The algorithm starts with determining $[\bar{x}]_K$ for any given real number $x$, where $\bar{x}$ is the integer nearest to $x$. $[\bar{x}]_K$ may be determined by exhaustive search or greedy algorithm.

Step 1) Find $[\bar{x}]_K$ by exhaustive search or by greedy algorithm.

Step 2) if $[\bar{x}]_K = x$, let $[x] = [\bar{x}]_K$; go to Step 7; otherwise, go to Step 3.

Step 3) If $[\bar{x}]_K > x$, let $[x] = [\bar{x}]_K$ and $i = -1$; otherwise, let $[x] = [\bar{x}]_K$ and $i = 1$.

Step 4) Let $x' = x$.

Step 5) Update $x' = x' + i$. If $[x']_K = [\bar{x}]_K$, repeat Step 5; otherwise, go to Step 6.

Step 6) If $i > 0$, let $[x] = [\bar{x}]_K$; otherwise, let $[x] = [\bar{x}]_K$.

Step 7) Stop.

In the above algorithm, the most time consuming operation is to find $[\bar{x}]_K$. If the subexpression basis set is static, i.e., if it is not growing or updated during the optimization, a look-up table can be created to store $[\bar{x}]_K$ for all integer $x$ within a given range. Then, using the above algorithm, the upper bound and lower bound of any continuous numbers can be obtained efficiently.

### B. Selection of Basis Set

Higher order basis set includes more subexpression bases requiring more adders to generate these bases, but may result in requiring less subexpression terms to approximate a given set of filter coefficients leading to a reduction in the total number of adders for the entire filter. Therefore, in the implementation, the objective is to determine the basis set that would result in the minimum number of adders required to implement the entire filter. Since the subexpressions 1, and 1 do not require any adders, they are automatic members of all basis sets.

Further investigations show that, for a given order of basis set, generally, contiguous set supports a denser discrete coefficient grid space. For example, if basis, $\{+1, \pm 3, \pm 5, \ldots, \pm 17\}$ are available, $\{\pm 1, \pm 3\}, \{\pm 1, \pm 5\}, \{\pm 1, \pm 7\}, \{\pm 1, \pm 9\}, \{\pm 1, \pm 15\}$, and $\{\pm 1, \pm 17\}$ are all first-order basis sets, but only $\{\pm 1, \pm 3\}$ is a contiguous set. For 10 bits integers, corresponding to integers between 0 and 1 023 (inclusive), the number of integers that are supported by these basis sets with more than 2 subexpression terms are 232, 220, 221, 207, 206 and 190, respectively. Thus, the basis set $\{\pm 1, \pm 3\}$ which supports 232 integers between 0 and 1 023 (inclusive) has the densest discrete coefficient grid space.

Table I lists the number of integers that are supported by contiguous first-, second-, third-, and fourth-order basis set, as well as the most number of integers that can be supported using any other first-, second-, third-, and fourth-order basis sets. The integer ranges from 0 to 1 023 and $K = 2$. From Table I, it is seen that for all the order of basis sets, the contiguous sets support the most number of integers.

Base on this observation, the contiguous basis set is a good choice for constructing the subexpression space. Furthermore, the greedy algorithm can be used to find the nearest discrete value to any number where the basis set is the contiguous basis set.
C. Number of Subexpression Terms

The statistics for the number of subexpression terms used by each coefficient is such that larger magnitude coefficients use more subexpression terms than smaller magnitude coefficients. A given frequency response specification can be met by either increasing the number of subexpression terms per coefficient or by enlarging the number of basis set.

Experience shows that, for the minimum number of adders used, the number of subexpression terms per coefficient should not be more than two; the size of the basis set should be enlarged to meet specifications instead of increasing the number of subexpression terms per coefficient beyond two.

When the maximum number of subexpression terms is limited to not more than 2, generally, the required order of the basis set depends on the coefficient effective wordlength (EWL). If the effective wordlength is not larger than 9, usually, second-order basis set is a good choice to minimize the total number of adders. For effective wordlength larger than 9, each additional bit requires about 3 more subexpression bases. Thus, the suggested order of basis set is

\[
\begin{cases}
  2, & \text{for EWL} \leq 9 \\
  2 + (\text{EWL} - 9) \times 3, & \text{for EWL} > 9.
\end{cases}
\]

IV. Numerical Examples

In this section, several FIR filters are designed using the proposed optimization techniques. The qualities of the results, in terms of the number of adders used to generate the coefficients, are compared with those obtained using two-stage approaches.

A. Example 1

This example is taken from example 2 of [2] (and labeled as L2 in Table IV). The specification is a lowpass filter with filter length \( N = 63 \), passband and stopband edges \( \omega_p = 0.2 \pi, \omega_s = 0.28 \pi \), passband and stopband ripples \( \delta_p = 0.028 \) and \( \delta_s = 0.001 \), and the effective wordlength is 11 bits excluding sign bit.

The subexpression space for the design is constructed based on a seventh-order contiguous basis set \( \{\pm 1, \pm 3, \pm 5, \pm 7, \pm 9, \pm 11, \pm 13, \pm 15\} \). In the 32 distinct coefficients, the 13 coefficients with the largest magnitude in the continuous design are allocated with a maximum of 2 subexpression terms each, whereas the remaining are assigned with a maximum of 1 term each. A total of not more than 45 subexpression terms will be used. The depth-first branch and bound MILP produces a set of coefficients meeting the given specifications using only 42 subexpression terms with 10 coefficients requiring 2 subexpression terms. Therefore, the total number of adders used to produce the coefficients is 17, which can be broken down into 7 adders for constructing the basis set and 10 adders for combining the bases into coefficient values. This design yields a 25% saving in the number of adders when compared with the previously obtained best results [7]–[11]. The filter coefficients expressed in the sum of subexpression terms are listed in Table II.

B. Example 2

Three other filters, labeled as \( S1, S2, \) and \( L3 \) are also designed. \( S1 \) and \( S2 \) are the filters originally given in Example 1 and 2 of [24], whereas \( L3 \) is the filter originally given in Example 3 of [2]. The filter specifications for \( S1 \) and \( S2 \) were not explicitly given in [24]. From the coefficient values given in [24], the frequency responses of the filters were evaluated; the specifications were recovered and tabulated in Table III for easy reference. The total number of adders used to produce the subexpression coefficients as well as the subexpression basis set for each design are listed in Table IV. The best results obtained in the previous literatures are tabulated for comparison. The filter

<p>| TABLE II |
| FILTER SPECIFICATIONS FOR S1 AND S2 [24] |</p>
<table>
<thead>
<tr>
<th>( N )</th>
<th>( \omega_p )</th>
<th>( \omega_s )</th>
<th>( \delta_p )</th>
<th>( \delta_s )</th>
<th>EWL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>25</td>
<td>0.3\pi</td>
<td>0.5\pi</td>
<td>0.0157</td>
<td>0.0066</td>
</tr>
<tr>
<td>S2</td>
<td>60</td>
<td>0.042\pi</td>
<td>0.14\pi</td>
<td>0.012</td>
<td>0.001</td>
</tr>
</tbody>
</table>

* EWL: Effective WordLength (excluding sign bit)

<p>| TABLE III |
| TOTAL NUMBER OF ADDERS USED TO PRODUCE THE FILTER COEFFICIENTS, DENOTED AS ( M ) |</p>
<table>
<thead>
<tr>
<th>Filters</th>
<th>Basis set used</th>
<th>Best Published</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>{\pm 1, \pm 3, \pm 5}</td>
<td>9 [8]–[11]</td>
</tr>
<tr>
<td>S2</td>
<td>{\pm 1, \pm 3, \pm 5, \pm 7, \pm 9, \pm 11, \pm 13, \pm 15}</td>
<td>11 [8]–[10]</td>
</tr>
<tr>
<td>L1</td>
<td>{\pm 1, \pm 3, \pm 5, \pm 7, \pm 9, \pm 11, \pm 13, \pm 15}</td>
<td>14 [8]–[10]</td>
</tr>
<tr>
<td>L2</td>
<td>{\pm 1, \pm 3, \pm 5, \pm 7, \pm 9, \pm 11, \pm 13, \pm 15}</td>
<td>11 [8]–[10]</td>
</tr>
<tr>
<td>L3</td>
<td>{\pm 1, \pm 3, \pm 5}</td>
<td>9 [7]–[11]</td>
</tr>
</tbody>
</table>
coefficients of S1, S2 and L3 are listed in Table V in the form of sum of subexpression terms.

It should be noted that in filter S2, 7 adders are required to construct the basis set (see Table IV), and 13 adders are used to combine the bases into coefficient values (see Table V). Thus, the total number of adders required to generate the coefficient should be 20. However, it is observed from the coefficient values of S2 in Table V that coefficients \( h(13) \) and \( h(14) \) have the same value. Therefore, an adder to sum the subexpression term can be saved, resulting in a total of 19 adders required, as shown in the entries of S2 in Table IV.

The results in Table IV show that the proposed algorithm not only use less adders to meet the same specifications, but also may produce filter coefficients with smaller dynamic ranges, as in the case of filter S2.

<table>
<thead>
<tr>
<th>TABLE V</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMPULSE RESPONSE OF SUBEXPRESSION COEFFICIENT FILTERS</td>
</tr>
<tr>
<td>OF S1, S2 AND L3</td>
</tr>
</tbody>
</table>

Filter \( S1, h(n) = h(24 - n) \) for \( 13 \leq n \leq 24 \)

Passband gain 485.2682

Impulse Response \( \times 512 \):

\[
\begin{align*}
 h(12) &= 3 \times 2^6 - 1 \times 2^3 \\
 h(11) &= 5 \times 2^3 \\
 h(10) &= 3 \times 2^3 \\
 h(9) &= -3 \times 2^3 \\
 h(8) &= -1 \times 2^3 \\
 h(7) &= -3 \times 2^3 \\
 h(6) &= 1 \times 2^3 \\
 h(5) &= 5 \times 2^3 \\
 h(4) &= -1 \times 2^3 \\
 h(3) &= -1 \times 2^3 \\
\end{align*}
\]

Filter \( S2, h(n) = h(59 - n) \) for \( 30 \leq n \leq 59 \)

Passband gain 10945.3361

Impulse Response \( \times 8192 \):

\[
\begin{align*}
 h(29) &= 7 \times 2^7 + 5 \times 2^3 \\
 h(28) &= 7 \times 2^5 + 1 \times 2^3 \\
 h(27) &= 13 \times 2^6 + 5 \times 2^3 \\
 h(26) &= 3 \times 2^8 + 1 \times 2^3 \\
 h(25) &= 11 \times 2^6 + 9 \times 2^3 \\
 h(24) &= 9 \times 2^6 + 11 \times 2^3 \\
 h(23) &= 15 \times 2^5 + 1 \times 2^3 \\
 h(22) &= 3 \times 2^7 + 1 \times 2^3 \\
 h(21) &= 9 \times 2^5 + 15 \times 2^3 \\
 h(20) &= 11 \times 2^4 + 1 \times 2^3 \\
 h(19) &= 3 \times 2^5 \\
 h(18) &= 7 \times 2^3 \\
 h(17) &= -13 \times 2^3 \\
 h(16) &= -1 \times 2^3 \\
 h(15) &= -11 \times 2^3 \\
\end{align*}
\]

Filter \( L3, h(n) = h(35 - n) \) for \( 18 \leq n \leq 35 \)

Passband gain 816.8370

Impulse Response \( \times 1024 \):

\[
\begin{align*}
 h(17) &= 3 \times 2^6 + 1 \times 2^3 \\
 h(16) &= 5 \times 2^5 \\
 h(15) &= 3 \times 2^5 \\
 h(14) &= 1 \times 2^5 \\
 h(13) &= -1 \times 2^4 \\
 h(12) &= -5 \times 2^3 \\
 h(11) &= -1 \times 2^3 \\
 h(10) &= -3 \times 2^3 \\
 h(9) &= -3 \times 2^4 \\
 h(8) &= 5 \times 2^3 \\
 h(7) &= 5 \times 2^3 \\
 h(6) &= 1 \times 2^6 \\
 h(5) &= -1 \times 2^6 \\
 h(4) &= -5 \times 2^3 \\
 h(3) &= -3 \times 2^2 \\
 h(2) &= -3 \times 2^3 \\
 h(1) &= -1 \times 2^6 \\
 h(0) &= 1 \times 2^6 \\
\end{align*}
\]

V. LONG FILTER DESIGN IN SUBEXPRESSION SPACE

For long filters, MILP may require unacceptably long computer time [2]. An approach to overcome this difficulty is to optimize the coefficients in multi-steps. The coefficients are split into two or more groups, say Group A and Group B. First, the coefficients in Group A are optimized in discrete subexpression space, while the others are optimized in continuous space. Next, the coefficients in Group A are fixed at the optimized discrete values and the coefficients in Group B are then optimized in subexpression space.

Such multi-step optimization tremendously accelerates the MILP search; the price to be paid for is the degradation of the resulting filter’s performance. Nevertheless, the multi-step optimization introduces an additional benefit, which may compensate for its disadvantage.

It is noted from the examples in Section V that the optimized discrete coefficient which is allocated with more than 1 subexpression terms may be a value which is not in the basis set. However, such new value cannot serve as a basis to produce other coefficients.

In multi-step optimization, the discrete values produced in earlier steps can serve as bases in later optimization steps. Thus, the supported discrete space is enlarged. For a given basis set, the supported discrete space is sparser in larger magnitude region than in smaller magnitude region. Therefore, intuitively, the coefficients with smaller magnitude in a filter should be optimized in the discrete space earlier than those with larger magnitude. Thus, the new values produced together with the original bases comprise a higher order basis set; this higher order basis set supports denser discrete space in larger magnitude region. Therefore, the optimization of coefficients with larger magnitude will have high chances to obtain better results.

An example taken from example 1 of [2] is used to illustrate the efficiency of this multi-step approach in optimizing long filters in subexpression space.

The filter specification is a 121-tap highpass filter with stop-band and passband edges at 0.747π and 0.85π; the stopband and passband ripples are not larger than 0.0001 and 0.0057, respectively; the effective wordlength of filter coefficients is 14 bits excluding the sign bit.

From our experience, using a 3 GHz CPU desktop PC, the computer time required by MILP becomes unacceptable for filter length longer than 70. In this example, the filter coefficients are split into three groups: group A includes coefficients \( h(0) \) to \( h(29) \); group B includes coefficients from \( h(30) \) to \( h(34) \), and group C includes coefficient from \( h(34) \) to \( h(40) \). The initial basis set is a seventh-order contiguous set, i.e.,

\[ SS_1 = \{ \pm1, \pm3, \pm5, \pm7, \pm9, \pm11, \pm13, \pm15 \} \]

The reason to have group B with only 5 coefficients is that the magnitude of most of the coefficients in group A are small. Many of those only need one subexpression term to represent when they are optimized in the subexpression space based on the initial basis set. There are only a small number of new discrete values being generated to form new bases. Group B is introduced for the purpose of producing more bases.

The optimization procedure runs as follows. First, the coefficients of group A are optimized in discrete subexpression space constructed based on basis set \( SS_1 \), while the coefficients of group B and C are optimized in continuous space. The 11 co-

\[ SS_1 = \{ \pm1, \pm3, \pm5, \pm7, \pm9, \pm11, \pm13, \pm15 \} \]
coefficients with the largest magnitude in the continuous design in group A are allocated with a maximum of two subexpression terms each, whereas the others are allocated with a maximum of one term each. MILP produces a set of coefficients using only 38 subexpression terms with 8 coefficients requiring 2 subexpression terms. Therefore, in addition to the adders used for constructing the basis set, the number of adders used to produce the coefficients of group A is 8. Meanwhile, the 8 coefficient values using 2 subexpression terms are repeatedly divided by 2 until becoming odd numbers. The 8 odd numbers are 29, 31, 57, 59, 89, 113, 129, 145; and they will serve as new basis values in the next step of the optimization. Thus, the basis set for the second step optimization is a noncontiguous set

\[
SS_2 = \{±1, ±3, ±5, ±7, ±9, ±11, ±13, ±15, ±29, ±31, ±57, ±59, ±89, ±113, ±129, ±145\}
\]

with order 15.

In the second step, the coefficients in group B are optimized in discrete subexpression space constructed based on basis set SS_2, while the coefficients of group A are fixed to the optimized discrete values. The coefficients in group C are optimized in continuous space. Each coefficient in group B is allocated with a maximum of 2 subexpression terms. The resulted coefficients in group B uses 9 subexpression terms, which corresponds to 4 adders. Four new values are produced which can be used in the next step of the optimization. They are 29, 31, 89, and 113, which are the quotients of the two-term coefficients by 2 repeatedly until the quotients are odd numbers. Thus, the basis set for the last step of the optimization becomes a 19th-order set

\[
SS_3 = \{±1, ±3, ±5, ±7, ±9, ±11, ±13, ±15, ±29, ±31, ±53, ±57, ±59, ±89, ±113, ±129, ±145, ±183, ±271\}.
\]

In the last step of the optimization, the coefficients in group C are optimized in discrete subexpression space constructed based on basis set SS_3, while the coefficients of group A and B are fixed to the optimized discrete values obtained in the earlier optimization steps. Each coefficient in group C is allocated with a maximum of 2 subexpression terms. The resulted coefficients in group C uses 51 subexpression terms, which corresponds to 25 adders.

After this 3-step optimization, all coefficients are discrete values in the subexpression space constructed based on basis set SS_3. The total number of adders used is 44, which can be broken down into 7 adders for constructing the initial basis set SS_1, 8, and 25 adders for combining the bases into coefficient values for coefficients in group A, B, and C, respectively. The resulting discrete coefficient filter has a stopband ripple of 0.0009704 and a passband ripple of 0.0056, which meet the given specifications. The total number of adders required to realize the coefficients is 18% less than that of the best results published; the number of adders required for the proposed algorithm as well as that for the best published are listed in Table IV for comparison. The filter coefficients obtained using the proposed algorithm are listed in Table VI.

VI. COEFFICIENT CRITICAL PATH LENGTH

In transposed direct form implementation, the “adder step” of a coefficient is the number of adders along the critical path in the realization of the coefficient. The maximum adder step of all coefficients determines the throughput rate of the filter. Therefore, it is desirable to limit the maximum adder step. Using our proposed technique, it is possible to limit the maximum adder step by choosing suitable basis set and limiting the maximum subexpression terms allowed to each coefficient.

For example, the maximum adder step of the coefficients of filter S2 given in Table V is 3, occurring in coefficients

<table>
<thead>
<tr>
<th>Passband gain 59327.7019</th>
<th>Impulse Response ×65536</th>
</tr>
</thead>
<tbody>
<tr>
<td>h(60) = 13 × 2^10 - 1 × 2^1</td>
<td>h(44) = -15 × 2^6 - 1 × 2^2</td>
</tr>
<tr>
<td>h(59) = 3 × 2^12 + 29 × 2^1</td>
<td>h(43) = 129 × 2^2 + 5 × 2^1</td>
</tr>
<tr>
<td>h(58) = 89 × 2^0 + 9 × 2^10</td>
<td>h(42) = 9 × 2^2 + 1 × 2^2</td>
</tr>
<tr>
<td>h(57) = -183 × 2^2 - 145 × 2^5</td>
<td>h(41) = -271 × 2^1 - 11 × 2^0</td>
</tr>
<tr>
<td>h(56) = 183 × 2^3 + 9 × 2^2</td>
<td>h(40) = 89 × 2^2 + 1 × 2^0</td>
</tr>
<tr>
<td>h(55) = 89 × 2^0 + 5 × 2^8</td>
<td>h(39) = -271 × 2^1 + 5 × 2^1</td>
</tr>
<tr>
<td>h(54) = -11 × 2^8 + 13 × 2^3</td>
<td>h(38) = -9 × 2^4 - 5 × 2^0</td>
</tr>
<tr>
<td>h(53) = 5 × 2^9 - 5 × 2^1</td>
<td>h(37) = 129 × 2^1 + 1 × 2^2</td>
</tr>
<tr>
<td>h(52) = -89 × 2^4 + 53 × 2^0</td>
<td>h(36) = -31 × 2^4 + 3 × 2^0</td>
</tr>
<tr>
<td>h(51) = -53 × 2^1</td>
<td>h(35) = 59 × 2^2 + 3 × 2^0</td>
</tr>
<tr>
<td>h(50) = 271 × 2^0 + 59 × 2^4</td>
<td>h(34) = -125 × 2^1</td>
</tr>
<tr>
<td>h(49) = 59 × 2^1 - 53 × 2^5</td>
<td>h(33) = -53 × 2^0</td>
</tr>
<tr>
<td>h(48) = 145 × 2^3 + 15 × 2^1</td>
<td>h(32) = -145 × 2^1</td>
</tr>
<tr>
<td>h(47) = -89 × 2^2 + 1 × 2^0</td>
<td>h(31) = -183 × 2^1</td>
</tr>
<tr>
<td>h(46) = -31 × 2^4 + 5 × 2^0</td>
<td>h(30) = 271 × 2^0</td>
</tr>
<tr>
<td>h(45) = 31 × 2^5 - 5 × 2^1</td>
<td>h(29) = -9 × 2^2</td>
</tr>
</tbody>
</table>

h(28) = -1 × 2^7
h(27) = 15 × 2^4
h(26) = -57 × 2^2
h(25) = 29 × 2^2
h(24) = 1 × 2^5
h(23) = -145 × 2^0
h(22) = 89 × 2^1
h(21) = -129 × 2^0
h(20) = 1 × 2^6
h(19) = 1 × 2^6
h(18) = -59 × 2^1
h(17) = 113 × 2^0
h(16) = -31 × 2^1
h(15) = 3 × 2^1
h(14) = 15 × 2^2
h(13) = 5 × 2^4
h(12) = 1 × 2^6
TABLE VII  

<table>
<thead>
<tr>
<th>Subexpression Coefficient Filter</th>
<th>Impulse Response</th>
<th>Passband gain 10738.8435</th>
<th>256 and 512</th>
</tr>
</thead>
<tbody>
<tr>
<td>h(27), h(25), h(24), h(20) and h(12)</td>
<td>7 x 2^6 + 1 x 2^4</td>
<td>7 x 2^6 + 3 x 2^3</td>
<td>5 x 2^6 + 1 x 2^4</td>
</tr>
<tr>
<td>h(27) = 3 x 2^6 + 7 x 2^3</td>
<td>7 x 2^6 + 3 x 2^3</td>
<td>7 x 2^6 + 3 x 2^6 + 5 x 2^9</td>
<td>5 x 2^6 + 1 x 2^4</td>
</tr>
<tr>
<td>h(26) = 3 x 2^6 + 3 x 2^2</td>
<td>5 x 2^7 + 3 x 2^6</td>
<td>5 x 2^7 + 1 x 2^6</td>
<td>9 x 2^6 - 1 x 2^6</td>
</tr>
<tr>
<td>h(23) = 15 x 2^5 - 7 x 2^2</td>
<td>8 x 2^2 + 3 x 2^3</td>
<td>15 x 2^5 - 7 x 2^2</td>
<td>9 x 2^5 - 1 x 2^4</td>
</tr>
<tr>
<td>h(22) = 3 x 2^6 - 15 x 2^2</td>
<td>2 x 2^2 + 1 x 2^6</td>
<td>3 x 2^6 - 15 x 2^2</td>
<td>2 x 2^2 + 1 x 2^4</td>
</tr>
<tr>
<td>h(21) = 17 x 2^4 - 3 x 2^3</td>
<td>6 x 2^2 + 1 x 2^3</td>
<td>17 x 2^4 - 3 x 2^3</td>
<td>6 x 2^2 + 1 x 2^3</td>
</tr>
<tr>
<td>h(20) = 5 x 2^5 + 1 x 2^4</td>
<td>5 x 2^5 + 1 x 2^4</td>
<td>5 x 2^5 + 1 x 2^4</td>
<td>5 x 2^5 + 1 x 2^4</td>
</tr>
<tr>
<td>h(19) = 3 x 2^5</td>
<td>4 x 2^2</td>
<td>3 x 2^5</td>
<td>4 x 2^2</td>
</tr>
<tr>
<td>h(18) = 7 x 2^2</td>
<td>3 x 2^2</td>
<td>7 x 2^2</td>
<td>3 x 2^2</td>
</tr>
<tr>
<td>h(17) = 3 x 2^3</td>
<td>2 x 2^2</td>
<td>3 x 2^3</td>
<td>2 x 2^2</td>
</tr>
<tr>
<td>h(16) = 31 x 2^1</td>
<td>1 x 2^3</td>
<td>31 x 2^1</td>
<td>1 x 2^3</td>
</tr>
<tr>
<td>h(15) = 5 x 2^4 - 3 x 2^2</td>
<td>0 x 2^2</td>
<td>5 x 2^4 - 3 x 2^2</td>
<td>0 x 2^2</td>
</tr>
</tbody>
</table>

VII. CONCLUSION

Common subexpression sharing in filter coefficients tremendously reduces the number of adders in FIR filter implementations. In this paper, subexpression spaces constructed based on subexpression basis set is presented. A greedy algorithm is proposed to efficiently find the best approximation of a number in a given subexpression space. MILP is used to optimize filter coefficients in subexpression spaces. Examples show that the number of adders used to synthesize the filter coefficients are significantly reduced when comparing with those obtained using the traditional two-stage optimization approach. Conventionally, MILP is notorious for its high computation load for the design of high-order filters. In this paper, a multi-step optimization approach is tailored to circumvent the difficulty of optimizing long filter in subexpression space. The algorithm proposed in this paper also showed its flexibility for designing filters with other constraints such as the maximum adder steps of coefficients when filters are implemented in transposed direct form structure.

REFERENCES


Ya Jun Yu (S’99–M’05) received both the B.Sc. and M.Eng. degrees in biomedical engineering from Zhejiang University, Hangzhou, China, in 1994 and 1997, respectively, and the Ph.D. degree in electrical and computer engineering from the National University of Singapore, Singapore, in 2004.

Since 2005, she has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, where she is currently an Assistant Professor. From 1997 to 1998, she was a Teaching Assistant with Zhejiang University. She joined the Department of Electrical and Computer Engineering, National University of Singapore as a Post Master Fellow in 1998 and remained in the same department as a Research Engineer until 2004. In 2002, she was a Visiting Researcher at the Tampere University of Technology, Tampere, Finland and the Hong Kong Polytechnic University, Hong Kong, China. She joined the Temasek Laboratories at Nanyang Technological University as a Research Fellow in 2004. Her research interests include digital signal processing and VLSI circuits and systems design.

Yong Ching Lim (S’79–M’82–SM’92–F’00) received the A.C.G.I. and B.Sc. degrees in 1977 and the D.I.C. and Ph.D. degrees in 1980, all in electrical engineering, from Imperial College, University of London, U.K.

Since 2003, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, where he is currently a Professor. From 1980 to 1982, he was a National Research Council Research Associate in the Naval Postgraduate School, Monterey, CA. From 1982 to 2003, he was with the Department of Electrical Engineering, National University of Singapore. His research interests include digital signal processing (DSP) and VLSI circuits and systems design.

Dr. Lim was a recipient of the 1996 IEEE Circuits and Systems Society’s Guillemin–Cauer Award, the 1990 IREE (Australia) Norman Hayes Memorial Award, 1977 IEE (U.K.) Prize and the 1974–77 Siemens Memorial (Imperial College) Award. He served as a lecturer for the IEEE Circuits and Systems Society under the distinguished lecturer program from 2001 to 2002 and as an associate editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS from 1991 to 1993 and from 1999 to 2001. He has also served as an Associate Editor for Circuits, Systems and Signal Processing from 1993 to 2000. He served as the Chairman of the DSP Technical Committee of the IEEE Circuits and Systems Society from 1998 to 2000. He served in the Technical Program Committee’s DSP Track as the Chairman in IEEE ISCAS’97 and IEEE ISCAS’00 and as a Co-chairman in IEEE ISCAS’99. He is the General Chairman for IEEE APCCAS’06. He is a member of Eta Kappa Nu.