Variable Fractional Delay FIR Filters with Sparse Coefficients

W.-S. Lu
Dept. of Electrical & Computer Engineering
University of Victoria
Victoria, Canada

T. Hinamoto
Graduate School of Engineering
Hiroshima University
Hiroshima, Japan
Outline

• Design Problem and Related Work
• Significance
• Design Method at a Glance
• A Design Example
• Concluding Remarks
1. Design Problem and Related Work

Design Problem

Find a VFD FIR digital filter of order \((N, K)\)

\[
H(z, p) = \sum_{n=0}^{N} a_{n}(p)z^{-n}, \quad a_{n}(p) = \sum_{k=0}^{K} a_{nk}p^{k}
\]

with \(0 \leq p \leq 1\) and sparse \(A = \{a_{ij}\} \in R^{(N+1) \times (K+1)}\) that optimally approximates a desired frequency response

\[
H_{d}(\omega, p) = e^{-j\omega(D+p)}
\]

in a weighted least-squares sense.

\[
J(A) = \frac{1}{2} \int_{0}^{\pi} \int_{0}^{1} W(\omega, p) \left| H(\omega, p) - H_{d}(\omega, p) \right|^{2} dpd\omega
\]

subject to: \(\text{sparsity}(A) = N_z\)
Related Work

- Linear programming algorithms for sparse FIR filters (T. Baran, D. Wei, and A.V. Oppenheim, March 2010)
- $l_1$-minimization algorithms for sparse 1-D and 2-D FIR filters (W.-S. Lu and T. Hinamoto, 2010, 2011)
2. Significance

• Implementing a VFD filter in Farrow model is costly as each coefficient of a VFD filter is a polynomial rather than a scalar. Hence VFD digital filters with sparse coefficients are of interest because the sparsity implies reduced implementation complexity (cost), hence real-time application potential.

• VFD filters are *not* sparse in general.
3. Design Method at a Glance

Define

\[ \omega = \begin{bmatrix} 1 & e^{-j\omega} & \cdots & e^{-jN\omega} \end{bmatrix}^T \] and \[ p = \begin{bmatrix} 1 & p & \cdots & p^K \end{bmatrix}^T \]

and assume a separable weighting function \( W(\omega, p) = W_1(\omega)W_2(p) \), then up to a constant one can write

\[
J(A) = \frac{1}{2} \int_0^\pi \int_0^1 W(\omega, p) \left| H(\omega, p) - H_d(\omega, p) \right|^2 dpd\omega \\
= \frac{1}{2} \text{tr}(PA^T\Omega A) - \text{tr}(SA)
\]

where

\[
P = \int_0^1 W_2(p)pp^T dp, \quad \Omega = \text{Re} \left[ \int_0^\pi W_1(\omega)\bar{\omega}\bar{\omega}^T d\omega \right]
\]

\[
S = \int_0^1 W_2(p)p\omega_p^T dp, \quad \omega_p^T = \text{Re} \left[ \int_0^\pi W_1(\omega)\omega^T e^{j\omega(D+p)} d\omega \right]
\]
To deal with large condition numbers of $P$ and $\Omega$, Cholesky decompositions $P = P_1^T P_1$, $\Omega = \Omega_1^T \Omega_1$ are used, and up to a constant $J(A)$ can be written as

$$J(a) = \frac{1}{2} \| \Gamma a - y \|_2^2, \quad \Gamma = P_1 \otimes \Omega_1, \quad y = \Gamma^{-T} s$$

where $a$ and $s$ are the vectors generated by concatenating the columns of $A$ and $S$, and $\otimes$ is the Kronecker product.

**Design Phase 1**

- To identify an index set of the most appropriate locations in $a$ to be set to zero in order to satisfy a target sparsity.
- This is done subject to maintaining closeness of $H(e^{i\omega}, p)$ to $H_d(\omega, p)$. 
• The target sparsity is achieved by introducing sparsity promoting $l_1$-norm of $a$ into an objective function:

$$\minimize_a \mu \|a\|_1 + \frac{1}{2} \|\Gamma a - y\|_2^2 \quad (1)$$

where $\|a\|_1 = \sum_{i=0}^{n} |a_i|$. Problem (1) is a convex, for which many fast algorithms are available. E.g. using FISTA, (1) can be solved with a small number of iterations.

**Input:** Data $\Gamma$ and $y$, parameter $\mu$ and iteration number $M$.

**Step 1.** Compute $A_0 = \Gamma^{-1} y$, $a_0 = A_0(\cdot)$. Set $b_1 = a_0$, $t_1 = 1$, and $m = 1$.

**Step 2.** Compute $a_m = S_{\mu/L}\{b_m - (1/L)\Gamma^T(\Gamma b_m - y)\}$,
where $S_{\alpha}(u) = \text{sgn}(u) \cdot \max\{|u| - \alpha, 0\}$

**Step 3.** Update $t_{m+1} = (1 + \sqrt{1 + 4t_m^2})/2$

**Step 4.** Update $b_{m+1} = a_m + ((t_m - 1)/t_{m+1})(a_m - a_{m-1})$.

**Step 5.** If $m < M$, set $m = m + 1$ and repeat from Step 2; otherwise stop and output $a_m$ as solution $\hat{a}$. 
• Hard thresholding is applied to vector $\hat{a}$ with an appropriate value of threshold $\varepsilon^*$ so that the length of the index set $I^* = \{i, |\hat{a}(i)| < \varepsilon^*\}$ equals to target sparsity $N_z$. The index set $I^*$ is a key ingredient of design phase 2.

Design Phase 2

• To find a coefficient matrix $A$ that minimizes the WLS error $J(A)$ subject to the sparsity constraint. This part of the design is carried out by solving the convex problem

$$\min_a J(a) = \frac{1}{2} \|\Gamma a - y\|_2^2$$

subject to: $a(i) = 0 \text{ for } i \in I^*$

• By simply substituting the constraints into the objective function, the above problem becomes an unconstrained least square problem whose solution is given by
\[ a^*(I^*) = \Gamma_s^{-1} y \quad \text{and} \quad a^*(I^*) = 0 \]

where \( I^* \) denotes the set of index not in \( I^* \) and \( \Gamma_s \) is composed of those columns of \( \Gamma \) with indices in set \( I^* \).

4. A Design Example

The Design Problem

- Design a VFD FIR filter of order \( N = 65 \) and \( K = 7 \) with cutoff \( \omega_c = 0.9\pi \).
- Design performance is evaluated in terms of maximum error
  \[ e_{\text{max}} = \max \{e(\omega, p), 0 \leq \omega \leq 0.9\pi, 0 \leq p \leq 1\} \]
  with
  \[ e(\omega, p) = 20\log_{10} |H(\omega, p) - H_d(\omega, p)| \]
  and L2-error
\[ e_2 = \left[ \int_0^{0.9\pi} \int_0^1 |H(\omega, p) - H_d(\omega, p)|^2 \, dp\, d\omega \right]^{1/2} \]

- The weighting function was set to \( W(\omega, p) = W_1(\omega)W_2(p) \) with \( W_2(p) = 1 \) for \( p \) in \([0, 1]\) and
  \[
  W_1(\omega) = \begin{cases} 
  1 & \text{for } \omega \in [0, 0.88\pi) \\
  3 & \text{for } \omega \in [0.88\pi, 0.8994\pi) \\
  0 & \text{for } \omega \in [0.88\pi, \pi]
  \end{cases}
  \]

- The target sparsity was set to \( N_z = 198 \) which means a 37.5% of coefficients were set to zero. To achieve this sparsity, the two key parameters in phase-1 were set to \( \mu = 10^{-5}, \epsilon^* = 10^{-3} \). It took 60 FISTA iterations for the algorithm in phase 1 to converge.
• Phase 2 of the design then produced an optimal $A^*$ with sparsity($A^*$) = 198. Below are the numerical evaluation results:
  • maximum error $e_{\text{max}} = 0.0021$.
  • L2-error $e_2 = -75.25$ dB.

===== Comparison 1 =====

• Compare with an equivalent nonsparse VFD filter of order ($N = 65, K = 4$) with the same specifications:
  • maximum error $e_{\text{max}} = 0.0609$.
  • L2-error $e_2 = -45.28$ dB.
• Profile of frequency response error

$$|e(\omega, p)| \text{ over } 0 \leq \omega \leq 0.9\pi, 0 \leq p \leq 1: \text{ the sparse VFD filter:}$$
• Profile of frequency response error

\[ |e(\omega, p)| \quad \text{over} \quad 0 \leq \omega \leq 0.9\pi, \; 0 \leq p \leq 1: \quad \text{the nonsparse VFD filter:} \]
Comparison 2

- To justify phase 1 of the design, we compare the above result with the following: We design a conversional (nonsparse) VFD filter of order \(N = 65, K = 7\), then applied hard thresholding to generate exactly 198 locations that may be considered appropriate to set to zero. We then went on to carry out phase 2 to yield a sparse VFD filter. It was found that
  - maximum error \(e_{\text{max}} = 0.0025\) (vs 0.0021 with phase 1)
  - L2-error \(e_2 = -73.41\) dB (vs \(-75.25\) dB with phase 1).
• Profile of frequency response error

\[ |e(\omega, p)| \text{ over } 0 \leq \omega \leq 0.9\pi, 0 \leq p \leq 1 : \text{ the sparse VFD filter without phase 1:} \]
Comparison 3

- Fractional delay over $0 \leq \omega \leq 0.9\pi$, $0 \leq p \leq 1$. The sparse filter:
• Fractional delay over $0 \leq \omega \leq 0.9\pi, 0 \leq p \leq 1$. The equivalent nonsparse filter:
5. Concluding Remarks

- A two-phase technique for the WLS design of VFD FIR filters subject to a target coefficient sparsity constraint has been proposed.
- The design algorithm is easy to implement and computationally efficient because it is based on $l_1 - l_2$ convex optimization.
- The performance of the filter appears to be satisfactory compared with its nonsparse counterpart.