Design of two-channel low-delay IIR nonuniform-division filter banks using L_1 error criteria

J.-H. Lee and W.-H. Chung

Abstract: The design of a two-channel nonuniform-division filter (NDF) bank with infinite impulse response (IIR) analysis/synthesis filters and low group delay in the sense of L_1 error criteria is considered. The problem formulation results in a nonlinear optimisation problem. Based on a variant of Karmarkar's algorithm, the optimisation problem is solved through a frequency sampling and iterative approximation technique to find the tap coefficients and the reflection coefficients for the numerator and the denominator of the IIR analysis filters. An efficient stabilisation procedure ensures that the reflection coefficients lie in $(-1, 1)$. Simulation results are provided for illustration and comparison.

1 Introduction

Uniform-subband decomposition by using a quadrature mirror filter (QMF) bank is not an appropriate scheme to match the requirements for the subband coding of speech and audio signals. The most appropriate decomposition must consider the critical bands of the ear. It has been mentioned in [1] that these critical bands have nonuniform bandwidths and cannot be easily constructed by the conventional tree structure based on two-channel QMF banks. Hence, it is worth exploiting the design problem of two-channel nonuniform-division filter (NDF) banks.

The conditions on the analysis and synthesis filters of NDF banks, as well as methods to design them, are given in [1]. Because of the difficulty in solving the design problem with nonlinear constraints, a structure and design procedure for pseudo-QMF banks was presented in [2]. The main drawback of this method is that FIR filters with complex coefficients are required by the resulting NDF bank to reduce the aliasing distortion. Recently, one of the authors considered a structure for two-channel NDF banks and proposed design methods for optimally designing NDF banks based on L_1 and L_2 error criteria in [3] and [4], respectively.

Designing an NDF bank has been widely considered in [1–4]. However, linear-phase FIR analysis filters were employed for developing these design techniques. Although using linear-phase FIR analysis filters leads to a favourable problem formulation for design, the system delay k_d of the designed NDF bank is determined by the lengths of FIR filters used and, hence, the long overall system delay of an NDF bank with linear-phase FIR analysis filters may make practical applications prohibitive. It is well known that an IIR filter requires a lower order than an FIR filter under the same stopband energy. For a given passband group delay, IIR NDFs have better passband ripple/stopband attenuation and/or lower computational complexity than FIR NDFs under the same stopband energy for the analysis and synthesis filters. Several approaches dealing with the design problem of two-channel IIR uniform-division filter banks with conventional filter structures have been reported in [5–7]. Recently, the results of designing a two-channel low-delay IIR NDF bank using the L_2 error criteria have been reported in [8].

In this paper, we consider the design of two-channel low-delay IIR NDF banks in the sense of L_1 error criteria. Utilising a lattice structure for the denominators of the IIR analysis filters, a design technique using a modified dual-affine scaling (MDAS) variant of Karmarkar's algorithm [9] in conjunction with an approximation scheme is presented for efficiently solving the resulting design problem that is basically a nonlinear optimisation problem. During the design process, the proposed technique finds the tap coefficients for the numerators and the reflection coefficients for the denominators of the IIR analysis filters simultaneously. It ensures the stability of the designed IIR analysis filters by incorporating an efficient stabilisation procedure to make the magnitude of each reflection coefficient within -1 and $+1$. Computer simulations show that the IIR NDF banks designed by using the proposed technique provide satisfactory performance, though the designed results are only local solutions (instead of the global optimum).

2 Two-channel low-delay NDF banks

Consider the two-channel NDF bank with the architecture given in [3] which is shown in Fig. 1. The analysis low-pass and high-pass filters are designated by $H_0(z)$ and $H_1(z)$, respectively, while the synthesis low-pass and high-pass filters are designated by $F_0(z)$ and $F_1(z)$, respectively. $B_0(z)$ and $B_1(z)$ are two low-pass filters responsible for achieving aliasing-free operation during the rational decimation and interpolation. It can be shown that using the modulations of multiplying $\exp(jn\pi)$ in the high-pass subband channel leads to the favourable result that $B_1(z)$ can be a low-pass filter with real coefficients. The desired magnitude responses

IEE Proceedings online no. 20020485

DOI: 10.1049/ip-vis:20020485

Paper first received 20th July 2001 and in revised form 5th April 2002

The authors are with the Department of Electrical Engineering, National Taiwan University, Taipei, 106, Taiwan, Republic of China

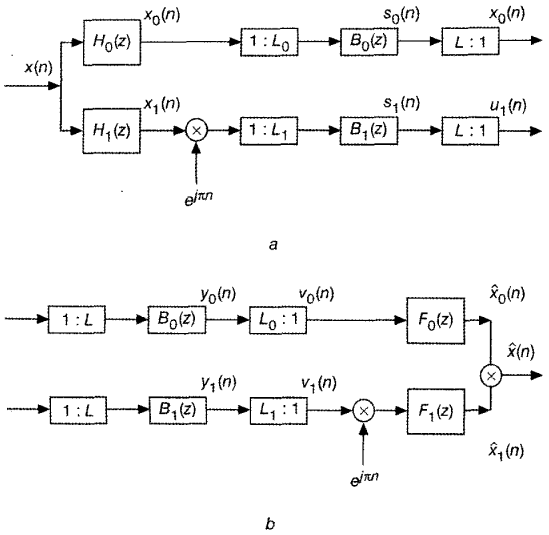

Fig. 1 *Two-channel NDF bank system structure*
a Analysis system
b Synthesis system

for the analysis filters $H_0(z)$ and $H_1(z)$ with passband widths equal to $L_0\pi/L$ and $L_1\pi/L$, respectively, are shown in Fig. 2; where $L = L_0 + L_1 \cdot \omega_p$ and ω_s denote the related band-edge frequencies satisfying $\omega_p + \omega_s = 2\pi L_0/L$.

Assume that the associated magnitude responses are set to

$$\begin{cases} B_0(\omega) = 1, \text{ for } \omega \in \left[0, \dfrac{\omega_s}{L_0}\right] \text{ and} \\ B_0(\omega) = 0, \text{ for } \omega \in \left[\dfrac{2\pi - \omega_s}{L_0}, \pi\right] \\ B_1(\omega) = 1, \text{ for } \omega \in \left[0, \dfrac{\pi - \omega_p}{L_1}\right] \text{ and} \\ B_1(\omega) = 0, \text{ for } \omega \in \left[\dfrac{\pi + \omega_p}{L_1}, \pi\right] \end{cases}$$

Further, assume that $H_0(z)$ and $H_1(z)$ have zero stopband response. As shown in [4], the input/output relationship of the NDF bank in the frequency domain is given by

$$\begin{aligned} \hat{X}(e^{j\omega}) = {} & \frac{e^{-j\omega G_0}}{LL_0}[X(e^{j\omega})H_0(e^{j\omega}) + X(e^{j\omega}W_L^{L_0})H_0(e^{j\omega}W_L^{L_0}) \\ & + X(e^{j\omega}W_L^{-L_0})H_0(e^{j\omega}W_L^{-L_0})]F_0(e^{j\omega}) + \frac{e^{-j\omega G_1}}{LL_1} \\ & \times [X(e^{j\omega})H_1(e^{j\omega}) + X(e^{j\omega}W_L^{L_1})H_1(e^{j\omega}W_L^{L_1}) \\ & + X(e^{j\omega}W_L^{-L_1})H_1(e^{j\omega}W_L^{-L_1})]F_1(e^{j\omega}) \end{aligned} \quad (1)$$

where G_0 and G_1 are the resulting group delays of the upper and lower channels, respectively. $W_L = \exp(-j2\pi/L)$. Substituting $L = L_0 + L_1$, $F_0(e^{j\omega}) = H_0(e^{j\omega})$ and $F_1(e^{j\omega}) = -H_1(e^{j\omega})$ into (1) yields

$$\begin{aligned} \hat{X}(e^{j\omega}) = {} & T(e^{j\omega})X(e^{j\omega}) + V_1(e^{j\omega})X(e^{j\omega}W_L^{L_0}) \\ & + V_2(e^{j\omega})X(e^{j\omega}W_L^{L_1}) \end{aligned} \quad (2)$$

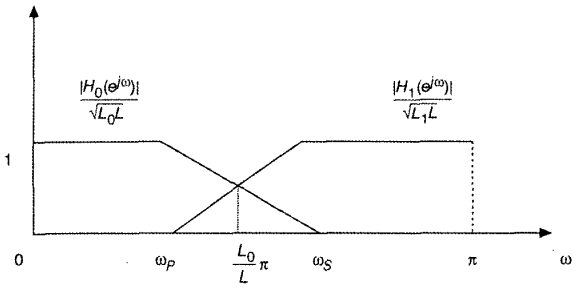

Fig. 2 *Desired magnitude responses for analysis filters*

where:

$$T(e^{j\omega}) = \frac{e^{-j\omega G_0}}{LL_0}H_0^2(e^{j\omega}) - \frac{e^{-j\omega G_1}}{LL_1}H_1^2(e^{j\omega}) \quad (3)$$

$$\begin{aligned} V_1(e^{j\omega}) = {} & \frac{e^{-j\omega G_0}}{LL_0}H_0(e^{j\omega})H_0(e^{j\omega}W_L^{L_0}) \\ & - \frac{e^{-j\omega G_1}}{LL_1}H_1(e^{j\omega})H_1(e^{j\omega}W_L^{L_0}) \end{aligned} \quad (4)$$

$$\begin{aligned} V_2(e^{j\omega}) = {} & \frac{e^{-j\omega G_0}}{LL_0}H_0(e^{j\omega})H_0(e^{j\omega}W_L^{L_1}) \\ & - \frac{e^{-j\omega G_1}}{LL_1}H_1(e^{j\omega})H_1(e^{j\omega}W_L^{L_1}) \end{aligned} \quad (5)$$

The first term of (2) represents the response of a linear shift-invariant system $T(e^{j\omega})$ with input $X(e^{j\omega})$, while the other two terms represent the resulting aliasing distortion. Therefore, perfect reconstruction for a group delay k_d requires the following conditions:

PR 1: $T(e^{j\omega})$ must be equal to $e^{-jk_d\omega}$ for all ω.
PR 2: The magnitude $V_1(\omega)$ of $V_1(e^{j\omega})$ must be zero, i.e. $V_1(\omega) = 0$, for all ω.
PR 3: The magnitude $V_2(\omega)$ of $V_2(e^{j\omega})$ must be zero, i.e. $V_2(\omega) = 0$, for all ω.

For simplicity, assuming $G_0 = G_1$, we can neglect the phase term $e^{-j\omega G_0}$ in (3) and express $T(e^{j\omega})$ as

$$\hat{T}(e^{j\omega}) = \frac{H_0^2(e^{j\omega})}{LL_0} - \frac{H_1^2(e^{j\omega})}{LL_1} \quad (6)$$

Furthermore, we note from (4), (5) and $L = L_0 + L_1$ that the aliasing distortion can be eliminated if

$$\begin{aligned} \hat{V}(e^{j\omega}) = {} & \frac{1}{LL_0}H_0(e^{j(\omega-\omega_p-\omega_s)})H_0(e^{j\omega}) \\ & - \frac{1}{LL_1}H_1(e^{j(\omega-\omega_p-\omega_s)})H_1(e^{j\omega}) \end{aligned} \quad (7)$$

equals zero for $\omega_p \le \omega \le \omega_s$. Following (6) and (7), we can reformulate the conditions required for perfect reconstruction as follows:

$$\begin{aligned} \hat{T}(e^{j\omega}) &= e^{-jk_d\omega}, \text{ for } 0 \le \omega \le \pi \\ H_0(e^{j\omega}) &= 0, \text{ for } \omega_s \le \omega \le \pi \\ H_1(e^{j\omega}) &= 0, \text{ for } 0 \le \omega \le \omega_p \\ \frac{1}{LL_0}H_0(e^{j\omega})H_0(e^{j(\omega-\omega_p-\omega_s)}) &= \frac{1}{LL_1}H_1(e^{j(\omega-\omega_p-\omega_s)})H_1(e^{j\omega}), \\ & \text{for } \omega_p \le \omega \le \omega_s \end{aligned} \quad (8)$$

Hence, the design problem of the IIR NDF banks of Fig. 1 is finding such IIR analysis filters, namely $H_0(z)$

and $H_1(z)$, that the conditions listed in (8) can be approximately met.

3 Design of two-channel IIR NDF banks in the L_1 sense

3.1 Problem formulation

Here, we consider the design of the two-channel IIR NDF as shown in Fig. 1. Let the low-pass analysis filter be an IIR filter with order M_0/N_0 (i.e. M_0 zeros and N_0 poles) and transfer function $H_0(z)=A_0(z)/B_{N_0}(z)$, where the numerator $A_0(z)$ is an M_0th-order polynomial with tap coefficient vector $\boldsymbol{a}_0=[a_{00}, a_{01}, \ldots, a_{0M_0}]^T$, and $B_{N_0}(z)$ is an N_0th-order FIR lattice filter with reflection coefficient vector $\boldsymbol{k}_0=[k_{01}, k_{02}, \ldots, k_{0N_0}]^T$. The superscript T denotes the transpose operation. Similarly, let the high-pass analysis filter be an IIR filter with order M_1/N_1 and transfer function $H_1(z)=A_1(z)/B_{N_1}(z)$, where the numerator $A_1(z)$ is an M_1th-order polynomial with tap coefficient vector $\boldsymbol{a}_1=[a_{10}, a_{11}, \ldots, a_{1M_1}]^T$, and $B_{N_1}(z)$ is an N_1th-order FIR lattice filter with reflection coefficient vector $\boldsymbol{k}_1=[k_{11}, k_{12}, \ldots, k_{1N_1}]^T$. Fig. 3 shows the system structure for $B_{N_0}(z)$ and $B_{N_1}(z)$ which can be obtained from the following recursive formula [10]:

$$\begin{aligned} B_0(z) &= Q_0(z) = 1 \\ B_n(z) &= B_{n-1}(z) + k_n z^{-1} Q_{n-1}(z) \\ Q_n(z) &= k_n B_{n-1}(z) + z^{-1} Q_{n-1}(z) \end{aligned} \tag{9}$$

Hence, the overall design task is to find the required tap and reflection coefficients $\{a_{im}, k_{in}\}$ for the stable IIR filters $H_i(z)$, $i=0, 1$, such that the conditions shown in (8) must be satisfied. Accordingly, we consider that the overall error function E to be minimised in the L_1 sense can be expressed as the following weighted sum of four terms:

$$E = \delta_a + \alpha\delta_1 + \beta\delta_0 + \gamma\delta_t \tag{10}$$

where $\delta_a=\int_{\omega=0}^{\pi}|e^{-jk_d\omega}-\hat{T}(e^{j\omega})|d\omega$, $\delta_1=\int_{\omega=0}^{\omega_p}|H_1(e^{j\omega})|d\omega$, $\delta_0=\int_{\omega=\omega_s}^{\pi}|H_0(e^{j\omega})|d\omega$ and $\delta_t=\int_{\omega_p}^{\omega_s}|\hat{V}(e^{j\omega})|d\omega$. Moreover, α, β and γ represent the relative weights between the four error terms. We note from (10) that the overall error function E is a highly nonlinear function of the tap and reflection coefficients. Therefore, minimising E directly leads to a very highly nonlinear programming problem in addition to the stability problem for $H_0(z)$ and $H_1(z)$.

3.2 Proposed design technique

Let $\{\omega_1=0, \omega_2, \ldots, \omega_I=\omega_p, \omega_{I+1}, \ldots, \omega_{I+J}, \omega_{I+J+1}=\omega_s, \ldots, \omega_{I+J+K}=\pi\}$ be a grid of equidistant frequencies distributed in the range of $\omega=0$ to $\omega=\pi$ for evaluating the magnitude response of the NDF bank and the related error terms defined as above. Moreover, assume that the set S_1 contains the grid points $\{\omega_1=0, \omega_2, \ldots, \omega_I=\omega_p\}$, S_2 contains the grid points $\{\omega_{I+1}, \omega_{I+2}, \ldots, \omega_{I+J}\}$ and S_3 contains the grid points $\{\omega_{I+J+1}=\omega_s, \omega_{I+J+2}, \ldots, \omega_{I+J+K}=\pi\}$. Then, we rewrite (10) as the following approximation:

$$\hat{E} = \hat{\delta}_a + \alpha\hat{\delta}_1 + \beta\hat{\delta}_0 + \gamma\hat{\delta}_t \tag{11}$$

where $\hat{\delta}_a=\sum_{\omega\in S_1\cup S_2\cup S_3}|e^{-jk_d\omega}-\hat{T}(e^{j\omega})|$, $\hat{\delta}_1=\sum_{\omega\in S_1}|H_1(e^{j\omega})|$, $\hat{\delta}_0=\sum_{\omega\in S_3}|H_0(e^{j\omega})|$ and $\hat{\delta}_t=\sum_{\omega\in\omega_p\cup S_2\cup\omega_s}|\hat{V}(e^{j\omega})|$. Next, we construct $\boldsymbol{b}_1^T=[\mathbf{1}_{I+J+K}^T\ \boldsymbol{p}_I^T\boldsymbol{q}_K^T\ \boldsymbol{r}_{J+2}^T]$ with the $(I+J+K)\times 1$ vector $\mathbf{1}_{I+J+K}=[1,\ldots,1]^T$, the $I\times 1$ vector $\boldsymbol{p}_I=[\alpha,\ldots,\alpha]^T$, the $K\times 1$ vector $\boldsymbol{q}_K=[\beta,\ldots,\beta]^T$ and the $(J+2)\times 1$ vector $\boldsymbol{r}_{J+2}=[\gamma,\ldots,\gamma]^T$. Moreover, let $\boldsymbol{e}_l=[\boldsymbol{e}_{a,l}^T\boldsymbol{e}_{1,l}^T\boldsymbol{e}_{0,l}^T\boldsymbol{e}_{t,l}^T]^T$ represent the error vector computed at the lth iteration, where $\boldsymbol{e}_{a,l}$ denotes the $(I+J+K)\times 1$ error vector due to $|e^{-jk_d\omega}-\hat{T}^l(e^{j\omega})|$ computed for $\omega\in S_1\cup S_2\cup S_3$, $\boldsymbol{e}_{1,l}^T$ the $I\times 1$ error vector due to $|H_1^l(e^{j\omega})|$ computed for $\omega\in S_1$, $\boldsymbol{e}_{0,l}$ the $K\times 1$ error vector due to $|H_0^l(e^{j\omega})|$ computed for $\omega\in S_3$ and $\boldsymbol{e}_{t,l}$ the $(J+2)\times 1$ error vector due to $|\hat{V}^l(e^{j\omega})|$ computed for $\omega\in\omega_p\cup S_2\cup\omega_s$, respectively, where $H_i^l(e^{j\omega})$, for $i=0$, 1, and $\hat{T}^l(e^{j\omega})$ are the frequency responses for the analysis filters and the corresponding filter bank response computed at the lth iteration during the design process, respectively. Let $\Delta H_i^l(e^{j\omega})$ and $\Delta\hat{T}^l(e^{j\omega})$ be the deviations for $H_i^l(e^{j\omega})$ and $\hat{T}^l(e^{j\omega})$ corresponding to the filter coefficient perturbation performed at the lth iteration. Accordingly, $\hat{V}^l(e^{j\omega})$ and $\Delta\hat{V}^l(e^{j\omega})$ are the response and the deviation related to (7). Thus, the design problem of using the L_1 criteria to find the optimal coefficients can be formulated as follows:

$$\begin{aligned} &\text{Minimise } \boldsymbol{b}_1^T\boldsymbol{e}_{l+1} \\ &\text{subject to } \begin{cases} |\boldsymbol{T}_d - \boldsymbol{T}^{l+1}| \le \boldsymbol{e}_{a,l+1} \\ |\boldsymbol{H}_1^{l+1}| \le \boldsymbol{e}_{1,l+1} \\ |\boldsymbol{H}_0^{l+1}| \le \boldsymbol{e}_{0,l+1} \\ |\hat{\boldsymbol{V}}^{l+1}| \le \boldsymbol{e}_{t,l+1} \end{cases} \end{aligned} \tag{12}$$

where $\boldsymbol{T}_d=[e^{-jk_d\omega_1}, e^{-jk_d\omega_2}, \ldots, e^{-jk_d\omega_{I+J+K}}]^T$, $\boldsymbol{T}^l=[\hat{T}^l(e^{j\omega_1}), \hat{T}^l(e^{j\omega_2}), \ldots, \hat{T}^l(e^{j\omega_{I+J+K}})]^T$, $\boldsymbol{H}_1^l=[H_1^l(e^{j\omega_1}), H_1^l(e^{j\omega_2}), \ldots, H_1^l(e^{j\omega_I})]^T$, $\boldsymbol{H}_0^l=[H_0^l(e^{j\omega_{I+J+1}}), H_0^l(e^{j\omega_{I+J+2}}), \ldots, H_0^l(e^{j\omega_{I+J+K}})]^T$ and $\hat{\boldsymbol{V}}^l=[\hat{V}^l(e^{j\omega_I}), \hat{V}^l(e^{j\omega_{I+1}}), \ldots, \hat{V}^l(e^{j\omega_{I+J+1}})]^T$. Moreover, the notation of $|\boldsymbol{Y}|\le\boldsymbol{Z}$ denotes that the absolute value of each entry in the matrix (or vector) $\boldsymbol{Y}$ is not greater than the value of the corresponding entry in the matrix (or vector) $\boldsymbol{Z}$. To deal with the nonlinear

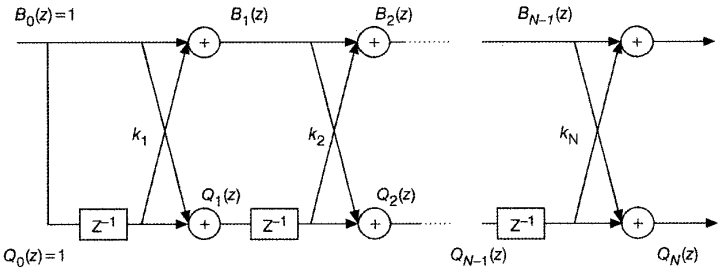

Fig. 3 *Lattice structure for an Nth-order FIR lattice filter*

constraints as shown in (12), we reformulate them as follows:

$$
\begin{aligned}
0 \le\ & |\hat{T}^l(e^{j\omega}) - e^{-jk_d\omega}|^2 - 2\mathrm{Re}\{[|T^l(e^{j\omega}) - e^{-jk_d\omega}]^* \\
& \times \Delta\hat{T}^{l+1}(e^{j\omega})\} + |\Delta\hat{T}^{l+1}(e^{j\omega})|^2 \\
\le\ & |\hat{T}^l(e^{j\omega}) - e^{-jk_d\omega}|^2, \text{ for } \omega \in S_1 \cup S_2 \cup S_3 \\
0 \le\ & |H_1^l(e^{j\omega})|^2 - 2\mathrm{Re}\{H_1^{l*}(e^{j\omega})\Delta H_1^{l+1}(e^{j\omega})\} \\
& + |\Delta H_1^{l+1}(e^{j\omega})|^2 \le |H_1^l(e^{j\omega})|^2, \text{ for } \omega \in S_1 \\
0 \le\ & |H_0^l(e^{j\omega})|^2 - 2\mathrm{Re}\{H_0^{l*}(e^{j\omega})\Delta H_0^{l+1}(e^{j\omega})\} \\
& + |\Delta H_0^{l+1}(e^{j\omega})|^2 \le |H_0^l(e^{j\omega})|^2, \text{ for } \omega \in S_3 \\
0 \le\ & |\hat{V}^l(e^{j\omega})|^2 - 2\mathrm{Re}\{\hat{V}^{l*}(e^{j\omega})\Delta\hat{V}^{l+1}(e^{j\omega})\} \\
& + |\Delta\hat{V}^{l+1}(e^{j\omega})|^2 \le |\hat{V}^l(e^{j\omega})|^2, \text{ for } \omega \in \omega_p \cup S_2 \cup \omega_s
\end{aligned} \tag{13}
$$

where the superscript '$*$' and '$\mathrm{Re}\{x\}$' denote the complex conjugate and the real part of x, respectively. Assume that the perturbations performed on the filter coefficients are significantly small at each iteration; then the second terms due to deviations in (13) can be neglected. As a result, the constraints in (13) can be approximated by the following expressions:

$$
\begin{aligned}
0 \le\ & |\hat{T}^l(e^{j\omega}) - e^{-jk_d\omega}|^2 - 2\mathrm{Re}\{[\hat{T}^l(e^{j\omega}) - e^{-jk_d\omega}]^* \\
& \times \Delta\hat{T}^{l+1}(e^{j\omega})\} \\
\le\ & |\hat{T}^l(e^{j\omega}) - e^{-jk_d\omega}|^2, \text{ for } \omega \in S_1 \cup S_2 \cup S_3 \\
0 \le\ & |H_1^l(e^{j\omega})|^2 - 2\mathrm{Re}\{H_1^{l*}(e^{j\omega})\Delta H_1^{l+1}(e^{j\omega})\} \\
\le\ & |H_1^l(e^{j\omega})|^2, \text{ for } \omega \in S_1 \\
0 \le\ & |H_0^l(e^{j\omega})|^2 - 2\mathrm{Re}\{H_0^{l*}(e^{j\omega})\Delta H_0^{l+1}(e^{j\omega})\} \\
\le\ & |H_0^l(e^{j\omega})|^2, \text{ for } \omega \in S_3 \\
0 \le\ & |\hat{V}^l(e^{j\omega})|^2 - 2\mathrm{Re}\{\hat{V}^{l*}(e^{j\omega})\Delta\hat{V}^{l+1}(e^{j\omega})\} \\
\le\ & |\hat{V}^l(e^{j\omega})|^2, \text{ for } \omega \in \omega_p \cup S_2 \cup \omega_s
\end{aligned} \tag{14}
$$

Next, several useful matrices are defined as follows: $\boldsymbol{U}_a$ is an $(I+J+K)\times 1$ vector with the mth entry given by $U_a(m) = e^{-jk_d\omega_m} - \hat{T}^l(e^{j\omega_m})$, $\omega_m \in S_1 \cup S_2 \cup S_3$, $m = 1, 2, \ldots, I+J+K$, $\mathrm{diag}\{\boldsymbol{U}_a\}$ is a diagonal matrix with the mth diagonal entry given by $U_a(m)$, $\boldsymbol{U}_1$ is an $I\times 1$ vector with the mth entry given by $U_1(m) = -\mathrm{H}_1^l(e^{j\omega_m})$, $\omega_m \in S_1$, $m = 1, 2, \ldots, I$, $\mathrm{diag}\{\boldsymbol{U}_1\}$ is a diagonal matrix with the mth diagonal entry given by $U_1(m)$, $\boldsymbol{U}_0$ is an $J\times 1$ vector with the mth entry given by $U_0(m) = -H_0^l(e^{j\omega_m})$, $\omega_m \in S_3$, $m = I+J+1, \ldots, I+J+K$, $\mathrm{diag}\{\boldsymbol{U}_0\}$ is a diagonal matrix with the mth diagonal entry given by $U_0(m)$, $\boldsymbol{U}_t$ is an $(J+2)\times 1$ vector with the mth entry given by $U_t(m) = -\hat{V}^l(e^{j\omega_m})$, $\omega_m \in \omega_p \cup S_2 \cup \omega_s$, $m = I+1, \ldots, I+J$, $\mathrm{diag}\{\boldsymbol{U}_t\}$ is a diagonal matrix with mth diagonal entry given by $U_t(m)$, the coefficient increment vector $\boldsymbol{v} = [\Delta\boldsymbol{k}_0, \Delta\boldsymbol{a}_0, \Delta\boldsymbol{k}_1, \Delta\boldsymbol{a}_1]^T = [\Delta k_{01}, \Delta k_{02}, \ldots, \Delta k_{0N_0}, \Delta a_{00}, \Delta a_{01}, \ldots, \Delta a_{0M_0}, \Delta k_{11}, \Delta k_{12}, \ldots, \Delta k_{1N_1}, \Delta a_{10}, \Delta a_{11}, \ldots, \Delta a_{1M_1}]^T$, the gradient matrices $\boldsymbol{H}_a = [\nabla\hat{T}^l(e^{j\omega_m})]$ for $\omega_m \in S_1 \cup S_2 \cup S_3$, $\boldsymbol{H}_p = [\nabla\mathrm{H}_1^l(e^{j\omega_m})]$ for $\omega_m \in S_1$, $\boldsymbol{H}_s = [\nabla H_0^l(e^{j\omega_m})]$ for $\omega_m \in S_3$ and $\boldsymbol{H}_t = [\nabla\hat{V}^l(e^{j\omega_m})]$ for $\omega_m \in \omega_p \cup S_2 \cup \omega_s$, where ∇ represents the gradient operator $[\partial/\partial k_{01}, \partial/\partial k_{02}, \ldots, \partial/\partial k_{0N_0}, \partial/\partial a_{00}, \partial/\partial a_{01}, \ldots, \partial/\partial a_{0M_0}, \partial/\partial k_{11}, \partial/\partial k_{12}, \ldots, \partial/\partial k_{1N_1}, \partial/\partial a_{10}, \partial/\partial a_{11}, \ldots, \partial/\partial a_{1M_1}]$. Using the above matrices, we can re-express the constraints approximated by (14) as follows:

$$
\begin{aligned}
0 &\le \mathrm{diag}\{\boldsymbol{U}_a\}^*\boldsymbol{U}_a - 2\mathrm{Re}\{\mathrm{diag}\{\boldsymbol{U}_a\}^*\boldsymbol{H}_a\}\boldsymbol{v} \le \boldsymbol{e}_{a,l+1}^2 \\
0 &\le \mathrm{diag}\{\boldsymbol{U}_1\}^*\boldsymbol{U}_1 - 2\mathrm{Re}\{\mathrm{diag}\{\boldsymbol{U}_1\}^*\boldsymbol{H}_p\}\boldsymbol{v} \le \boldsymbol{e}_{1,l+1}^2 \\
0 &\le \mathrm{diag}\{\boldsymbol{U}_0\}^*\boldsymbol{U}_0 - 2\mathrm{Re}\{\mathrm{diag}\{\boldsymbol{U}_0\}^*\boldsymbol{H}_s\}\boldsymbol{v} \le \boldsymbol{e}_{0,l+1}^2 \\
0 &\le \mathrm{diag}\{\boldsymbol{U}_t\}^*\boldsymbol{U}_t - 2\mathrm{Re}\{\mathrm{diag}\{\boldsymbol{U}_t\}^*\boldsymbol{H}_t\}\boldsymbol{v} \le \boldsymbol{e}_{t,l+1}^2
\end{aligned} \tag{15}
$$

where $\boldsymbol{e}^2$ denotes the vector with each entry equal to the square of the corresponding entry in the vector $\boldsymbol{e}$, and $\boldsymbol{0}$ the zero vector with appropriate size.

Based on the results obtained by (12) and (15), the discrete approximation for the original L_1 design problem can therefore be reformulated as follows:

$$
\begin{aligned}
&\text{Maximise } \boldsymbol{b}^T\boldsymbol{w} \\
&\text{Subject to } \boldsymbol{A}^T\boldsymbol{w} \le \boldsymbol{c}
\end{aligned} \tag{16}
$$

where the vectors $\boldsymbol{b} = [\boldsymbol{0}^T - \boldsymbol{b}_1^T]^T$, $\boldsymbol{w} = [(\boldsymbol{v}^{l+1})^T(e_{l+1})^T]^T$ and $\boldsymbol{c} = [-(\mathrm{diag}\{\boldsymbol{U}_a\}^*\boldsymbol{U}_a)^T - (\mathrm{diag}\{\boldsymbol{U}_1\}^*\boldsymbol{U}_1)^T - (\mathrm{diag}\{\boldsymbol{U}_0\}^*\boldsymbol{U}_0)^T - (\mathrm{diag}\{\boldsymbol{U}_t\}^*\boldsymbol{U}_t)^T \ (\mathrm{diag}\{\boldsymbol{U}_a\}^*\boldsymbol{U}_a)^T \ (\mathrm{diag}\{\boldsymbol{U}_1\}^*\boldsymbol{U}_1)^T \ (\mathrm{diag}\{\boldsymbol{U}_0\}^*\boldsymbol{U}_0)^T \ (\mathrm{diag}\{\boldsymbol{U}_t\}^*\boldsymbol{U}_t)^T]^T$, and the matrix $\boldsymbol{A}$ is given as follows:

$$
\boldsymbol{A} = \begin{bmatrix}
-\boldsymbol{F}_a^T & -\boldsymbol{F}_1^T & -\boldsymbol{F}_0^T & -\boldsymbol{F}_t^T & \boldsymbol{F}_a^T & \boldsymbol{F}_1^T & \boldsymbol{F}_0^T & \boldsymbol{F}_t^T \\
-\boldsymbol{1}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T \\
\boldsymbol{0}^T & -\boldsymbol{1}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T \\
\boldsymbol{0}^T & \boldsymbol{0}^T & -\boldsymbol{1}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T \\
\boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & -\boldsymbol{1}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T & \boldsymbol{0}^T
\end{bmatrix}
$$

where $\boldsymbol{F}_a = 2\mathrm{Re}\{\mathrm{diag}\{\boldsymbol{U}_a\}^*\boldsymbol{H}_a\}$, $\boldsymbol{F}_1 = 2\mathrm{Re}\{\mathrm{diag}\{\boldsymbol{U}_1\}^*\boldsymbol{H}_1\}$, $\boldsymbol{F}_0 = 2\mathrm{Re}\{\mathrm{diag}\{\boldsymbol{U}_0\}^*\boldsymbol{H}_0\}$ and $\boldsymbol{F}_t = 2\mathrm{Re}\{\mathrm{diag}\{\boldsymbol{U}_t\}^*\boldsymbol{H}_t\}$. Equation (16) represents a standard linear programming problem in dual form. In the following, we present an iterative process based on the dual affine-scaling variant of Karmarkar's algorithm (the DAS algorithm) [9] for solving (16).

Introducing slack variables to the formulation of (16), we have

$$
\begin{aligned}
&\text{Maximise } \boldsymbol{b}^T\boldsymbol{w} \\
&\text{Subject to } \boldsymbol{A}^T\boldsymbol{w} + \boldsymbol{z} = \boldsymbol{c}, \boldsymbol{z} \ge \boldsymbol{0}
\end{aligned} \tag{17}
$$

where $\boldsymbol{z}$ is the vector containing slack variables. Assume that we have an interior feasible solution $\boldsymbol{w}^0$ which satisfies $\boldsymbol{A}^T\boldsymbol{w}^0 + \boldsymbol{z}^0 = \boldsymbol{c}$ and $\boldsymbol{z}^0 > 0$ at the initial stage. With these initial solutions $\boldsymbol{w}^0$ and $\boldsymbol{z}^0$, it has been shown in [11] that an appropriate scaling operation must be performed in order to update $\boldsymbol{w}$ and $\boldsymbol{z}$ such that the objective function $\boldsymbol{b}^T\boldsymbol{w}$ can be improved at a faster rate. In [9], it was proposed to scaled the slack variables as follows:

$$
\hat{\boldsymbol{z}} = \boldsymbol{D}_z^{-1}\boldsymbol{z} \tag{18}
$$

where $\boldsymbol{D}_z = \mathrm{diag}\{\boldsymbol{z}\}$. Substituting (18) into (17), we obtain

$$
\begin{aligned}
&\text{Maximise } \boldsymbol{b}^T\boldsymbol{w} \\
&\text{Subject to } \boldsymbol{A}^T\boldsymbol{w} + \boldsymbol{D}_z\hat{\boldsymbol{z}} = \boldsymbol{c}, \hat{\boldsymbol{z}} \ge 0
\end{aligned} \tag{19}
$$

Let the set of the feasible solutions for (17) be given by

$$
\boldsymbol{W} = \{\boldsymbol{w} \in \boldsymbol{R}^r \,|\, \boldsymbol{A}^T\boldsymbol{w} \le \boldsymbol{c}\} \tag{20}
$$

where $r = N_0 + N_1 + M_0 + M_1 + 2(I+J+K+2)$. Then, the set of the feasible scaled slack vectors for (19) is given by

$$
\hat{\boldsymbol{Z}} = \{\hat{\boldsymbol{z}} \in \boldsymbol{R}^{4(I+J+K+1)} \,|\, \exists \boldsymbol{w} \in \boldsymbol{W},\ \boldsymbol{A}^T\boldsymbol{w} + \boldsymbol{D}_z\hat{\boldsymbol{z}} = \boldsymbol{c}\} \tag{21}
$$

From (21), it is easy to show that the corresponding $\boldsymbol{w}$ in $\boldsymbol{W}$ for a given scaled slack vector $\hat{\boldsymbol{z}}$ in $\hat{\boldsymbol{Z}}$ is given by

$$\boldsymbol{w}(\hat{\boldsymbol{z}}) = (\boldsymbol{A}\boldsymbol{D}_z^{-2}\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{D}_z^{-1}(\boldsymbol{D}_z^{-1}\boldsymbol{c} - \hat{\boldsymbol{z}}) \tag{22}$$

and the one-to-one relationship between the feasible directions $\boldsymbol{f}_w$ in $\boldsymbol{W}$ and $\boldsymbol{f}_{\hat{z}}$ in $\hat{\boldsymbol{Z}}$ is given by

$$\boldsymbol{f}_{\hat{z}} = -\boldsymbol{D}_z^{-1}\boldsymbol{A}^T\boldsymbol{f}_w \tag{23}$$

Based on (19) and (22), the feasible direction $\boldsymbol{f}_{\hat{z}}$ can be obtained by computing the gradient of the objective function $\boldsymbol{b}^T\boldsymbol{W}$ with respect to $\hat{\boldsymbol{z}}$ as follows:

$$\boldsymbol{f}_{\hat{z}} = \nabla_{\hat{z}}(\boldsymbol{b}^T(\boldsymbol{w}(\hat{\boldsymbol{z}})) = -\boldsymbol{D}_z^{-1}\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{D}_z^{-2}\boldsymbol{A}^T)^{-1}\boldsymbol{b} \tag{24}$$

Comparing (23) and (24), we obtain

$$\boldsymbol{f}_w = (\boldsymbol{A}\boldsymbol{D}_z^{-2}\boldsymbol{A}^T)^{-1}\boldsymbol{b} \tag{25}$$

After determining $\boldsymbol{f}_w$ from (25), we note that updating $\boldsymbol{w}$ can be carried out as follows:

$$\boldsymbol{w}^{l+1} = \boldsymbol{w}^l + \alpha\boldsymbol{f}_w = \begin{bmatrix} \boldsymbol{v}^l \\ \boldsymbol{e}_l^2 \end{bmatrix} + \alpha\begin{bmatrix} \boldsymbol{f}_v \\ \boldsymbol{f}_{e^2} \end{bmatrix} \tag{26}$$

if a suitable step size α is found, where $\boldsymbol{w}^l$, $\boldsymbol{v}^l$ and $\boldsymbol{e}_l^2$ represent the $\boldsymbol{w}$, $\boldsymbol{v}$ and $\boldsymbol{e}^2$ obtained after the $(l-1)$th iteration, respectively, during the optimisation process. We use this $\boldsymbol{f}_v$, which is the subvector containing the first $r - 2(I+J+K+1)$ entries of the feasible direction $\boldsymbol{f}_w$, as the true descent direction for updating the coefficient increment vector $\boldsymbol{v}$ of the optimisation problem in (19). To find a suitable step size α analytically instead of numerically for updating $\boldsymbol{w}$, we propose an efficient method by considering the feasibility of using the updated slack variable $\boldsymbol{z}^l + \alpha\boldsymbol{f}_z$. First, from (18) and (24), we have the feasible direction for the unscaled slack variable as follows:

$$\mathbf{f}_z = -\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{D}_z^{-2}\boldsymbol{A}^T)^{-1}\boldsymbol{b} = -\boldsymbol{A}^T\boldsymbol{f}_w \tag{27}$$

Then, based on the fact that the updated slack vector $\boldsymbol{z}$ must be a vector with all entries greater than or equal to zero, a suitable step size α can be obtained by taking the most appropriate feasible step in the direction of $\boldsymbol{f}_z$ as follows:

$$\alpha = \rho \times \min\left\{-\frac{(\boldsymbol{z}^l)_j}{(\boldsymbol{f}_z)_j} \,|\, (\boldsymbol{f}_z)_j < 0\right\} \tag{28}$$

where $0 < \rho < 1$ is in general chosen experimentally. $(\boldsymbol{y})_j$ denotes the jth entry of the vector $\boldsymbol{y}$ and $\min\{x\}$ the minimum value of x. Based on (26), the formula for updating $\boldsymbol{v}$ after the lth iteration is then given by

$$\boldsymbol{v}^{l+1} = \boldsymbol{v}^l + \alpha\boldsymbol{f}_v \tag{29}$$

3.2.1 Determination of initial guesses for $H_0(z)$ and $H_1(z)$:

Our design experience shows that using better initial guesses for $H_0(e^{j\omega})$ and $H_1(e^{j\omega})$ usually provides better design results. To initiate the design process, we propose a procedure for determining appropriate initial guesses $H_0^0(z)$ and $H_1^0(z)$ for $H_0(z)$ and $H_1(z)$, respectively. According to the principle of a two-channel NDF bank, we define two desired frequency responses $D_0(e^{j\omega})$ and $D_1(e^{j\omega})$ as

$$D_0(e^{j\omega}) = \begin{cases} \sqrt{LL_0}e^{-j(k_d/2)\omega}, & \text{for } \omega \in [0, \omega_p] \\ \sqrt{\dfrac{LL_0}{2}}e^{-j(k_d/2)\omega}, & \text{for } \omega = \dfrac{\omega_p + \omega_s}{2} \\ 0, & \text{for } \omega \in [\omega_s, \pi] \end{cases} \tag{30}$$

$$D_1(e^{j\omega}) = \begin{cases} 0, & \text{for } \omega \in [0, \omega_p] \\ j\sqrt{\dfrac{LL_1}{2}}e^{-j(k_d/2)\omega}, & \text{for } \omega = \dfrac{\omega_p + \omega_s}{2} \\ j\sqrt{LL_1}e^{-j(k_d/2)\omega}, & \text{for } \omega \in [\omega_s, \pi] \end{cases} \tag{31}$$

Two FIR filters $G_0(z)$ and $G_1(z)$ with orders $2N_0$ and $2N_1$ are designed to optimally approximate $D_0(e^{j\omega})$ and $D_1(e^{j\omega})$, respectively, using conventional least-squares error criteria. Let the resulting filter coefficients be given by $\{h_{00}, h_{01}, \ldots, h_{0(2N_0)}\}$ and $\{h_{10}, h_{11}, \ldots, h_{1(2N_1)}\}$, respectively. Through the use of the model reduction algorithm presented in [12], we find two IIR filters where the numerator and denominator are of order N_0 (corresponding to $G_0(z)$) and of order N_1 (corresponding to $G_1(z)$), respectively. Assume that the IIR filters have denominators $C_0(z)$ and $C_1(z)$ with coefficients $\{c_{00} = 1, c_{01}, \ldots, c_{0N_0}\}$ and $\{c_{10} = 1, c_{11}, \ldots, c_{1N_1}\}$, respectively. Then, the initial lattice systems $B_{N_0}^0$ and $B_{N_1}^0$ with reflection coefficients $\{k_{01}^0, k_{02}^0, \ldots, k_{0N_0}^0\}$ and $\{k_{11}^0, k_{12}^0, \ldots, k_{1N_1}^0\}$ corresponding to $C_0(z)$ and $C_1(z)$, respectively, can be found since there exists a one-to-one correspondence between $\{c_{i0}, c_{i1}, \ldots, c_{iN_i}\}$ and $\{k_{i1}^0, k_{i2}^0, \ldots, k_{iN_i}^0\}$ [13] for $i = 0, 1$. Finally, the best solution for the corresponding initial numerator $A_i^0(z)$, $i = 0, 1$, can be obtained by solving the following optimisation problem

Minimise

$$\left| D_i(e^{j\omega}) - \frac{A_i^0(e^{j\omega})}{B_{N_i}^0(e^{j\omega})} \right|^2, \quad \text{for all } \omega \in [0, \omega_p] \cup \frac{\omega_p + \omega_s}{2} \cup [\omega_s, \pi] \tag{32}$$

For evaluating the related error functions given by (14), we again take a set of discrete frequency points linearly distributed over $S = [0, \omega_p] \cup (\omega_p + \omega_s)/2 \cup [\omega_s, \pi]$. Let $S_d = S_1 \cup (\omega_p + \omega_s)/2 \cup S_3 = \{\omega_1 = 0, \omega_2, \ldots, \omega_I = \omega_p, \omega_{I+1} = (\omega_p + \omega_s)/2, \omega_{I+2} = \omega_s, \ldots, \omega_{I+K+1} = \pi\}$ be the dense grid of frequency points and $\boldsymbol{P}_i$ be a complex $(I+K+1) \times (M_i+1)$ matrix with its (m, n)th element given by

$$P_i(m, n) = \frac{e^{-j(n-1)\omega_m}}{B_{N_i}^0(e^{j\omega_m})}, \quad 1 \le m \le I+K+1, \quad 1 \le n \le M_i + 1 \tag{33}$$

and $\boldsymbol{d}_i$ be a complex $(I+K+1) \times 1$ vector with its mth element given by

$$d_i(m) = D_i(e^{j\omega_m}), \quad 1 \le m \le I+K+1 \tag{34}$$

Then, the initial coefficient vector $\boldsymbol{a}_i^0 = [a_{i0}^0, a_{i1}^0, \ldots, a_{iM_i}^0]^T$ of $A_i^0(z)$ optimal in the least-squares sense for (32) is found by minimising $|\boldsymbol{P}_i\boldsymbol{a}_i^0 - \boldsymbol{d}_i|^2 = (\boldsymbol{P}_i\boldsymbol{a}_i^0 - \boldsymbol{d}_i)^H(\boldsymbol{P}_i\boldsymbol{a}_i^0 - \boldsymbol{d}_i)$, where the superscript H denotes the complex conjugate transpose. Clearly, this leads to the optimal solution given by $\boldsymbol{a}_i^0 = \{\mathrm{Re}\{\boldsymbol{P}_i^H\boldsymbol{P}_i\}\}^{-1}\ [\mathrm{Re}\{\boldsymbol{P}_i^H\boldsymbol{d}_i\}]$, for $i = 0, 1$. After finding the appropriate initial guesses $H_i^0(z) = A_i^0(z)/B_{N_i}^0$ and setting the initial coefficient increment vector $\boldsymbol{v}^0$ to a zero vector, we present an iterative procedure step by step for computing $A_i(z)$ and $B_{N_i}(z)$, for $i = 0, 1$, during the design process.

3.2.2 Iterative procedure:

Step 1: At the lth iteration, let the analysis filters be $H_0^l(e^{j\omega}) = A_0^l(e^{j\omega})/B_{N_0}^l(e^{j\omega})$ and $H_1^l(e^{j\omega}) = A_1^l(e^{j\omega})/B_{N_1}^l(e^{j\omega})$. Compute the gradient matrices $\boldsymbol{H}_a = [\nabla\hat{T}^l(e^{j\omega_m})]$ for $\omega_m \in S_1 \cup S_2 \cup S_3$, $\boldsymbol{H}_p = [\nabla H_1^l(e^{j\omega_m})]$ for $\omega_m \in S_1$, $\boldsymbol{H}_s = [\nabla H_0^l(e^{j\omega_m})]$ for $\omega_m \in S_3$ and $\boldsymbol{H}_t = [\nabla\hat{V}^l(e^{j\omega_m})]$ for

$\omega_m \in \omega_p \cup S_2 \cup \omega_s$. Moreover, find the corresponding $\boldsymbol{U}_a$, $\boldsymbol{U}_1$, $\boldsymbol{U}_0$ and $\boldsymbol{U}_t$. Set $k=0$.

Step 2: Compute $\boldsymbol{z}=\boldsymbol{c}-\boldsymbol{A}^T\boldsymbol{w}=[\boldsymbol{z}_1^T\boldsymbol{z}_2^T\boldsymbol{z}_3^T\boldsymbol{z}_4^T\boldsymbol{z}_5^T\boldsymbol{z}_6^T\boldsymbol{z}_7^T\boldsymbol{z}_8^T]^T$, where $\boldsymbol{z}_1=-\text{diag}\{\boldsymbol{U}_a\}^*\boldsymbol{U}_a+2\text{Re}\{\text{diag}\{\boldsymbol{U}_a\}^*\boldsymbol{H}_a\}\boldsymbol{v}^k+\boldsymbol{e}_{a,l}$, $\boldsymbol{z}_2=-\text{diag}\{\boldsymbol{U}_1\}^*\boldsymbol{U}_1+2\text{Re}\{\text{diag}\{\boldsymbol{U}_1\}^*\boldsymbol{H}_1\}\boldsymbol{v}^k+\boldsymbol{e}_{1,l}$, $\boldsymbol{z}_3=-\text{diag}\{\boldsymbol{U}_0\}^*\boldsymbol{U}_0+2\text{Re}\{\text{diag}\{\boldsymbol{U}_0\}^*\boldsymbol{H}_0\}\boldsymbol{v}^k+\boldsymbol{e}_{0,l}$, $\boldsymbol{z}_4=-\text{diag}\{\boldsymbol{U}_t\}^*\boldsymbol{U}_t+2\text{Re}\{\text{diag}\{\boldsymbol{U}_t\}^*\boldsymbol{H}_t\}\boldsymbol{v}^k+\boldsymbol{e}_{t,l}$, $\boldsymbol{z}_5=\text{diag}\{\boldsymbol{U}_t\}^*\boldsymbol{U}_a-2\text{Re}\{\text{diag}\{\boldsymbol{U}_a\}^*\boldsymbol{H}_a\}\boldsymbol{v}^k$, $\boldsymbol{z}_6=\text{diag}\{\boldsymbol{U}_1\}^*\boldsymbol{U}_1-2\text{Re}\{\text{diag}\{\boldsymbol{U}_1\}^*\boldsymbol{H}_1\}\boldsymbol{v}^k$, $\boldsymbol{z}_7=\text{diag}\{\boldsymbol{U}_0\}^*\boldsymbol{U}_0-2\text{Re}\{\text{diag}\{\boldsymbol{U}_0\}^*\boldsymbol{H}_0\}\ \boldsymbol{v}^k$, $\boldsymbol{z}_8=\text{diag}\{\boldsymbol{U}_t\}^*\boldsymbol{U}_t-2\text{Re}\{\text{diag}\{\boldsymbol{U}_t\}^*\boldsymbol{H}_t\}\boldsymbol{v}^k$.

Step 3: From (25), compute the search vector $\boldsymbol{f}_w=(\boldsymbol{A}\boldsymbol{D}_z^{-2}\boldsymbol{A}^T)^{-1}\boldsymbol{b}=[\boldsymbol{f}_v^T\boldsymbol{f}_{e^2}^T]^T$. Accordingly find the search vector $\boldsymbol{f}_z$ from (27) and the feasible step size α from (28).

Step 4: Using (29), we update the increment vector and obtain $\boldsymbol{v}^{k+1}$ after the kth iteration.

Step 5: Compute the ratio $|(\boldsymbol{b}^T\boldsymbol{w}^k-\boldsymbol{b}^T\boldsymbol{w}^{k+1})/\boldsymbol{b}^T\boldsymbol{w}^k|$. If the ratio is larger than a preset positive real number η_1, then set $k=k+1$ and go to step 2. Otherwise, set $\boldsymbol{v}^{l+1}=\boldsymbol{v}^{k+1}$ and go to step 6.

Step 6: Perform a line search by using the Nelder–Meade simplex algorithm of [14] to find the best step size t with $0<t<1$ to update the numerator and denominators of $H_0(z)$ and $H_1(z)$ such that the error function shown by (11) reaches its minimum:

$$\hat{E}^{l+1}=\sum_{\omega_m\in S_1\cup S_2\cup S_3}|e^{-jk_d\omega_m}-\hat{T}^{l+1}(e^{j\omega_m})|+\alpha\sum_{\omega_m\in S_1}|H_1^{l+1}(e^{j\omega_m})|+\beta\sum_{\omega_m\in S_3}|H_0^{l+1}(e^{j\omega_m})|+\gamma\sum_{\omega_m\in\omega_p\cup S_2\cup\omega_s}|\hat{V}^{l+1}(e^{j\omega_m})| \tag{35}$$

subject to the constraints of $\max|k_{ij}^l+t\Delta k_{ij}|<k_{max}$, for $i=0$, $j=1, 2, \ldots, N_0$, and $i=1$, $j=1, 2, \ldots, N_1$, where $H_0^{l+1}(e^{j\omega})$ and $H_1^{l+1}(e^{j\omega})$ have tap and reflection coefficient vectors given by $\boldsymbol{a}_0^{l+1}=[a_{00}^l+t\Delta a_{00},\ a_{01}^l+t\Delta a_{01}, \ldots, a_{0M_0}^l+t\Delta a_{0M_0}]^T$, $\boldsymbol{k}_0^{l+1}=[k_{01}^l+t\Delta k_{01},\ k_{02}^l+t\Delta k_{02}, \ldots, k_{0N_0}^l+t\Delta k_{0N_0}]^T$, $\boldsymbol{a}_1^{l+1}=[a_{10}^l+t\Delta a_{10}, a_{11}^l+t\Delta a_{11}, \ldots, a_{1M_1}^l+t\Delta a_{1M_1}]^T$ and $\boldsymbol{k}_1^{l+1}=[k_{11}^l+t\Delta k_{11}, k_{12}^l+t\Delta k_{12}, \ldots, k_{1N_1}^l+t\Delta k_{1N_1}]^T$, respectively, and $\hat{T}^{l+1}(e^{j\omega})$ and $\hat{V}^{l+1}(e^{j\omega})$ are the corresponding filter bank response and aliasing distortion. Moreover, k_{max} is a preset maximal absolute value and must be less than 1 for the reflection coefficients in order to ensure the stability of the designed IIR NDF bank.

Step 7: Using the updated coefficient increment vector $\boldsymbol{v}^{l+1}$, we compute the corresponding

$$H_0^{l+1}(e^{j\omega})=\frac{A_0^{l+1}(e^{j\omega})}{B_{N_0}^{l+1}(e^{j\omega})},\qquad H_1^{l+1}(e^{j\omega})=\frac{A_1^{l+1}(e^{j\omega})}{B_{N_1}^{l+1}(e^{j\omega})}$$

and $\boldsymbol{e}_{l+1}$.

Step 8: Compute the ratio $|(\boldsymbol{b}_1^T\boldsymbol{e}_l-\boldsymbol{b}_1^T\boldsymbol{e}_{l+1})/\boldsymbol{b}^T\boldsymbol{e}_l|$. If the ratio is larger than a preset positive real number η_2, then set $l=l+1$ and go to step 1. Otherwise, we terminate the design process.

Remarks: There are two situations where the gradient matrices $\boldsymbol{H}_t$, $\boldsymbol{H}_p$, $\boldsymbol{H}_s$ and $\boldsymbol{H}_a$ may degenerate.

Case 1: The columns of $\boldsymbol{H}_t$, $\boldsymbol{H}_p$, $\boldsymbol{H}_s$ and $\boldsymbol{H}_a$ are not linearly independent. Then, the optimal solution for (16) will not be unique. To find an appropriate optimal solution, we construct matrices $\boldsymbol{G}_t$, $\boldsymbol{G}_p$, $\boldsymbol{G}_s$ and $\boldsymbol{G}_a$ by choosing the independent columns from $\boldsymbol{H}_t$, $\boldsymbol{H}_p$, $\boldsymbol{H}_s$ and $\boldsymbol{H}_a$, and a vector $\boldsymbol{u}$ by choosing the components of $\boldsymbol{v}$ corresponding to the independent columns. Then use $\boldsymbol{G}_t$, $\boldsymbol{G}_p$, $\boldsymbol{G}_s$, $\boldsymbol{G}_a$ and $\boldsymbol{u}$ to replace $\boldsymbol{H}_t$, $\boldsymbol{H}_p$, $\boldsymbol{H}_s$, $\boldsymbol{H}_a$ and $\boldsymbol{v}$ in (16). On the other hand, if only $\boldsymbol{H}_t$ (or $\boldsymbol{H}_p$ or $\boldsymbol{H}_s$ or $\boldsymbol{H}_a$) has columns not linearly independent, then we construct a matrix $\boldsymbol{G}_t$ (or $\boldsymbol{G}_p$ or $\boldsymbol{G}_s$ or $\boldsymbol{G}_a$) by replacing the elements of those columns that are not linearly independent with zero elements from $\boldsymbol{H}_t$ (or $\boldsymbol{H}_p$ or $\boldsymbol{H}_s$ or $\boldsymbol{H}_a$). Then use $\boldsymbol{G}_t$ (or $\boldsymbol{G}_p$ or $\boldsymbol{G}_s$ or $\boldsymbol{G}_a$) to replace $\boldsymbol{H}_t$ (or $\boldsymbol{H}_p$ or $\boldsymbol{H}_s$ or $\boldsymbol{H}_a$) in (16) to obtain an appropriate optimal solution.

Case 2: At the lth iteration, the ith reflection coefficient k_i may have the absolute value equal to k_{max}. To tackle this difficulty, we construct a vector $\boldsymbol{u}$ by eliminating Δk_i of the vector $\boldsymbol{v}$ and four matrices $\boldsymbol{G}_t$, $\boldsymbol{G}_p$, $\boldsymbol{G}_s$ and $\boldsymbol{G}_a$ by eliminating the columns of $\boldsymbol{H}_t$, $\boldsymbol{H}_p$, $\boldsymbol{H}_s$ and $\boldsymbol{H}_a$, corresponding to Δk_i, respectively. Then, we use $\boldsymbol{G}_t$, $\boldsymbol{G}_p$, $\boldsymbol{G}_s$, $\boldsymbol{G}_a$ and $\boldsymbol{u}$ to replace $\boldsymbol{H}_t$, $\boldsymbol{H}_p$, $\boldsymbol{H}_s$, $\boldsymbol{H}_a$ and $\boldsymbol{v}$ in (16).

4 Simulation results

In this section, we present simulation results of designing two-channel IIR NDF banks for illustration. These designs are performed on a personal computer with a Pentium CPU using MATLAB programming language. To perform the design process for all design examples, the spacing for two adjacent frequency points in $[0, \pi]$ is set to $\pi/299$, i.e. the number of grid points taken in $[0, \pi]$ is 300, including ω_p and ω_s. The presented design examples are, in fact, related to the optimal design of two-channel IIR NDF banks with low group delay. The performance for the designed filter bank is evaluated in terms of the peak reconstruction error (PRE), the normalised peak stopband ripple (NPSR) and the stopband error energies (SEE) of the designed $H_0(z)$ and $H_1(z)$. They are given by

$$\text{PRE(dB)}=\max\{|20\log_{10}|\hat{T}(\omega_i)||\}\ \text{for}\ \omega_i\in[0,\pi]$$

$$\text{NPSR}_0\text{(dB)}=-20\log_{10}\left(\max\frac{|H_0(\omega_i)|}{\sqrt{LL_0}}\right)\ \text{for}\ \omega_i\in[\omega_s,\pi]$$

$$\text{NPSR}_1\text{(dB)}=-20\log_{10}\left(\max\frac{|H_1(\omega_i)|}{\sqrt{LL_1}}\right)\ \text{for}\ \omega_i\in[0,\omega_p]$$

$$\text{SEE}_0=\sum_{\omega_i\in S_3}|H_0(\omega_i)|^2,\ \text{SEE}_1=\sum_{\omega_i\in S_1}|H_1(\omega_i)|^2$$

The maximum variation of $\text{GD}(\hat{T}(e^{j\omega_i}))$,

$$\text{MVGD}=\max_{\omega_i\in[0,\pi]}|\text{GD}(\hat{T}(e^{j\omega_i}))-k_d|\text{(samples)}$$

The maximum variation of passband $\text{GD}(H_0(e^{j\omega_i}))$,

$$\text{MVPGD}_0=\max_{\omega_i\in[0,\omega_p]}|\text{GD}(H_0(e^{j\omega_i}))-k_d/2|\text{(samples)}$$

The maximum variation of passband $\text{GD}(H_1(e^{j\omega_i}))$,

$$\text{MVPGD}_1=\max_{\omega_i\in[\omega_s,\pi]}|\text{GD}(H_1(e^{j\omega_i}))-k_d/2|\text{(samples)}$$

and the maximum variation of filter bank response,

$$\text{MVFBR}=\max_{\omega_i\in[0,\pi]}|e^{-j\omega_i k_d}-\hat{T}(e^{j\omega_i})|$$

where $\text{GD}(x)$ denotes the group delay of x.

Example 1: The design specifications used are the same as those of example 2 given in [8] and are listed as follows: the group delay k_d is 29, $\omega_p=0.12\pi$ and $\omega_s=0.28\pi$, $L_0=1$ and $L_1=4$, $M_0=13$ and $N_0=14$, and $M_1=N_1=17$. Hence, the number of independent coefficients for the design is 63. The value of ρ and k_{max} required by (28) and (35) are set to 0.1 and 0.98, respectively.

Table 1: Significant design results for design example 1

	IIR NDF bank	FIR NDF bank	IIR NDF bank on L_2 [8]
Filter order	13/14, 17/17	49, 49	13/14, 17/17
No. of coefficients	63	100	63
PRE, DB	0.0086	0.0183	0.0176
MVGD	0.0587	0.0809	0.0703
$NPSR_0$, dB	40.62	39.66	40.88
$NPSR_1$, dB	42.11	40.63	40.66
$MVPGD_0$	0.0075	0.0286	0.0091
$MVPGD_1$	0.0187	0.0199	0.0247
SEE_0	2.97×10^{-3}	4.00×10^{-3}	4.05×10^{-3}
SEE_1	5.24×10^{-3}	7.60×10^{-3}	7.42×10^{-3}
Max. variation of filter bank response	1.18×10^{-3}	2.60×10^{-3}	2.10×10^{-3}
No. of iterations	527	238	41

Table 2: Tap and reflection coefficients for design example 1

i	a_{oi}	k_{oi}	a_{1i}	k_{1i}
0	−0.00284396842596		0.00098569514329	
1	0.02107418338718	−0.85867144175667	0.00296755136857	−0.81601544762663
2	−0.07481434954937	0.93089115660287	0.00453136101016	0.86359343843395
3	0.17488031325743	−0.89975082116660	0.00662730146165	0.00042393418844
4	−0.30825708250530	0.88908873074420	0.00623436576810	0.11068465274787
5	0.43719272004115	−0.89591623447826	0.00286464198576	0.10981418221419
6	−0.51607256399714	0.89148506572985	−0.00510509678221	0.10352590517617
7	0.51416981237393	−0.86152169249230	−0.01693293851708	0.09838880361666
8	−0.42958084918359	0.78201889322278	−0.03015868644344	0.09329256952786
9	0.29701273335629	−0.63717415951998	−0.03901970017885	0.08616457602272
10	−0.16282725395564	0.56748828713168	−0.03440558319699	0.07572921726086
11	0.06736808025591	−0.57468024106798	−0.00207815513623	0.06211428970049
12	−0.01825155562653	0.42581314584237	0.08691866125004	0.04675521774424
13	0.00262077131619	−0.14586255149631	0.33650173940870	0.03173514020921
14		0.01041482556264	1.87314592226429	0.01898136265320
15			−6.17662677471384	0.00965586483405
16			5.79249202800981	0.00390268942021
17			−1.80846418983885	0.00103185011984

Moreover, the values of α, β, γ, η_1 and η_2 are set to 0.015, 0.015, 5×10^{-8}, 10^{-2} and 10^{-6}, respectively, for the design using the proposed technique. The design results with FIR filters using the proposed technique and the results of using the technique [8] to design the same example are also presented for comparison. In the FIR case, $H_0(z)$ and $H_1(z)$ have orders equal to 49. Table 1 shows the significant design results for the designed NDF banks. The resulting tap and reflection coefficients of the IIR NDF bank designed by the proposed technique are listed in Table 2. Figs. 4 and 5 plot the corresponding magnitude responses and group delays of $H_0(e^{j\omega})/\sqrt{(LL_0)}$ and $H_1(e^{j\omega})/\sqrt{(LL_1)}$, respectively. Figs. 6 and 7 depict the magnitude response and the group delay response of $\hat{T}(e^{j\omega})$, respectively, for the IIR NDF bank designed by the proposed technique. Finally, the magnitude response of $e^{-jk_d\omega} - \hat{T}(e^{j\omega})$ is shown in Fig. 8. From the presented design results, we note that the proposed technique provides very satisfactory performance.

Example 2: We use the following design specifications: the group delay k_d is 19; $\omega_p = 0.3\pi$ and $\omega_s = 0.5\pi$; $L_0 = 2$ and $L_1 = 3$; $M_0 = N_0 = 10$ and $M_1 = N_1 = 11$. In the FIR case, $H_0(z)$ and $H_1(z)$ have orders equal to 32 and 33, respectively. Accordingly, the numbers of independent coefficients for the IIR and FIR NDF banks are 44 and 67, respectively. The values of ρ and k_{max} are the same as those of example 1. Moreover, the values of α, β, γ, η_1 and η_2 are set to 0.012, 0.012, 0.055, 10^{-2} and 3.5×10^{-7}, respectively, for the design using the proposed technique. The design results with FIR filters using the proposed technique and the results of using the technique [8] to design the same example are also presented for comparison. Table 3 shows the significant design results for the designed NDF banks. The resulting tap and reflection coefficients of the IIR NDF bank obtained after CPU time = 112 s using the proposed technique are given in Table 4. Figs. 9 and 10 plot the corresponding magnitude responses and group delays of $H_0(e^{j\omega})/\sqrt{(LL_0)}$ and $H_1(e^{j\omega})/\sqrt{(LL_1)}$,

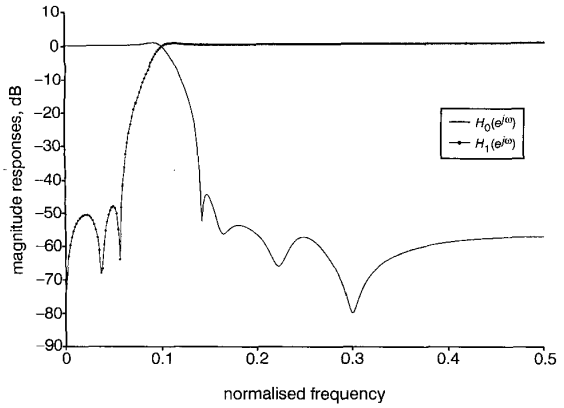

Fig. 4 *Magnitude responses of designed analysis filters*

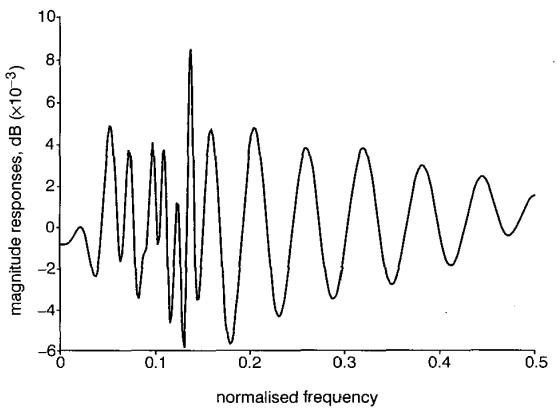

Fig. 6 *Magnitude response of $\hat{T}(e^{j\omega})$*

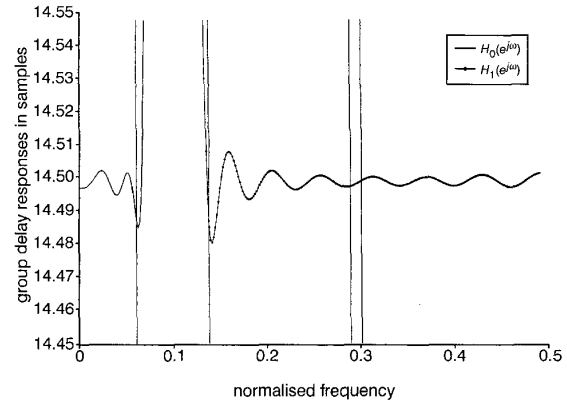

Fig. 5 *Group delay responses of analysis filters*

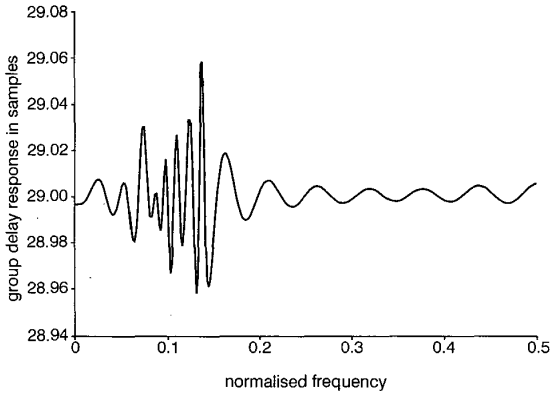

Fig. 7 *Group delay response of $\hat{T}(e^{j\omega})$*

respectively. Figs. 11 and 12 depict the magnitude response and the group delay response of $\hat{T}(e^{j\omega})$, respectively, for the IIR NDF bank designed by the proposed technique. Finally, the magnitude response of $e^{-jk_d\omega} - \hat{T}(e^{j\omega})$ is shown in Fig. 13. To show the convergence property of the proposed technique, Table 5 lists the value of the objective function (11) against the iteration number and Fig. 14 depicts the corresponding convergence plot. From the presented design results, we note again that the proposed technique provides very satisfactory performance.

Table 3: Significant design results for design example 2

	IIR NDF bank	FIR NDF bank	IIR NDF bank on L_2 [8]
Filter order	10/10, 11/11	32, 33	10/10, 11/11
No. of coefficients	44	67	44
PRE, dB	0.0141	0.0172	0.0148
MVGD	0.0555	0.0535	0.0583
$NPSR_0$, dB	32.02	31.07	32.21
$NPSR_1$, dB	32.03	29.88	32.07
$MVPGD_0$	0.0149	0.0113	0.0158
$MVPGD_1$	0.0226	0.0131	0.0230
SEE_0	3.35×10^{-2}	4.80×10^{-2}	3.33×10^{-2}
SEE_1	5.82×10^{-2}	7.52×10^{-2}	5.68×10^{-2}
Max. variation of filter bank response	2.22×10^{-3}	2.20×10^{-3}	2.27×10^{-3}
No. of iterations	25	51	37

Table 4: Tap and reflection coefficients for design example 2

i	a_{oi}	k_{oi}	a_{li}	k_{li}
0	0.00840280852248		−0.00687997871021	
1	−0.03972551284230	−0.52474169869307	−0.02743281043562	−0.10792777435909
2	0.07104206765306	0.77850267598451	−0.04953643573062	0.64929487935185
3	−0.05515331560924	−0.72643825913675	−0.03922455092752	0.47021734110726
4	0.02945729780658	0.67700633123663	0.02260921825689	0.40390034618481
5	−0.05825096332504	−0.63734769119002	0.08034363531087	0.35423085640786
6	0.02878009263438	0.58276091261142	0.01221363431628	0.31470903504061
7	0.08734916697783	−0.45656489202335	−0.22451992481268	0.26989578865561
8	0.16610765633258	0.24696936484061	−0.33011181755577	0.20440951536865
9	0.11958231494381	−0.06993178256951	1.37462171587862	0.12312487776655
10	0.08882380824677	0.00615175958300	−1.44444556334549	0.05197988875961
11			0.62349392295269	0.01195672802207

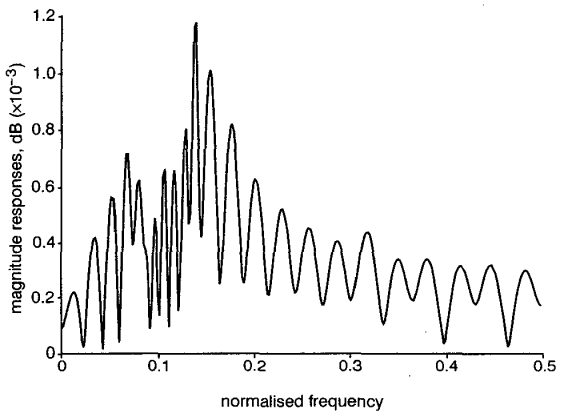

Fig. 8 *Magnitude response of $e^{-jk_d\omega} - \hat{T}(e^{j\omega})$*

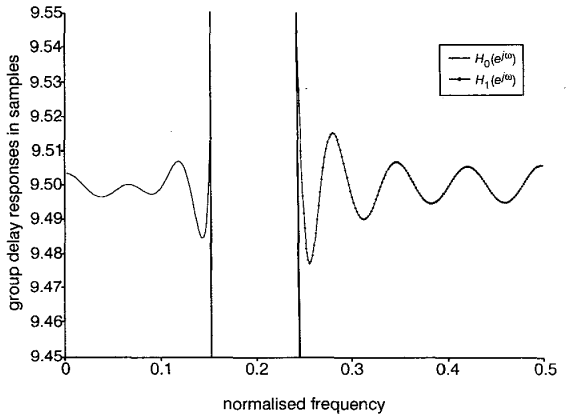

Fig. 10 *Group delay responses of designed analysis filters for example 2*

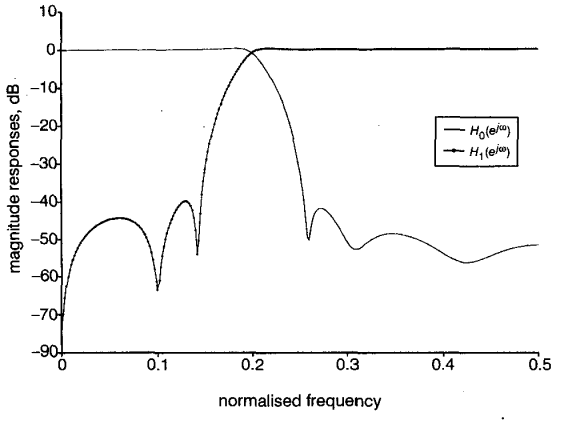

Fig. 9 *Magnitude responses of designed analysis filters for example 2*

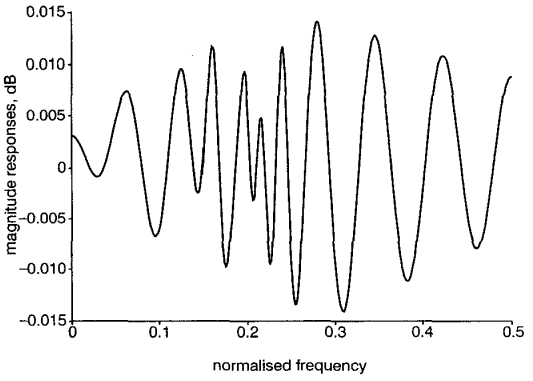

Fig. 11 *Magnitude response of $\hat{T}(e^{j\omega})$ for example 2*

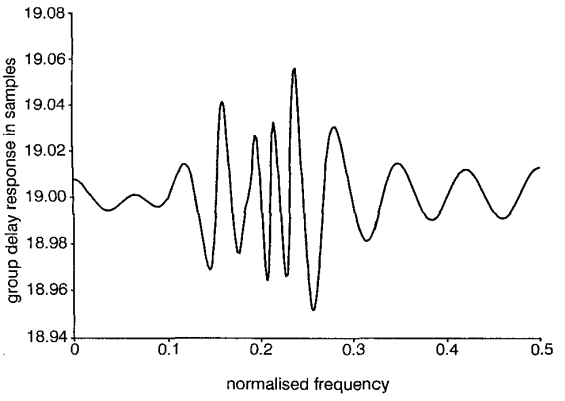

Fig. 12 *Group delay response of $\hat{T}(e^{j\omega})$ for example 2*

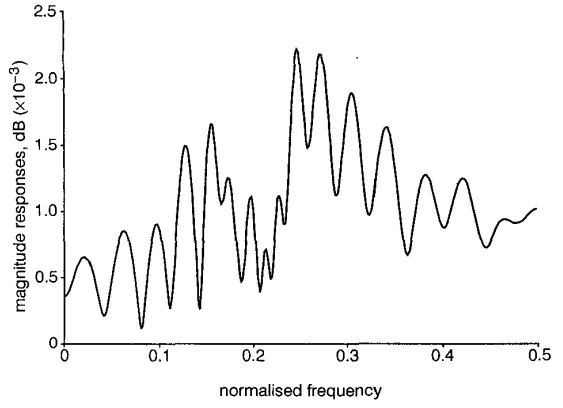

Fig. 13 *Magnitude response of $e^{-jk_d\omega} - \hat{T}(e^{j\omega})$ for example 2*

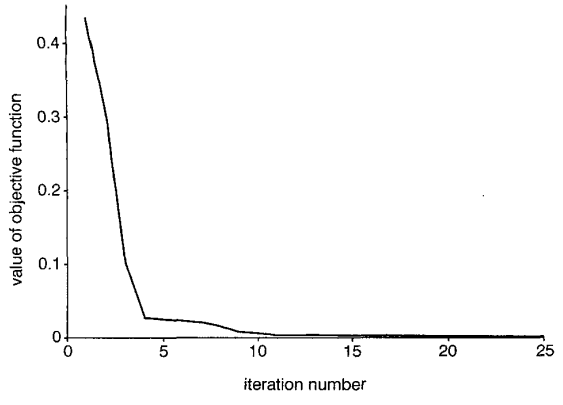

Fig. 14 *Convergence plot for example 2*

5 Conclusion

Employing recursive analysis filters with a lattice denominator for designing IIR NDF banks under the L_1 error criteria has been presented. An approximation scheme has been proposed to solve the resulting highly nonlinear programming problem. The stability of the designed recursive analysis filters is ensured by incorporating an efficient stabilisation procedure. Therefore, at each iteration, the proposed design technique adjusts the tap and reflection coefficients for the analysis filters to reduce the resulting design error and keep the designed IIR analysis filters stable. Simulation results have shown the effectiveness of the proposed technique.

Table 5: Convergence property for design example 2

Iteration number	Value of objective function (11)
0 (initial)	0.6567676534
1	0.4327604785
2	0.3066314681
3	0.1018665326
4	0.0265105776
5	0.0242037392
6	0.0218806623
7	0.0195353138
8	0.0143261386
9	0.0066852773
10	0.0046182609
11	0.0023880977
12	0.0020821327
13	0.0019039882
14	0.0018358464
15	0.0017263049
16	0.0016922422
17	0.0016855382
18	0.0016449093
19	0.0016177912
20	0.0016097093
21	0.0015664104
22	0.0015611703
23	0.0015534602
24	0.0015302981
25	0.0015302980

6 Acknowledgment

This work was supported by the National Science Council under grant NSC89-2213-E002-162.

7 References

1 NAYEBI, K., BARNWELL, T.P., and SMITH, M.J.T.: 'Nonuniform filter banks: a reconstruction and design theory', *IEEE Trans. Signal Process.*, 1993, **41**, pp. 1114–1127
2 WADA, S.: 'Design of nonuniform division multirate FIR filter banks', *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, 1995, **42**, pp. 115–121
3 LEE, J.-H., and HUANG, S.-C.: 'Design of two-channel nonuniform-division maximally decimated filter banks using L_1 criteria', *IEE Proc., Vis. Image Signal Process.*, 1996, **143**, pp. 79–83
4 LEE, J.-H., and TANG, D.-C.: 'Optimal design of two-channel nonuniform-division FIR filter banks with -1, 0, and $+1$ coefficients', *IEEE Trans. Signal Process.*, 1999, **47**, pp. 422–432
5 EKANAYAKE, M.M., and PREMARATNE, K.: 'Two-channel IIR QMF banks with approximately linear-phase analysis and synthesis filters', *IEEE Trans. Signal Process.*, 1995, **43**, pp. 2313–2322
6 TAY, D.B.H.: 'Design of causal stable IIR perfect reconstruction filter banks using transformation of variables', *IEE Proc., Vis. Image Signal Process.*, 1998, **145**, pp. 287–292
7 OKUDA, M., FUKUOKA, T., IKEHARA, M., and TAKAHASHI, S.-I.: 'The design of two-channel perfect reconstruction IIR filter banks with causality', *Electron. Commun. Jpn.*, 1998, **81**, pp. 22–31
8 LEE, J.-H., and NIU, I.-C.: 'Design of two-channel IIR nonuniform division filter banks with arbitrary group delay', *IEE Proc., Vis. Image Signal Process.*, 2000, **147**, (6), pp. 534–542

9 ADLER, I., KARMARKAR, N., RESENDE, M.G.C., and VEIGA, G.: 'An implementation of Karmarkar's algorithm for linear programming', *Math. Program.*, 1989, **44**, pp. 297–235
10 LIM, Y.C.: 'On the synthesis of IIR digital filters derived from AR lattice network', *IEEE Trans. Acoust. Speech Signal Process.*, 1984, **32**, pp. 741–749
11 HILLER, F.S., and LIEBERMAN, G.J.: 'Introduction to mathematical programming' (McGraw-Hill, New York, 1991)
12 BELICZYNSKI, B., KALE, I., and CAIN, G.D.: 'Approximation of FIR by IIR digital filters: an algorithm based on balanced model reduction', *IEEE Trans. Signal Process.*, 1992, **40**, pp. 532–542
13 OPPENHEIM, A.V., and SCHAFER, R.W.: 'Discrete-time signal processing' (Prentice-Hall, Englewood Cliffs, New Jersey, 1989)
14 NELDER, J.A., and MEADE, R.: 'A simplex method for function minimization', *Comput. J.*, 1965, **7**, pp. 308–313