EFFECTS OF INPUT DATA CORRELATION ON THE CONVERGENCE OF BLIND ADAPTIVE EQUALIZERS

James P. LeBlanc¹ Kutluyl Doğançay² Rodney A. Kennedy² C. Richard Johnson, Jr.¹

¹School of Electrical Engineering, Cornell University, Ithaca, NY USA

²Dept. of Systems Engineering, RSISE, The Australian National University, Canberra ACT 0200, Australia

ABSTRACT

A variety of blind equalization algorithms exist. These algorithms, which draw on some theoretical justification for the demonstration or analysis of their purportedly ideal convergence properties, almost invariably rely on the input data being independent and identically distributed (i.i.d.). In contrast, in this paper we show that input correlation can have a marked effect on the character of algorithm convergence. We demonstrate that under suitable input data correlation and channels: *i*) undesirable local minima present in the i.i.d. case are absent for certain correlated sources implying ideal global convergence for some situations and, *ii*) the most commonly employed practical algorithm can exhibit ill-convergence to closed-eye minima even under the popular single spike initialization when an eye-opening equalizer parameterization is possible.

1. INTRODUCTION

In high performance telecommunication systems, coded modulated signals are employed and the receiver has to typically perform the twin tasks of equalization and decoding. The need for equalization arises from the dispersion of the transmitted symbols (intersymbol interference) over a band-limited channel at a high data rate. It is clear that the input symbols will display *some* degree of correlation (reflected in the second and higher order statistics) due to the redundancy introduced in the coding process which is designed to mitigate the effects of noise. Therefore, it may be unrealistic, as is done in many theoretical equalization studies, to assume the data samples can be drawn from an i.i.d. process.

In this paper we make two contributions to the investigation into the effects of input data correlation on the convergence behavior of blind equalization algorithms and pose a number of important open questions which naturally arise from this work. The first contribution deals with the detrimental effects of correlation where a well behaved system in the i.i.d. case displays ill-convergence when the data sequence is generated by a particular Markov source. The second contribution deals with an extreme case of the beneficial effects of correlation where all stable undesirable local minima present in the i.i.d. case for a well studied example disappear for a range of input data correlations. These results point to the care that should be taken when considering the most suitable blind algorithm for a data source of a known correlation. Contrarywise, these results illustrate that there may be advantages in carefully designing the input correlation, perhaps via coding, to assist in the blind equalization identification process.

¹Supported by NSF Grant ECS-9022785

²Partially supported by CRC for Robust and Adaptive Systems and The Australian Research Council

In particular, we use the Constant Modulus Algorithm (CMA), also known as the Godard Algorithm [3], to illustrate these correlation effects, although it is not deemed that such effects are confined strictly to the CMA algorithm.

Consider the channel/equalization system as in Figure 1. For the examples presented here, the channel is modelled as a three tap moving average channel with impulse response $C = [c_0 \ c_1 \ c_2]^T$, and the equalizer consists of a two tap moving average filter given by $D = [d_0 \ d_1]^T$. The source alphabet members are considered binary (i.e. $s_k \in \{1, -1\}$). In this case the CMA adaptation algorithm which uses a gradient descent on the CMA cost function ($J = \frac{1}{4}E\{(1 - y^2)^2\}$) becomes,

$$D_{k+1} = D_k + \mu y(1 - y^2)X_k \quad (1)$$

where X_k represents the equalizer's regressor at time k ($X_k = [r_k \ r_{k-1}]^T$) and μ is the step-size.

Figure 1. Block Diagram

2. MARKOV SOURCE MODEL

Consider the source to be modelled as a finite state Markov process with the probability transition matrix Ξ . This implies that the source is correlated (for the particular parameters used herein). The element $\Xi_{(i,j)}$ is the transition probability from state j to state i . The states are indexed by the binary encoding of last n states, where a +1 source element corresponds to a 1 and a -1 source element corresponds to a 0 (e.g. state #3 for the case $n = 4$ represents $s_{k-4} = -1, s_{k-3} = -1, s_{k-2} = 1, s_{k-1} = 1$, and is encoded as 0011 $\rightarrow$ 3). Here then state #3 (0011) has a probability of 0.5 to become state #6 (0110) and 0.5 probability to become state #7 (0111), hence $\Xi_{(6,3)} = 0.5$, and $\Xi_{(7,3)} = 0.5$. (Note that the first matrix entry is $\Xi_{(0,0)}$, not $\Xi_{(1,1)}$.) For one of the example Markov sources used, labelled "source 1" (with spectrum as shown in Figure 8), we have the transition matrix Ξ_1 given by

$$\Xi_1 = \begin{bmatrix} 0.25 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0.75 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 & 0 & 0.8 & 0 & 0 \\ 0 & 0.2 & 0 & 0 & 0 & 0.2 & 0 & 0 \\ 0 & 0 & 0.2 & 0 & 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0.8 & 0 & 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0.75 \\ 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0.25 \end{bmatrix}. \quad (2)$$

3. STATIONARY POINT CONDITION

According to averaging theory, the stationary points of the instantaneous CMA cost used in (1) are the same as those of the average cost [1]. From $y_k = X_k^T D$ we obtain an expression describing algorithmic stationary points,

$$\frac{\partial J}{\partial D} = E \{ (1 - X^T D X^T D) (X X^T D) \} = 0 \quad (3)$$

which may be written as,

$$\begin{aligned} E \left\{ (r_k d_0 + r_{k-1} d_1)^2 \begin{bmatrix} r_k^2 & r_k r_{k-1} \\ r_{k-1} r_k & r_{k-1}^2 \end{bmatrix} \begin{bmatrix} d_0 \\ d_1 \end{bmatrix} \right\} \\ = E \left\{ \begin{bmatrix} r_k^2 & r_k r_{k-1} \\ r_{k-1} r_k & r_{k-1}^2 \end{bmatrix} \begin{bmatrix} d_0 \\ d_1 \end{bmatrix} \right\}. \end{aligned} \quad (4)$$

Assuming that the received signal is fourth order weakly stationary, the fourth and second moments of the received signal may be represented by,

$$\begin{aligned} R_{ijk} &= E \{ r_n r_{n-i} r_{n-j} r_{n-k} \} \\ R_i &= E \{ r_n r_{n-i} \}. \end{aligned} \quad (5)$$

Now the matrix equation above can be written as,

$$\begin{bmatrix} R_{000} d_0^3 + 3R_{001} d_0^2 d_1 + 3R_{011} d_0 d_1^2 + R_{111} d_1^3 \\ R_{001} d_0^3 + 3R_{011} d_0^2 d_1 + 3R_{111} d_0 d_1^2 + R_{000} d_1^3 \end{bmatrix} = \begin{bmatrix} R_0 d_0 + R_1 d_1 \\ R_1 d_0 + R_0 d_1 \end{bmatrix}. \quad (6)$$

4. CHANGING NUMBER OF STATIONARY POINTS

The multi-modality of the CMA error surface resulting from similar channel-equalizer examples (but i.i.d. source) has been shown in [2]. From these earlier presentations of 2 tap equalizers, it has been seen that there generally are 2 minima exhibiting globally best performance, 2 minima exhibiting lesser performance (with 4 saddle points and a local maximum at the origin). It is appropriate to ask if this is always the case. Computations show that this is not always true, the number of stationary points on the error surface can change.

In Figure 2, a Markov source resulting from Ξ_1 given in (2) is filtered through a channel with impulse response [1.0 0.3 0.09]. The resulting received signal is then used to compute numerical estimates of the necessary correlation quantities (R_{000} , R_{001} , etc). These values (see Tables 1 and 2) are then substituted into (6) allowing computation of the roots of each of the two equations.

Each point on the curves in Figure 2 represents an equalizer parameterization that is a solution to some row of (6), in other words, a parameterization where $\frac{\partial J}{\partial d_i} = 0$ for some i . The intersections of these curves then correspond to the simultaneous solutions of these equations, hence stationary points of the CMA algorithm. Note that for this received signal correlation (due to correlated source 1, and channel

Figure 2. Channel/Correlation with 9 Stationary Points

response C), we have the 9 stationary points. Evaluation of the stability of these stationary points allows classification of each point as either a local maximum, local minimum or saddle point, as labelled in the figure.

Another case is examined in which the same source is passed through a different channel with $c = [1 \ 0.8 \ 0.64]$. The resulting received signal has different correlations, hence different polynomial coefficients (Tables 1 and 2). In this case the number of stationary points (shown by the intersections) is only 5, as in Figure 3. Thus, the error surface has only global minima (2 minima of same performance due to symmetry). The error surface has been deformed in a way as to change the number of minima. It is interesting to note that this correlation allows *any* equalizer initialization to converge to an open-eye equalizer setting.

Figure 3. Channel/Correlation with 5 Stationary Points, Open-Eye Minima

In the case of a two-tap equalizer, Morse theory provides useful insight into the relationship between the number of stationary points of different type. For a given function $f(\mathbf{x}) : \mathbf{R}^2 \rightarrow \mathbf{R}$ where $f(\mathbf{x})$ is smooth, has isolated extrema and $f(\mathbf{x}) \rightarrow \infty$ when $\|\mathbf{x}\| \rightarrow \infty$, we have the following relations for its stationary points:

$$\begin{aligned} \text{No. maxima} &\geq 0 \\ \text{No. saddles} &\geq \text{no. maxima} \\ \text{No. minima} &= \text{no. saddles} - \text{no. maxima} + 1. \end{aligned}$$

In our case, $J(D)$ satisfies the requirements for $f(\mathbf{x})$. What is more, it has only one maximum at the origin [4]. Thus

the number of minima is equal to the number of saddles. Two cubic equations can have at most 9 simultaneous solutions. Also, the number of minima has to be an even number. Hence, the *only* other possibility is to have 5 solutions with 2 minima.

5. CORRELATION AND CHANNEL COMBINATION PRODUCES ONLY CLOSED-EYE STATIONARY POINTS

A more disturbing example is found below. Here a different correlated source (source “2” given by Ξ_2 as in (7)) which has spectrum as shown in Figure 8 and correlations found in Tables 1 and 2. When used in conjunction with the $[1 \ 0.8 \ 0.64]$ channel this source produces a CMA error surface possessing only 2 minima with the added twist that both of these minima are closed eye (Figure 4). In this situation the source correlation has produced an error surface with only global closed-eye minima. Notice that this occurs in spite of the existence of a fairly good eye-opening equalizer parameterization. Figure 5 confirms this analysis by initializing the equalizer at an approximate channel inverse. With this “omniscient” open-eye initialization, the CMA adaptation then brings the equalizer to a closed-eye a parameterization.

It should be pointed out that a small perturbation in source correlation (achieved by slightly changing the elements of Ξ_2 , to produce Ξ_3 as in (7) and (8)) produces again the familiar CMA error surface with 9 stationary points. The curves intersect at a parameterization allowing eye-opening equalization (Figure 6). Here, the trajectories of a CMA simulation display this behavior in Figure 7.

While we have seen a large number of cases where the number of stationary points can be changed between 5 and 9 due to source correlation, we have chosen this particular example to demonstrate that small changes in the source correlation can produce markedly different behavior of the CMA algorithm. Regarding the similarity of the spectra of the source signals and the differing convergent behavior, it is suggested that spectral measurements alone may not be beneficial in flagging algorithmic (mis)convergence.

Figure 4. Channel/Correlation with 5 Stationary Point, Closed-Eye Minima

6. CONCLUSIONS

Two novel results regarding the effects of data correlation on the convergence behavior of blind algorithms have been indicated —correlation can aid convergence by reducing the number of local minima —or correlation can foil an algorithm that works well in the i.i.d. case. These results point to the conclusion that some theory dealing with the

Figure 5. Tap Trajectories due to Source “2”

Figure 6. Channel/Correlation with 9 Stationary Point, Open-Eye Minima

correlation input data case is badly needed and a number of open problems need to be addressed. Equally significant, these results put in question a majority of the known results in the theory of blind equalization which based their analysis and discussion on the assumption that the data sequence is i.i.d. since for many practical situations this is a property that does not hold.

Figure 7. Tap Trajectories due to Source “3”

7. APPENDIX

$$\Xi_i = \begin{bmatrix} \alpha & . & . & . & . & . & \beta & . & . & . & . & . \\ \beta & . & . & . & . & . & \alpha & . & . & . & . & . \\ . & \alpha & . & . & . & . & \beta & . & . & . & . & . \\ . & \beta & . & . & . & . & \alpha & . & . & . & . & . \\ . & . & \gamma & . & . & . & . & \gamma & . & . & . & . \\ . & . & \delta & . & . & . & . & \delta & . & . & . & . \\ . & . & . & \beta & . & . & . & . & \gamma & . & . & . \\ . & . & . & \alpha & . & . & . & . & \delta & . & . & . \\ . & . & . & . & \delta & . & . & . & \beta & . & . & . \\ . & . & . & . & \gamma & . & . & . & . & \alpha & . & . \\ . & . & . & . & . & \delta & . & . & . & . & \beta & . \\ . & . & . & . & . & \gamma & . & . & . & . & . & \alpha \\ . & . & . & . & . & . & \alpha & . & . & . & . & . \\ . & . & . & . & . & . & \beta & . & . & . & . & . \\ . & . & . & . & . & . & \alpha & . & . & . & . & . \\ . & . & . & . & . & . & \beta & . & . & . & . & . \\ . & . & . & . & . & . & \alpha & . & . & . & . & . \end{bmatrix}$$

where “.” represents 0 and with,

$$\text{for } \Xi_2, \quad \alpha = 0.1 \quad \beta = 0.9 \quad \gamma = 0.95 \quad \delta = 0.05 \quad (7)$$

$$\text{for } \Xi_3, \quad \alpha = 0.0 \quad \beta = 1.0 \quad \gamma = 0.90 \quad \delta = 0.10 \quad (8)$$

Figure 8. Spectra of Various Sources

Figure 9. Example Channel Frequency Response

REFERENCES

- [1] C. R. Johnson, Jr., S. Dasgupta and W. A. Sethares, "Averaging analysis of local stability of a real constant modulus algorithm adaptive filter," *IEEE Trans. on ASSP*, vol. ASSP-36, pp. 900-910, June 1988.
- [2] C.R. Johnson, Jr., "Admissibility in Blind Adaptive Channel Equalization" *IEEE Control Systems*, vol. 11, pp.3-15, 1991.

Source	Channel	R_0	R_1
#1	[1 0.8 0.64]	1.2504	0.2931
#2	''	0.9505	0.6953
#3	''	1.0423	0.7713
#1	[1 0.3 0.09]	0.8441	-0.1626
#2	''	0.8698	-0.0160
#3	''	0.8824	0.0043

Table 1. 2nd Order Corr. Coef. for Source/Channel Combinations

Source	Channel	R_{000}	R_{001}	R_{011}	R_{111}
#1	[1 0.8 0.64]	4.1262	2.1145	1.9808	1.5664
#2	''	2.1284	1.4460	1.2661	1.2954
#3	''	2.7400	1.9253	1.6674	1.6838
#1	[1 0.3 0.09]	0.9518	0.2733	0.7423	-0.0531
#2	''	1.0176	0.4011	0.6674	-0.1242
#3	''	1.0497	0.4420	0.7156	-0.0648

Table 2. 4th Order Corr. Coef. for Source/Channel Combinations

- [3] D.N. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems", *IEEE Trans. on Commun.*, vol. COM-28, pp. 1867-1875, 1988
- [4] Z. Ding, R.A. Kennedy, "On the Whereabouts of Local Minima for Blind Adaptive Equalizers", *IEEE Trans. on Circuits and Systems-II: Analog and Digital Signal Processing* vol. 39, no. 2, Feb. 1992.
- [5] S. Haykin, *Blind Deconvolution*, Prentice Hall, Englewood Cliffs, NJ, 1994
- [6] A. Benveniste, M. Goursat, "Blind Equalizers", *IEEE Trans. on Comm.*, vol. COM-32, pp. 871-882, 1980