CLOSED FORM EXPRESSION OF EMSE FOR BUSSGANG EQUALIZATION WITH SPATIO-TEMPORAL DIVERSITY

A. Touzni, T. J. Endres, R.A. Casas, C. H. Strolle

atouzni,endres,raul,cstrolle@nxtwavecomm.com
Nxtwave Communications Inc., Langhorne, PA 19047

ABSTRACT

Bussgang algorithms are a class of simple stochastic gradient type solutions for blind channel equalization. In this contribution we investigate the degradation of performance in the source estimation resulting from the stochastic jitter around the stationary points of Bussgang algorithms (which correspond to the equalizers of the channel). More precisely, we derive a closed form approximation of the Excess Mean Square Error (EMSE) defined as the variance of the jitter which depends on the effects of the channel characteristics, the source distribution and the non-linearity used in the update equation. The analysis is performed in a context of spatio-temporal channel diversity involving a single-input/multiple outputs data model.

1. INTRODUCTION

The increasing demand for wireless digital communications services over propagation channels with delay spread much greater than the symbol period places strict performance requirments on today's receivers. Adaptive equalizers usually operate in a blind mode (or non-data aided) and must simultaneously mitigate severe multipath and track time variations. As a result the adaptation process is usually not halted in a stochastic gradient descent implementation and a residual error component known as Excess Mean Square Error (EMSE) is introduced due to stochastic jitter. For long equalizer structures with a large number of adaptive coefficients, the EMSE can be prohibitive.

Our contribution studies the EMSE associated with Bussgang algorithms in the context of equalization with spatio-temporal diversity. The class of Bussgang algorithms includes, for example, Godard's algorithm (or the constant modulus algorithm (CMA)) [7] and Sato's agloritm [13]. Spatio-temporal diversity is characterized either by multiple receivers and/or an over-sampling of the received analog data at a rate greater than the symbol rate [9]. Spatio-temporal diversity with additional assumptions, such as sufficient equalizer length and no channel noise, allows perfect source symbol recovery. We use this property to derive a closed-form expression of the EMSE for Bussgang algorithms. Our result is a generalization of the EMSE for LMS derived by Macchi [11] and the EMSE characterization for CMA given more recently by Fijalkow [5].

The paper is organized as follows. Section 2 describes our system model for the digital communications link, Section 3 the MSE at steady state associated with the adaptation process, and Section 4 provides a closed form estimation for the EMSE of the Bussgang class. Section 5 provides some computer examples for CMA and Sato type algorithm, and Section 6 contains concluding remarks.

This work has been accomplished in part at the ENSEA, France with the collaboration of Prof. I. Fijalkow.

2. SYSTEM MODEL

2.1. Data model

We consider the data model that follows,

$$\begin{aligned} \underline{x}(n) &= [h_1(z) \cdots h_p(z)]^t s(n) \\ &= H \underline{s}(n) \end{aligned} \quad (1)$$

where $\underline{x}(n)$ is a length-N vector of observations of a linear single-input, p-output data model (where $p > 1$), $\underline{s}(n)$ is a length-M vector of source symbols. The degree of the sub-channels $h_k(z)$ is finite. We assume that the coefficients of $h_k(z)$ can be stacked in a matrix H of dimensions $N \times M$ (where $N > M$). The matrix H and the source $\underline{s}(n)$ are unknown and real-valued. The source $\{s(n)\}$ is an i.i.d. sequence of symbols belonging to a known alphabet.

2.2. Bussgang algorithms

The equalizer used is a linear, multichannel, adaptive FIR filter which produces estimates of the delayed symbols $s(n-m)$ given by the output

$$y(n) = \underline{f}(n)^t \underline{x}(n). \quad (2)$$

The equalizer vector follows the update rule,

$$\underline{f}(n+1) = \underline{f}(n) - \mu\, (y(n) - \psi(y(n)))\, \underline{x}(n). \quad (3)$$

Here, μ is a small positive constant known as the stepsize of the algorithm, and $\psi(.)$ is a non-linear, differentiable, odd function [1] which, roughly speaking, can be seen as a soft-decision function of the estimated symbol. The vector $\underline{f}_*$ denotes the point of convergence of the stochastic adaptive Bussgang algorithm.

Most popular blind adaptive equalization algorithms based on a gradient descent optimization technique can be described by the update rule (3). For example, the Constant Modulus Algorithm (CMA) [7], [14] is a special case of a Bussgang algorithm for which

$$\psi(y(n)) = \frac{\mathbb{E}\{s^2\}}{\mathbb{E}\{s^4\}}\, y^3(n) \quad (4)$$

where $\mathbb{E}\{.\}$ is the mathematical symbol for statistical expectation.

Another example is given by the maximum likelihood (ML) estimation of a signal $s(n) = \pm 1, \pm 3, \cdots, \pm(P-1)$ (PAM-P modulation) for which,

$$\psi(y(n)) \simeq \frac{\sum_{k=1}^{P-1} (2k-1)\, e^{-\frac{(2k-1)^2}{2\rho}} \sinh(\frac{(2k-1)}{\rho} y(n))}{\sum_{k=1}^{P-1} e^{-\frac{(2k-1)^2}{2\rho}} \cosh(\frac{(2k-1)}{\rho} y(n))} \quad (5)$$

where ρ is defined by $\rho = \frac{\sqrt{SRI}}{\sigma}$ and $SRI = \frac{E\{s^2\}}{\sigma^2}$ denotes the signal to residual Inter-Symbol Interference (ISI) ratio (σ^2 is the variance of the ISI generated by the signal subspace minus the dominant tap). For an ML estimation of a PAM-2 source, the expression (5) can be simplified to $\psi(y(n)) = E\{s^2\}/E\{|s|\}\times \tanh(\rho\, y(n))$. In particular when $\rho \to +\infty$, i.e., when the ISI goes to zero, we have $\psi(y(n)) \to sign(y(n))\times \mathbb{E}\{s^2\}/\mathbb{E}\{|s|\}$ which is precisely the non-linear function introduced by Sato in [13]. Other examples of Bussgang type algorithms, of a non-exhaustive list, are given in [12], [8].

Finally, the choice $\psi(y(n)) = s(n-m)$ can be seen as a borderline case where roughly speaking the soft decision $\psi(.)$ is substituted by a known transmitted symbol. In this case (3) is the LMS algorithm which is the stochastic gradient descent optimization of the Mean Square Error (MSE) criterion $\mathbb{E}\{(y(n) - s(n-m))^2\}$.

3. MEAN SQUARE ERROR AT STEADY STATE

We investigate the performance of the estimator (3) with the Mean Square Error (MSE) measure defined as follows,

$$\Delta\mathcal{E}(\psi) = \mathbb{E}\left\{ \left(\underline{f}(n)^t \underline{x}(n) - s(n-m)\right)^2 \right\} \tag{6}$$

where we assume that $\underline{f}(n)$ evolves around the stationary point $\underline{f}_*$. This measure characterizes the behavior of the Bussgang algorithm at steady state .

More precisely we are interested in the estimation of the contribution $\overline{\Delta\mathcal{E}}$ of $\Delta\mathcal{E}(\psi)$ that measures the jitter of $\underline{f}(n)$. We rewrite (6) as,

$$\begin{aligned} \Delta\mathcal{E}(\psi) &= \underbrace{\mathbb{E}\left\{ \left(\underline{f}_*^t \underline{x}(n) - s(n-m)\right)^2 \right\}}_{\Delta\mathcal{E}_0} + \\ &\quad \underbrace{\mathbb{E}\left\{ \left(\underline{f}(n) - \underline{f}_*\right)^t \underline{x}(n)\underline{x}(n)^t \left(\underline{f}(n) - \underline{f}_*\right) \right\}}_{\overline{\Delta\mathcal{E}}} \end{aligned} \tag{7}$$

where $\Delta\mathcal{E}_0$ denotes the (static) MSE associated with the solution $\underline{f}_*$ and where $\overline{\Delta\mathcal{E}}$ is defined as the Excess MSE (EMSE) resulting from stochastic jitter.

Note that,

$$\Delta\mathcal{E}_0(\underline{f}_*) \geq \text{MMSE} = \Delta\mathcal{E}_0(\underline{f}_{mmse}) \tag{8}$$

where $\underline{f}_{mmse} = \mathbb{E}\{\underline{x}(n)\underline{x}(n)^t\}^{-1}\, \mathbb{E}\{\underline{x}(n)s(n-m)\}$ is the equalizer minimizing the MSE criterion defined as the Minimum MSE (MMSE) equalizer. Calculation of the error $\Delta\mathcal{E}_0(\underline{f}_*)$ has been analyzed by several authors, mainly for the MSE and CM criteria under different channel and noise conditions (see [6], [16] and [15] for instance).

The second contribution of (7) $\overline{\Delta\mathcal{E}}$ measures the perturbation of the solution $\underline{f}_*$ resulting from the non-vanishing update error term $\mu(y(n) - \psi(y(n)))\underline{x}(n)$ of (3). The stepsize is usually non-zero to facilitate tracking of time varying channels. The EMSE term can be approximated, for small μ, as (see [11])

$$\begin{aligned} \overline{\Delta\mathcal{E}} &= trace\mathbb{E}\left\{ \left(\underline{f}(n) - \underline{f}_*\right) \left(\underline{f}(n) - \underline{f}_*\right)^t \underline{x}(n)\underline{x}(n)^t \right\} \\ &\simeq \mathbb{E}\left\{ \|\underline{f}(n) - \underline{f}_*\|^2 \right\} trace\mathbb{E}\left\{ \underline{x}(n)\underline{x}(n)^t \right\} \end{aligned} \tag{9}$$

Reference [2] shows that the set of vectors $\underline{f}(n)$ given by (3) are asymptotically unbiased estimators of $\underline{f}_*$ with variance,

$$\mathbb{E}\left\{ \|\underline{f}(n) - \underline{f}_*\|^2 \right\} = \mu\, trace(\Sigma^*) \tag{10}$$

where Σ^* is defined as the unique positive solution of the Lyapunov equation defined by,

$$G(\underline{f}_*)\Sigma^* + \Sigma^* G(\underline{f}_*)^t + R(\underline{f}_*) = 0 \tag{11}$$

where $G(\underline{f}_*) = \nabla \underline{f}_* \mathbb{E}\{(y(n) - \psi(y(n)))\}$ is the gradient of the update error term of (3) with respect to $\underline{f}(n)$ taken at the point $\underline{f}_*$. The matrix $R(\underline{f}_*)$ is the covariance of the update error term, also taken at $\underline{f}_*$.

The next section introduces further assumptions on the data model to give an explicit expression for the EMSE.

4. CLOSED FORM EXPRESSION OF EMSE

In this section we give a closed-form expression for the EMSE, $\overline{\Delta\mathcal{E}}$, of Bussgang algorithms as a function of the channels $h_k(z)$ and the non-linearity $\psi(.)$.

Our result relies on the main assumption of left-invertibility of the matrix H in the data model (1). Since the matrix H is taller than it is wide ($N > M$), the condition of left-invertibility is guaranteed if the subchannels $h_k(z)$ do not share common roots and if the equalizer provides sufficient length.

Under these assumptions the EMSE in (7) can be written as,

$$\overline{\Delta\mathcal{E}} = \mathbb{E}\{s^2\}\, \mathbb{E}\left\{ \|\underline{q}(n) - \underline{e}_{m+1}\|^2 \right\} \tag{12}$$

where $\underline{q}(n)$ denotes the combined impulse response of the channel and equalizer defined as $\underline{q}(n) = H^t \underline{f}(n)$ and where $\underline{e}_{m+1}$ denotes a single spike-vector where the unity value is located at the $(m+1)$-th position (*i.e.*, $\underline{f}_*^t H \underline{s}(n) = \underline{e}_{m+1}^t \underline{s}(n) = s(n-m)$). Since $\Delta\mathcal{E}_0$ goes to zero in this case, the EMSE equals the variance of the source estimation (*i.e.* the MSE),

$$\overline{\Delta\mathcal{E}} = \mathbb{E}\left\{ (\hat{s}(n) - s(n-m))^2 \right\} \tag{13}$$

where $\hat{s}(n) = \underline{q}(n)^t \underline{s}(n)$. The EMSE expression in (13) is found by solving the equation (11), which for the Bussgang class can be written as,

$$HG(\underline{e}_{m+1})H^t\Sigma^* + \Sigma^* HG(\underline{e}_{m+1})^t H^t - HR(\underline{e}_{m+1})H^t = 0 \tag{14}$$

where the matrices $G(\underline{e}_{m+1})$ and $R(\underline{e}_{m+1})$ are respectively given by,

$$G(\underline{e}_{m+1}) = \mathbb{E}\left\{ \psi'(\underline{e}_{m+1}^t \underline{s}(n))\underline{s}(n)\underline{s}(n)^t \right\} - \mathbb{E}\{s^2\}\, I_M \tag{15}$$

and,

$$\begin{aligned} R(\underline{e}_{m+1}) &= \sum_{n=-\infty}^{+\infty} \mathbb{E}\left\{ \left(\underline{e}_{m+1}^t \underline{s}(n) - \psi(\underline{e}_{m+1}^t \underline{s}(n))\right) \underline{s}(n) \right. \\ &\quad \left. \underline{s}(0)^t \left(\underline{e}_{m+1}^t \underline{s}(0) - \psi(\underline{e}_{m+1}^t \underline{s}(0))\right) \right\}. \end{aligned} \tag{16}$$

where I_M is the $M \times M$ identity matrix.

The expression of the EMSE term is summarized in the proposition that follows. The proof is given in the Appendix.

Proposition 1. *Under the assumptions:* i*) the matrix H is left-invertible, and* ii*) the stepsize μ is small enough, a closed form expression of the EMSE for Bussgang algorithms in (3) is given by,*

$$\overline{\Delta\mathcal{E}}_{Buss} \simeq \frac{\mathbb{E}\left\{(s-\psi(s))^2\right\}}{\mathbb{E}\left\{\psi'(s)\right\}-1}\,\Delta\mathcal{E}_{LMS} \qquad (17)$$

where,

$$\Delta\mathcal{E}_{LMS} = \mu\,\frac{N}{2}\mathbb{E}\{s^2\}\sum_{k=1}^{p}\int_{-\pi}^{+\pi}|h_k(e^{i\omega})|^2\,d\omega \qquad (18)$$

is the EMSE of the LMS algorithm [11].

Note that the EMSE of the LMS algorithm $\Delta\mathcal{E}_{LMS}$ is defined as the product of three terms: the step size μ, the length of the equalizer N and the power of the received signal expressed as a function of the channels $h_k(z)$. Notice that the EMSE increases linearly with the length of the equalizer and the step-size. The effect of the non-linearity $\psi(.)$ in (17) is contained in the ratio $\mathbb{E}\left\{(s-\psi(s))^2\right\}/(\mathbb{E}\{\psi'(s)\}-1)$. The gain induced by this ratio depends on the non-linearity and the source distribution. The numerator stems directly from the contribution of the non vanishing error $(y(n)-\psi(y(n))$ of the algorithm (3). The contribution of the denominator depends on the derivative of the non-linear function $\psi(.)$ (recall that it is assumed that $\psi(.)$ is differentiable).

5. EXAMPLES

Next, we use the result of Proposition 1 to calculate the EMSE for two important Bussgang algorithms: the CMA and a Sato-type algorithm. We first study the specific form of (17) for each of these algorithms. A numerical validation of the EMSE formula (17) is then given.

CM criterion [7]: $\psi(s) = \frac{\mathbb{E}\{s^2\}}{\mathbb{E}\{s^4\}}\,s^3$.

$$\frac{\mathbb{E}\left\{(s-\psi(s))^2\right\}}{\mathbb{E}\left\{\psi'(s)\right\}-1} = \frac{\mathbb{E}\{s^2\}}{\mathbb{E}\{s^4\}}\,\frac{\left(\frac{\mathbb{E}\{s^6\}}{\mathbb{E}\{s^2\}^3}-\kappa_s^2\right)}{3-\kappa_s}\,\mathbb{E}\{s^2\}^2 \qquad (19)$$

where $\kappa_s = \frac{\mathbb{E}\{s^4\}}{\mathbb{E}\{s^2\}^2}$ denotes the kurtosis of the source. The contribution of the EMSE in formula (19) above matches the EMSE result for CMA in [5]. According to the update equation (3), the relation between the step size μ_{cm} used in [5] and μ, is $\mu_{cm} = \mu\frac{\mathbb{E}\{s^4\}}{\mathbb{E}\{s^2\}}$. As mentioned in [5] the EMSE increases (goes toward infinity) when the source is close to a Gaussian distribution (*i.e.,* when $\kappa_s \to 3$). Note that $3-\kappa_s > 0$ for a sub-gaussian source. The sub-gaussianity of the source is indeed a necessary condition to guarantee a stable equalizer setting for CMA [4].

SATO type criterion [13]: $\psi(s) = \frac{\mathbb{E}\{s^2\}}{\mathbb{E}\{|s|\}}\,\tanh(\alpha s)$.

$$\frac{\mathbb{E}\left\{(s-\psi(s))^2\right\}}{\mathbb{E}\left\{\psi'(s)\right\}-1} = \frac{\mathbb{E}\left\{(s-\frac{\mathbb{E}\{s^2\}}{\mathbb{E}\{|s|\}}tanh(\alpha s))^2\right\}}{\alpha\frac{\mathbb{E}\{s^2\}}{\mathbb{E}\{|s|\}}(1-\mathbb{E}\{tanh(\alpha s)^2\})-1}. \qquad (20)$$

When α is big enough we have $|\mathbb{E}\{\psi'(s)\}-1| \to 1$. For the Sato algorithm the Gaussianity of the source is not a determining factor in the enhancement of the EMSE as it is for the CMA.

Next we compare the EMSE furnished in (17) and the experimental EMSE $\hat{\overline{\Delta\mathcal{E}}}_{Buss} = \frac{1}{T}\sum_{n=1}^{T}(\hat{s}(n)-s(n-m))^2$, with

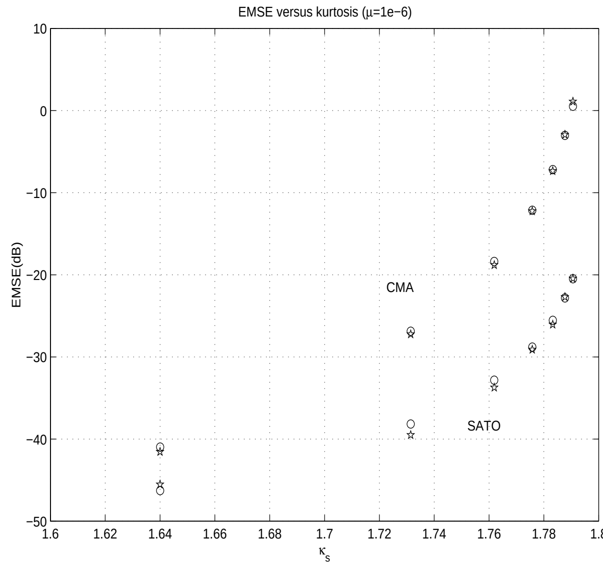

Fig. 1. EMSE of CMA ($\psi(s) = \frac{\mathbb{E}\{s^2\}}{\mathbb{E}\{s^4\}}\,s^3$) and Sato-type algorithm ($\psi(s) = \frac{\mathbb{E}\{s^2\}}{\mathbb{E}\{|s|\}}\,\tanh(\alpha s)$ with $\alpha = 5$) versus source kurtosis.

$T = 5\times 10^5$, computed at the steady-state solution of the CMA and Sato type algorithm. Both simulations use the same (arbitrary) channel defined by $h_1(z) = 1+0.3z^{-1}-0.2z^{-2}+0.6z^{-3}$, $h_2(z) = -0.2+0.7z^{-1}+0.7z^{-2}+0.1z^{-3}$. Figure 1 plots the EMSE for different source kurtoses κ_s corresponding to a PAM-P modulation where P goes from 4 to 16. The theoretical EMSE $\overline{\Delta\mathcal{E}}_{Buss}$ is plotted with o's and the experimental EMSE $\hat{\overline{\Delta\mathcal{E}}}_{Buss}$ is plotted with $\star$'s. We can see that for a small stepsize (here $\mu = 1e^{-6}$) the theoretical and experimental results match well.

6. CONCLUSION

In this contribution we provide a closed form expression for the Excess Mean Square Error (EMSE) of Bussgang algorithms under a channel invertibility condition. Under the assumption of perfect channel inversion and no noise, we show that the EMSE for Bussgang type algorithms is proportional to the EMSE of the LMS algorithm and generalizes results in the open literature. Our result can be used to provide a better approximation of the MSE for blind equalization algorithms and can therefore be used to derive practical design guidelines. For example, for a given source and class of channels, our EMSE expression can be used to better select the equalizer length and stepsize. Our analysis can also be applied to quantized versions of Bussgang algorithms, for example like [3] which aims to reduce complexity for implementation. An extended analysis in the case of noisy and/or non invertible channel conditions is in preparation.

APPENDIX

Next we prove result (17) in Proposition 1. To solve the Lyapunov equation (14) we need to compute the matrices $G(\underline{e}_{m+1})$ and $R(\underline{e}_{m+1})$.

Estimation of $G(\underline{e}_{m+1})$:

The Hessian $G(\underline{e}_{m+1})$ is a diagonal matrix for which the diagonal entries are given by,

$$\left(G(\underline{e}_{m+1})\right)_{i,i} = \begin{cases} \mathbb{E}\{\psi^{'}(s)s^2\} - \mathbb{E}\{s^2\} & \text{for } i = m \\ \mathbb{E}\{\psi^{'}(s)\}\mathbb{E}\{s^2\} - \mathbb{E}\{s^2\} & \text{elsewhere} \end{cases} \tag{21}$$

for $1 \leq i \leq N$. A compact expression is,

$$G(\underline{e}_{m+1}) = (\mathbb{E}\{\psi^{'}(s)\}\mathbb{E}\{s^2\} - \mathbb{E}\{s^2\})\, I_N + \epsilon_1\, \underline{e}_{m+1}\underline{e}_{m+1}^t \tag{22}$$

where $\epsilon_1 = \mathbb{E}\{\psi^{'}(s)s^2\} - \mathbb{E}\{\psi^{'}(s)\}\mathbb{E}\{s^2\}$.

Estimation of $R(\underline{e}_{m+1})$:

We evaluate first the contributions of the (i,j)-th entries of the expectation of the formula (16). We get,

$$\begin{aligned}
&\Big(\mathbb{E}\big\{\left(\underline{e}_{m+1}^t\underline{s}(n) - \psi(\underline{e}_{m+1}^t\underline{s}(n))\right)\underline{s}(n) \\
&\quad \underline{s}(0)^t\left(\underline{e}_{m+1}^t\underline{s}(0) - \psi(\underline{e}_{m+1}^t\underline{s}(0))\right)^t\big\}\Big)_{i,j} \\
&= \mathbb{E}\{s^2\}^2(\delta(i-m)\delta(j-m) + \delta(n)\delta(n+j-i) \\
&+\delta(n-m+j)\delta(n+m-i)) + \mathbb{E}\{s^4\}\delta(n)\delta(m-j) \\
&\delta(m-i) + \mathbb{E}\{\psi(s)^2\}\mathbb{E}\{s^2\}(\delta(n)\delta(n+j-i) + \\
&\delta(n-m+j)\delta(n+m-i)) + E[s^4]\delta(n)\delta(m-j)\delta(m-i) \\
&+\mathbb{E}\{\psi(s)s\}\mathbb{E}\{s^2\}((\delta(n)\delta(n+j-i) + \delta(m-i)\delta(m-j) \\
&+\delta(n-m+j)\delta(n-i+m)) + \mathbb{E}\{\psi(s)s^3\}\delta(n)\delta(m-j) \\
&\delta(j-i) + \mathbb{E}\{\psi(s)s\}\mathbb{E}\{s^2\}(\delta(n)\delta(n-i+j) + \\
&\delta(n+m-i)\delta(n-m+j) + \delta(m-i)\delta(m-j)) \\
&+\mathbb{E}\{\psi(s)s^3\}\delta(n)\delta(i-j)\delta(m-i).
\end{aligned} \tag{23}$$

Summation of the terms above leads to the expression that follows for the expression of the entries (i,j)-th of the covariance matrix $R(\underline{e}_{m+1})$,

$$\begin{aligned}
&R\left(\underline{e}_{m+1}\right)\big)_{i,j} = \\
&\sum_{n\in\mathbb{Z}} \big\{\left(\mathbb{E}\{s^2\}^2 + \mathbb{E}\{\psi(s)^2\}\mathbb{E}\{s^2\} - 2\mathbb{E}\{\psi(s)s\}\mathbb{E}\{s^2\}\right) \\
&\delta(n)\delta(n+j-i) + \left(\mathbb{E}\{s^4\} + \mathbb{E}\{\psi(s)^2s^2\} - 2\mathbb{E}\{\psi(s)s^3\}\right) \\
&\delta(n)\delta(m-j)\delta(m-i)\big\}
\end{aligned} \tag{24}$$

for which the only non-zero contributions are given at $n = 0$. The matrix $R(\underline{e}_{m+1})$ is therefore a diagonal matrix of the form,

$$\left(R(\underline{e}_{m+1})\right)_{i,i} = \begin{cases} \mathbb{E}\{s^2(s-\psi(s))^2\} & \text{for } i = m \\ \mathbb{E}\{s^2\}\mathbb{E}\{(s-\psi(s))^2\} & \text{elsewhere} \end{cases} \tag{25}$$

A compact expression, is given by,

$$R(\underline{e}_{m+1}) = \mathbb{E}\{s^2\}\mathbb{E}\{(s-\psi(s))^2\}\, I_N + \epsilon_2\, \underline{e}_{m+1}\underline{e}_{m+1}^t \tag{26}$$

where $\epsilon_2 = \mathbb{E}\{s^2(s-\psi(s))^2\} - \mathbb{E}\{s^2\}\mathbb{E}\{(s-\psi(s))^2\}$.

Solution of Lyapunov equation:

The Lyapunov equation (14) is given by,

$$HH^tG(\underline{e}_{m+1})\Sigma^* + \Sigma^*G(\underline{e}_{m+1})^tH^tH - HR(\underline{e}_{m+1})H^t = 0. \tag{27}$$

Under the assumption $|\epsilon_1| < \epsilon$ and $|\epsilon_2| < \epsilon$ where ϵ is small enough, we get,

$$HH^t\Sigma^* + \Sigma^*HH^t - \frac{\mathbb{E}\{s^2\}\mathbb{E}\{(s-\psi(s))^2\}}{\mathbb{E}\{\psi'(s)\}\mathbb{E}\{s^2\} - \mathbb{E}\{s^2\}}HH^t = 0 \tag{28}$$

which admits an obvious (unique) solution given by,

$$\Sigma^* = \frac{1}{2}\,\frac{\mathbb{E}\left\{(s-\psi(s))^2\right\}}{\mathbb{E}\left\{\psi'(s)\right\} - 1}HH^t \tag{29}$$

thus,

$$\begin{aligned}
\overline{\Delta\mathcal{E}} &\simeq \mu\,\mathbb{E}\{s^2\}\,trace(\Sigma^*) \\
&\simeq \frac{\mu}{2}\frac{\mathbb{E}\left\{(s-\psi(s))^2\right\}}{\mathbb{E}\left\{\psi'(s)\right\} - 1}\mathbb{E}\{s^2\}\,trace(HH^t).
\end{aligned}$$

Finally, to conclude the proof, we notice that $trace(HH^t)$ can be expressed as a function of the sub-channels $h_k(z)$ as follows,

$$trace(HH^t) = N\mathbb{E}\{s^2\}\sum_{k=1}^{p}\int_{-\pi}^{+\pi}|h_k(e^{i\omega})|^2\,d\omega. \tag{30}$$

7. REFERENCES

[1] S. Bellini, "Bussgang techniques for blind deconvolution and equalization", in *Blind Deconvolution* (S. Haykin, ed.), pp. 8-52, Englewood Cliffs, New Jersey : Prentice Hall, 1994.

[2] A.Beneveniste, M. Métivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*, Springer-Verlag, 1990.

[3] T.J. Endres, S.N. Hulyalkar, C.H. Strolle and T.A. Schaffer, "Low complexity and low latency implementation of the Godard/ CMA update," to appear *IEEE Transactions on Communications*, February 2001.

[4] I. Fijalkow, F. Lopez de Victoria, and C.R. Johnson Jr., "Adaptive fractionally spaced blind equalization", *Proc. IEEE DSP-94 Workshop on Digital Signal Processing*, pp. 257-260, Yosemite, CA, October 1994.

[5] I. Fijalkow, C. E. Manlove, C. R. Johnson, Jr., "Adaptive fractionally spaced blind CMA equalizaton: Excess MSE," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 56-66, Jan. 1998.

[6] I. Fijalkow, A. Touzni, J.R. Treichler, "Fractionally-spaced equalization using CMA: robustness to channel noise and lack of disparity", *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 56-66, January 1997.

[7] D.N. Godard, "Self-recovering equalization and carrier tracking in two-dimentional data communiation systems," *IEEE Transactions on Communications*, vol. 28, pp. 1867-1875, Nov. 1980.

[8] R. Godfrey and F. Rocca, "Zero memory non-linear deconvolution", *Geophysical Prospecting*, vol. 29, pp. 189-228, 1981

[9] H. Liu, G. Xu, L. Tong, T. Kailath, "Recent developments in blind channel equalization: from cyclostationarity to subspaces," *Signal Processing*, vol. 50, pp. 83-99, 1996.

[10] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Transactions on Automatic Control*, vol. 22, no. 4, 1977

[11] O. Macchi, *Advances in adaptive filtering, in Digital Communications*, Elsevier Publishing, North-Holland, 1986.

[12] G. Picchi and G. Prati, "Blind equalization and carrier recovery using stop-and-go decsision directed algorithm", *IEEE Transactions on Communications*, vol. 35, pp. 877-887, Sept. 1987

[13] Y. Sato, "A method of self recovering equalization for multilevel amplitude modulation system," *IEEE Transactions on Communications*, vol. 23, pp. 679-682, 1975.

[14] J.R. Treichler, and B.G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 2, pp. 459-472, April 1983.

[15] P. Schniter and C.R. Johnson, Jr., "Bounds for the MSE performance of constant modulus estimators," *IEEE Transactions on Information Theory*, November 2000.

[16] H. H. Zeng, L. Tong, and C.R. Johnson, "Relasionships between the constant modulus and Wiener receivers," *IEEE Transactions on Information Theory*, vol. 44, pp. 1523-1538, July 1998.