Joint Design of Interpolation Filters and Decision Feedback Equalizers

Mu-Huo Cheng, *Member, IEEE,* and Tsai-Sheng Kao

***Abstract*—This paper presents an algorithm to design jointly interpolation filters and decision feedback equalizers in the sense of minimum mean-square error such that the joint capacity which is neglected in conventional design is explored to improve the receiver performance. The algorithm comprises an iteration of two alternating simple quadratic minimizing operations and ensures convergence. A simulation example for the raised-cosine channel demonstrates that via this approach an improvement over the conventional design can be achieved.**

***Index Terms*—Alternating coordinates minimization algorithm, decision feedback equalizers, interpolation filters.**

I. Introduction

In a digital baseband communication receiver, a timing recovery system is used to compensate for the timing offset between the transmitted data and the received sample while an equalizer serves to balance the channel effect for reducing the intersymbol interference (ISI). The timing recovery system is commonly realized by a timing offset estimator combined with either a voltage control oscillator (VCO) or an interpolation filter [1] and the commonly used equalizer is the decision-feedback equalizer (DFE) [2]. It is known that in the receiver, the timing recovery and the equalizer do not work independently of each other and the interaction has been studied [3]–[6]. The single-sideband AM digital communication system is studied in [3], [4] to jointly design an analog timing loop for carrier recovery and a finite impulse response (FIR) equalizer in the receiver. In [5], [6], a single adaptive fractionally-spaced FIR filter is used to realize the functions of both the timing recovery and the equalizer. In the present paper, we consider the digital baseband communication systems with a receiver containing a timing recovery system as well as a DFE and concentrate on the joint design of the interpolation filter and the DFE.

In convention, the interpolation filter and equalizer are designed separately: the interpolation filter is designed assuming the channel is known and fixed [7], [8] and the DFE is designed assuming the timing offset has been completely compensated [2]. The reason for designing each independently is mainly the simplicity because the joint design of both requires to solve a nonlinear optimization problem. The price, however, is that the joint capability is sacrificed. Investigating closely this problem, we observe that the complexity for solution of joint design does not seem so formidable. While the design of the interpolation filter and DFE independently requires only solving a quadratic minimization problem each, an algorithm for solving the joint design, presented in this paper, requires only an iteration of two quadratic minimizing operations. Therefore, the capacities of the interpolation filter and DFE can be further employed for improving the receiver performance. Specifically, we formulate together the interpolation filter and the DFE to minimize a mean-square error (MSE) and present an algorithm for solution. The algorithm comprises only an iteration of two simple quadratic minimizations and thus is simple to realize; it also ensures convergence and the convergence solution, by choosing a proper initial estimate, guarantees better than those obtained from conventional designs. A simulation example for the raised-cosine channel is performed to illustrate the design and the performance improvement.

Paper approved by C.-L. Wang, Editor for Wireless Spread Spectrum of the IEEE Communications Society. Manuscript received January 6, 2004; revised October 6, 2004. This work was supported in part by the National Science Council, Taiwan, R.O.C., under Grant NSC92-2213-E-009-084.

The authors are with the Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan, R.O.C. (e-mail: mhcheng@mail.nctu.edu.tw; u8812817@cc.nctu.edu.tw).

Digital Object Identifier 10.1109/TCOMM.2005.849689

II. Problem Formulation

The received signal of a digital baseband communication system can be expressed as

$$x(t) = \sum_{k=-\infty}^{\infty} d_k h(t - kT) + n(t) \tag{1}$$

where d_k is the transmitted data symbol with period T, $h(t)$ is the cascaded impulse response of the transmission filter, the channel, and the receiver filter, and $n(t)$ is an additive noise which may be white or colored depending on applications. Assume baud-rate sampling with a normalized sampling timing offset represented by μ, the received sample is given by

$$x_k(\mu) = \sum_{i=-\infty}^{\infty} d_{k-i} h((i-\mu)T) + n_k = \sum_{i=-\infty}^{\infty} d_{k-i} h_i(\mu) + n_k \tag{2}$$

where $x_k(\mu) = x((k-\mu)T)$, $h_i(\mu) = h((i-\mu)T)$, and n_k is the noise sample. We also assume that the timing offset μ is uniformly distributed within the range $[-0.5, 0.5]$, as is commonly done. Note that the baud-rate sampling is assumed here for simplicity; the interpolation filter with a higher sampling rate can be similarly formulated but requires further mechanism for down-sampling processing.

Fig. 1 depicts an equivalent discrete-time model of a digital baseband communication receiver; the receiver consists of a timing recovery system, a DFE, and a detector. The timing recovery system includes a timing offset estimator and an interpolation filter. Like conventional designs, the timing offset estimator is assumed to obtain correctly the timing offset μ and the detector obtains correct decision, i.e., $\hat{d}_k = d_k$. The purpose of this paper is to design the interpolation filter and DFE such that the mean square of the error between the transmitted data and the DFE output is minimized.

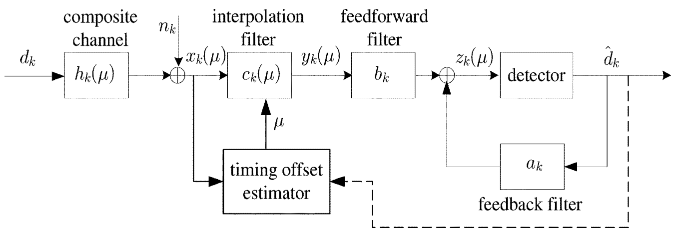

Fig. 1. Equivalent discrete-time model of a digital baseband communication receiver.

A. MSE Criterion

As usual, an FIR interpolation filter with coefficients $c_k(\mu)$ is used to compensate for the timing offset [9], yielding its output sample $y_k(\mu)$

$$y_k(\mu) = \sum_{i=-L_1}^{L_2} c_i(\mu) x_{k-i}(\mu) \tag{3}$$

where integers L_1 and L_2 indicate the lengths of noncausal and causal parts of the interpolation filter. Note that to make the interpolation realizable we need an extra delay $D \geq L_1$ which may be arisen either from the physical channel delay or from the artificially included delay memory on the received data. Each coefficient is usually characterized by a polynomial of degree M in μ

$$c_k(\mu) = \sum_{m=0}^{M} f_{k,m} \mu^m. \tag{4}$$

Farrow [10] has proposed an efficient structure to realize such an interpolation filter and thus $f_{k,m}$'s are also called the Farrow coefficients [8]. The DFE including a feedforward filter of order K_1 and a decision feedback filter of order K_2 is then used to combat the ISI, yielding the output $z_k(\mu)$,

$$z_k(\mu) = \sum_{i=0}^{K_1} b_i y_{k-i}(\mu) + \sum_{i=1}^{K_2} a_i d_{k-i}. \tag{5}$$

Note that the assumption of correct decisions has been used to replace $\hat{d}_{k-i}$ by d_{k-i}. The MSE criterion J, therefore, is given by

$$J = \mathrm{E}[d_k - z_k(\mu)]^2 \tag{6}$$

where the expectation operation $\mathrm{E}[\cdot]$ is taken with respect to the randomness of the input data, the noise sample and the timing offset μ.

It is more convenient to express the MSE in the frequency domain. Let $H(\omega, \mu)$, $C(\omega, \mu)$, $B(\omega)$, and $A(\omega)$ denote, respectively, the frequency responses of the composite channel, the interpolation filter, the feedforward filter, and the decision feedback filter; that is $H(\omega, \mu) = \sum_{n=-\infty}^{\infty} h_n(\mu) e^{-jn\omega}$, $C(\omega, \mu) = \sum_{n=-L_1}^{L_2} c_n(\mu) e^{-jn\omega}$, $B(\omega) = \sum_{n=0}^{K_1} b_n e^{-jn\omega}$, and $A(\omega) = \sum_{n=1}^{K_2} a_n e^{-jn\omega}$. Then, via the Parseval's theorem [2], the MSE in frequency domain can be derived

$$\begin{aligned} J = \frac{1}{2\pi} \int_{-0.5}^{0.5} \int_{-\pi}^{\pi} [D(\omega)|1 - H(\omega, \mu) C(\omega, \mu) B(\omega) \\ - A(\omega)|^2 + N(\omega) |C(\omega, \mu) B(\omega)|^2]\, d\omega d\mu \end{aligned} \tag{7}$$

$$= \int_{-0.5}^{0.5} J_\mu d\mu \tag{8}$$

where $D(\omega)$ is the power spectrum density (PSD) of d_k and $N(\omega)$ is the PSD of n_k. Note that J_μ in (8) is the MSE of a given fixed μ, which will be used later to illustrate performance difference between various designs.

The frequency response of the interpolation filter can be represented in a more compact form using (4) [11]

$$C(\omega, \mu) = \boldsymbol{f}^T (\boldsymbol{\mu} \otimes \boldsymbol{\omega}_c) \tag{9}$$

where $\boldsymbol{f} = [f_{-L_1,0}, \ldots, f_{L_2,0}, \ldots, f_{-L_1,M}, \ldots, f_{L_2,M}]^T$, $\boldsymbol{\mu} = [1, \mu, \ \ldots, \mu^M]^T$, $\boldsymbol{\omega}_c = [e^{j\omega L_1}, \ \ldots, \ 1, \ldots, \ e^{-j\omega L_2}]^T$, the superscript T denotes the transpose operation and the notation $\otimes$ represents the right Kronecker product [12]. Similarly, the frequency responses of the feedforward and decision feedback filters can be represented in a vector form

$$B(\omega) = \boldsymbol{b}^T \boldsymbol{\omega}_b, \quad A(\omega) = \boldsymbol{a}^T \boldsymbol{\omega}_a \tag{10}$$

where $\boldsymbol{b} = [b_0, b_1, \ldots, b_{K_1}]^T$, $\boldsymbol{a} = [a_1, a_2, \ldots, a_{K_2}]^T$, $\boldsymbol{\omega}_b = [1, e^{-j\omega}, \ldots, e^{-j\omega K_1}]^T$, and $\boldsymbol{\omega}_a = [e^{-j\omega}, \ e^{-j\omega 2} \ldots, e^{-j\omega K_2}]^T$.

Substituting (9) and (10) into (7), we obtain the MSE J as a nonlinear function of the interpolation filter coefficients $\boldsymbol{f}$ and the DFE parameters $\boldsymbol{\theta} = [\boldsymbol{b}^T, \boldsymbol{a}^T]^T$. The nonlinear optimization approaches [13] can be used for solution but are complicated. In this paper, the alternating coordinates minimization (ACM) [14] algorithm is applied for solution such that simple realization is obtained. Before discussing the detail of the algorithm, note that since the interpolation filter and the feedforward filter are cascaded, a constant factor redundancy thus exists between $\boldsymbol{f}$ and $\boldsymbol{b}$. Hence an extra constraint $f_{0,0} = 1$ is imposed to remove this redundancy. The optimization problem, therefore, is given by

$$\mathrm{Min}_{\boldsymbol{f}, \boldsymbol{\theta}} \, J \quad \text{subject to } f_{0,0} = 1. \tag{11}$$

III. ACM Algorithm for Optimal Joint Design

The ACM algorithm for solving this optimization problem involves iterations of two alternating optimizing operations; in the pth iteration, the first operation solves $\boldsymbol{\theta}^{(p)}$ of (11) given $\boldsymbol{f} = \boldsymbol{f}^{(p-1)}$, and then the second operation solves $\boldsymbol{f}^{(p)}$ of (11) given $\boldsymbol{\theta} = \boldsymbol{\theta}^{(p)}$ which is obtained from the first operation. The iteration continues until the convergence of $\boldsymbol{f}^{(p)}$ and $\boldsymbol{\theta}^{(p)}$. Each optimizing operation, shown below, only requires solving a simple quadratic optimization and thus its solution is unique. Also the two operations solve the coefficients $\boldsymbol{\theta}^{(p)}$, $\boldsymbol{f}^{(p)}$ alternatingly, the obtained MSE J is therefore guaranteed nonincreasing in every iteration. Since the MSE J is nonnegative and thus bounded from below, the ACM algorithm always converges. The derivations of two optimizing operations are described in the following subsections.

A. First Optimizing Operation: Solve $\boldsymbol{\theta}^{(p)}$ of (11) Given $\boldsymbol{f} = \boldsymbol{f}^{(p-1)}$

Since $\boldsymbol{f}$ is given and fixed, the constraint is naturally satisfied and $C(\omega,\mu)$, for a given ω and μ, is a fixed scalar; the MSE, after substituting (10) into (7), turns into a quadratic function of $\boldsymbol{\theta}$, shown in (12)–(13) at the bottom of the page, where $\boldsymbol{\omega}_\theta = [\boldsymbol{\omega}_b^T, \boldsymbol{\omega}_a^T]^T$

$$\boldsymbol{R} = \begin{bmatrix} H(\omega,\mu)C(\omega,\mu)\boldsymbol{I}_{K_1+1} & \mathbf{0}_{(K_1+1)\times K_2} \\ \mathbf{0}_{K_2\times(K_1+1)} & \boldsymbol{I}_{K_2} \end{bmatrix}$$

$$\boldsymbol{Q} = \begin{bmatrix} C(\omega,\mu)\boldsymbol{I}_{K_1+1} & \mathbf{0}_{(K_1+1)\times K_2} \\ \mathbf{0}_{K_2\times(K_1+1)} & \mathbf{0}_{K_2\times K_2} \end{bmatrix}$$

with $\boldsymbol{I}_m$ representing the identity matrix of dimension m and $\mathbf{0}_{m\times n}$ the $m \times n$ zero matrix. The solution $\boldsymbol{\theta}^{(p)}$ can be obtained by setting the gradient vector of J in (13) with respect to $\boldsymbol{\theta}$ to zero and rearranging, yielding

$$\boldsymbol{\theta}^{(p)} = \boldsymbol{\Omega}_f^{-1}\boldsymbol{v}_f \tag{14}$$

where

$$\boldsymbol{\Omega}_f = \frac{1}{2\pi}\int_{-0.5}^{0.5}\int_{-\pi}^{\pi}\mathrm{Re}[D(\omega)\boldsymbol{R}\boldsymbol{w}_\theta\boldsymbol{w}_\theta^H\boldsymbol{R}^H + N(\omega)\boldsymbol{Q}\boldsymbol{w}_\theta\boldsymbol{w}_\theta^H\boldsymbol{Q}^H]d\omega d\mu \tag{15}$$

and

$$\boldsymbol{v}_f = \frac{1}{2\pi}\int_{-0.5}^{0.5}\int_{-\pi}^{\pi} D(\omega)\mathrm{Re}[\boldsymbol{R}\boldsymbol{\omega}_\theta]d\omega d\mu \tag{16}$$

with $\mathrm{Re}[\cdot]$ representing the real part of a variable. The subscript f in $\boldsymbol{\Omega}_f$, $\boldsymbol{v}_f$ indicates that they are evaluated given a fixed interpolation filter $\boldsymbol{f}$. Note that the matrix $\boldsymbol{\Omega}_f$ is symmetric and some of its submatrices have a Toeplitz form; these properties can be used to simplify the matrix evaluation and are not elaborated further for brevity.

B. Second Optimizing Operation: Solve $\boldsymbol{f}^{(p)}$ of (11) Given $\boldsymbol{\theta} = \boldsymbol{\theta}^{(p)}$

Note that the given $\boldsymbol{\theta}^{(p)}$ is obtained from the previous optimizing operation. Since $\boldsymbol{\theta}$ is known, $A(\omega)$ and $B(\omega)$ can be evaluated and hence the optimization problem (11) is turned into a simple constraint quadratic optimization problem

$$\mathrm{Min}_{\boldsymbol{f}}\ J \quad \text{subject to } f_{0,0} = 1 \tag{17}$$

where (18), shown at the bottom of the page. Express the constraint as $\boldsymbol{f}^T\boldsymbol{i}_c = f_{0,0} = 1$ where $\boldsymbol{i}_c$ is a vector whose (L_1+1)th component is unity and whose other components are zero. Then, the solution $\boldsymbol{f}^{(p)}$ can be derived using the Lagrange multiplier technique [13], yielding

$$\boldsymbol{f}^{(p)} = \boldsymbol{\Omega}_\theta^{-1}\left(\boldsymbol{v}_\theta + \frac{1-\boldsymbol{i}_c^T\boldsymbol{\Omega}_\theta^{-1}\boldsymbol{v}_\theta}{\boldsymbol{i}_c^T\boldsymbol{\Omega}_\theta^{-1}\boldsymbol{i}_c}\boldsymbol{i}_c\right) \tag{19}$$

where (20)–(21), shown at the bottom of the page and

$$\boldsymbol{v}_\theta = \frac{1}{2\pi}\int_{-0.5}^{0.5}\int_{-\pi}^{\pi} D(\omega)\mathrm{Re}\left[H(\omega,\mu)B(\omega)(1-A^*(\omega)) \times(\boldsymbol{\mu}\otimes\boldsymbol{\omega}_c)\right]d\omega d\mu \tag{22}$$

with the superscript $*$ standing for the complex conjugate operation.

The algorithm starts with an initial guess $\boldsymbol{f}^{(0)}$ and iteratively performs the above two optimizations until convergence. Numerically, the algorithm terminates when the ratio of MSE improvement over MSE in previous iteration, $|J^{(l)} - J^{(l-1)}|/J^{(l-1)}$, is less than a predetermined small value ϵ.

Note that even the ACM algorithm ensures convergence, like most nonlinear optimization algorithms, it is only a suboptimal algorithm as it may converge to a local minimum. Therefore, a sensible initial estimate may be required. One good initial estimate is to take the $\boldsymbol{f}$ obtained from the conventional approach and normalizes it to obtain $f_{0,0} = 1$. The convergence solution using this initial estimate, because of the nonincreasing MSE of the algorithm, is ensured to result in a lower MSE than that by the conventional design. Another good initial estimate

$$J = \frac{1}{2\pi}\int_{-0.5}^{0.5}\int_{-\pi}^{\pi}\left[D(\omega)|1 - H(\omega,\mu)C(\omega,\mu)\boldsymbol{b}^T\boldsymbol{\omega}_b - \boldsymbol{a}^T\boldsymbol{\omega}_a|^2 + N(\omega)|C(\omega,\mu)\boldsymbol{b}^T\boldsymbol{\omega}_b|^2\right]d\omega d\mu \tag{12}$$

$$= \frac{1}{2\pi}\int_{-0.5}^{0.5}\int_{-\pi}^{\pi}\left[D(\omega)|1 - \boldsymbol{\theta}^T\boldsymbol{R}\boldsymbol{\omega}_\theta|^2 + N(\omega)|\boldsymbol{\theta}^T\boldsymbol{Q}\boldsymbol{\omega}_\theta|^2\right]d\omega d\mu \tag{13}$$

$$J = \frac{1}{2\pi}\int_{-0.5}^{0.5}\int_{-\pi}^{\pi}[D(\omega)|1 - H(\omega,\mu)B(\omega)\boldsymbol{f}^T(\boldsymbol{\mu}\otimes\boldsymbol{\omega}_c) - A(\omega)|^2 + N(\omega)|B(\omega)\boldsymbol{f}^T(\boldsymbol{\mu}\otimes\boldsymbol{\omega}_c)|^2]d\omega d\mu. \tag{18}$$

$$\boldsymbol{\Omega}_\theta = \frac{1}{2\pi}\int_{-0.5}^{0.5}\int_{-\pi}^{\pi}|B(\omega)|^2[D(\omega)|H(\omega,\mu)|^2 + N(\omega)]\mathrm{Re}[(\boldsymbol{\mu}\otimes\boldsymbol{\omega}_c)(\boldsymbol{\mu}\otimes\boldsymbol{\omega}_c)^H]d\omega d\mu \tag{20}$$

$$= \frac{1}{2\pi}\int_{-0.5}^{0.5}\int_{-\pi}^{\pi}|B(\omega)|^2[D(\omega)|H(\omega,\mu)|^2 + N(\omega)][(\boldsymbol{\mu}\boldsymbol{\mu}^H)\otimes\mathrm{Re}(\boldsymbol{\omega}_c\boldsymbol{\omega}_c^H)]d\omega d\mu \tag{21}$$

is $\boldsymbol{f}^{(0)} = \boldsymbol{i}_c$, i.e., $f_{00}^{(0)} = 1$ and all other components are zero. The interpolation filter corresponding to this initial estimate is just a pure unity gain filter, hence the first operation will obtain a DFE without the intervention of interpolation filter.

IV. Demonstration Example

One design example for standard raised-cosine channel is given to illustrate the advantage gained through the joint design approach. The channel impulse response with the symbol rate normalized as $T = 1$ is known to be $h(t) = \sin(\pi t)\cos(\beta\pi t)/(\pi t(1 - 4\beta^2 t^2))$ with $\beta \in [0, 1]$ as the roll-off factor. Since the channel has been ideally equalized, no equalizer is needed. For illustration, however, we assume that a first-order DFE ($K_1 = 0$, $K_2 = 1$) is used. Since the interpolation filter commonly operates at a high data rate and the delay time μ, in practice, is varying with time, as discussed in [1], the filter order is normally short and the degree of polynomial to characterize the coefficients is also low. Hence, we choose six taps ($L_1 = 2$, $L_2 = 3$) interpolation FIR filter with each coefficient characterized by a polynomial of degree 3 ($M = 3$). Assume the input data are white such that its PSD $D(\omega) = 1$ for all ω. Generally, the noise is colored because of the receiving filter, but for simplicity, it is also assumed white. The raised-cosine channels of $\beta = 0.2$ with the output signal-to-noise ratios (SNRs) set to 15, 20, 25, and 30 dB, respectively, are used in simulations. The conventional approach first designs the DFE for minimizing the MSE assuming exact sampling time and then designs the interpolation filter for minimizing J in (7). The joint approach normalizes the interpolation filter obtained via the conventional approach and uses it as the initial data, then the iteration terminates when the ratio of MSE improvement is less than $\epsilon = 10^{-5}$.

For example, for the output SNR of 20 dB, the DFE via the conventional approach yields $b_0 = 0.9906$, $a_1 = 0$, the interpolation filter is then designed, yielding the minimum J of -13.51 dB. The joint design, in this case, obtains the MSEs at each iteration which is shown in Fig. 2; the algorithm takes 120 iterations to converge and the convergence MSE equals -17.63 dB. Therefore, the performance gain of 4.12 dB is achieved. Note that the obtained MSEs with respect to iteration, as expected, are nonincreasing. A faster convergence speed, of course, can be obtained if a larger ϵ is given. We have also tested the simulated annealing algorithm for solution and observed that its convergence solution is the same as that obtained via the ACM algorithm. Hence, the obtained design in this example is optimal. To further illustrate the difference between the conventional approach and the joint method, Fig. 3 depicts the J_μ defined in (8) of both methods for μ increasing from -0.5 to 0.5 with the step size of 0.1 for various SNRs. The conventional approach obtains good performance only when the timing offset is small; the joint design, however, achieves lower and more uniform J_μ, resulting in a smaller MSE J. The MSEs (J) obtained from conventional and joint design methods under different output SNRs are listed in Table I. Note that the improvement, as shown from the table, increases as the SNR is increasing. When the SNR equals 30 dB, the improvement in MSE attains 10.7 dB; the improvement, however, is only about 1.28 dB for SNR of 15 dB. These results explain that because the compensation of the timing offset does not reduce the effect of noise, the joint design has less room for improvement when the noise power is larger. Hence, the joint design obtains better improvement for higher SNR of the received signal. This simulation, therefore, demonstrates that the joint design may significantly improve the MSE performance over the conventional approach.

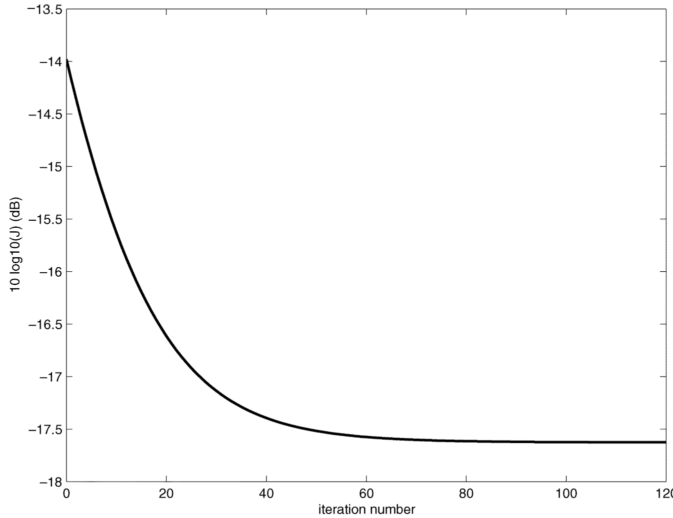

Fig. 2. Obtained MSEs J at each iteration for SNR of 20 dB.

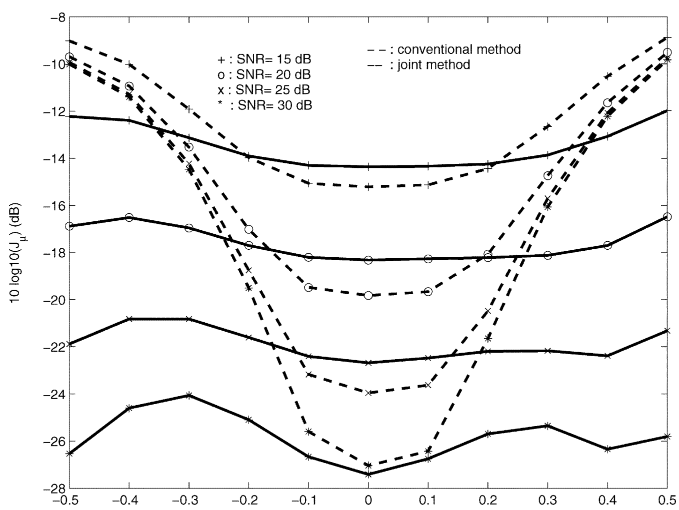

Fig. 3. J_μ versus the timing offset μ for the joint method and conventional method in different output SNRs.

TABLE I

MSEs of Conventional and Joint Methods for Raised-Cosine Channel With Various Output SNRs

SNR(dB)	15	20	25	30
conventional method, MSE(dB)	-12.23	-13.98	-14.74	-15.02
joint method, MSE(dB)	-13.51	-17.63	-21.88	-25.72
improvement (dB)	1.28	3.65	7.14	10.7

V. SUMMARY

In this paper, we present an algorithm to design both the interpolation filter and the DEF such that the joint capability is explored to improve the performance of a communication receiver. The algorithm is simple to realize and ensures convergence; the convergence solution, for a proper initial estimate, guarantees better than that obtained from the conventional design. This approach exploits the joint capacity which is neglected in the conventional design and achieves the performance improvement without increasing the complexity of either the interpolation filter or the DFE.

REFERENCES

[1] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers, Synchronization: Channel Estimation, and Signal Processing*. New York: Wiley, 1998.

[2] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[3] R. W. Chang, "Joint optimization of automatic equalization and carrier acquisition for digital communcation," *Bell Syst. Tech. J.*, vol. 49, pp. 1069–1104, Jul.-Aug. 1970.

[4] R. W. Chang and R. Srinivasagopalan, "Carrier recovery for data communication systems with adaptive equalization," *IEEE Trans. Commun.*, vol. 28, no. 8, pp. 1142–1153, Aug. 1980.

[5] B. R. Ghosh and D. G. Messerschmitt, "Asynchronous adaptive equalization in voiceband data modems," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Mar. 1992, pp. 485–488.

[6] D. J. Artman, S. Chari, and R. P. Gooch, "Joint equalization and timing recovery in a fractionally-spaced equalizer," in *Proc. IEEE 26th Asilomar Conf. Signals, Systems and Computers*, Oct. 1992, pp. 25–29.

[7] D. Kim, M. J. Narasimha, and D. C. Cox, "Design of optimal interpolation filter for symbol timing recovery," *IEEE Trans. Commun.*, vol. 45, no. 7, pp. 877–884, Jul. 1997.

[8] G. Watkins, "Optimal farrow coefficients for symbol timing recovery," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 381–383, Sep. 2001.

[9] L. Erup, F. M. Gardner, and R. A. Harris, "Interpolation in digital modems- Part II: Implementation and performance," *IEEE Trans. Commun.*, vol. 41, no. 6, pp. 998–1008, Jun. 1993.

[10] C. W. Farrow, "A continuously variable digital delay element," in *Proc. IEEE Int. Symp. Circuits and Systems*, Jun. 1988, pp. 2641–2645.

[11] T. B. Deng, "Discretization-free design of variable fractional-delay FIR digital filters," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 48, no. 6, pp. 637–644, Jun. 2001.

[12] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed: Academic Press, 1985.

[13] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. New York: Wiley, 1987.

[14] I. Csiszár and G. Tusnády, "Information geometry and alternating minimization procedures," *Statist. Decisions*, pp. 205–237, 1984.