Hardware Efficient Digital Filter Design by Multimode Mean Field Annealing

Per Persson, Sven Nordebo, and Ingvar Claesson

***Abstract*—The multimode mean field annealing (MM-MFA) approach is introduced as an efficient tool to design digital filters with discrete coefficients, each implemented as a sum of signed power-of-two terms and additional general hardware constraints. As an application example demonstrating the potential of the new method, we consider the design of a linear phase nonrecursive finite impulse response (FIR) filter with a least squares criterion and minimum number of power-of-two terms.**

***Index Terms*—Combinatorial optimization, digital filter design.**

I. Introduction

Digital filters with coefficients made up of a minimum number of signed power-of-two (SPT) terms are desirable but difficult to design [1]–[3]. From a combinatorial point of view, such a design can be considered a search for optimality over a state space consisting of 3^{BN} possible outcomes of BN trinary variables (with values 1, 0, or -1), where B is the coefficient wordlength, and N is the number of coefficients.

Simulated annealing [4], [5] and the mean field annealing (MFA) algorithm [6]–[8] are efficient tools to solve combinatorial problems. We will introduce a generalized form of MFA capable of handling trinary variables.

Multimode mean field annealing (MM-MFA) works by mapping the BN discrete state variables on to a new set of BN continuous state variables by applying a probabilistic interpretation. During the course of optimization, the solution to the continuous-valued problem is obtained from a deterministic expression. By utilizing the properties of a two-dimensional sigmoid-like function emerging naturally from the probabilistic interpretation, the continuous-valued state variables are gradually forced to become either 0, 1, or -1, while still representing a valid solution.

II. Problem Formulation

Consider the design of a type 1 linear phase nonrecursive finite impulse response (FIR) filter with impulse response $h(n)$, even symmetry $h(n) = h(N-1-n)$, and odd length N. The frequency response is given by [9]

$$H(\omega) = \sum_{n=0}^{N-1} h(n)e^{-j\omega n} = e^{-j\omega M} A(\omega) \quad (1)$$

$$A(\omega) = \sum_{m=0}^{M} a_m \phi_m(\omega) = \mathbf{a}^T \phi(\omega) \quad (2)$$

where

$A(\omega)$	real amplitude (with sign);
M	$= (N-1)/2$;
$\mathbf{a}$	$(M+1)\times 1$ vector of coefficients $a_m = h(M-m)$;
$\phi(\omega)$	vector of basis functions $\phi_m(\omega)$ where $\phi_0(\omega) = 1$;
$\phi_m(\omega)$	$= 2\cos(m\omega)$ for $m = 1, \ldots, M$.

The design error E_0 is defined by

$$E_0 = \sum_{k=1}^{K} w_k^2 (A(\omega_k) - A_d(\omega_k))^2 \quad (3)$$

where $A_d(\omega_k)$ is the desired response defined on a set of frequency points $\omega_k \in [0, \pi]$ for $k = 1, \ldots, K$, and w_k is a weighting factor. In the case of infinite precision coefficients a_m, the optimum solution, denoted E_{LS}, is known.

A hardware-efficient coding of the finite precision coefficients is the following sum of signed power-of-two terms

$$a_m = \sum_{b=1}^{B} s_{mB+b} 2^{-b} \quad (4)$$

where the trinary variables $s_i \in \{-1, 0, 1\}$ have been introduced, see for example [3]. Hence, in terms of the trinary variables the response $A(\omega)$ can be written

$$A(\omega) = \mathbf{s}^T \psi(\omega) \quad (5)$$

where s is an $L \times 1$, $L = (M+1)B$ vector of variables s_i, and $\psi(\omega) = \phi(\omega) \otimes [2^{-1} \ \cdots \ 2^{-B}]^T$. By employing (5), we may now rewrite (3) in terms of s

$$E_0(\mathbf{s}) = \mathbf{s}^T \mathbf{R} \mathbf{s} - 2\mathbf{p}^T \mathbf{s} + c \quad (6)$$

where

$$\mathbf{R} = \sum_{k=1}^{K} w_k^2 \psi(\omega_k) \psi^T(\omega_k) \quad (7)$$

$$\mathbf{p} = \sum_{k=1}^{K} w_k^2 \psi(\omega_k) A_d(\omega_k) \quad (8)$$

$$c = \sum_{k=1}^{K} w_k^2 A_d^2(\omega_k). \quad (9)$$

In order to minimize the number of SPT terms corresponding to $s_i \neq 0$, a penalty term proportional to $\mathbf{s}^T\mathbf{s}$ is added to (6), giving

$$E(\mathbf{s}) = \mathbf{s}^T \left(\mathbf{R} + \alpha \frac{E_{\mathrm{LS}}}{L} \cdot \mathbf{I} \right) \mathbf{s} - 2\mathbf{p}^T \mathbf{s} + c \quad (10)$$

where E_{LS}/L is a normalizing factor, while $\alpha > 0$ is a weighting factor determining the tradeoff between filter design error and total number of SPT terms.

Manuscript received October 19, 2000. This work was sponsored by the Swedish Knowledge Foundation. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. P. Milanfar.

The authors are with the Department of Telecommunications and Signal Processing, Blekinge Institute of Technology, S-372 25 Ronneby, Sweden.

Publisher Item Identifier S 1070-9908(01)05240-3.

1070–9908/01$10.00 © 2001 IEEE

We emphasize that the type 1 linear phase FIR filter structure [9] and the quadratic design criterion chosen above are used merely as examples to demonstrate the potential of the new algorithm. Other filter structures, design criteria, and hardware costs may be treated similarly.

III. Multimode Mean Field Annealing

Consider the set $\mathcal{S}$ of 3^L states defined by

$$\mathcal{S} = \{\mathbf{s} \mid s_i \in \{-1, 0, 1\},\ i = 1, \ldots, L\} \tag{11}$$

and the global cost $E(\mathbf{s})$ associated with each state according to (10). From statistical mechanics, it is well known [10] that for a system in thermal equilibrium at temperature T, the probability of a particular state s is given by the Boltzmann distribution

$$P(\mathbf{s}) = \frac{e^{-E(\mathbf{s})/T}}{\sum_{\mathcal{S}} e^{-E(\mathbf{s})/T}}. \tag{12}$$

Under the condition that the system is kept in thermal equilibrium, it is easy to show that (12) gives

$$\lim_{T \to 0} P(\mathbf{s}^{\text{opt}}) = 1 \tag{13}$$

where $\mathbf{s}^{\text{opt}}$ is the state with $E(\mathbf{s}^{\text{opt}}) < E(\mathbf{s} \neq \mathbf{s}^{\text{opt}})$ [11].

Furthermore, the expectation value of s_i denoted by $\langle s_i \rangle$ for a system in thermal equilibrium is [10]

$$\langle s_i \rangle_T = \frac{\sum_{\mathcal{S}} s_i e^{-E(\mathbf{s})/T}}{\sum_{\mathcal{S}} e^{-E(\mathbf{s})/T}}. \tag{14}$$

For most problems, (14) is impossible to compute, but a good approximation can be found by rewriting (14) using conditional expectations over $s_i = \{-1, 0, 1\}$ as

$$\langle s_i \rangle_T = \frac{\sum_{\mathcal{S}} e^{-E(\mathbf{s})/T} \dfrac{\sum_{s_i=-1,0,1} s_i e^{-\frac{1}{T}E(\mathbf{s})|_{s_i}}}{\sum_{s_i=-1,0,1} e^{-\frac{1}{T}E(\mathbf{s})|_{s_i}}}}{\sum_{\mathcal{S}} e^{-E(\mathbf{s})/T}}. \tag{15}$$

Noting that

$$\langle s_i \rangle_T \equiv \left\langle \frac{\sum_{s_i=-1,0,1} s_i e^{-\frac{1}{T}E(\mathbf{s})|_{s_i}}}{\sum_{s_i=-1,0,1} e^{-\frac{1}{T}E(\mathbf{s})|_{s_i}}} \right\rangle \tag{16}$$

the assumption is made that (16) can be approximated with

$$\langle s_i \rangle_T \approx \frac{\sum_{s_i=-1,0,1} s_i e^{-\frac{1}{T}E(\langle \mathbf{s} \rangle)|_{s_i}}}{\sum_{s_i=-1,0,1} e^{-\frac{1}{T}E(\langle \mathbf{s} \rangle)|_{s_i}}} \tag{17}$$

where the expectation operator $\langle \cdot \rangle$ has been moved inside the argument of $E(\cdot)$.

By introducing the set of continuous variables $\mathcal{V} = \{\mathbf{v} \mid v_i \in [-1, 1], i = 1, \ldots, L\}$ where $v_i = \langle s_i \rangle$, (17) can be formulated as a continuous-valued fixed point equation

$$v_i = \frac{\sum_{s_i=-1,0,1} s_i e^{-\frac{1}{T}E(\mathbf{v})|_{v_i=s_i}}}{\sum_{s_i=-1,0,1} e^{-\frac{1}{T}E(\mathbf{v})|_{v_i=s_i}}} \tag{18}$$

provided that $E(\mathbf{v})$ is defined for $\mathbf{v} \in \mathcal{V}$.

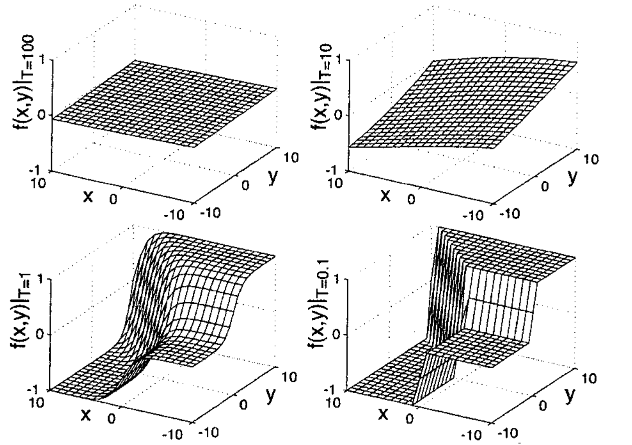

Fig. 1. Trigmoid function plotted for $T = 100$ (top left), $T = 10$ (top right), $T = 1$ (bottom left) and $T = 0.1$ (bottom right).

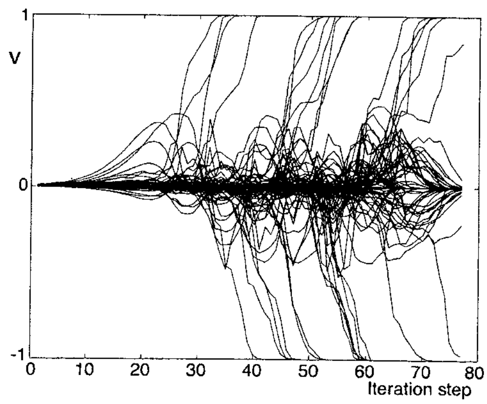

Fig. 2. The plot shows how the components of v evolve over the optimization course. In this case, $N = 31$ and $B = 10$.

To visualize the properties of (18), the following variable substitution is made

$$x \equiv E(\mathbf{v})|_{v_i=1} - E(\mathbf{v})|_{v_i=0} \tag{19}$$

$$y \equiv E(\mathbf{v})|_{v_i=-1} - E(\mathbf{v})|_{v_i=0} \tag{20}$$

introducing the two-dimensional (2-D) *trigmoid* function

$$f(x, y)|_T = \frac{e^{-\frac{x}{T}} - e^{-\frac{y}{T}}}{1 + e^{-\frac{x}{T}} + e^{-\frac{y}{T}}} \tag{21}$$

which is plotted for different values of T in Fig. 1.

For infinite T, the trivial fixed point is $v_i = 0$, but as $T \to 0, v_i \to s_i$ and (13) imply that $\mathbf{s} \to \mathbf{s}^{\text{opt}}$, provided that the approximation error in (17) is sufficiently small. In practice, the starting temperature should be chosen such that $v_i \approx 0$ in (18) regardless of $E(\mathbf{v})$, and the algorithm is terminated when the effects on $E(\mathbf{v})$ of rounding v_i to the nearest integer are negligible.

The evolution of variables v_i as T is decreased is shown in Fig. 2 for a run corresponding to the design example below with $N = 31$ and $B = 10$. At each iteration T is lowered by a factor of 0.8, starting from $T = 100$.

The MM-MFA algorithm can be described as follows.

1) Set initial $v_i = 0$, $i = 1 \ldots L$ and temperature T.
2) Calculate v_i, $i = 1 \ldots L$ from (18).

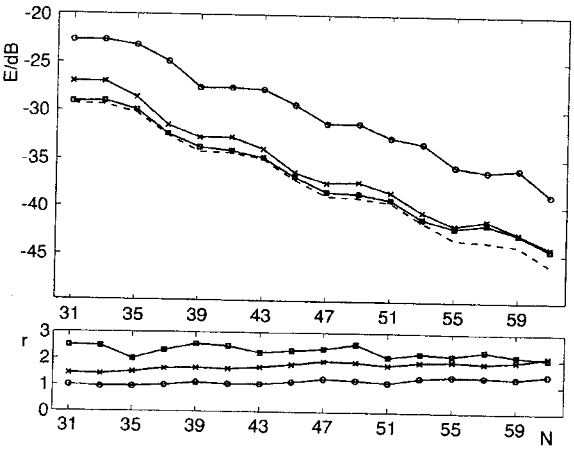

Fig. 3. Upper plot shows the infinite precision design error $E_{\rm LS}$ (dashed) and the design error E_0 computed for the MM-MFA solutions with $\alpha = 2$ (□), α =20 (×), and $\alpha = 200$ (o). The lower plot shows the average number of signed power-of-two terms per coefficient r.

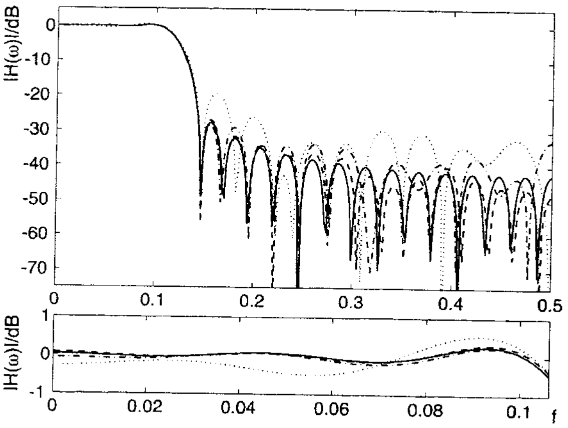

Fig. 4. Amplitude response of filters with $N = 37$ for the infinite precision case (solid) and for the MM-MFA case with $\alpha = 2$, (dashed), $\alpha = 20$, (dash-dotted) and $\alpha = 200$, (dotted). The upper plot shows the full frequency scale, and the lower plot shows the passband only.

3) Decrease the temperature geometrically.
4) Repeat steps 2–3 until $\frac{1}{L}\sum_i (v_i - \text{Round}(v_i))^2 < \varepsilon_{\text{stop}}$.

Since the number of iterations is independent of problem size, the computational complexity of the MM-MFA algorithm q, expressed as the number of times the cost function (10) is evaluated, is $q \sim L$.

IV. Design Examples

The MM-MFA algorithm was applied to the problem presented in [1], [2], a low-pass filter with passband [0 0.106 25], stopband [0.143 75 0.5], and weighting factor

$$w(\omega) = \begin{cases} 1, & 0 \le \omega < 0.103 \\ \sqrt{3}, & 0.103 \le \omega \le 0.106\,25 \\ \sqrt{8}, & 0.143\,75 \le \omega < 0.15 \\ 1, & 0.15 \le \omega \le 0.5 \end{cases} \qquad (22)$$

The filter length N (odd) ranged from 31 to 61, $B = 10$, and $\alpha = 2, 20$, and 200. The filters thus obtained were compared to the infinite precision LS-solution $E_{\rm LS}$ by computing $E_0 = E(\mathbf{s})|_{\alpha=0}$, and the results are plotted versus N in the upper plot of Fig. 3. Since α is a soft constraint, the exact number of SPT terms is not known beforehand, but the average number per coefficient $r = \mathbf{s}^T\mathbf{s}/(M+1)$ is plotted against N in the lower plot of Fig. 3.

In Fig. 4 the amplitude response of filters corresponding to the results in Fig. 3 for the case $N = 37$ are shown.

V. Summary and Conclusions

The multimode mean field annealing (MM-MFA) algorithm is introduced as a means of designing digital filters with hardware constraints. Examples are given for a linear phase FIR-filter, implemented as sums of signed power-of-two terms, with increasing demands on the hardware aspect as measured by the number of SPT terms. The tradeoff between filter characteristics and hardware constraints is imposed in a soft manner by parameter $\alpha \ge 0$.

The strength of the MM-MFA algorithm lies in the fact that it operates directly on the hardware level of the filter, which means that hardware constraints can be smoothly incorporated into the design. Even though the examples given are for linear phase FIR-filters, the algorithm is clearly not limited to such cases. It is our opinion that it can be turned into a useful tool to design general digital filters, with or without hardware constraints, for which no general method exists today.

References

[1] J.-J. Shyu and Y.-C. Lim, "A new approach to the design of discrete coefficient FIR digital filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 310–314, Jan. 1995.

[2] S. S. Rao and A. Ramasubrahmanyan, "Design of discrete coefficient FIR filters by simulated evolution," *IEEE Signal Processing Lett.*, vol. 3, pp. 137–140, May 1996.

[3] Y. C. Lim, R. Yang, D. Li, and J. Song, "Signed power-of-two term allocation scheme for the design of digital filters," *IEEE Trans. Circuits Syst. II*, vol. 46, pp. 577–584, May 1999.

[4] N. Benvenuto, M. Marchesi, and A. Uncini, "Applications of simulated annealing for the design of special digital filters," *IEEE Trans. Signal Processing*, vol. 40, pp. 323–332, Feb. 1992.

[5] J. Radecki, J. Konrad, and E. Dubois, "Design of multidimensional finite-wordlength FIR and IIR filters by simulated annealing," *IEEE Trans. Circuits Syst. II*, vol. 42, pp. 424–431, June 1995.

[6] J. J. Hopfield and D. W. Tank, "Neural computation of decisions in optimization problems," *Biol. Cybern.*, vol. 52, pp. 141–152, 1985.

[7] C. Peterson and B. Söderberg, "A new method for mapping optimization problems onto neural networks," *Int. J. Neural Syst.*, vol. 1, pp. 3–22, Jan. 1989.

[8] M. Ohlsson, C. Peterson, and B. Söderberg, "Neural networks for optimization problems with inequality constraints: The knapsack problem," *Neural Comput.*, vol. 5, pp. 331–339, Mar. 1993.

[9] T. W. Parks and C. S. Burrus, *Digital Filter Design*. New York: Wiley, 1987.

[10] G. Parisi, *Statistical Field Theory*. Reading, MA: Perseus, 1998.

[11] E. Aarts and J. Korst, *Simulated Annealing and Boltzmann Machines*. New York: Wiley, 1989.