A brief comparison between Affine Projection Algorithm and Normalized Least Mean Squares Algorithm on adaptive channel equalizer designs

110064511 陳冠銘, 110064533 陳劭珩, 110064535 陳逸衍, 110064544 楊哲寬

***Abstract*—利用仿射投影演算法 (Affine Projection Algorithm)** 於適應性等化器設計以消除通道雜訊之應用。**APA** 跟其他適應性濾波器演算法，如 **LMS, NLMS, RLS...** 相比，最大的優點在於能夠快速地收斂至最佳解，因此在對於需要快速得到最佳解的例子，是非常的有效。

***Index Terms*—LMS, NLMS, Affine projection algorithm, adaptive filter, adaptive equalizer.**

I. Introduction

適應性濾波器演算法中最廣爲人知的方法爲 LMS [1] 是一種基於 LMMSE 以及梯度下降機制所設計的濾波器，其最主要關鍵的設計就是利用瞬間錯誤來得到更新方程式，然而 LMS 有一個很大的缺點，就是存在非均勻收斂的現象。爲了能夠更進一步改善上述問題，NLMS 而後被提出來 [2]，而 NLMS 僅是對更新方程式做些微改變，就能夠有效彌補 LMS 的缺點，因此 NLMS 也是一個被廣爲使用的演算法。雖然 NLMS 可以有效解決非均勻收斂問題，但是 NLMS 與 LMS 同樣遇到一個應用上的困難，即當設定越小的步距參數，LMS 與 NLMS 演算法收斂速度會越慢，對於應用在需要小步距去解決的複雜問題來説，不是個很好的選擇，因此 [3][4] 提出一個基於仿射投影所設計的演算法 (Affine projection algorithm, 簡稱爲 APA)，而 APA 最大的特點在於它可以在複雜問題中更快收斂，更適合應用在一些有需要快速收斂需求的複雜問題上。

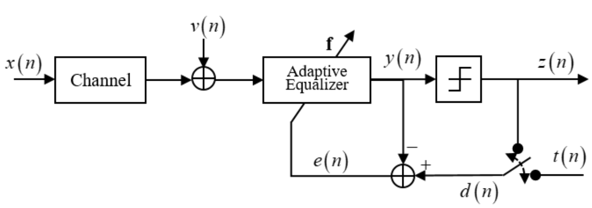


Fig. 1. 等化器方塊圖

此次的報告是基於 APA 所設計的演算法來設計適應性等化器，詳細的步驟如 Fig.1所示。等化器在通訊系統中是爲了要來補償，因爲受到了通道以及雜訊的影響，因此爲了要復原傳送的信號，我們可以先傳送一段接收端已知的信號，叫做訓練序列，當訓練序列傳送時會受到環境影響，而到接收端時，因爲是一段已知的信號，因此接收端可以經過一個適應性濾波器將信號做補償，盡可能的除去通道的影響而還原回原始信號。而圖中的濾波器係數爲、經過偵測後的信號爲。因此此次報告是要用 APA 之適應性等化器來將信號還原，另外也使用了 NLMS 以及 APA-SE 來比較性能。

接下來將會依照下列區塊介紹: section II爲 APA 演算法説明及推導，section III爲各等化器設計的模擬結果，section IV爲結論及工作分配。

II. Algorithm

在演算法的部分，第一部分解釋爲何 APA 可以比 NLMS 還要更快收斂；第二部分則是使用基於最佳化理論的方式來證明出 APA 的更新方程式；第三部分則是證明 APA 的收斂行爲。而在以下演算法中，符號 $(\cdot)^T$ 代表轉置矩陣，Π 爲仿射子空間、P_Π 代表仿射投影到 Π。

A. 空間迭代收斂證明

要如何證明在相同的迭代次數下，NLMS 濾波器係數與收斂的結果相差一定會大於 APA 濾波器係數與收斂的結果，也就是

$$\| \mathbf{f}^o - \mathbf{f}_{n+1}^{(1)} \| \geq \| \mathbf{f}^o - \mathbf{f}_{n+1}^{(2)} \| \tag{1}$$

如 Fig.2(a)所示，假設濾波器最佳解 $\mathbf{f}^o$ 剛好在平面 Π_{n-1} 及 Π_n 的交界處。其中新的濾波器係數 $\mathbf{f}_{n+1}$ 在 NLMS 中表示爲 $\mathbf{f}_{n+1}^{(1)}$，在 APA 中爲 $\mathbf{f}^o - \mathbf{f}_{n+1}^{(2)}$，且 $0 < \alpha < 2$，NLMS 是直接垂直投影到平面 Π_n，而 APA 則是投影到平面 Π_{n-1} 及 Π_n。因爲 $P_{\Pi_n}\mathbf{f}_n - \mathbf{f}_n$ 與仿射子空間 Π_n，$P_{\Pi_n}\mathbf{f}_n - \mathbf{f}_n$ 與 $\Pi_n \cap \Pi_{n-1}$ 正交。$P_{\Pi_n \cap \Pi_{n-1}}\mathbf{f}_n - \mathbf{f}_n$ 與 $\Pi_n \cap \Pi_{n-1}$ 也會正交。藉由這些關係，可以得到

$$P_{\Pi_n \cap \Pi_{n-1}}\mathbf{f}_n - P_{\Pi_n}\mathbf{f}_n = P_{\Pi_n \cap \Pi_{n-1}}\mathbf{f}_n - \mathbf{f}_n - (P_{\Pi_n}\mathbf{f}_n - \mathbf{f}_n) \tag{2}$$

而(2)會與 $\Pi_n \cap \Pi_{n-1}$ 正交，因此，向量 $P_{\Pi_n \cap \Pi_{n-1}}\mathbf{f}_n - P_{\Pi_n}\mathbf{f}_n$ 會與 $P_{\Pi_n \cap \Pi_{n-1}}\mathbf{f}_n - \mathbf{f}^o$ 正交，依照此關係，可得

$$\| P_{\Pi_n}\mathbf{f}_n - \mathbf{f}^o \| \geq \| P_{\Pi_n \cap \Pi_{n-1}}\mathbf{f}_n - \mathbf{f}^o \| \tag{3}$$

再來假設

$$\begin{aligned} a &= \| \mathbf{f}^o - \mathbf{f}_n \|, b = \| \mathbf{f}^o - P_{\Pi_n}\mathbf{f}_n \| \\ c &= \| \mathbf{f}^o - P_{\Pi_n \cap \Pi_{n-1}}\mathbf{f}_n \|, d = \| \mathbf{f}_n - P_{\Pi_n}\mathbf{f}_n \| \\ e &= \| \mathbf{f}_n - P_{\Pi_n \cap \Pi_{n-1}}\mathbf{f}_n \| \end{aligned} \tag{4}$$

如 Fig. 2(b)所示。而在 Fig. 2(b)中可以看到 $\mathbf{f}^o$ 與 $\mathbf{f}_{n+1}^{(1)}$ 跟 $\mathbf{f}^o$ 與 $\mathbf{f}_{n+1}^{(2)}$ 的距離可以藉由直角三角關係是得到，如下

$$\begin{aligned} \| \mathbf{f}^o - \mathbf{f}_{n+1}^{(1)} \|^2 &= b^2 + \| \mathbf{f}_{n+1}^{(1)} - P_{\Pi_n}\mathbf{f}_n \|^2 \\ &= b^2 + (1-\alpha)^2 d^2 \\ &= b^2 + (1-\alpha)^2 \left(a^2 - b^2\right) \\ &= \alpha (2-\alpha) b^2 + (1-\alpha) a^2 \end{aligned} \tag{5}$$

同樣

$$
\begin{aligned}
\| \mathbf{f}^o - \mathbf{f}_{n+1}^{(2)} \|^2 &= c^2 + \| \mathbf{f}_{n+1}^{(2)} - P_{\Pi_n \cap \Pi_{n-1}} \mathbf{f}_n \|^2 \\
&= c^2 + (1-\alpha)^2 e^2 \\
&= c^2 + (1-\alpha)^2 \left(a^2 - c^2\right) \\
&= \alpha (2-\alpha) c^2 + (1-\alpha) a^2 \qquad (6)
\end{aligned}
$$

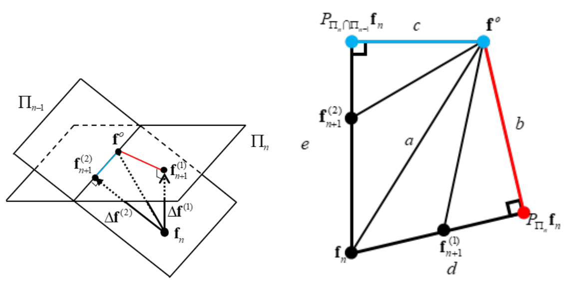
然後 $b > c$、$\alpha (2-\alpha) \geq 0$ 將(5)與(6)組合，可以得到

$$\| \mathbf{f}^o - \mathbf{f}_{n+1}^{(1)} \| - \| \mathbf{f}^o - \mathbf{f}_{n+1}^{(2)} \| = \alpha (2-\alpha) \left(b^2 - c^2\right) \qquad (7)$$

因爲兩者相減所剩下的值恆大於等於零，因此可以得到

$$\| \mathbf{f}^o - \mathbf{f}_{n+1}^{(1)} \| \geq \| \mathbf{f}^o - \mathbf{f}_{n+1}^{(2)} \| \qquad (8)$$

所以證明了 APA 會比 NLMS 還要快收斂。



(a) (b)

Fig. 2. (a) 收斂投影圖 (b) 收斂平面圖

B. 由最佳化理論證明 APA 的更新方程式

一般證明更新方程式的方式有很多種，爲了能夠看出 NLMS 以及 APA 的相似之處，在此使用最佳化理論 [5] 來證明 NLMS 以及 APA，而 NLMS 的方式如下

$$\min_{\mathbf{f}} \quad \| \mathbf{f} - \mathbf{f}_n \|^2 \qquad (9)$$

$$\text{s.t.} \quad d(n) - \mathbf{x}_n^T \mathbf{f} = 0 \qquad (10)$$

其中式(9)及(10)代表的是要讓實際濾波器係數與估計的濾波器係數相差越小越好，而且要在錯誤爲零的條件下達到，接著使用拉格朗日乘數可以得到 NLMS 的更新方程式，爲

$$\mathbf{f}_{n+1} = \mathbf{f}_n + \alpha e(n) \frac{\mathbf{x}_n}{c + \mathbf{x}_n^T \mathbf{x}_n} \qquad (11)$$

其中 c 爲非常小的正實數。其中

$$\mathbf{x}_n = \left[x(n), x(n-1), \ldots, x(n-M+1)\right]^T \qquad (12)$$

$$\mathbf{f}_n = \left[f(n), f(n-1), \ldots, f(n-M+1)\right]^T \qquad (13)$$

M 代表濾波器長度。

而 APA 的證明與 NLMS 極爲相似，差在限制除了限制當下時間點的誤差外，也考慮了 $L-1$ 個時間點的誤差，如下所示

$$\min_{\mathbf{f}} \quad \| \mathbf{f} - \mathbf{f}_n \|^2 \qquad (14)$$

$$
\begin{aligned}
\text{s.t.} \quad & d(n) - \mathbf{x}_n^T \mathbf{f} = 0 \\
& d(n-1) - \mathbf{x}_{n-1}^T \mathbf{f} = 0 \\
& \vdots \\
& d(n-L+1) - \mathbf{x}_{n-L+1}^T \mathbf{f} = 0 \qquad (15)
\end{aligned}
$$

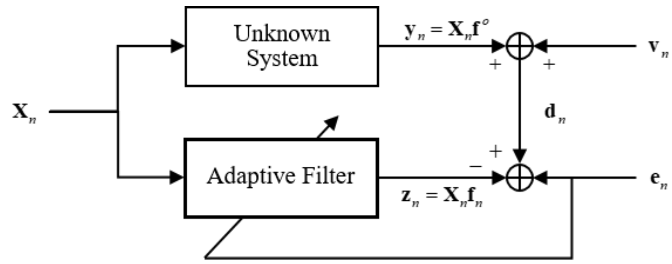


Fig. 3. 收斂平面圖

L 代表投影誤差長度。利用拉格朗日乘數，可以得到

$$
\begin{aligned}
\mathcal{L}(\mathbf{f}, \lambda_0, \lambda_1, \ldots, \lambda_{L-1}) &= \| \mathbf{f} - \mathbf{f}_n \|^2 \\
&+ \sum_{i=0}^{L-1} \lambda_i \left[d(n-i) - \mathbf{x}_{n-i}^T \mathbf{f}\right] \qquad (16)
\end{aligned}
$$

並對(16)中 $\mathbf{f}$ 偏微分令其等於零得到

$$2\mathbf{f} - 2\mathbf{f}_n - \sum_{i=0}^{L-1} \lambda_i \mathbf{x}_{n-i} = 0 \qquad (17)$$

整理得

$$
\begin{aligned}
\mathbf{f} &= \mathbf{f}_n + \sum_{i=0}^{L-1} \frac{\lambda_i}{2} \mathbf{x}_{n-i} \\
&= \mathbf{f}_n + \mathbf{X}_n^T \boldsymbol{\lambda} \qquad (18)
\end{aligned}
$$

其中，

$$\mathbf{X}_n = \left[\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{n-L+1}\right]^T \qquad (19)$$

$$\boldsymbol{\lambda} = \left[\tfrac{\lambda_0}{2}, \tfrac{\lambda_1}{2}, \ldots, \tfrac{\lambda_{L-1}}{2}\right]^T \qquad (20)$$

並對(18)等號兩邊的左邊乘上 $\mathbf{X}_n$，若假設(19)每列爲線性獨立，且 $\mathbf{X}_n \mathbf{X}_n^T$ 爲非奇異矩陣，可以得到

$$
\begin{aligned}
\mathbf{X}_n \mathbf{X}_n^T \boldsymbol{\lambda} &= \mathbf{e}_n \\
\boldsymbol{\lambda} &= \left(\mathbf{X}_n \mathbf{X}_n^T\right)^{-1} \mathbf{e}_n \qquad (21)
\end{aligned}
$$

將(21)帶入(18)中，可得

$$
\begin{aligned}
\mathbf{f} &= \mathbf{f}_n + \mathbf{X}_n^T \left(\mathbf{X}_n \mathbf{X}_n^T\right)^{-1} \mathbf{e}_n \\
\mathbf{e}_n &= \mathbf{d}_n - \mathbf{X}_n \mathbf{f}_n \qquad (22)
\end{aligned}
$$

其中

$$\mathbf{d}_n = \left[d(n), d(n-1), \ldots, d(n-L+1)\right]^T \qquad (23)$$

因此 APA 的更新方程式爲

$$
\begin{aligned}
\mathbf{f}_{n+1} &= \mathbf{f}_n + \alpha \mathbf{X}_n^T \left(\mathbf{X}_n \mathbf{X}_n^T\right)^{-1} \mathbf{e}_n \\
&= \mathbf{f}_n + \alpha \mathbf{X}_n^{\dagger} \mathbf{e}_n \qquad (24)
\end{aligned}
$$

其中 $(\cdot)^{\dagger}$ 代表虛擬反矩陣。再來，爲了簡化更新方程式，我們將 $\mathbf{e}_n$ 取 sign，得到新的更新方程式爲

$$\mathbf{f}_{n+1} = \mathbf{f}_n + \alpha \mathbf{X}_n^T \left(\mathbf{X}_n \mathbf{X}_n^T\right)^{-1} sgn(\mathbf{e}_n) \qquad (25)$$

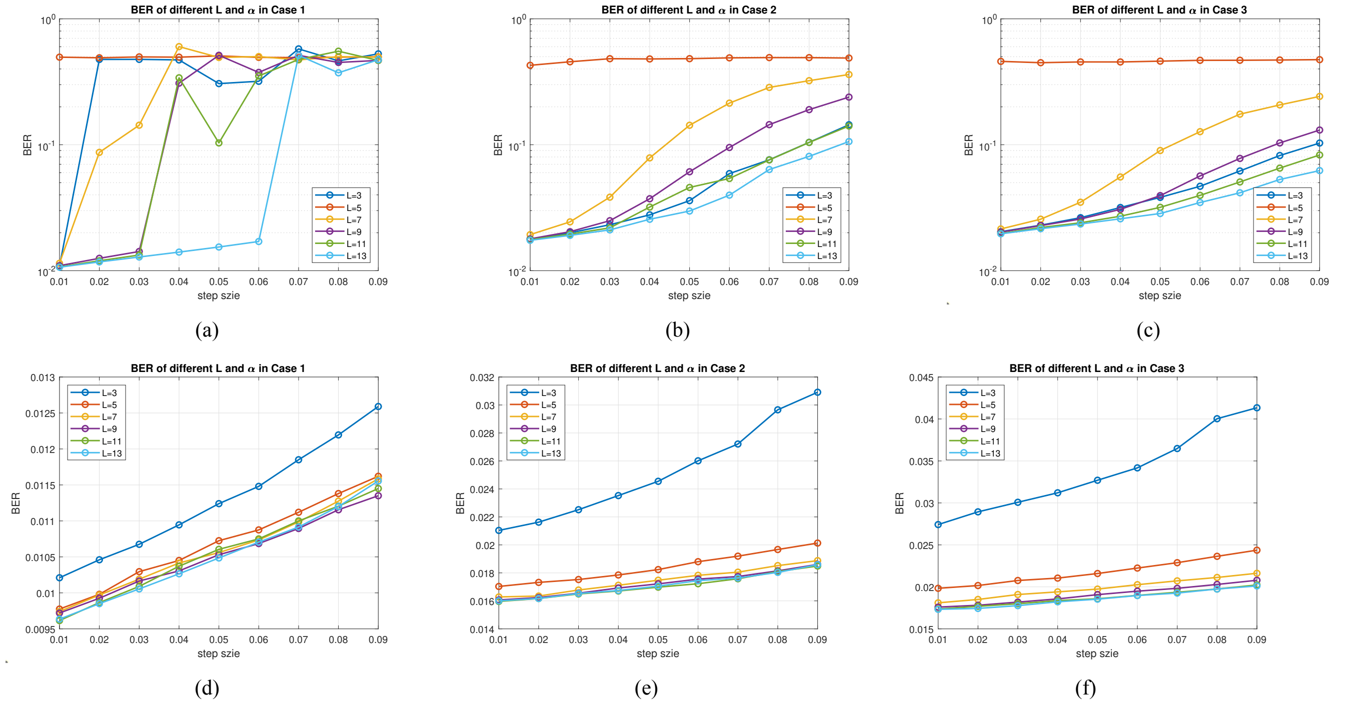


Fig. 4. (a) APA 在 static channel 的 BER (b) APA 在 quasi-static channel 的 BER (c) APA 在 time varying channel 的 BER (d) NLMS 在 static channel 的 BER (e) NLMS 在 quasi-static channel 的 BER (f) NLMS 在 time varying channel 的 BER

C. APA 收斂行爲

要證明收斂行爲，我們先假設是一個系統識別的問題，Fig. 3所示。

其中，輸入信號 $\mathbf{X}_n$ 爲廣義穩態信號、$\mathbf{v}_n$ 爲 AWGN、且 $\mathbf{X}_n$ 與 $\mathbf{v}_n$ 相互獨立。未知系統中的係數爲 $\mathbf{f}^o$，更新方程式如下

$$\begin{aligned}\mathbf{f}_{n+1} &= \mathbf{f}_n + \alpha\mathbf{X}_n^T\left(\mathbf{X}_n\mathbf{X}_n^T\right)^{-1}\mathbf{e}_n \\ &= \mathbf{f}_n + \alpha\mathbf{X}_n^T\left(\mathbf{X}_n\mathbf{X}_n^T\right)^{-1}\left(\mathbf{d}_n - \mathbf{X}_n\mathbf{f}_n\right) \qquad (26)\end{aligned}$$

而 $\mathbf{d}_n = \mathbf{X}_n^T\mathbf{f}_n^o + \mathbf{v}_n$，帶回(26)式如下

$$\mathbf{f}_{n+1} = \mathbf{f}_n + \alpha\mathbf{X}_n^T\left(\mathbf{X}_n\mathbf{X}_n^T\right)^{-1}\left(\mathbf{X}_n^T\mathbf{f}_n^o + \mathbf{v}_n - \mathbf{X}_n\mathbf{f}_n\right) \qquad (27)$$

將(27)簡化可得

$$\begin{aligned}\widetilde{\mathbf{f}}_{n+1} &= \left(\mathbf{I} - \alpha\mathbf{X}_n^T\left(\mathbf{X}_n\mathbf{X}_n^T\right)^{-1}\mathbf{X}_n\right)\widetilde{\mathbf{f}}_n \\ &\quad - \alpha\mathbf{X}_n^T\left(\mathbf{X}_n\mathbf{X}_n^T\right)^{-1}\mathbf{v}_n \qquad (28)\end{aligned}$$

(28)中第二項因爲 $\mathbf{X}_n$ 與 $\mathbf{v}_n$ 相互獨立因此可以直接省略，此外 $\widetilde{\mathbf{f}}_{n+1} = \mathbf{f}_o - \mathbf{f}_n$。

假設 $\mathbf{A}_n = \mathbf{X}_n^T\left(\mathbf{X}_n\mathbf{X}_n^T\right)^{-1}\mathbf{X}_n$，而藉由 (28)可以證明在 $0 < \alpha < 2$ 的條件下，$\widetilde{\mathbf{f}}$ 將會收斂到 0。首先，將 (28)兩邊同取期望值，並且做化簡可得

$$E\{\widetilde{\mathbf{f}}_{n+1}\} = \left(\mathbf{I} - \alpha E\{(\mathbf{A}_n)\}\right)E\{\widetilde{\mathbf{f}}_n\} \qquad (29)$$

再來將 $E\{\mathbf{A}_n\}$ 做特徵分解如下

$$E\{\widetilde{\mathbf{f}}_{n+1}\} = \left(\mathbf{I} - \alpha\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T\right)E\{\widetilde{\mathbf{f}}_n\} \qquad (30)$$

(30)中的 $\mathbf{Q}$ 爲特徵向量，其對應到的特徵值爲 $\mathbf{\Lambda} = diag\left(\lambda_1, \lambda_2, \ldots, \lambda_k\right)$，再來假設 $\widehat{\mathbf{f}}_n = \mathbf{Q}^T E\{\widetilde{\mathbf{f}}_n\}$ 帶回上式，會得到

$$\mathbf{Q}^T E\{\widetilde{\mathbf{f}}_{n+1}\} = \mathbf{Q}^T E\{\widehat{\mathbf{f}}_n\} - \alpha\mathbf{Q}^T\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T E\{\widetilde{\mathbf{f}}_n\} \qquad (31)$$

$$\begin{aligned}\widehat{\mathbf{f}}_{n+1} &= \widehat{\mathbf{f}}_n - \alpha\mathbf{\Lambda}\widehat{\mathbf{f}}_n \\ &= \left(1 - \alpha\mathbf{\Lambda}\right)\widehat{\mathbf{f}}_n \\ \widehat{f}_n\left(j\right) &= \left(1 - \alpha\lambda_j\right)^n \widehat{f}_0\left(j\right) \qquad (32)\end{aligned}$$

而其中 $0 < \lambda_j < n$，因此當 $0 < \alpha < 2$ 時，其括號內值都會小於 0，因此可以得知 $\widetilde{\mathbf{f}}_n$ 將會收斂至 0。

III. Simulations

由 Fig. 4 所示可以看到，根據我們嘗試不同 filter length L 與不同 step size α 的模擬結果，無論是 NLMS 或是 APA，選擇越高的 filter length L 與越小的 step size α 所得到的 BER 會越低。由上述模擬結果推估，filter length L 大約在 11、13 左右時效能會逐漸收斂。因此我們最後使用 $L = 13$。若設備與時間許可的話，理論上更高的 filter length L 應能達到更好的結果。至於最終演算法選擇，由於 NLMS 的效能稍微比 APA 更佳 (即 BER 較低)，因此在這個通道雜訊消除的問題應用上，我們會選用 NLMS 來完成。

IV. Conclusion

LMS 是利用瞬間的錯誤來得到更新方程式，LMS 有一個很大的缺點，就是存在非均勻收斂的現象。爲了能夠更進一步改善上述問題，NLMS 而後被提出來，雖然 NLMS 可以有效解

決非均勻收斂問題，但是 NLMS 與 LMS 同樣遇到一個應用上的困難，即當設定越小的步距參數，LMS 與 NLMS 演算法收斂速度會越慢，對於應用在需要小步距去解決的複雜問題來說，不是個很好的選擇，因此利用仿射投影演算法 (Affine projection algorithm) 去優化，而 APA 最大的優點在於它可以在複雜問題中更快收斂，更適合應用在一些有需要快速收斂需求的複雜問題上。

我們這次對 NLMS 與 APA 這兩種演算法進行簡單的分析與比較。APA 的優勢在於收斂速度較快，不過它的演算法複雜度也較高。因此，需要考量問題需求及硬體資源、時間成本等因素，來判斷選擇何種演算法更合適。如果愈解決的問題相當複雜且有快速收斂的要求，APA 可能是合適的選擇；而如果當 BER 比收斂速度更加重要時，則是 NLMS 比 APA 更加合適。從模擬結果來看，雖然 APA 收斂速度快，但相比之下其 BER 表現相較 NLMS 在簡單的雜訊消除問題下並沒有顯著的優勢，故 APA 在等化器演算法設計中並不是最佳解，但若只看 MSE，因爲 APA 收斂速度快的緣故，是較 NLMS 更佳的演算法。

最後是工作分配，本組同學都積極參與期末專案實作，雖然投入時間不同但大家整體貢獻都差不多，謝謝老師及助教學長。

學號	姓名	工作內容	貢獻度
110064511	陳冠銘	APA 演算法數學推導, 及物理意義分析	25%
110064533	陳劭珩	Latex 撰寫, 上台報告, 演算法程式實作	25%
110064535	陳逸衍	NLMS 程式實作, 圖表, 及投影片製作	25%
110064544	楊哲寬	RLS 程式實作, 資料蒐集, 及程式除錯	25%

TABLE I
ASP Term Project 第 6 組工作分配表

References

[1] B. Widrow and M. E. Hoff, "Adaptive switching circuits," *7960 IRE WESCON Conv. Record*, pp. 96–104, Aug. 1960.

[2] T. Hsia, "Convergence analysis of lms and nlms adaptive algorithms," in *ICASSP '83. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 8, 1983, pp. 667–670.

[3] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electronics and Communications in Japan (Part I: Communications)*, vol. 67, no. 5, pp. 19–27, 1984.

[4] K. Ozeki, "Theory of affine projection algorithms for adaptive filtering," in *Springer*, 2016.

[5] H. S., "Adaptive filter theory," in *Pearson*, 2014.