Fractionally-Spaced Equalization for Time-Varying Channels

M.-L. Alberi, I. Fijalkow
Equipe de Traitement des Images et du Signal, URA CNRS 2235
ETIS / ENSEA, 6 av. du Ponceau, 95014 Cergy-Pontoise Cedex, France
Tel : 01 30 73 66 10, Fax : 01 30 73 66 27
alberi@ensea.fr,fijalkow@ensea.fr

ABSTRACT

The good convergence tracking properties of spatio-temporal equalizers are pointed out and analyzed when there is effective diversity. The analysis is illustrated in the case of a frequency offset between the baud and sampling clocks that induces important time-variations.

1 Introduction

Equalization is a crucial point of the digital receiver when the propagation is affected by intersymbol interference. It is usually studied assuming the channel is invariant during the processing. However, this assumption is not always met for high-rate transmissions, and to our best knowledge no analysis is available in this case. Abrupt changes occur in radio-mobile communications when the mobile enters into a new street, the equalizer should at this point start a new convergence procedure. Channel drift may be due to Doppler effect or to a frequency offset between the transmitter baud clock and sampling clock, see [3]. Good equalizer tracking abilities are crucial for both abrupt changes and drifts in the channel transfer function.

Simulations of transverse equalization show an increased convergence speed when using spatial and/or temporal diversity, i.e., either an array of sensors or a sampling rate greater than the transmitting baud rate. Such behavior has been reported for gradient descent algorithms such as the Least Mean Square (LMS) [2], as well as for the Constant Modulus Algorithm, [5]. However, this may sound contradictory with the results in [10], where possible deterioration of the convergence speed is noticed when using temporal diversity. Therefore, we intend to give an analytical explanation of these behaviors based on the observation that improved or deteriorated performances seem to be related to the channel disparity, i.e., the effective diversity, see [6].

We know that spatio-temporal diversity sometimes allows improved mean asymptotic performances using a finite impulse response transverse equalizer. In order to address the effect of spatio-temporal diversity on the convergence speed and tracking capabilities, we study the input / output mean square error during the transient convergence to the mean asymptotic setting. For the study to be fair, we need also to look at the behavior when asymptotic convergence in mean is achieved, following the results of [7] in the conventional case.

2 Spatio-Temporal Equalization

2.1 Spatio-Temporal model

The spatial or temporal diversity of a factor ($L > 1$) is modelized by a single input / multiple outputs system ([8]).

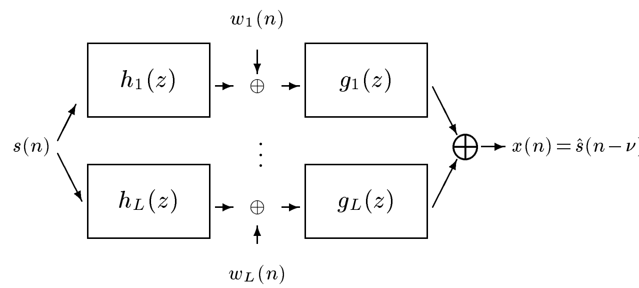


Figure 1: Noisy Multi-Channel / Equalizer scheme

The signal received by the spatial-temporal equalizer is composed by a L-dimensional observation :

$$\mathbf{y}(n) = \sum_{k=0}^{Q} \mathbf{h}_k s(n-k) + \mathbf{w}(n)$$

where $s(n)$ is the input scalar i.i.d. sequence with variance unity and $\mathbf{h}(z) = \sum_{k=0}^{Q} \mathbf{h}_k z^{-k}$ represents the multichannel transfer function (a degree Q polynomial vector of length L). We denote σ^2 the variance of the spatially and temporally white noise, $\mathbf{w}(n)$.

The transverse equalizer with transfer function (a degree $N-1$ polynomial vector of length L) $\mathbf{g}(z) = \sum_{k=0}^{N-1} \mathbf{g}_k z^{-k}$ is meant to restore a good estimation of the input source with a delay ν:

$$x(n) = \sum_{k=0}^{N-1} \mathbf{g}_k^{\top} \mathbf{y}(n-k) = \hat{s}(n-\nu)$$

Under **channel invertability conditions** ([8]) (i.e., no common root to all the components of $\mathbf{h}(z)$ and $N \geq Q$) perfect equalization can be achieved, with finite impulse response (FIR) equalizer, in the absence of noise. It implies that when there is no noise, the equalizer output, $x(n)$, can match exactly the delayed input $s(n-\nu)$. However in the noisy case, an infinite impulse response is required. Still, equalization with a FIR filter can achieve good performances in terms of input / output Minimum Mean Square Error (MMSE) when the amount of noise is not too high, [6].

2.2 Performances Measure

Equalization performances are measured by the input / output Mean Square Error (MSE),

$$MSE(n) = E[|x(n) - s(n-\nu)|^2]$$

Spatio-temporal equalization is analyzed using the reference adaptive algorithm, Least Mean Square algorithm (LMS),

$$\mathbf{g}(n+1) = \mathbf{g}(n) - \mu(x(n) - s(n-\nu))\mathbf{Y}(n)^*, \; x(n) = \mathbf{g}(n)^\top \mathbf{Y}(n)$$

where $\mathbf{g} = (\mathbf{g}_0^\top, \cdots, \mathbf{g}_{N-1}^\top)^\top$ is the equalizer impulse response and $\mathbf{Y}(n) = (\mathbf{y}(n)^\top, \cdots, \mathbf{y}(n-N+1)^\top)^\top$ is the regression vector. MSE is characterized in the sequel for both transient and asymptotical contexts.

3 Asymptotic performances

When using an adaptive algorithm, asymptotic convergence is achieved in mean. We denote $\mathbf{g}_{opt}$ this mean value. Using LMS, $\mathbf{g}_{opt}$ corresponds to achieving MMSE. Still, the equalizer taps jitter around their mean values, inducing some Excess Mean Square Error (EMSE). The total Mean Square Error (MSE),

$$MSE = MMSE + EMSE$$

characterizes the asymptotic, or steady-state, performances.

3.1 MMSE

For small noise level, an expression of MMSE is given in [6]. In the case of spatio-temporal effective diversity,

$$MMSE = \sigma^2 \delta_\nu^\top (\mathcal{T}(\mathbf{h})^* \mathcal{T}(\mathbf{h})^\top)^{-1} \delta_\nu$$

where $\mathcal{T}(\mathbf{h})$ is the channel convolution matrix and δ_ν has all but its ν^{th} component are equal to 0. In the conventional case, the MMSE (denoted $MMSE_0$) is the sum of an incompressible term due to channel non-invertability, and of a term proportional to the noise power σ^2. $MMSE$ is always lower than $MMSE_0$.

3.2 EMSE

In [7], the EMSE expression is approximated for a stationary channel,

$$EMSE = MMSE \, \frac{\mu N P_{\mathbf{y}}}{2}$$

with $P_{\mathbf{y}} = E[\mathbf{y}(n)^\top \mathbf{y}(n)^*]$, the received power and μ is the algorithm step-size (μ is small). The derivation makes this calculus valid whether there is spatio-temporal diversity or not, so that we can use it in the sequel. Moreover, we assume for the analysis that the power received on each subchannel is the same, so that $P_{\mathbf{y}} = L P_{y_0}$ where P_{y_0} is the received power on each subchannel.

3.3 Comparing spatio-temporal and conventional equalizers

Comparisons between adaptive algorithms often consider a similar EMSE value (by tuning accordingly the step-sizes), [4]. The convergence speeds are then compared. In order to make a fair comparison between spatio-temporal diversity equalization ($L \geq 2$) and the conventional case ($L = 1$), we consider in the conventional case a $N_0 = NL$-long equalizer. This would imply in our case, $\mu MMSE = \mu_0 MMSE_0$ where μ_0 is the conventional algorithm step-size. Therefore, when diversity allows a much smaller MMSE ($MMSE << MMSE_0$), μ_0 should be much smaller than μ, making the comparison quite unfair. Moreover, the improvement of the asymptotic performances doesn't ensure of good tracking performances. Therefore, we will sometimes consider equal step-sizes, and sometimes equal MMSE, in the sequel.

4 Transient behavior with diversity

In order to study the algorithm convergence, we will analyze the evolution of the deviation $V(n) = \mathbf{g}(n) - \mathbf{g}_{opt}$, defined as the difference between the equalizer at the instant n and the optimal equalizer. We assume, for sake of easy calculus, that the observation vector $\mathbf{y}(n)$ is zero-mean and temporally white even if this assumption doesn't reflect the reality. We suppose also that $|\mathbf{Y}(n)|^2 = \mathrm{E}\left[|\mathbf{Y}(n)|^2\right]$.

4.1 Mean and Mean Square deviation

The initial and the final convergence of the deviation is respectively characterized by the mean deviation and by the mean square deviation. However the analysis of the convergence of this two terms is made by an equivalent and indirectly manner in using the transformation, $Q(n) = \Gamma^\top V(n)$. Γ is an orthogonal matrix corresponding to the eigendecomposition of the observation covariance matrix $\mathcal{R}$ which is expressed as:

$$\mathcal{R} = E[\mathbf{Y}(\mathbf{n})^* \mathbf{Y}(\mathbf{n})^\top] = \mathcal{T}(\mathbf{h})^* \mathcal{T}(\mathbf{h})^\top + \sigma^2 \mathbf{I}$$

The eigen-decomposition of $\mathcal{R}$ is such as (see [8]):

$$\Gamma^\top \mathcal{R} \Gamma = diag(\lambda_1, \ldots, \lambda_{N+Q}, \underbrace{\sigma^2, \ldots, \sigma^2}_{NL-(N+Q)})$$

where the $(N+Q)$ largest eigenvalues $\lambda_1, \ldots, \lambda_{N+Q}$ are associated to the signal subspace ξ_S and the $(NL-(N+Q))$ remaining eigenvalues, equal to σ^2, correspond to its orthogonal complement, the noise subspace ξ_W.

4.2 Influent eigenvalues

We can establish an exact evaluation of the quantities $\mathrm{E}[Q(n)]$ and $\mathrm{E}\left[|Q(n)|^2\right]$ according to their values at the instant $(n-1)$, following [9] and [7]. $\mathrm{E}[Q(n)] =$

$$diag(\underbrace{\ldots, (1-\mu\lambda_l), \ldots}_{N+Q}, \underbrace{\ldots, (1-\mu\sigma^2), \ldots}_{NL-(N+Q)}) \mathrm{E}[Q(n-1)] \quad (1)$$

The variance of each component, $\mathrm{E}\left[|q_i(n)|^2\right] = ()$

$$\begin{aligned} (1-\mu(2-\mu N P_{\mathbf{y}})\lambda_i)\mathrm{E}\left[|q_i(n-1)|^2\right], i &= 1, ..., N+Q, \\ (1-\mu(2-\mu N P_{\mathbf{y}})\sigma^2)\mathrm{E}\left[|q_i(n-1)|^2\right], i &= N+Q+1, ..., NL \end{aligned}$$

First, we consider that the coefficients of the equalizer have been initialized with $g(0) = (0 \ldots 0)^\top$. So that

$$Q(0) = -\Gamma^\top g_{opt} = -[a_1 \ldots a_{NL}] \quad (2)$$

In absence of noise ($\sigma^2 = 0$), the spatio-temporal equalizer g_{opt} inverts strictly the channel. It is called Zero-Forcing equalizer, is orthogonal to the subspace noise ξ_W and belongs to the signal subspace ξ_S. $\Gamma^\top g_{opt} = [a_1 \dots a_{N+Q} 0 \dots 0]$. (1) and (4.2) become :

$$\mathrm{E}\left[Q(n)\right] = [\underbrace{\dots,(1-\mu\lambda_l)^n \mathrm{E}\left[w_l(0)\right],\dots,}_{N+Q} \underbrace{\dots,0,\dots}_{NL-(N+Q)}]^\top \quad (3)$$

and

$$\mathrm{E}\left[|Q(n)|^2\right] = \sum_{i=1}^{N+Q} [1-\mu(2-\mu N P_{\mathbf{y}})\lambda_i]\mathrm{E}\left[|q_i(n-1)|^2\right] \quad (4)$$

The $(NL-(N+Q))$ last terms of $\mathrm{E}\left[Q(n)\right]$ and $\mathrm{E}\left[|Q(n)|^2\right]$, associated with $\sigma^2 = 0$, are zero. Only the eigenvalues linked to ξ_S contributes to the convergence of $\mathrm{E}\left[Q(n)\right]$ and $\mathrm{E}\left[|Q(n)|^2\right]$.

In presence of a small amount of noise (σ^2), the equalizer achieves a trade-off between the subspace ξ_S and ξ_W, but is still very close to the noise-free expression, [6]. Therefore, the contribution of ξ_S in the expression of g_{opt} is most important than that of ξ_W. The $(NL-(N+Q))$ last terms of $\mathrm{E}\left[Q(n)\right]$ and $\mathrm{E}\left[|Q(n)|^2\right]$, associated with σ^2, aren't anymore zero but their evolution remain negligible with respect to the $(N+Q)$ first taps.

When $g(0) \neq [0\dots0]^\top$, the $NL-(N+Q)$ last terms of $\mathrm{E}\left[Q(n)\right]$ and $\mathrm{E}\left[|Q(n)|^2\right]$, associated with σ^2, aren't zero. Also the smallest eigenvalues have detrimental influence in the transient convergence of the algorithm. Since we would like to have the non-influent taps equal to 0, an initialization to the $NL-(N+Q)$ last coefficients of the equalizer at 0 is the most sensible thing to do. This setting is assumed in the sequel.

4.3 Importance of disparity

We have just seen that the transient behavior of the LMS depends mostly on the eigenvalues in the signal subspace. So that a measure of convergence rate (similar to that in [7]) could be given by the effective condition number defined as,

$$\rho_{eff} = \frac{\lambda_{max(Signal)}}{\lambda_{min(Signal)}}$$

$\lambda_{max(Signal)}$ and $\lambda_{min(Signal)}$ are respectively the maximum and the minimum values of signal subspace eigenvalues, which are also the eigenvalues of $\mathcal{T}(\mathbf{h})^\top\mathcal{T}(\mathbf{h})^*$. ρ_{eff} is therefore the condition number of $\mathcal{T}(\mathbf{h})^\top\mathcal{T}(\mathbf{h})^*$,

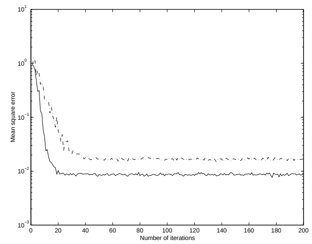
$$\rho_{eff} = \|\mathcal{T}(\mathbf{h})^\top\mathcal{T}(\mathbf{h})^*\|\|(\mathcal{T}(\mathbf{h})^\top\mathcal{T}(\mathbf{h})^*)^{-1}\|$$

ρ_{eff} is also inversely proportional to the determinant of $\mathcal{T}(\mathbf{h})^\top\mathcal{T}(\mathbf{h})^*$ ([6]) given by,

$$det(\mathcal{T}(\mathbf{h})^\top\mathcal{T}(\mathbf{h})^*) = \sum_{k<l} K_{k,l} \prod_{i,j} |z_k^i - z_l^j|^2$$

where $K_{k,l}$ is some polynomial bounded function of the subchannels k and l and $(z_k^i)_{i=1,..Q}$ denotes the set of roots of subchannel k.

This implies that ρ_{eff} should be a good way to measure the effect of channel disparity on transient behavior as well as on asymptotic behavior. An example is provided in 2 to show how LMS performances deteriorate when there is a lack of channel disparity. Equalization was performed for a SNR=20 dB, $N = Q + 3$ and an average of 100 realizations of LMS with step-size equal to 0.055. Channel ($\mathbf{h}_1$), which owns disparity, has $\rho_{eff} = 3.08$ and channel ($\mathbf{h}_2$), which lacks disparity, has $\rho_{eff} = 233.9$. As expected, the speed of convergence and the asymptotic performances are worse when ρ_{eff} is larger.



channel $\mathbf{h}_1$ in full line - channel $\mathbf{h}_2$ in dotted line

Figure 2: Equalization error versus iterations

5 Comparison

We have just shown that spatio-temporal equalizers have good transient convergence performances as well as asymptotic performances when there is effective channel diversity. In the case of temporal diversity, as considered in [10], there is often lack of disparity. The remaining questions are: "does the spatio-temporal equalizer converge faster than the conventional equalizer ? Or does it deteriorates when there is lack of disparity ?"

We will next compare the evolution of the MSE versus the number of iterations in both cases, in order to answer this question because more precise results require (study of the signal subspaces eigenvalues isn't suffisant and difficult analytical). The case of spatio-temporal effective diversity (a) is modelized by a channel $\mathbf{h}(z)$ whose L subchannels have no common zeros. The conventional case (b), modelized by $h(z)$ is equivalent to a spatio-temporal channel which subchannels are all identical. To do so, we used results given in [9] and [7], in order to extend the expression of the conventional $MSE(n)$ to the case of spatio-temporal diversity.

The $MSE(n)$ can be generalized to both cases as:
$MSE(n) = (1-\mu NLP_{\mathbf{y}})^2 \; MSE(n-1) + 2\mu NLP_{\mathbf{y}} \; MMSE$

We note that the exponentially decreasing term $(1 - \mu NLP_{\mathbf{y}})$ has the same value in the case (a) and (b), only the MMSE changes.

5.1 Equal step-size

The convergence speeds are first compared with an identical step-size, i.e ($\mu N P_{\mathbf{y}} = \mu_0 N_0 P_{\mathbf{y_0}}$).

It means that the convergence speed of the spatio-temporal and conventional equalizers should be similar as a first approximation.

The curves in figure 3 confirm the approximation. The simulations is achieved with L=2 for SNR=30dB. The common zeros between the case (a) and the case (b) are $(0.35, -0.14, 0.02)$ and $(3.35, -3.54, -15)$ are the zeros of the second subchannel in the case of diversity.

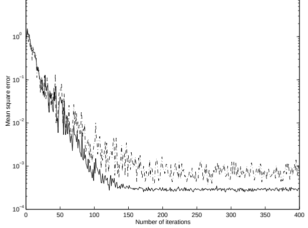


Figure 3: MSE versus iterations: with diversity (in full line), conventional (in dotted line)

We conclude that with equal step-size, spatio-temporal and conventional equalizers converge approximately with the same speed, whether there is effective diversity or not.

5.2 Equal EMSE

Equal EMSE implies in our case, $\mu MMSE = \mu_0 MMSE_0$. Therefore, when diversity allows a much smaller MMSE ($MMSE << MMSE_0$), μ_0 should be much smaller than μ. This should induce an decrease of the convergence speed of the conventional equalizer.

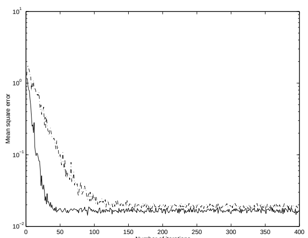


Figure 4: MSE versus iterations: with diversity (in full line), conventional (in dotted line)

At equal asymptotic performances, spatio-temporal equalizer converge faster than the conventional's one equalizer.

6 Case of time-varying channel

Finally, we are going to illustrate the tracking properties of the equalizer with temporal diversity and the conventional equalizer in the case of a time-varying channel. Good convergence properties should result in good tracking capabilities. Tracking is a trade-off between convergence speed, μ large enough to be able to to update the correct inversion at each instant ([2], and a small EMSE, small μ ([7]). The channel variation is induced by a drift due to a frequency offset in the sampling clock: the transmitter rate is $1/T$ and the receiver sampling rate $2/(T+\epsilon)$. The channel is assumed to be ideal and the frequency offset results in $\epsilon/T = 10^{-2}$, which is larger than realistic values. The emitted data is an i.i.d and binary sequence and $SNR = 20dB$. The LMS is run with a step-size equal to 0.05, processing 100 realizations of a sequence of 500T, with different values of the equalizer length N. As we can see on Figure 5, the FSE behave in a globally way best face to an high offset than the conventional equalizer. Between $N = 6$ and $N = 15$, the performances of the FSE are better until reach an optimal length which induce the smallest mean square error. Out of this interval, the performances of the both equalizer are identical and worst. Below $N = 6$, the squared error is very important because the equalizer is unable to remove the ISI induced by the frequency offset. Above $N = 15$, the squared error becomes large because of the residual stochastic jitter due to the high value of the step-size and of the noise, see [7]. In this case, temporal diversity implies good tracking properties and an improved robustness to a very high level of sampling clock frequency offset.

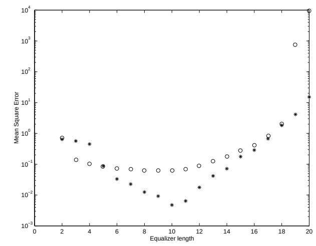


Figure 5: MSE versus length: with diversity ($*$), conventional (o)

7 Conclusion

We have shown in this paper that spatio-temporal effective diversity improves a lot the asymptotic performances and in a more subtle way the transient behavior. Although calculus and simulations were made for LMS, we can extend this analysis to other adaptive algorithms in order to evaluate the contribution of the spatio-temporal diversity.

References

[1] M.L. Alberi, I. Fijalkow, *Caractérisation de l'égalisabilité spatio-temporelle : cas d'un canal variant dans le temps* in Proc. GRETSI-97, September 1997.

[2] M.L. Alberi, I. Fijalkow, *Fractionally-Spaced Equalization of Time-Varying Mobile Communications* in Proc. ICASSP-98, Mai 1998.

[3] J.D. Behm et al., *Characterization of empirically-derived time-varying microwave channel models with implications for blind equalization and identification*, in Proc. ASILOMAR-97, November 1997.

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

[5] Blind Equalization Research Group (C.R. Johnson Jr., I. Fijalkow, L. Tong, J.R. Treichler), *The core of CMA behavior theory*, to be published in Blind Deconvolution II, edited by S. Haykin.

[6] I. Fijalkow, *Multichannel equalization lower bound: a function of channel noise and disparity*, in Proc. SSAP-96, June 1996.

[7] O. Macchi, Adaptive processing: The least mean squares approach with applications in transmission, Wiley & Sons, 1995.

[8] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, *Subspace methods for the blind identification of multichannel FIR filters*, IEEE Trans. on SP, February 1995.

[9] G. Ungerboeck, *Theory on the Speed of Convergence in Adaptive Equalizers for Digital Communications*, IBM J. Res. Dev., 856-864, november 1972.

[10] G. Ungerboeck, *Fractional tap-Spacing Equalizers and Consequences for Clock Recovery in Data Modems*, IEEE Trans. on Comm., august 1976.