The Design of Sparse and Non-Sparse FIR Filters using Linear Complementarity Problem Approach

Muhammad Muzammal Naseer, Kamal Harb, Ahmad Nuseirat
Electrical Engineering Department
KFUPM/UOHB University, Saudi Arabia
Email: 1- muz.pak@gmail.com; 2- kamal_harb@hotmail.com; 3- anuseirat@anu.edu.jo

***Abstract*—In this article, the problem of linear phase finite impulse response (FIR) filter design is reconsidered as a linear complementarity problem (LCP) with a weighting strategy. The LCP is not an optimization technique because there is no objective function to optimize; however, quadratic programming, one of the applications of LCP, can be used to find an optimal solution for the 1D FIR filter. Quadratic programs are an extremely important source of applications of LCP; in fact, several algorithms for quadratic programs are based on LCP. It has been shown that, by selecting proper weights, the LCP approach is capable of producing equiripple response. Since length of the impulse response of discrete time filters is often an indicator of computational cost, an algorithm is proposed that iteratively thins the impulse response of a non-sparse filter. The resulting LCP has been solved by a computationally effective Lemke's algorithm. Different examples are presented to illustrate the efficiency of the proposed methods.**

***Keywords*−*Finite Impulse Response (FIR)*; *Lemke's algorithm*; *Linear Comlimentarity Problem (LCP)*; *Quadratic Program(QP)*.**

I. INTRODUCTION

The linear phase finite impulse response (FIR) filters are essential in many applications and there are many well documented techniques in literature for designing such filters [1]–[10]. In [3], Vaidyanathan et. al designed the linear phase FIR filters by minimizing a quadratic measure of the error in the passband and stopband. In [4], Medlin et. al introduced the Lagrange multiplier method to design FIR filters for multirate applications. In [5], M. H. Er and C. K. Siew presented the FIR filter design problem as a quadratic program (QP) with quadratic constraints. Rabiner [2] used the theory of linear programming to design discrete time FIR filters with equiripple response. Nuseirat [1] studied the design problem using the LCP approach.

The classical LCP has been explicitly stated by Du val in 1940 [11]. The LCP is not an optimization technique as there is no objective function to optimize, however, QP, one of the applications of LCP, can be used to find an optimal solution for the linear phase FIR filter. QPs are extremely important source of application for the LCP. There are several highly effective algorithms for solving QPs that are based on the LCP [9] [11].

In filter design problems, it is required to optimize a desired frequency response by minimizing an error norm, which can be measured in L_2 or L_∞ norm. L_2 is known for its poor performance especially at band edges $(Gibb'sPhenomenon)$. In [1], the author tried to improve the performance by simply removing the transition region from the error measure but the presented examples show that a compromise has to be made at one of the band edges.

We revisit the problem with a weighting strategy and show that QP converted into LCP is capable of producing equiripple response filters comparable with Park and McClellan [6] who used polynomial interpolation technique to solve for the desired filters. The resulted LCP-QP is solved by the most robust Lemke's algorithm. Based on pivoting, Lemke is a direct algorithm. It is computationally very effective as no matrix inversion is needed [1].

In many applications, the number of arithmetic operations indicate the cost of implementation, thus reducing the length of impulse response that is designing sparse filters is beneficial not only in terms of computational cost but also in hardware and energy consumption [13] [14]. Sparse filters offer opportunity to omit the arithmetic operations associated with zero-valued coefficients. In this article, a simple iterative algorithm is proposed to reduce the number of coefficients of a non-sparse filter.

This paper is organized as follows: Section II presents the problem formulation, and shows the effect of LCP-QP with weights on the design problem. Section III describes an iterative algorithm to thin the impulse response. Section IV presents the discussion and simulation results. Finally, we conclude in Section V.

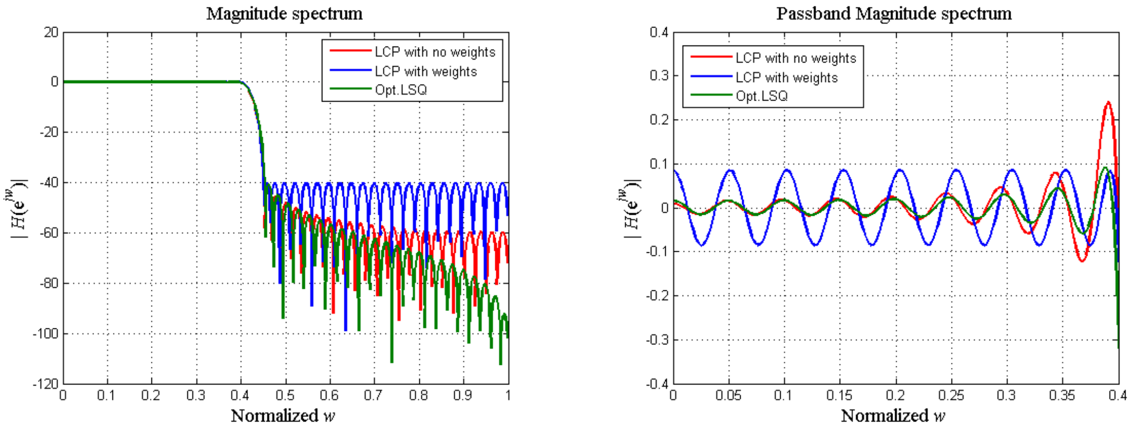
II. PROBLEM FORMULATION

For simplicity of presentation, consider the frequency response of type I FIR filter given by [1] [13] [14]:

$$H(w) = \sum_{n=0}^{N-1} h(n)cos(\omega n), \tag{1}$$

where $\{ h_n\}, n = 0, ..., N-1$ is the impulse response. Discretizing ω as $\omega_k | 1 \leq k \leq L$, the frequency response in (1) can be written in the following matrix form:

$$\mathbf{H} = \mathbf{Ch}, \tag{2}$$



(a) a (b) b

Figure 1. Performance comparison of LCP with and without weights compared with the optimal least square method. It is clear that with appropriate weights LCP approach is capable of producing FIR filters with equiripple response.

where

$$\mathbf{C} = \begin{bmatrix} 1 & \cos(w_1) & . & . & . & \cos(Nw_1) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & \cos(w_m) & . & . & . & \cos(Nw_m) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & \cos(w_L) & . & . & . & \cos(Nw_L) \end{bmatrix}$$

and

$$\mathbf{h} = \begin{bmatrix} h[0] & h[1] & . & . & . & h[N-1] & h[N] \end{bmatrix}^T$$

Since LCP is solved for a positive vector, the impulse response has to be written as a difference of two non-negative vectors [1]:

$$\mathbf{h} = \mathbf{h}^+ - \mathbf{h}^-. \tag{3}$$

Using (2) and (3), the frequency response of a discrete time filter becomes [1]:

$$\mathbf{H} = \mathbf{Bd}, \tag{4}$$

where $\mathbf{B} = \begin{bmatrix} \mathbf{C} & -\mathbf{C} \end{bmatrix}$ and $\mathbf{d} = \begin{bmatrix} \mathbf{h}^+ \\ \mathbf{h}^- \end{bmatrix}$. For a desired frequency response $\mathbf{D}$, error vector will be [1]:

$$\mathbf{E} = \mathbf{Bd} - \mathbf{D}. \tag{5}$$

The problem of finding $\mathbf{h}$ can be formulated by minimizing the squared error [1]:

$$\begin{aligned} &\text{Minimize} \quad \varepsilon(\mathbf{d}) = \mathbf{E}^\tau \mathbf{E} \\ &\text{Subject to} \\ &\qquad\qquad |\mathbf{W} \times \mathbf{E}| \leq \delta, \end{aligned} \tag{6}$$

where $\mathbf{W}$ is a strictly positive weighting vector and δ, is the tolerance scheme. The objective function in (6) can be written as:

$$\begin{aligned} \varepsilon(\mathbf{d}) &= \mathbf{E}^\tau \mathbf{E}, \\ &= (\mathbf{Bd} - \mathbf{D})^\tau (\mathbf{Bd} - \mathbf{D}), \\ &= (\mathbf{d}^\tau \mathbf{B}^\tau - \mathbf{D}^\tau)(\mathbf{Bd} - \mathbf{D}), \\ &= \mathbf{d}^\tau \mathbf{B}^\tau \mathbf{Bd} - \mathbf{d}^\tau \mathbf{B}^\tau \mathbf{D} - \mathbf{D}^\tau \mathbf{Bd} + \mathbf{D}^\tau \mathbf{D}, \\ &= \mathbf{d}^\tau \mathbf{B}^\tau \mathbf{Bd} - 2\mathbf{d}^\tau \mathbf{B}^\tau \mathbf{D} + \mathbf{D}^\tau \mathbf{D}. \\ \varepsilon(\mathbf{d}) &= \frac{1}{2} \mathbf{d}^\tau \mathbf{Qd} - \mathbf{d}^\tau \mathbf{R} + \mathbf{D}^\tau \mathbf{D}, \end{aligned} \tag{7}$$

where $\mathbf{d} \in R^{2N+2}$, $\mathbf{Q} = \mathbf{2B}^\tau \mathbf{B}$ is symmetric and semi-positive definite matrix and $\mathbf{R} = \mathbf{2B}^\tau \mathbf{D}$. The linear constraints in problem (6) can be written in a compact form as follows:

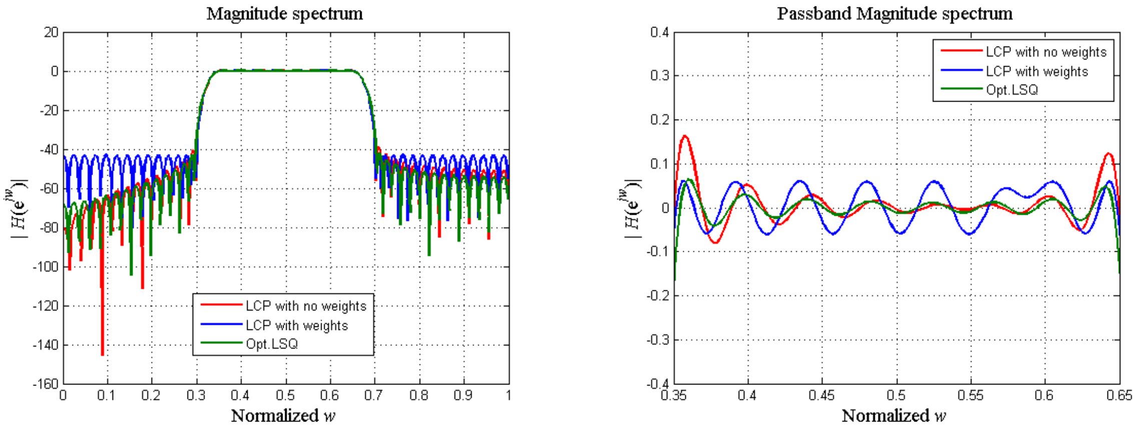
$$\begin{aligned} &|\mathbf{W} \times \mathbf{E}| \leq \delta \\ &|\mathbf{Bd} - \mathbf{D}| \leq \frac{\delta}{\mathbf{W}} \\ &\underbrace{\begin{bmatrix} \mathbf{B} \\ -\mathbf{B} \end{bmatrix}}_{\mathbf{A}} \mathbf{d} \leq \underbrace{\begin{bmatrix} \mathbf{D} + \frac{\delta}{\mathbf{W}} \\ -\mathbf{D} + \frac{\delta}{\mathbf{W}} \end{bmatrix}}_{\mathbf{b}}. \end{aligned} \tag{8}$$

In compact form, the above minimization problem becomes [1]:

$$\begin{aligned} &\text{Minimize} \quad f(\mathbf{d}) = \frac{1}{2} \mathbf{d}^\tau \mathbf{Qd} - \mathbf{d}^\tau \mathbf{R} + \mathbf{D}^\tau \mathbf{D} \\ &\text{Subject to} \\ &\qquad\qquad \mathbf{Ad} \leq \mathbf{b} \\ &\qquad\qquad \mathbf{d} \geq 0. \end{aligned} \tag{9}$$

The Kuhn-Tucker necessary conditions for the above QP in (9) are that there must exist vectors $\mathbf{u} \in R^{2N+2}$, $\mathbf{v} \in R^{2L}$, $\lambda \in R^{2L}$ such that [1] [9] [11]:

$$\begin{aligned} &-\mathbf{R} + \mathbf{Qd} + \mathbf{A}^\tau \lambda - \mathbf{u} = 0, \\ &\mathbf{Ad} + \mathbf{v} = \mathbf{b}, \\ &\mathbf{u} \geq 0, \;\; \mathbf{v} \geq 0, \;\; \mathbf{d} \geq 0, \;\; \lambda \geq 0, \;\; \mathbf{u}^\tau \mathbf{d} = 0, \;\; \mathbf{v}^\tau \lambda = 0. \end{aligned} \tag{10}$$



(a) a

(b) b

Figure 2. Performance comparison of LCP with and without weights compared with the optimal least square method. Again with appropriate weights, the LCP approach produces equiripple response.

Clearly, this can be written as:

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \equiv \begin{pmatrix} -\mathbf{R} \\ \mathbf{b} \end{pmatrix} + \begin{pmatrix} \mathbf{Q} & \mathbf{A}^{\tau} \\ -\mathbf{A} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d} \\ \lambda \end{pmatrix},$$
$$\mathbf{u} \geq 0, \quad \mathbf{v} \geq 0, \quad \mathbf{d} \geq 0, \quad \lambda \geq 0, \quad \mathbf{u}^{\tau}\mathbf{d} = 0, \quad \mathbf{v}^{\tau}\lambda = 0. \tag{11}$$

Thus, the minimization problem in (6) represents a LCP. In a compact form, the LCP for the problem in (6) can be written as [1]:

$$\begin{aligned} &\mathbf{z} - \mathbf{M}\mathbf{w} = \mathbf{q}, \\ &\mathbf{z} \geq 0, \quad \mathbf{w} \geq \mathbf{0}, \quad \mathbf{z}^{\tau}\mathbf{w} = \mathbf{0}, \end{aligned} \tag{12}$$

where $\mathbf{M} = \begin{bmatrix} \mathbf{Q} & \mathbf{A}^{\tau} \\ -\mathbf{A} & 0 \end{bmatrix}$,$\mathbf{z} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$,$\mathbf{w} = \begin{bmatrix} \mathbf{d} \\ \lambda \end{bmatrix}$ and $\mathbf{q} = \begin{bmatrix} -\mathbf{R} \\ \mathbf{b} \end{bmatrix}$. If $\mathbf{Q}$ is semi positive definite then $\mathbf{M}$ is also a semi positive semi-definite matrix [1] [11].

LCP Solver:

There are two main families of algorithms are available to solve for the LCP(q,M) (12): *a*) direct algorithms and *b*) indirect algorithms . In this research, the most robust and direct Lemke's algorithm [1] [9] [11] is used to solve LCP (12).

A. Design Examples

In this section, a set of design examples are provided. The objective here is to show various FIR filter designs via LCP with and without weights compared with optimal least square method. Figure 1 shows the design of 79th-order linear phase lowpass FIR filter. The tolerance scheme for the passband $[0, 0.4\pi]$ and stopband $[0.45\pi, \pi]$ is 0.02 [1].

Another example of the 87th-order linear phase band-pass FIR filter design is shown in Figure 2. Tolerance schemes for the passband $[0.35\pi, 0.65\pi]$ and stopband $[0, 0.3\pi]$, $[0.7\pi, \pi]$ are 0.04 and 0.06 respectively [1].

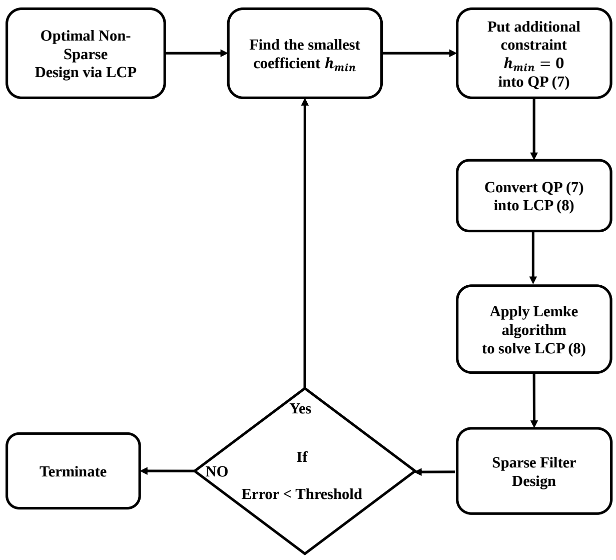


Figure 3. LCP based successive thinning algorithm to design sparse FIR digital filter.

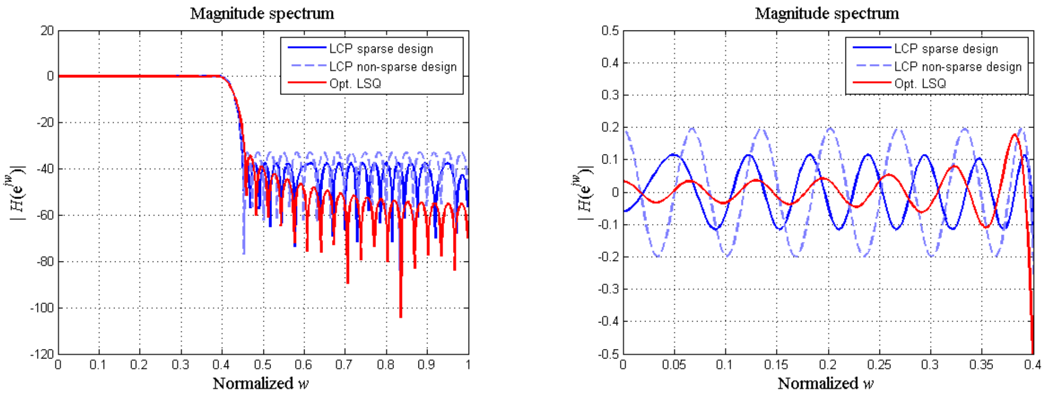
Design examples show that with the appropriate weighting strategy LCP method [1] can lead to FIR filters with equiripple response.

III. Sparse Filter Design using LCP

Since in many applications, the number of arithmetic operations indicate the cost of implementation, thus reducing the length of impulse response that is designing sparse filters is beneficial not only in term of computational cost but also in hardware and energy consumption [13] [14]. To test the performance of LCP in context of spares filter designing, we proposed a simple algorithm shown in Figure 3 that iteratively thins the impulse response of

TABLE I. PERFORMANCE OF LCP IN THE CONTEXT OF SPARSE AND NON-SPARSE FILTER DESIGN

No. of coefficients	Non-zero weights	zero weights	Max.pass-band error [dB]	Min.stop-band attenuation [dB]
79	79	0	0.0852	-40.2
79	61	18	0.1148	-37.57
61	61	0	0.1956	-32.89



(a) a

(b) b

Figure 4. Performance comparison of sparse and non-sparse filters with 61 non-zero co-coefficients. This can be observed that sparse FIR filter designed by LCP offer advantage, when compared to its equivalent non-sparse filter designed by LCP and optimal least square method.

a non-sparse filter. For successive thinning of the impulse response, usually we start with higher no. of coefficients than required [13] [14].

Application of Successive thinning algorithm:

The successive thinning algorithm shown in Figure 3 is applied on the 79th order non-sparse filter shown in Figure 4 with passband $[0, 0.4\pi]$ and stop-band $[0.45\pi, \pi]$. Figure 4 and Table I compare the sparse and non-sparse FIR filters designed by LCP approach.

IV. Discussion Results

In [1], the author used quadratic measure to design digital filters via LCP without weights and tried to minimize the effect of Gibb's phenomenon by ignoring the transition band and simply removing it from the error measure. Thus, the presented examples show that a compromise has to be made at one of the band edges. However, the FIR examples presented in sections II and III showed the efficiency of the proposed weighting strategy.

Moreover, it has been observed that, the LCP-solvers like Lemke's algorithm are very sensitive to the frequency grid. Different passband to stopband frequency grid (p/s-fg) ratio can lead to different solutions. The denser the frequency grid in passband compared to the stop-band, the smaller the error in passband but at a cost of increased error in stopband and vice versa. To show the effect of frequency grid ratio, a 95th order FIR filter with passband $[0, 0.11\pi]$ and stopband $[0.15\pi, \pi]$ designed by LCP is shown in Figure 5.

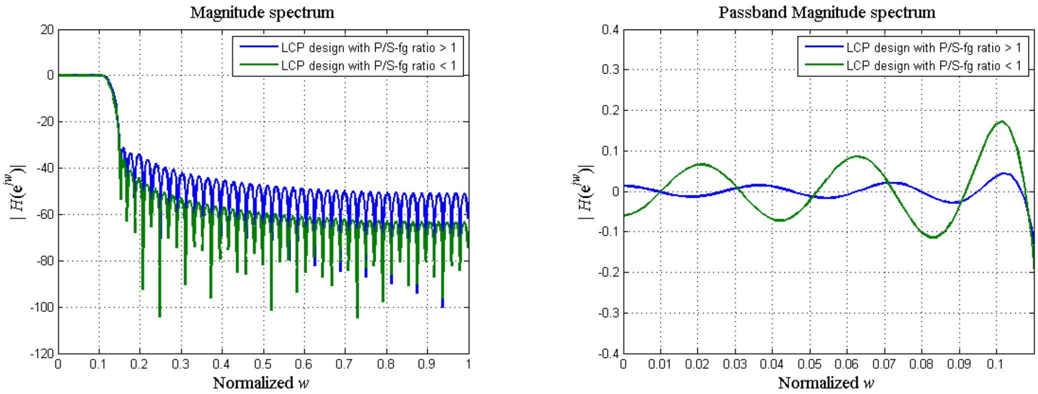
V. Conclusion and Future Work

In this paper, the problem of linear phase FIR filter design is reconsidered as a LCP with a weighting strategy. The LCP is not an optimization technique because there is no objective function to optimize; thus, the design problem of the linear phase 1D FIR filter is formulated via quadratic programming, and then the equivalent semidefinite LCP form is obtained by applying the Karush Kuhn Tucker conditions. One advantage of LCP is its well-developed theory because there are a number of algorithms available to solve a particular LCP. In the case of an FIR filter, the resulted semidefinite LCP is solved by the most robust Lemkes algorithm. It is shown with simulations that with a proper weighting strategy LCP can lead to equiripple solution for 1D FIR filters. In addition, a simple but effective algorithm is presented to design sparse FIR filters. Sparse filters designed by the proposed successive thinning algorithm outperform the non-sparse filters with equal number of non-zero coefficients.

Future work is in progress to extend the LCP technique in order to design two dimensional FIR filter.

References

[1] A. M. Al-Fahed Nuseirat, *"Design of linear phase FIR filters via linear complementarity problem (LCP) approach,"* Interna-



(a) a

(b) b

Figure 5. Performance of LCP designed FIR filters without weights. Effect of passband to stopband frequency grid (p/s-fg) ratio can be observed on Magnitude spectrum.

tional Journal of Circuit Theory and Applications Sept. 2005, vol. 33, no. 5, pp. 353-364.

[2] L. R. Rabiner, *"The design of finite impulse response digital filters using linear programming techniques,"* Bell System Technical Journal Aug. 1972, vol. 51, no.6, pp. 1117-1198.

[3] P. P. Vaidyanathan and T. Q. Nguyen, *"Eigenfilters: a new approach to least-square FIR filter design and applications including Nyquist filters,"* IEEE Transactions on Circuits and Systems, Jan. 1987, vol. 34, no. 1, pp. 11-23.

[4] G. W. Medlin, J. W. Adams, and C.T. Leondes, *"Lagrange multiplier approach to the design of FIR filters for multirate applications,"* IEEE Transactions on Circuits and Systems, Aug. 2002, vol. 35, no.10, pp. 1210-1219.

[5] M. H. Er and C. K. Siew, *"Design of FIR filters using quadratic programming approach,"* IEEE Transactions on Circuits and Systems II Aug. 2002, vol. 42, no. 3, pp. 217-220.

[6] T. W. Parks and J. H. McClellan,*"Chebyshev approximation for nonrecursive digital filters with linear phase,"* IEEE Trans.Circuit Theory, Mar. 1972, vol. 19, pp. 189-194.

[7] V. R. Algazi, M. Suk, and R. Chong-Suck, *"Design of almost minimax FIR filters in one and two dimensions by WLS technique,"* IEEE Transactions on Circuits and Systems Jan. 2003, vol. 33, no. 6, pp. 590-596.

[8] M. Lang, I. W. Selesnick, and C. S. Burrus, *"Constrained least squares design of 2-D FIR filters,"* IEEE Transactions on Signal Processing, May 1996, vol. 44, no. 5, pp. 1234-1241.

[9] K. G. Murty, Linear Complementarity: Linear and Nonlinear Programming, 1997.

[10] A. Chandra and S. Chattopadhyay, *"Design optimization of powers-of-two FIR filter using self-organizing random immigrants GA,"* International Journal of Electronics, 2015, vol. 102, no. 1, pp. 127-140.

[11] O. L. Mangasarian and J. S. Pang, *"The Linear Complementarity Problem,"* SIAM Journal on Matrix Analysis and Applications, 1995, Vol. 16, no. 2, pp. 359-368.

[12] A. Fischer, *"A Newton-Type Method for Positive-Semidefinite Linear Complementarity Problems,"* Journal of optimization theory and applications, September 1995, vol. 86, no. 3, pp. 585-608.

[13] T. Baran, D. Wei, and A. V. Oppenheim, *"Linear Programming Algorithms for Sparse Filter Design,"* IEEE Transaction on signal processing, March 2010, no. 3, vol. 58, pp. 1605-1617.

[14] D. Wei, C. K. Sestok, and A. V. Oppenheim, *"Sparse Filter Design Under a Quadratic Constraint:Low-Complexity Algorithms,"* IEEE Transaction on signal processing, Feb. 2013, no. 4, vol. 61, pp. 857-870.