

Adaptive MMSE Equalizer with Optimum Tap-length and Decision Delay

Yu Gong, Xia Hong and Khalid F. Abu-Salim

School of Systems Engineering
The University of Reading, Reading
RG6 6AY, UK

E-mail: {y.gong, x.hong, k.f.abusalem}@reading.ac.uk

Tel: +44 (0) 118 378 8581, Fax: +44 (0) 118 378 8583

Abstract—In a classic adaptive MMSE equalizer, the structure (or the number of the free coefficients) and the decision delay are often fixed either to some compromise values or simply from experience, making it hard to achieve the best potential performance. In this paper, we propose a novel method to jointly adapt the tap-length and decision delay so that the best performance can be achieved with minimum tap-length and decision delay. The proposed approach needs little extra complexity. It describes an innovative way to implement the adaptive equalizer where not only the structure is optimized but also the decision delay is self tuned.

I. INTRODUCTION

The adaptive minimum mean square error (MMSE) equalizer as a classic approach has been widely used in communications. In a traditional design, the structure (or the tap-length) and decision delay of the equalizer are normally fixed, making it sub-optimum in most cases. While the tap-length refers to the number of free tap coefficients, the decision delay determines which symbol is detected at the current time.

Both the tap-length and decision delay can significantly affect the performance of the equalizer. On the one hand, if the decision delay is fixed, the MMSE of the equalizer is a monotonic decreasing function of the tap-length, but the MMSE improvement due to length increase always becomes insignificant when the tap-length is large enough. There exists an optimum tap-length that best balances the performance and complexity. Length adaptation algorithms can be used to find the optimum tap-length of the equalizer with the decision delay being fixed. Among various length adaptation algorithms (e.g. [1]–[4]), of particular interest is the fractional tap-length (FT) algorithm due to its robustness and simplicity [4], in which the length adaptation is based on a pseudo fractional tap-length whose integer part gives the true tap-length. The performance of the FT algorithm can be further improved with convex combining of two FT algorithms [5].

On the other hand, for a given tap-length, there exists an optimum decision delay Δ that minimizes the MMSE. The optimum Δ depends on specific channels. As will be shown later, Δ should be set as 0 and $N - 1$ for minimum and maximum phase channels respectively, where N is the tap-length. In general, even with accurate channel state information, it is still not straightforward to obtain the optimum Δ which normally lies between 0 zero and $N - 1$. This can be even more complicated if the tap-length N is also to be determined

since the optimum Δ is different for different N . In general, the best MMSE is achieved when $N \rightarrow \infty$.

In classic applications, the decision delay is often chosen based on either the pre-assumption of the channel or simply from experience (e.g. [6, Section 9.7]). This normally requires the tap-length to be long enough such that there exist a large number of “appropriate” Δ which have the similar MMSE to that for the best Δ . Then the chance that the selected Δ happens to be one of those “appropriate” Δ can be high. Obviously such approaches have no guarantee to have the optimum Δ . Moreover, too large the tap-length contradicts the requirement that the tap-length should remain as small as possible without sacrificing the performance.

Delay selection is an important issue in various equalization approaches including blind equalization [7], decision feedback equalization (DFE) [8] and multichannel equalization [9]–[11]. The first decision delay adaptation algorithm for a fixed length linear equalizer was proposed in [12]. This algorithm may lose convergence in some cases because it is based on tracking the mean squared error (MSE) for different decision delays which are not always reliable. A joint length and delay optimization algorithm was proposed in [13] for the DFE equalizer but cannot be applied in the linear adaptive equalizer. Up to now, the only joint tap-length N and decision delay Δ optimization for the linear adaptive equalizer was derived in [14], in which a linear pre-whitener is applied prior to the equalizer. The pre-whitener makes the equivalent channel be maximum-phase so that the two dimensional search for the optimum N and Δ reduces to one dimensional search by letting $\Delta = N - 1$. This approach is far from ideal: First, since the decision delay is fixed as large as $N - 1$, it brings too much delay in the equalization in many cases. For instance, for the minimum phase channel, the optimum Δ should be 0. Secondly, the equalizer has to be longer as now both the channel and pre-whitener need to be “equalized”. Finally, the length of the pre-whitener needs to be adapted as well, which may diverge the overall system if it is not well handled.

Therefore, it is of great interest to have a robust and simple way to jointly adapt the tap-length and decision delay so that the best MMSE performance of the equalizer can be achieved with minimum tap-length and decision delay. As generally a two dimensional search problem, such joint approach is hard to realize adaptively.

In this paper, we first investigate how the tap-length and

decision delay affect the MMSE performance based on an equivalent model of the equalizer. We then propose to use forward and backward length approaches to jointly adapt the tap-length and decision delay. The proposed algorithm is the first robust yet simple approach in literature for joint tap-length and decision delay optimization. It describes an innovative way to implement the adaptive equalizer.

II. SYSTEM MODEL

The model of the adaptive equalization is illustrated in Fig. 1, where $x(n)$ is the transmitted information signal, $H(z)$ is the channel transfer function, $\eta(n)$ is the channel noise, $y(n)$ is the received signal, $z^{-\Delta}$ is the decision delay, $z(n)$ is the equalizer output, $e(n)$ is the error signal and $W(z)$ is the transfer function of the equalizer.

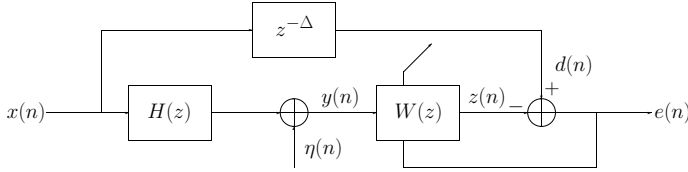


Fig. 1. Adaptive channel equalization

The MMSE equalizer [6] is obtained by minimizing the MMSE cost function

$$\xi = E|d(n) - \mathbf{w}^T(n)\mathbf{y}(n)|^2, \quad (1)$$

with respect to the tap-vector $\mathbf{w}(n)$, where $\mathbf{y}(n) = [y(n), \dots, y(n-N+1)]^T$ which is the tap-input vector, N is the tap-length and $d(n) = x(n-\Delta)$. The decision delay Δ determines which symbol is being detected at the current time n , or the current equalization output $z(n)$ is an estimate of $x(n-\Delta)$.

The maximum potential performance of an MMSE equalizer is achieved by the ideal equalizer with $N \rightarrow \infty$ extending from $-\infty$ to ∞ as (see [15]):

$$\mathbf{w}_\infty = [w_\infty(-\infty), \dots, w_\infty(i), \dots, w_\infty(\infty)]^T, \quad (2)$$

If the decision delay is Δ , the ideal equalizer is also delayed by Δ and is denoted as $\mathbf{w}_{\infty, \Delta}(n)$. Then (1) can be re-written as:

$$\xi = E|e_\infty(n) - (\tilde{d}(n) - \mathbf{w}^T(n)\mathbf{y}(n))|^2, \quad (3)$$

where $e_\infty(n) = d(n) - \mathbf{w}_{\infty, \Delta}^T(n)\mathbf{y}(n)$ and $\tilde{d}(n) = \mathbf{w}_{\infty, \Delta}^T(n)\mathbf{y}(n)$. Since $E|e_\infty(n)|^2$ is a constant and $E|e_\infty(n)\mathbf{y}(n)| = \mathbf{0}$ due to orthogonality principle [6], minimizing (3) is equivalent to minimizing

$$\tilde{\xi} = E|\tilde{d}(n) - \mathbf{w}^T(n)\mathbf{y}(n)|^2. \quad (4)$$

Therefore, the equalization can be regarded as the system modelling shown in Fig. 2, where $W_\infty(z)$ is the transfer function of $\mathbf{w}_\infty(n)$ and $\tilde{e}(n) = \tilde{d}(n) - \mathbf{w}^T(n)\mathbf{y}(n)$. The model in Fig. 2 is very useful in analyzing the tap-length and decision delay.

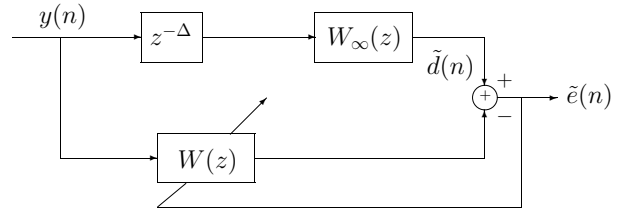


Fig. 2. The equivalent model for the MMSE equalizer.

III. OPTIMUM TAP-LENGTH AND DECISION DELAY

For an equalizer with fixed decision delay Δ , if the tap-length N is too large, some last tap coefficients will be close to zero. Then N can be decreased by removing these near-zero taps with little performance lost. If N is not large enough, all tap coefficients have significant values. Then it is highly possible that the MMSE performance can be improved with one or more taps so that N should be increased. Thus an *optimum* tap-length that best balances the MMSE performance and complexity can be defined as the minimum N satisfying

$$\xi_{N-K, \Delta}^{(f)} - \xi_{N, \Delta} \leq \mathcal{E}, \quad (5)$$

where $\xi_{N, \Delta}$ is the mean square error (MSE) with tap-length N and decision delay Δ , $\xi_{N-K, \Delta}^{(f)}$ is the forward trunked MSE with the last K tap coefficients being removed from the equalizer and \mathcal{E} is a small positive constant.

On another hand, for a given tap-length N , there exists a minimum Δ that minimizes the MMSE. As is shown in Fig. 2, the equalization is equivalent to using an FIR filter to model the unconstrained optimum equalizer $W_\infty(z)$ which is shifted by the decision delay.

The transfer function of the ideal equalizer can be expressed as

$$W_\infty(z) = \sum_{n=-\infty}^{-1} w_\infty(n)z^{-n} + \sum_0^{n=\infty} w_\infty(n)z^{-n}, \quad (6)$$

where the first and second summation are the non-causal and casual terms respectively. In fact, when the noise is small, we have $W_\infty(z) \approx 1/H(z)$ so that the non-causal and casual terms of $W_\infty(z)$ are obtained by inverting the maximum and minimum parts of the channel, which contain the zeros outside and inside the unit circle respectively.

Since the non-causal part cannot be modelled by the causal FIR equalizer no matter how long the tap-length is, the decision delay Δ is introduced to shift non-causal terms into the causal range. In practice, while $W_\infty(z)$ has infinite number of non-causal and casual terms, we always have $w_\infty(i) \rightarrow 0$ when $i \rightarrow \pm\infty$. Therefore, for an equalizer with tap-length N and decision delay Δ , the MMSE (or the equalization error) is contributed by, beside the noise, the non-causal terms of $z^{-\Delta}W_\infty(z)$ and the casual terms with time index larger than z^{-N} .

The optimum Δ that minimizes the MMSE depends on specific channels. For the minimum phase channel that all zeros of $H(z)$ are within the unit circle, $W_\infty(z)$ only contains

causal terms so that we should have $\Delta = 0$. For the maximum phase channel that all zeros of $H(z)$ are outside the unit circle, $W_\infty(z)$ only contains non-causal terms so that we should have $\Delta = N - 1$. In general, unfortunately, there is no fixed relationship between N and Δ . If the decision delay Δ is too large, too many non-causal terms of $W_\infty(z)$ are moved into the causal range so that some initial taps of $\mathbf{w}(n)$ are close to zero. Then the decision delay can be decreased to move these near-zero taps back to the non-causal range with little affect on the MMSE performance. On the other hand, if Δ is not large enough, some non-causal terms with significant values are still in the non-causal range so that the MMSE is too large. Then Δ should be increased to move these “significant” non-causal terms into the causal range. Based on these observation, the optimum decision delay is defined as the minimum Δ that satisfies

$$\xi_{N-K,\Delta}^{(b)} - \xi_{N,\Delta} \leq \mathcal{E}, \quad (7)$$

where $\xi_{N-K,\Delta}^{(b)}$ is the *backward* trunked MSE with the first K tap-coefficients being removed from the tap-vector.

Overall, the joint optimum tap-length and decision delay is defined as the the minimum N and Δ that satisfy (5) and (7) simultaneously.

IV. JOINT TAP-LENGTH AND DECISION DELAY ADAPTATION

While the tap-length and decision delay are jointly optimized by satisfying both (5) and (7), accordingly, they can be adapted by the forward and backward length adaptation as is shown below.

A. Forward length adaptation

At time n , we assume the decision delay is fixed at Δ which is obtained from the previous backward adaptation at time $(n - 1)$. The forward length adaptation is then applied with the fractional length algorithm to adapt the tap-length [4]. To be specific, letting n_f be the pseudo fractional tap-length which can take real positive values, the length adaptation rule is constructed as:

$$n_f(n+1) = (n_f(n) - \alpha) - \gamma \cdot \left[|e_{N(n)}(n)|^2 - |e_{N(n)-K}^{(f)}(n)|^2 \right], \quad (8)$$

where α and γ are small positive numbers, $e_{N(n)}(n)$ is the error signal of the equalizer with tap-length $N(n)$, $e_{N(n)-K}^{(f)}(n)$ is the forward trunked error signal with the last K tap coefficients being removed from the equalizer. α is the leaky factor and satisfies $\alpha \ll \gamma$. Initially we can have $n_f(0) = N(0)$.

The “true” tap-length $N(n)$ is adjusted only if the change of $n_f(n+1)$ is big enough as:

$$N(m+1) = \begin{cases} \lfloor n_f(n+1) \rfloor, & |N(n) - n_f(n+1)| \geq \delta \\ N(m), & \text{otherwise} \end{cases} \quad (9)$$

where δ is a positive threshold value and $\lfloor \cdot \rfloor$ rounds the embraced value to the nearest integer. Note that we use a different time index m for N because the length adjustment may happen in both forward and backward approaches at time n .

In the forward approach, the tap-length adjustment is applied at the “end” of the tap vector. Specifically, if $N(n+1)$ is increased by K , K zero taps are appended after the last tap, and otherwise the last K tap-coefficients are removed from the tap-vector. Or we have

$$\mathbf{w}(n) = \begin{cases} [w_1, \dots, w_{N(n)}, 0, \dots, 0]^T, & \delta_N(m) > 0 \\ [w_1, \dots, w_{N(n)-|\delta_N(m)|}]^T, & \delta_N(m) < 0 \end{cases} \quad (10)$$

where $\delta_N(m) = N(m+1) - N(m)$.

B. Backward length adaptation

While the forward length adaptation adjust the tap-length at the “end” of the tap-vector, based on the symmetry between the cost functions of (5) and (7), a backward approach can be applied to search for the optimum decision delay defined (7).

To be specific, if the initial tap-coefficients are large, the decision delay should be increased, for instance by δ_N , to shift more non-causal terms into to the casual range. At the same time, the tap-length should also be increased by padding δ_N zero taps before the first tap to prevent the last δ_N tap-coefficients from moving out of the equalization range. Similarly, if the decision delay is decreased by δ_N , the tap-length should also be decreased by removing the first δ_N from the equalizer. Therefore, the decision delay adjustment should be synchronized with the tap-length adjustment applying at the “beginning” of the tap-vector.

Similar to the forward approach, we define n_b as the fractional backward tap-length with the adaptation rule as:

$$n_b(n+1) = (n_b(n) - \alpha) - \gamma \cdot \left[|e_{N(n)}(n)|^2 - |e_{N(n)-K}^{(b)}(n)|^2 \right], \quad (11)$$

where $e_{N(n)-K}^{(b)}(n)$ is the backward trunked error signal of the equalizer with the first K tap-coefficients being removed. Initially we set $n_b(0) = n_f(0)$.

If the change of $n_f(n+1)$ is big enough, the tap-length N is adjusted as

$$N(m+1) = \begin{cases} \lfloor n_b(n+1) \rfloor, & |N(n) - n_b(n+1)| \geq \delta \\ N(m), & \text{otherwise} \end{cases} \quad (12)$$

Unlike the forward approach, the tap-length adjustment here is applied at the “beginning” of the equalizer as

$$\mathbf{w}(n) = \begin{cases} [0, \dots, 0, w_1, \dots, w_{N(n)}]^T, & \delta_N(m) > 0 \\ [w_{|\delta_N(m)|+1}, \dots, w_{N(n)}]^T, & \delta_N(m) < 0 \end{cases} \quad (13)$$

where $\delta_N(m) = N(m+1) - N(m)$.

The decision delay adjustment is synchronized with (12) as

$$\Delta(n+1) = \Delta(n) + \delta_N(m), \quad (14)$$

We must ensure $\Delta(n) \geq 0$ at all time to maintain causality as otherwise the equalizer is detecting future data.

C. The algorithm

The forward and backward length adaptation are applied alternatively until they converge to the optimum tap-length and decision delay. At any time n , the forward length adaptation is first applied to adapt the tap-length at the “end” of the tap-vector with the decision delay being fixed at what the backward approach obtained at time $n - 1$. After that, the backward length algorithm is applied to adapt the decision delay and adjust the tap-length at the “beginning” of the tap-vector. Since the tap-length adjustment can happen at both forward and backward approaches at time n , we must apply $n_f(n+1) = n_f + \delta_N(m)$ and $n_b(n+1) = n_b + \delta_N(m)$ before (8) and (11) respectively.

For stability, the tap-length and decision delay should be limited as $N_{min} \leq N(n) \leq N_{max}$ and $\Delta(n) \geq 0$ respectively, where N_{min} and N_{max} are set based on system requirements. It is clear that, during the adaptation, $\Delta(n)$ is always smaller than $N(n)$ so that $\Delta(n)$ is automatically upper limited. Accordingly, $n_f(n)$ and $n_b(n)$ adaptation must be limited to satisfy the constraints on the tap-length and decision delay. To be specific, since $n_f(n)$ only affects the tap-length, its constraint is set as same as that for $N(n)$:

$$N_{min} \leq n_f(n) \leq N_{max}. \quad (15)$$

It is known that $n_b(n)$ adaptation may change both the tap-length and decision delay. On the one hand, we must have $n_b(n) \leq N_{max} - N(n)$ to ensure $N(n) < N_{max}$. On the other hand, $n_b(n)$ should also be larger than N_{min} and $N(n) - \Delta(n)$ to ensure $N(n) \geq N_{min}$ and $\Delta(n) \geq 0$ respectively. Therefore, we have

$$\max(N_{min}, N(n) - \Delta(n)) \leq n_b(n) \leq N_{max} - N(n) \quad (16)$$

Because the forward and backward approaches are highly interactive, of particular importance for convergence is to make sure that (15) and (16) are checked after the length adaption (8) and (11) respectively at every time n .

The convergence analysis of the proposed algorithm can follow that for the original FT algorithm [4] but is more complicated as now there are two FT running in parallel. The details are not shown in this paper due to the space constraint.

V. NUMERICAL SIMULATIONS

Computer simulations are given to verify the proposed algorithm. Three channels are considered:

$$\begin{aligned} \text{minimum phase : } \mathbf{h}_1 &= [1 \ 0.4 \ 0.3 \ -0.1]^T \\ \text{maximum phase : } \mathbf{h}_2 &= [-0.1 \ -0.3 \ 0.4 \ 1]^T \\ \text{mixed phase : } \mathbf{h}_3 &= [-0.1 \ -0.3 \ 0.4 \ 1 \ 0.4 \ 0.3 \ -0.1]^T \end{aligned} \quad (17)$$

For comparison, we deliberately let \mathbf{h}_1 and \mathbf{h}_2 be nearly symmetric, and \mathbf{h}_3 be the combination of \mathbf{h}_1 and \mathbf{h}_2 .

Fig. 3 plots the MMSE with respect to the decision delay Δ for the mixed phase channel \mathbf{h}_3 , where the tap-length N is fixed at 5, 10, 15, 30 and ∞ respectively. It is clearly shown that for a fixed Δ , the MMSE is a non-increasing function of N . For a fixed N , on the other hand, there exists

an optimum Δ that minimizes the MMSE. Overall, the joint optimum tap-length and decision delay are the minimum N and Δ corresponding to the similar MMSE as the best MMSE, where the best MMSE is achieved when $N \rightarrow \infty$. It is shown in Fig. 3 that the optimum N and Δ for \mathbf{h}_3 are around 15 and 10 respectively.

For the minimum and maximum channels \mathbf{h}_1 and \mathbf{h}_2 , we can also plot the MMSE vs. N and Δ curves which are not shown here due to the space constraint.

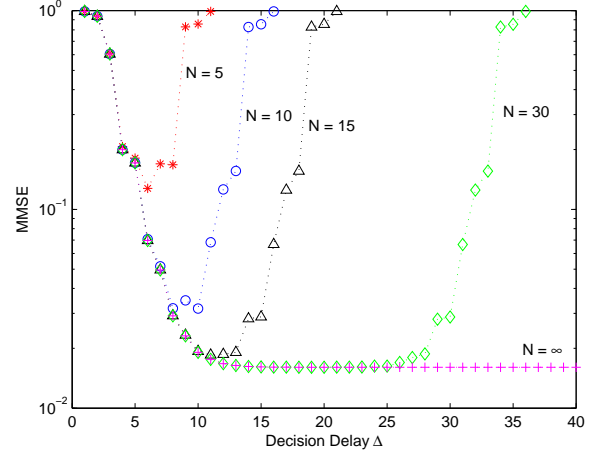


Fig. 3. MMSE vs Decision Delay for the channel \mathbf{h}_3 .

Fig. 4 shows the learning curves of the length adaptation. It is clear that the tap-lengths converges to around the optimum values. It is interesting to observe that, the length learning curves for the \mathbf{h}_1 and \mathbf{h}_2 converge to similar value because \mathbf{h}_1 and \mathbf{h}_2 are nearly symmetric. On the other hand, the mixed channel requires longer tap-length as it is the combination of the other two.

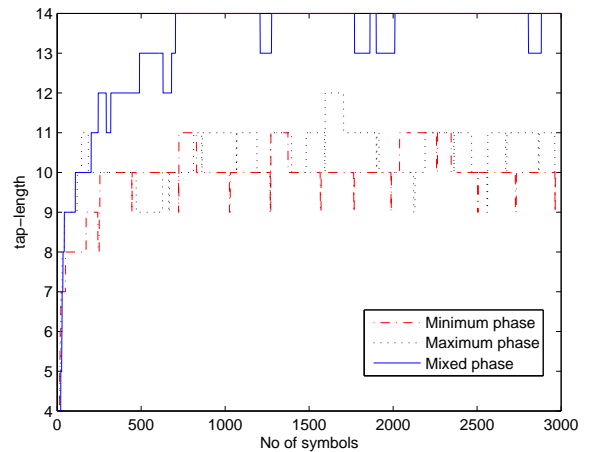


Fig. 4. Tap-length learning curves.

Fig. 5 plots the learning curves of the decision delay

adaptation, where it is shown that the decision delay for the minimum phase channel \mathbf{h}_1 is 3 and those for the maximum and mixed phase channels \mathbf{h}_2 and \mathbf{h}_3 are both around 10. This well matches our previous analysis: First, the decision delays for minimum and maximum phase channels should be around 0 and $N - 1$ respectively; Secondly, the decision delay is determined by the maximum phase part of the channel, if we note that the maximum phase part of \mathbf{h}_3 is just as same as \mathbf{h}_2 .

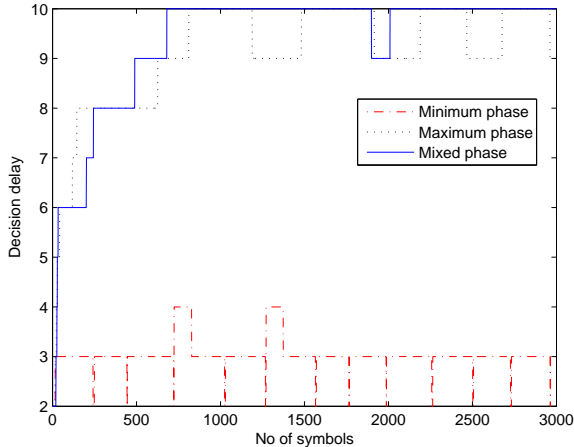


Fig. 5. Decision delay learning curves.

Finally, Fig. 6 shows the MSE learning curves for the three channels to indicate the convergence of the equalizer.

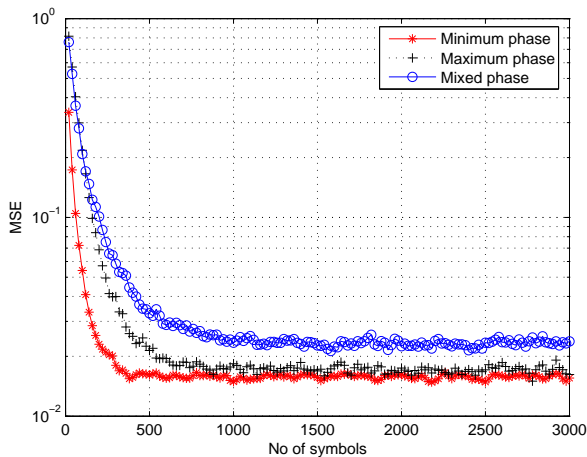


Fig. 6. MSE learning curves.

VI. CONCLUSION

This paper proposes a novel approach to jointly adapt the tap-length and decision delay of the adaptive equalizer so that the best MMSE performance can be achieved with minimum tap-length and decision delay. This is achieved by running

the forward and backward length adaption together at every iteration. Numerical results have been given to verify the algorithm. This is first joint tap-length and decision delay adaptive algorithm in literature with little extra complexity. It describes an innovative way to implement the adaptive equalizer with wide applications.

ACKNOWLEDGEMENT

This research was sponsored by the UK Engineering and Physical Sciences Research Council and DSTL under grant number EP/H012516/1.

REFERENCES

- [1] F. Riera-Palou, J. M. Noras, and D. G. M. Cruickshank, "Linear equalisers with dynamic and automatic length selection," *Electronic Letters*, vol. 37, no. 25, pp. 1553 – 1554, Dec. 2001.
- [2] Y. Gu, K. Tang, H. Cui, and W. Du, "LMS algorithm with gradient descent filter length," *IEEE Signal Processing letters*, vol. 11, no. 3, pp. 305 – 307, March 2004.
- [3] Y. Gong and C. F. N. Cowan, "Structure adaptation of linear MMSE adaptive filters," *IEE Proceedings - Vision, Image and Signal Processing*, vol. 151, no. 4, pp. 271 – 277, Aug. 2004.
- [4] —, "An LMS style variable tap-length algorithm for structure adaptation," *IEEE Trans. on Signal Processing*, vol. 53, no. 7, pp. 2400 – 2407, July 2005.
- [5] Y. Zhang and J. A. Chambers, "Convex combination of adaptive filters for variable tap-length LMS algorithm," *IEEE Signal Processing Letters*, vol. 13, no. 10, pp. 628–631, Oct. 2006.
- [6] S. Haykin, *Adaptive Filter Theory*. Prentice Hall, Englewood Cliffs, NJ, 1996.
- [7] R. Lopez-Valcarce and F. Perez-Gonzalez, "Efficient reinitialization of the prewhitened constant modulus algorithm," *IEEE communications letters*, vol. 5, no. 12, pp. 488 – 490, Dec. 1999.
- [8] P. A. Voois, I. Lee, and J. M. Cioffi, "The effect of decision delay in finite-length decision feedback equalization," *IEEE Trans. on Information Theory*, vol. 42, no. 2, pp. 618 – 621, March 1996.
- [9] A. P. Liavas, P. A. Regalia, and J. P. Delmas, "Least-squares channel equalization performance versus equalization delay," *IEEE Trans. on Signal Processing*, vol. 48, no. 6, pp. 1832 – 1835, June 2000.
- [10] H. Luo and R.-W. Liu, "Blind equalizers for multipath channels with best equalization delay," in *Proceedings of 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, 15-19 March 1999, pp. 2511 – 2514.
- [11] D. Gesbert, P. Duhamel, and S. Mayrargue, "Blind multichannel adaptive mmse equalization with controlled delay," in *Proceedings of 8th IEEE Signal Processing Workshop on Statistical Signal Array Processing (Cat. No.96TB10004)*, June 1996, pp. 172 – 175.
- [12] Y. Gong and C. F. N. Cowan, "Equalization with adaptive time lag," *IEE Proceedings - Communications*, vol. 152, no. 5, pp. 661 – 667, Oct. 2005.
- [13] —, "A self-constructed adaptive decision feedback equalizer," *IEEE Signal Processing Letters*, vol. 13, no. 3, pp. 169 – 172, March 2006.
- [14] Y. Gong, C. F. N. Cowan, and J. Chen, "An adaptive linear equalizer with optimum filter length and decision delay," in *SEVENTH international conference on mathematics in signal processing*, The Royal Agricultural College, Cirencester, UK, 18-20 December 2006.
- [15] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Application*. John Wiley & Sons, 1998.