

# Sparse FIR Filter Design Based on Signomial Programming

Maja Jurisic Bellotti, Mladen Vucic\*

Faculty of Electrical Engineering and Computing, University of Zagreb,  
Unska 3, HR-10000 Zagreb, Croatia  
mladen.vucic@fer.hr

**Abstract**—The goal of sparse FIR filter design is to minimize the number of nonzero filter coefficients, while keeping its frequency response within specified boundaries. Such a design can be formally expressed via minimization of  $l_0$ -norm of filter's impulse response. Unfortunately, the corresponding minimization problem has combinatorial complexity. Therefore, many design methods are developed, which solve the problem approximately, or which solve the approximate problem exactly. In this paper, we propose an approach, which is based on the approximation of the  $l_0$ -norm by an  $l_p$ -norm with  $0 < p < 1$ . We minimize the  $l_p$ -norm using recently developed method for signomial programming (SGP). Our design starts with forming an SGP problem that describes filter specifications. The optimum solution of the problem is then found by using iterative procedure, which solves a geometric program in each iteration. The filters whose magnitude responses are constrained in minimax sense are considered. The design examples are provided illustrating that the proposed method, in most cases, results in filters with higher sparsity than those of the filters obtained by recently published methods.

**Index Terms**—Digital filters; FIR filter design;  $l_p$ -norm; Signomial programming; Sparsity.

## I. INTRODUCTION

Finite impulse response (FIR) filters are used in many signal-processing applications. However, their implementation complexity is higher than the complexity of infinite-impulse-response filters with similar properties [1]. One technique for reducing the complexity is sparse-filter design [2]–[7].

The goal of sparse FIR filter design is to minimize the number of nonzero filter coefficients, while keeping its frequency response within specified boundaries. Such a design can be formally described as:

$$\begin{aligned} \min_{\mathbf{h}} \quad & \|\mathbf{h}\|_0, \\ \text{s. t.} \quad & \|E(\omega, \mathbf{h})\|_{2,\infty} \leq \delta, \quad \omega \in \Omega, \end{aligned} \quad (1)$$

where  $\mathbf{h} = [h_0, h_1, \dots, h_N]^T$  is filter's impulse response,  $E(\omega, \mathbf{h})$  is the approximation error,  $\delta > 0$  is an acceptable upper

bound of the approximation error, and  $\Omega$  is the union of frequency bands of interest. The measure  $\|\cdot\|_0$ , referred to as  $l_0$ -norm, counts the number of nonzero elements in  $\mathbf{h}$ , whereas  $\|\cdot\|_{2,\infty}$  denotes  $l_2$ -norm or  $l_\infty$ -norm, whichever is preferred in the design at hand. The  $l_0$ -norm is not a proper norm. However, we use this name for simplicity.

Additional criteria are sometimes added to (1), e.g., bounding the length of the impulse-response [8]–[10].

The problem in (1) is highly nonconvex due to the  $l_0$ -norm. Its exact solving requires combinatorial optimization as in [11]. However, several approximate methods have been developed. Many of them utilize  $l_0$ -norm and estimate the positions of zero coefficients by using greedy algorithms [12], [13], integer programming [2], [14], iterative second order cone programming [15], iterative shrinkage/thresholding [16], and genetic algorithms [17]. If  $l_0$ -norm in (1) is replaced by  $l_1$ -norm, the filter is obtained with many small coefficients [5]. Some of them can be set to zero and then the design can be repeated by varying only the rest. This idea is used in iterative  $l_1$  methods described in [5] and [18]–[20]. The approximation of  $l_0$ -norm by  $l_p$ -norm with  $0 < p < 1$  has been considered in [21]. Such an approximation relaxes the problem in (1). However, the problem remains nonconvex and it is still difficult to solve.

In this paper, we present a new method for the design of sparse symmetric FIR filters. The sparsity is promoted by using  $l_p$ -norm with  $0 < p < 1$ . To minimize this norm, we propose the signomial programming. We consider the filters whose magnitude responses are constrained in a minimax sense. We illustrate by examples that the proposed method results in filters with higher sparsity than those of the filters obtained by recently published methods.

The paper is organized as follows. Section II briefly describes the signomial programming and its relationship to the geometric programming. Section III presents the proposed method for the design of sparse symmetric FIR filters. Section IV discusses key implementation aspects of the presented design. Finally, Section V provides several filters obtained by using the proposed method and their comparison with the recently published sparse FIR filters.

## II. SIGNOMIAL PROGRAMMING

Signomial programming (SGP) is an optimization method that solves the problem [22]:

Manuscript received 11 June, 2019; accepted 24 January, 2020.  
This work was partly supported by Croatian Science Foundation under grants (No. IP-2014-09-2625 and No. IP-2019-04-4189), and by European Regional Development Fund under grant (No. KK.01.1.1.01.0009 (DATACROSS)).

$$\begin{aligned}
\min \quad & \sum_{j=1}^{m_0} c_{0j} \prod_{i=1}^n x_i^{a_{0ij}}, \\
\text{s. t.} \quad & \sum_{j=1}^{m_k} c_{kj} \prod_{i=1}^n x_i^{a_{kij}} \leq 1, \quad k = 1, 2, \dots, q, \\
& \sum_{j=1}^{m_k} c_{kj} \prod_{i=1}^n x_i^{a_{kij}} = 1, \quad k = q+1, q+2, \dots, r, \\
& x_i > 0, \quad i = 1, 2, \dots, n,
\end{aligned} \tag{2}$$

where  $\mathbf{x} = [x_1, x_2, \dots, x_n]^T \in \mathcal{R}^n$ ,  $c_{kj} \in \mathcal{R}$ ,  $a_{kij} \in \mathcal{R}$ , and  $m_k$ ,  $k = 0, 1, \dots, r$ , is the number of product terms in the objective function and in the constraints.

If  $c_{kj} > 0$  for all  $j$  and  $k$ , and if  $m_k = 1$  for  $q+1 \leq k \leq r$ , the problem in (2) reduces to a geometric programming (GP) problem. Such a problem can be easily transformed into a convex problem, which can be solved globally [22]. Otherwise, the problem in (2) is difficult to solve. However, in recent years, iterative methods have been developed (e.g., see [23] and the references therein), which efficiently solve SGP problems. Here, we apply the method proposed in [23]. This method uses transformations and convexifications to convert the SGP problem into a sequence of GP problems whose solutions converge to the solution of the original problem.

### III. SPARSE FILTER DESIGN

#### A. Problem Formulation

The design of sparse FIR filter whose magnitude response is constrained in a minimax sense is given by the optimization problem:

$$\begin{aligned}
\min_{\mathbf{h}} \quad & \|\mathbf{h}\|_0, \\
\text{s. t.} \quad & |H_a(\omega, \mathbf{h}) - H_d(\omega)| \leq \delta(\omega), \quad \omega \in \Omega,
\end{aligned} \tag{3}$$

where  $H_a(\omega, \mathbf{h})$  is filter's amplitude response,  $H_d(\omega)$  is desired amplitude response, and  $\delta(\omega)$  is maximum acceptable approximation error.

The  $l_0$ -norm in the objective function in (3) can be approximated with the  $l_p$ -norm as follows

$$\|\mathbf{h}\|_p = \left( \sum_{i=0}^N |h_i|^p \right)^{1/p}, \quad 0 < p < 1. \tag{4}$$

Furthermore, the constraint in (3) can be approximated by a set of constraints evaluated on a finite frequency grid  $\omega_q$ ,  $q = 1, 2, \dots, Q$ ,  $\omega_q \in \Omega$ . By assuming

$$\arg \min_{\mathbf{h}} \|\mathbf{h}\|_p = \arg \min_{\mathbf{h}} \|\mathbf{h}\|_p^p, \tag{5}$$

and denoting  $d_q = H_d(\omega_q)$  and  $\delta_q = \delta(\omega_q)$ , the optimization problem in (3) takes the form as expressed below:

$$\begin{aligned}
\min_{\mathbf{h}} \quad & \|\mathbf{h}\|_p^p, \\
\text{s. t.} \quad & |H_a(\omega_q, \mathbf{h}) - d_q| \leq \delta_q, \quad q = 1, 2, \dots, Q.
\end{aligned} \tag{6}$$

The problem in (6) describes the design of FIR filter of an

arbitrary type. For simplicity, we consider only the design of type-I FIR filters in further text, i.e., the filters with even-symmetric impulse responses and even filter orders. Such filters are uniquely described by  $K = N/2 + 1$  impulse response samples as in

$$\mathbf{z} = [z_1, z_2, \dots, z_K]^T = [h_{K-1}, h_K, \dots, h_N]^T. \tag{7}$$

By introducing  $\mathbf{z}$  as the optimization variable, the problem in (6) takes the form

$$\begin{aligned}
\min_{\mathbf{z}} \quad & |z_1|^p + 2|z_2|^p + \dots + 2|z_K|^p, \\
\text{s. t.} \quad & |H_a(\omega_q, \mathbf{z}) - d_q| \leq \delta_q, \quad q = 1, 2, \dots, Q,
\end{aligned} \tag{8}$$

where:

$$H_a(\omega_q, \mathbf{z}) = \mathbf{A}_q \mathbf{z}, \tag{9}$$

$$\mathbf{A}_q = [1 \quad 2\cos(\omega_q) \quad 2\cos(2\omega_q) \quad \dots \quad 2\cos((K-1)\omega_q)]. \tag{10}$$

#### B. SGP Problem

To write the above problem in a form of standard SGP program, simple manipulations are required. To ensure the optimization variables are positive, the substitution is made

$$\mathbf{y} = \mathbf{z} + \boldsymbol{\rho}, \tag{11}$$

where  $\boldsymbol{\rho} = [\rho, \rho, \dots, \rho]^T$  is a vector with  $K$  equal positive elements sufficiently large to ensure  $\mathbf{y} > \mathbf{0}$ . Using (11), the problem in (8) takes the form as:

$$\begin{aligned}
\min_{\mathbf{y}} \quad & |y_1 - \rho|^p + 2|y_2 - \rho|^p + \dots + 2|y_K - \rho|^p, \\
\text{s. t.} \quad & |H_a(\omega_q, \mathbf{y}) - d_q| \leq \delta_q, \quad q = 1, 2, \dots, Q, \\
& \mathbf{y} > \mathbf{0},
\end{aligned} \tag{12}$$

where

$$H_a(\omega_q, \mathbf{y}) = \mathbf{A}_q \mathbf{y} - \mathbf{A}_q \boldsymbol{\rho}. \tag{13}$$

The problem obtained is equivalent to:

$$\begin{aligned}
\min_{\mathbf{t}, \mathbf{y}} \quad & t_1^p + 2t_2^p + \dots + 2t_K^p, \\
\text{s. t.} \quad & y_k - \rho \leq t_k, \quad k = 1, 2, \dots, K, \\
& -y_k + \rho \leq t_k, \quad k = 1, 2, \dots, K, \\
& \mathbf{A}_q \mathbf{y} - (\mathbf{A}_q \boldsymbol{\rho} + d_q) \leq \delta_q, \quad q = 1, 2, \dots, Q, \\
& -\mathbf{A}_q \mathbf{y} + (\mathbf{A}_q \boldsymbol{\rho} + d_q) \leq \delta_q, \quad q = 1, 2, \dots, Q, \\
& \mathbf{t} > \mathbf{0}, \mathbf{y} > \mathbf{0}.
\end{aligned} \tag{14}$$

The problem in (14) is recognized as an SGP problem with  $2K$  variables and  $2(K+Q)$  inequality constraints.

#### C. Solving of SGP Problem

We solve the problem in (14) by using the optimization method presented in [23]. The method referred to is able to solve a general SGP problem shown in (2). However, our problem is simpler because it does not contain equality

constraints. Moreover, our objective function contains only positive terms, i.e., it is a posynomial.

For convenience, we describe here the optimization procedure from [23] tailored for SGP problem in (14). Let us first express the problem in a more compact form:

$$\begin{aligned} \min_{\mathbf{x}} \quad & x_1^p + 2x_2^p + \dots + 2x_K^p, \\ \text{s. t.} \quad & f_j(\mathbf{x}) = a_{0j} + \sum_{i=1}^{2K} a_{ij}x_i \leq 0, \quad j=1, 2, \dots, 2(K+Q), \\ & \mathbf{x} > \mathbf{0}, \end{aligned} \quad (15)$$

where

$$\mathbf{x} = [x_1, x_2, \dots, x_{2K}]^T = \begin{bmatrix} \mathbf{t} \\ \mathbf{y} \end{bmatrix}. \quad (16)$$

The functions  $f_j(\mathbf{x})$ ,  $j = 1, 2, \dots, 2(K+Q)$ , are obtained by simple rearranging of the constraints in (14). Note that we left the constraints in a form  $f_j(\mathbf{x}) \leq 0$  instead of transforming them into a common SGP form  $f_j(\mathbf{x}) \leq 1$ , because such a notation simplifies some of further expressions. The optimization procedure is given by the following algorithm:

```

Step 0: Choose feasible point  $\mathbf{x}^{(0)}$  and
        solution accuracy  $\varepsilon > 0$ .
Step 1: Set  $r = 0$ .
Step 2: Approximate original SGP
        problem with a GP problem
        around  $\mathbf{x}^{(r)}$ .
Step 3: Solve the GP problem to obtain
         $\mathbf{x}^{(r+1)}$ .
Step 4: If  $\|\mathbf{x}^{(r+1)} - \mathbf{x}^{(r)}\| > \varepsilon$ , set  $r =$ 
         $r + 1$  and go to step 2.
Step 5: Return  $\mathbf{x}_{\text{opt}} = \mathbf{x}^{(r+1)}$ .

```

The approximation in Step 2 is rather simple. Since the objective function in (15) is a posynomial, it remains unchanged. On the other hand, the constraints should be approximated by  $f_j^*(\mathbf{x}) \leq 1$ , where  $f_j^*(\mathbf{x})$ ,  $j = 1, 2, \dots, 2(K+Q)$  are also posynomials. Following the procedure described in [23], each constraint in (15) is first written in a form

$$f_j^+(\mathbf{x}) \leq f_j^-(\mathbf{x}), \quad (17)$$

where  $f_j^+(\mathbf{x})$  and  $f_j^-(\mathbf{x})$  collect the terms with  $a_{ij} > 0$  and  $a_{ij} < 0$ , respectively. Note that  $a_{0j}$  is also a part of  $f_j^-(\mathbf{x})$  provided  $a_{0j} < 0$ . The right-hand side of (17) is now approximated with only one positive term, i.e., with a monomial. It is performed by using arithmetic-geometric mean approximation [22], which, in our case, takes the form as

$$g_j(\mathbf{x}) = \prod_{i \in P_j} \left( \frac{\beta_{ij}(\mathbf{x})}{\gamma_{ij}(\mathbf{x}^{(r)})} \right)^{\gamma_{ij}(\mathbf{x}^{(r)})}, \quad (18)$$

where  $\beta_{ij}(\mathbf{x})$  is an  $i^{\text{th}}$  term in the  $j^{\text{th}}$  constraint,  $P_j \subset \{0, 1, \dots,$

$2K\}$  is a set of indexes  $i$ , which appear in  $f_j^-(\mathbf{x})$ , and

$$\gamma_{ij}(\mathbf{x}^{(r)}) = \frac{\beta_{ij}(\mathbf{x}^{(r)})}{f_j^-(\mathbf{x}^{(r)})}. \quad (19)$$

Using (18) and (19), we arrive at

$$f_j^+(\mathbf{x}) \leq g_j(\mathbf{x}). \quad (20)$$

Finally, the GP problem that approximates original SGP problem takes the form:

$$\begin{aligned} \min_{\mathbf{x}} \quad & x_1^p + 2x_2^p + \dots + 2x_K^p, \\ \text{s. t.} \quad & f_j^*(\mathbf{x}) = f_j^+(\mathbf{x}) \times g_j^{-1}(\mathbf{x}) \leq 1, \quad j=1, 2, \dots, 2(K+Q), \\ & \mathbf{x} > \mathbf{0}. \end{aligned} \quad (21)$$

The obtained GP problem can be easily solved in Step 3 by using an available solver for geometric programming, which can be found for example in GGPlab [24] and MOSEK [25]. We have used the latter.

Geometric and signomial programming require positive optimization variables. On the other hand, the sparsity criterion pushes some components in  $\mathbf{t}$  towards zero. The solvers handle this case correctly and return the optimum containing negligible values of such components. However, to shorten the design time, a constraint  $\mathbf{t} \geq \boldsymbol{\mu}$ , where  $\boldsymbol{\mu} = [\mu, \mu, \dots, \mu]^T$ , can be used in (14) instead of  $\mathbf{t} > \mathbf{0}$ . Clearly, a sufficiently small  $\mu$  should be chosen to promote sparsity. We use the value of  $\mu = 1e - 8$ . Note that  $\mathbf{t} \geq \boldsymbol{\mu}$  is already a GP constraint and it can be simply incorporated into (21) as in:

$$\begin{aligned} \min_{\mathbf{x}} \quad & x_1^p + 2x_2^p + \dots + 2x_K^p, \\ \text{s. t.} \quad & f_j^*(\mathbf{x}) \times g_j^{-1}(\mathbf{x}) \leq 1, \quad j=1, 2, \dots, 2(K+Q), \\ & \mu x_k^{-1} \leq 1, \quad k=1, 2, \dots, K, \\ & x_k > 0. \quad k=K+1, K+2, \dots, 2K. \end{aligned} \quad (22)$$

#### IV. PRACTICAL CONSIDERATIONS

The optimization method described in section III.C requires initial point that is feasible to the original SGP problem. An initial point,  $\mathbf{x}_0$ , consists of two parts,  $\mathbf{y}_0$ , representing the initial filter coefficients and  $\mathbf{t}_0$ , representing its sparsity. Therefore, the generation of an initial point starts with the design of initial filter. A nonsparse minimax filter approximates the desired response with a lower approximation error than does the sparse filter of the same order. Therefore, an initial filter suitable for the proposed design can be obtained by using Parks–McClellan algorithm. In our experiments, we generated the impulse response of the initial filter,  $\mathbf{h}_0$ , by using Matlab function *firpm*. We used unweighted design whenever possible. However, if unequal approximation errors are specified in various bands of interest, the initial filter can be generated using appropriate weighting factors. An initial  $\mathbf{z}$ , denoted by  $\mathbf{z}_0$ , is obtained from  $\mathbf{h}_0$  by using (7).

The value of  $\rho$  is not critical. Assuming that  $\mathbf{z}_0$  contains positive and negative components,  $\rho$  is obtained as

$$\rho = -2 \times \min(\mathbf{z}_0). \quad (23)$$

An initial  $\mathbf{y}$ , denoted by  $\mathbf{y}_0$ , is calculated from  $\mathbf{z}_0$  and  $\rho$  using (11). Finally, the initial sparsity is obtained as

$$\mathbf{t}_0 = 1.1 \times |\mathbf{z}_0| + \boldsymbol{\mu}, \quad (24)$$

where  $|\cdot|$  denotes absolute value calculated componentwise.

Figure 1 shows the number of nonzero coefficients of three different low-pass filters obtained for various values of  $p$ . The filters of the 60<sup>th</sup> order are chosen with passband edge  $\omega_p = 0.3\pi$ , stopband edge  $\omega_s = 0.5\pi$ , maximum passband error  $\delta_p = 0.001$  dB, and maximum stopband errors  $\delta_p = -60$  dB,  $\delta_p = -70$  dB, and  $\delta_p = -80$  dB. The optimization is run with  $\varepsilon = 1e - 8$ . Clearly, the number of nonzero coefficients decreases with the decrease of  $p$  reaching the value that is considered the optimum. Experiments show that  $p = 0.1$  is sufficiently low to obtain the optimum in most cases.

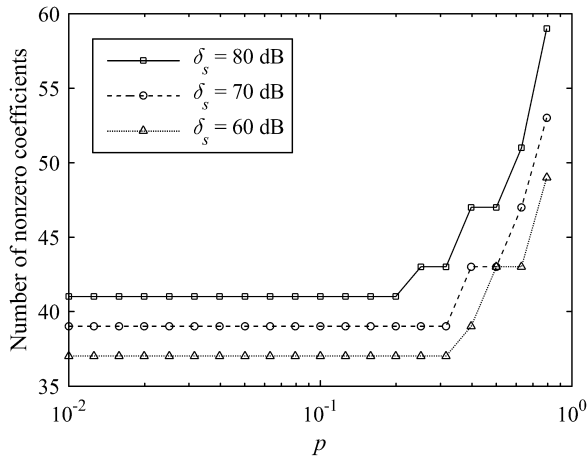


Fig. 1. Number of nonzero coefficients obtained for various values of  $p$  for 60<sup>th</sup> order low-pass filters with passband edge  $\omega_p = 0.3\pi$ , stopband edge  $\omega_s = 0.5\pi$ , maximum passband error  $\delta_p = 0.001$  dB, and maximum stopband errors  $\delta_s = -60$  dB,  $\delta_s = -70$  dB, and  $\delta_s = -80$  dB.

After a sparse filter satisfying the given requirements has been found, its magnitude response can be additionally refined. It is performed by running classic minimax design optimizing only nonzero coefficients whose positions are obtained in the sparse design. Thus, the final filter is obtained by solving the problem:

$$\begin{aligned} \min_{\delta, z_k, k \in C_x} \quad & \delta, \\ \text{s. t.} \quad & w_q |A_q \mathbf{z} - d_q| \leq \delta, \quad q = 1, 2, \dots, Q, \\ & z_k = 0, \quad k \in C_0, \end{aligned} \quad (25)$$

where  $C_0, C_x \subset \{1, 2, \dots, K\}$  contain the indexes of zero and nonzero components in  $\mathbf{z}$  and  $w_q = 1/\delta_q$ .

## V. DESIGN EXAMPLES

To illustrate the features of the proposed method, three examples are provided. In all of them,  $Q = 10N$  and  $\varepsilon = 1e - 8$  are used.

### A. Example 1

The first example illustrates the design of low-pass symmetric FIR filters with passband edge  $\omega_p = 0.3\pi$ , stopband edge  $\omega_s = 0.5\pi$ , maximum passband error  $\delta_p = 0.001$  dB, and maximum stopband errors  $\delta_s = -60$  dB,  $\delta_s = -65$  dB,  $\delta_s = -70$  dB,  $\delta_s = -75$  dB, and  $\delta_s = -80$  dB. The optimization is performed with  $p = 0.1$ . The initial filters are obtained by using Matlab function *firpm* with equally weighted passbands and stopbands.

Table I shows the obtained positions of nonzero coefficients. They are given for the right-hand side of the impulse responses as defined in (7). The total number of nonzero coefficients denoted by  $L_{NZ}$  is shown for our filters and for the filters presented in [19], which are obtained using the same specifications. Our filters exhibit higher sparsities in six cases, which are marked in bold, whereas in other cases the sparsities are equal to those in [19]. The obtained impulse responses are short. Moreover, they remain the same if filter orders increase. Figure 2 shows the right-hand side of a typical impulse response, which is obtained for  $N = 60$  and  $\delta_s = -75$  dB.

TABLE I. OPTIMUM POSITION AND NUMBER,  $L_{NZ}$ , OF NONZERO COEFFICIENTS OF PROPOSED FILTERS WITH PASSBAND  $\omega \in [0, 0.3\pi]$ , STOPBAND  $\omega \in [0.5\pi, \pi]$ , MAXIMUM PASSBAND ERROR  $\delta_p = 0.001$  DB, AND MAXIMUM STOPBAND ERROR  $\delta_s$  OBTAINED FOR INITIAL ORDER  $N$ , COMPARED WITH FILTERS WITH THE SAME SPECIFICATIONS PRESENTED IN [19].

Proposed			[19]	
$N$	$\delta_s$ , dB	Optimum position of nonzero coefficients	$L_{NZ}$	$L_{NZ}$
60	-60	xxxxx0xxxx0xxxx0xxxx0x0000x0000	37	37
70	-60	xxxxx0xxxx0xxxx0xxxx0x0000x000000000	37	37
80	-60	xxxxx0xxxx0xxxx0xxxx0x0000x00000000000000	37	37
60	-65	xxxxx0xxxx0xxxx0xxxx0x0000x0000	<b>37</b>	39
70	-65	xxxxx0xxxx0xxxx0xxxx0x0000x000000000	<b>37</b>	39
80	-65	xxxxx0xxxx0xxxx0xxxx0x0000x00000000000000	<b>37</b>	39
60	-70	xxxxx0xxxx0xxxx0xxxx0xx000x0000	39	39
70	-70	xxxxx0xxxx0xxxx0xxxx0xx000x000000000	39	39
80	-70	xxxxx0xxxx0xxxx0xxxx0xx000x00000000000000	39	39
60	-75	xxxxx0xxxx0xxxx0xxxx0xx000x0000	<b>39</b>	41
70	-75	xxxxx0xxxx0xxxx0xxxx0xx000x000000000	<b>39</b>	41
80	-75	xxxxx0xxxx0xxxx0xxxx0xx000x00000000000000	<b>39</b>	41
60	-80	xxxxx0xxxx0xxxx0xxxx0xxxx000000	41	41
70	-80	xxxxx0xxxx0xxxx0xxxx0xxxx0000000000	41	41
80	-80	xxxxx0xxxx0xxxx0xxxx0xxxx00000000000000	41	41

The  $l_{0.1}$ -norm approximates  $l_0$ -norm closely. Consequently, it forces the sparsity well (Fig. 3). Figure 3 shows the solution of SGP problem in logarithmic scale for the filter with  $N = 60$  and  $\delta_s = -75$  dB. It is clear that minimization of the  $l_{0.1}$ -norm pushes some coefficients towards  $\mu$ , whereas the rest of the coefficients take higher values. Such nearly bimodal distribution is similar to that expected from the  $l_0$ -norm, which differentiates only zero from nonzero coefficients. Clearly, the coefficients near  $\mu$  can be neglected. We consider a coefficient negligible if its

absolute value is less than  $1.01\mu$ . We use this threshold in all further examples.

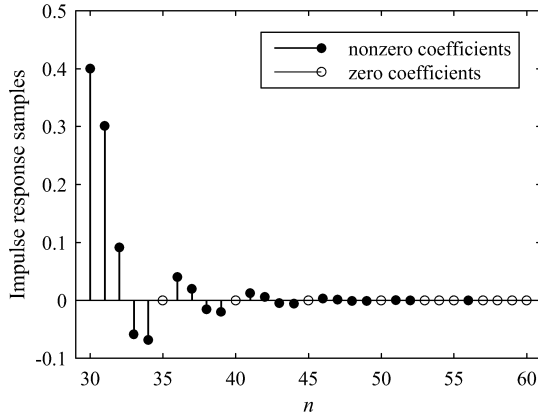


Fig. 2. Impulse response samples of filter with passband  $\omega \in [0, 0.3\pi]$ , stopband  $\omega \in [0.5\pi, \pi]$ , maximum passband error  $\delta_p = 0.001$  dB, and maximum stopband error  $\delta_s = -75$  dB obtained for initial order  $N = 60$ .

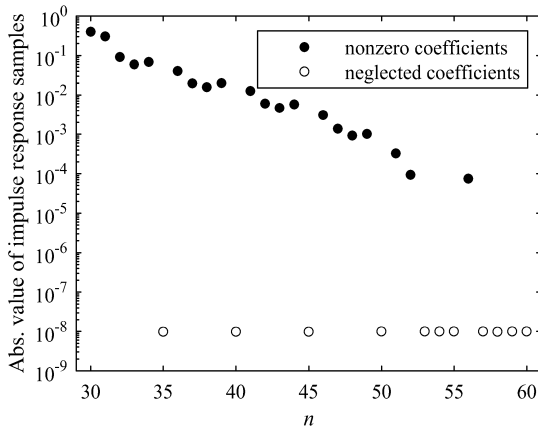


Fig. 3. Absolute values of impulse response samples after solving SGP problem for filter with passband  $\omega \in [0, 0.3\pi]$ , stopband  $\omega \in [0.5\pi, \pi]$ , maximum passband error  $\delta_p = 0.001$  dB, and maximum stopband error  $\delta_s = -75$  dB obtained for initial order  $N = 60$ .

As expected, all of the obtained filters have smaller passband ripple and higher stopband attenuation than the non-sparse minimax filters with the same number of nonzero coefficients. Such a behavior is illustrated in Fig. 4.

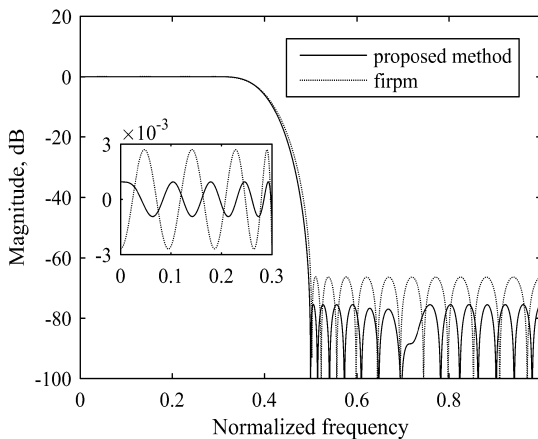


Fig. 4. Magnitude response of sparse low-pass filter obtained for passband  $\omega \in [0, 0.3\pi]$ , stopband  $\omega \in [0.5\pi, \pi]$ , maximum passband error  $\delta_p = 0.001$  dB, maximum stopband error  $\delta_s = -75$  dB and initial order  $N = 60$  having impulse response with 53 samples, of which only 39 are nonzero compared to magnitude response of non-sparse minimax filter with 39 coefficients.

The Figure shows the magnitude response of the sparse filter whose impulse response contains 53 samples, of which only 39 are nonzero compared to the magnitude response of the non-sparse minimax filter with 39 coefficients. Clearly, the improvement in magnitude is paid by an increase in the impulse response length.

### B. Example 2

For the second example, we have chosen low-pass filters with the same passband and stopband, as well as with the same maximum stopband errors as in the Example 1. However, the maximum passband error is increased to  $\delta_p = 0.5$  dB. Sparse filters obtained using these specifications were presented in [20]. Here, we repeat the design by using the proposed method. The optimization is performed with  $p = 0.1$ . Feasible initial filters are obtained by using Matlab function *firpm* with passband and stopband weighting factors equal to 1 and  $\delta_p/\delta_s$ , respectively.

The optimum positions and the total numbers of nonzero coefficients are shown in Table II. Compared to [20], the number of nonzero coefficients is improved for all filters, except for two cases, for which it remained unchanged. The positions of nonzero coefficients exhibit similar behavior to those in the Example 1. The responses are short, with the lengths that are not significantly influenced by the increase of the initial filter orders. The only exception is the filter with  $N = 80$  and  $\delta_s = -65$  dB, for which the result similar to the one with  $N = 70$  and  $\delta_s = -65$  dB is expected.

TABLE II. OPTIMUM POSITIONS AND NUMBER,  $L_{NZ}$ , OF NONZERO COEFFICIENTS OF PROPOSED FILTERS WITH PASSBAND  $\omega \in [0, 0.3\pi]$ , STOPBAND  $\omega \in [0.5\pi, \pi]$ , MAXIMUM PASSBAND ERROR  $\delta_p = 0.5$  DB, AND MAXIMUM STOPBAND ERROR  $\delta_s$  OBTAINED FOR INITIAL ORDER  $N$ , COMPARED WITH FILTERS WITH THE SAME SPECIFICATIONS PRESENTED IN [20].

Proposed		[20]
$N$	$\delta_s$ , dB	$L_{NZ}$
60	-60	25
70	-60	25
80	-60	25
60	-65	25
70	-65	25
80	-65	29
60	-70	25
70	-70	25
80	-70	25
60	-75	27
70	-75	27
80	-75	27
60	-80	29
70	-80	29
80	-80	29

### C. Example 3

The third example describes the design of various filters obtained using specifications in [9]. The paper referred to describes joint minimization of sparsity and filter order. Our method does not support such a design. However, we found the specifications in [9] interesting because these filters have high initial orders, steep transition bands, and low required approximation errors.

The first five columns in Table III contain the specifications of seven different filters from [9], based on which we perform our design. We used  $p = 0.05$  and we

