

On the equivalence between the Godard and Shalvi–Weinstein schemes of blind equalization

Phillip A. Regalia*

Département Signal et Image, Institut National des Télécommunications, 9, rue Charles Fourier, F-91011 Evry cedex, France

Received 11 September 1997; received in revised form 11 June 1998

Abstract

Certain equivalences between the Godard and Shalvi–Weinstein schemes have been previously noted under special circumstances. We present here a simple proof for real signals that an equivalence can be established assuming little more than stationarity to fourth order of the equalizer input; the exact nature of the input sequence proves otherwise irrelevant to the validity of the equivalence. The equivalence also carries over to complex signals, but subject to more restrictive circularity conditions. In a communication context, the equivalence implies that many performance issues, such as susceptibility to local minima, the ability (or lack thereof) to open the eye, or mean performance degradations due to channel noise and/or source correlation properties, are common to the two, even when applied with nonlinear channels. Our equivalence also indicates a simple modification to the Godard algorithm to render it applicable to leptokurtic sources. © 1999 Elsevier Science B.V. All rights reserved.

Zusammenfassung

Gewisse Äquivalenzen der Godard-Methode und der Shalvi–Weinstein-Methode wurden bereits früher unter speziellen Bedingungen festgestellt. Wir zeigen hier auf einfache Weise, daß bei reellwertigen Signalen eine Äquivalenz hergestellt werden kann, ohne viel mehr als Stationarität bis zur vierten Ordnung des Entzerrereingangssignals anzunehmen. Die detaillierten Eigenschaften des Eingangssignals sind ansonsten für die Gültigkeit der Äquivalenz unerheblich. Die Äquivalenz läßt sich weiters auf komplexwertige Signale übertragen, allerdings unter restriktiveren Zirkularitätsbedingungen. In einem übertragungstechnischen Kontext impliziert diese Äquivalenz, daß viele Aspekte der Leistungsfähigkeit – wie die Empfindlichkeit gegenüber lokalen Minima, die Fähigkeit (oder der Mangel derselben) das Augendiagramm zu öffnen oder die Verringerung der mittleren Leistungsfähigkeit durch Rauschen am Kanal und/oder Korrelation der Quellen – auf beide Verfahren zutreffen, sogar bei Anwendung auf nichtlineare Kanäle. Unsere Äquivalenz weist auch auf eine einfache Modifikation des Godard-Algorithmus hin, durch welche dieser auf leptokurtische Quellen anwendbar wird. © 1999 Elsevier Science B.V. All rights reserved.

Résumé

Certaines équivalences entre les méthodes de Godard et de Shalvi–Weinstein ont été remarquées dans le passé, sous certaines conditions. Pour des signaux réels, on présente ici une démonstration simple du fait qu'une équivalence peut être établie sous la seule hypothèse de stationnarité à l'ordre quatre de l'entrée de l'égaliseur; la nature exacte de la

* Tel.: + 33 1 60 76 46 31; fax: + 33 1 60 76 44 33; e-mail: regalia@int-evry.fr.

suite d'entrée n'est par ailleurs pas pertinente en ce qui concerne cette équivalence. L'équivalence s'étend aussi aux signaux complexes, mais sous des hypothèses plus restrictives de circularité. Dans un contexte de communications, cette équivalence implique que beaucoup de résultats sur les performances, comme la sensibilité aux minima locaux, la capacité (ou le manque de capacité) à ouvrir l'oeil, ou les dégradations de performances moyennes dues au bruit de canal et/ou aux propriétés de corrélation des sources, sont communs aux deux schémas, même lorsqu'ils sont appliqués à des canaux non linéaires. Notre équivalence indique aussi une simple modification à apporter dans l'algorithme de Godard pour le rendre applicable aux sources leptokurtiques. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Godard algorithm; Shalvi–Weinstein algorithm; Blind equalization; Constant modulus algorithm

1. Introduction

Recent years have witnessed considerable creativity in algorithm design and subsequent analysis techniques for blind equalization schemes. Those which have gained popularity can all claim, modulo various (and sometimes questionable) hypotheses, to properly equalize a communication channel without recourse to a training sequence. A more recent trend is to compare relative robustness issues which are of paramount importance in applications. In this light, our contribution is to show a simple proof that two often contrasted approaches, the Godard method [3] and the Shalvi–Weinstein method [10] are in fact equivalent in their mean performance measures.

That the two methods show some relations is already known. The original paper by Shalvi and Weinstein [10], in fact, shows that their criterion reduces to the Godard cost function under special circumstances: a linear channel, no noise, and an i.i.d. source with some known statistics. We show here a simple proof that an equivalence between the two schemes can be established for real signals assuming little more than stationarity to fourth order of the equalizer input. Any further assumptions, such as whether the source is i.i.d. versus correlated, whether the channel is linear versus nonlinear, or whether any noise present is additive versus multiplicative, have no bearing on the validity of the equivalence. The equivalence can of course be extended to complex signals as well, but subject to more restrictive circularity conditions on the received signal.

Rather than show the Shalvi–Weinstein criterion reduces to the Godard cost function under certain

conditions, as in [10], we show instead that a reduced error surface of the Godard cost function is but a deformation of the Shalvi–Weinstein criterion. The deformation in question is easily shown to be monotonic and order preserving. With the two algorithms applied to a common signal environment, the set of stationary points are thus the same, as is the classification of each stationary point (i.e., minimum/maximum, local/global, or saddle point). An immediate consequence is that many performance considerations, such as susceptibility to local minima, the ability (or lack thereof) to adequately open the eye, or mean performance degradations due to channel noise or source correlation properties, are fundamentally no different between the two methods. Our equivalence also indicates a simple modification to the Godard algorithm to render it applicable to leptokurtic source sequences, as the conventional form is known to be suitable only for platykurtic source sequences.

2. Problem structure

Let \mathbf{u}_n be the sequence of k -element column vectors which form the input to the equalizer, corresponding to k -fold oversampling or using k sensors; the baud-rate single-sensor case corresponds to $k = 1$. The equalizer coefficients are k -element column vectors $\mathbf{g}_0, \dots, \mathbf{g}_{L-1}$, in which the equalizer length L is for now arbitrary. The equalizer output is

$$y_n = \sum_{l=0}^{L-1} \mathbf{g}_l^T \mathbf{u}_{n-l} = \mathbf{G}^T \mathbf{U}_n,$$

in which superscript T denotes transposition, and where $\mathbf{G}^T = [\mathbf{g}_0^T \dots \mathbf{g}_{L-1}^T]$ and $\mathbf{U}_n^T = [\mathbf{u}_n^T \dots \mathbf{u}_{n-L+1}^T]$.

For convenience, we shall parametrize \mathbf{G} in polar form as

$$\mathbf{G} = r\bar{\mathbf{G}},$$

where $\bar{\mathbf{G}}$ has unit Euclidean norm ($\|\bar{\mathbf{G}}\| = 1$) and $r = \|\mathbf{G}\|$ thus sets the norm. The kL components of $\bar{\mathbf{G}}$ may be parametrized canonically in terms of $kL - 1$ rotation angles $\theta_1, \dots, \theta_{kL-1}$; setting $s_i = \sin \theta_i$ and $c_i = \cos \theta_i$, one possibility for real equalizers is

$$\bar{\mathbf{G}}^T = [s_1 \ s_2 c_1 \ s_3 c_2 c_1 \ \dots \ s_{kL-1} c_{kL-2} \dots c_1 \ c_{kL-1} \dots c_1].$$

The complex case can be handled by letting s_i take complex values of the form $s_i = e^{j\phi_i} |\sin \theta_i|$. For all choices of the rotation angles $\Theta = \{\theta_i\}$ (or $\Theta = \{\theta_i, \phi_i\}$ for the complex case), the resulting $\bar{\mathbf{G}}$ has unit norm and, conversely, an arbitrary unit-norm vector may be reached for an essentially unique choice of the rotation angles Θ (see, e.g., [7], [8] (pp. 337–339)).

The output of the unit-norm filter is denoted as

$$\bar{y}_n = \bar{\mathbf{G}}^T \mathbf{U}_n,$$

and its scaled version yields the equalizer output as $y_n = r\bar{y}_n$.

We assume only that the equalizer input $\{\mathbf{u}_n\}$ has zero mean and is stationary to fourth order. For real signals, any further hypotheses on the problem structure, e.g., whether the equalizer input is obtained from a linear versus nonlinear channel, whether any noise present is additive versus multiplicative, or whether the equalizer length is chosen correctly or not, have no bearing on the validity of the equivalence to be derived, and hence need not be invoked. The complex case, by contrast, will involve a circularity condition, which will be highlighted separately in the development to follow.

2.1. Shalvi–Weinstein algorithm

The Shalvi–Weinstein algorithm adjust the equalizer parameters so as to seek the extremal

point(s) of the objective function [10,12]

$$F_1(\Theta) = \frac{c_4(y_n)}{(E[y_n^2])^2} = \frac{c_4(\bar{y}_n)}{(E[\bar{y}_n^2])^2}, \tag{1}$$

where $c_4(y_n)$ is the fourth-order cumulant of the equalizer output, viz.

$$c_4(y_n) = E(|y_n|^4) - 2[E(|y_n|^2)]^2 - E(y_n^2)E[(y_n^*)^2]. \tag{2}$$

For real y_n this becomes $E(y_n^4) - 3[E(y_n^2)]^2$, while for circularly symmetric complex processes, for which $E(y_n^2) = 0$ [6], this becomes $E(y_n^4) - 2[E(y_n^2)]^2$. It is readily checked that the ratio of the fourth-order cumulant to the variance squared is invariant to the choice of the scale factor r (for $r \neq 0$). This is reflected in the second equality of Eq. (1), such that the objective function varies only with the rotation angles Θ .

2.2. Godard algorithm

The Godard algorithm adjusts the equalizer parameters in the negative gradient direction (to within stochastic estimation errors) of the following cost function [3,11]:

$$F_2(r, \Theta) = E[(\gamma - |y_n|^2)^2]. \tag{3}$$

In the special case where the equalizer input is a (multichannel) convolved version of some source sequence, call it $\{s_n\}$, the constant γ should be chosen as $\gamma = E(|s_n|^4)/E(|s_n|^2)^2$. Any error in γ , however, is benign, in that such an error is completely compensated by a corresponding error in r , with no influence on intersymbol interference (e.g., [5]). We assume only that γ is some positive constant, as its exact value has little bearing on what follows.

3. The main result

A device used in previous works [2,4] is to derive from Eq. (3) a reduced error surface, in which $F_2(r, \Theta)$ is first optimized with respect to the radial parameter r . The resulting expression is then studied with respect to remaining degrees of freedom contained in Θ .

With $y_n = r\bar{y}_n$, the cost function from Eq. (3) can be rewritten as

$$F_2(r, \Theta) = E[(\gamma - r^2|\bar{y}_n|^2)^2] \\ = \gamma^2 - 2\gamma r^2 E(|\bar{y}_n|^2) + r^4 E(|\bar{y}_n|^4). \quad (4)$$

The derivative with respect to r is

$$\frac{\partial F_2(r, \Theta)}{\partial r} = 4r[r^2 E(|\bar{y}_n|^4) - \gamma E(|\bar{y}_n|^2)].$$

Equating this to zero yields $r = 0$ and $r = \pm \sqrt{\gamma E(|\bar{y}_n|^2)/E(|\bar{y}_n|^4)}$ as solutions. The former is known [11] to yield a local maximum of F_2 , while the latter two solutions are (equally good) minima [2]. Note that in a communication context, the sign ambiguity on r is innocuous whenever differential source encoding is used, and subsequent expressions will involve only r^2 anyway. Setting thus

$$r_{\text{opt}}^2(\Theta) = \frac{\gamma E(|\bar{y}_n|^2)}{E(|\bar{y}_n|^4)},$$

in which we notationally emphasize that the optimal choice of r^2 depends on the angular parameters Θ , the cost function $F_2(r, \Theta)$ from Eq. (4) reduces to

$$\bar{F}_2(\Theta) \triangleq F_2(r, \Theta)|_{r^2_{\text{opt}}(\Theta)} \\ = \gamma^2 \left(1 - \frac{[E(|\bar{y}_n|^2)]^2}{E(|\bar{y}_n|^4)} \right) \quad (5) \\ = \gamma^2 \left(1 - \frac{1}{2 + \frac{c_4(\bar{y}_n)}{[E(|\bar{y}_n|^2)]^2} + \frac{E(\bar{y}_n^2)E[(\bar{y}_n^*)^2]}{[E(|\bar{y}_n|^2)]^2}} \right), \quad (6)$$

in which the fourth-order cumulant expression (2) may help the reader obtain the second line. We distinguish now two cases:

1. *Real equalizer output.* In this case, $E(\bar{y}_n^2) = E[(\bar{y}_n^*)^2] = E(|\bar{y}_n|^2)$, so that Eq. (6) becomes

$$\bar{F}_2(\Theta) = \gamma^2 \left(1 - \frac{1}{3 + F_1(\Theta)} \right). \quad (7)$$

This shows that the reduced error surface $\bar{F}_2(\Theta)$ of the Godard scheme is a simple deformation of the Shalvi–Weinstein criterion $F_1(\Theta)$. It should be clear that any minima of the reduced error surface $\bar{F}_2(\Theta)$ are minima of the full cost function (3), and that any maxima or saddle points of

the reduced error surface are saddle points of the full error surface. [The only local maximum of Eq. (3) is at $r = 0$, which is excluded in the construction of the reduced error surface; any other stationary points of the full error surface are minima with respect to r , and hence cannot appear as maxima with respect to Eq. (3)].

The deformation in question is monotonic, since from Eq. (7)

$$\frac{d\bar{F}_2(\Theta)}{dF_1(\Theta)} = \frac{\gamma^2}{[3 + F_1(\Theta)]^2} > 0. \quad (8)$$

Moreover, for any two parameter choices Θ_1 and Θ_2 , a simple calculation gives

$$\bar{F}_2(\Theta_1) - \bar{F}_2(\Theta_2) \\ = \gamma^2 \frac{F_1(\Theta_1) - F_1(\Theta_2)}{[3 + F_1(\Theta_1)][3 + F_1(\Theta_2)]}$$

This shows the deformation to be order preserving, i.e.,

$$\bar{F}_2(\Theta_1) > \bar{F}_2(\Theta_2) \Leftrightarrow F_1(\Theta_1) > F_1(\Theta_2), \quad (9)$$

once it is recognized that the constraint $\bar{F}_2(\Theta) < \gamma^2$ for all Θ [cf. Eq. (5)] implies, via Eq. (7), that the term $3 + F_1(\Theta)$ is positive-valued for all Θ .

It follows readily from Eq. (8) that stationary points in the Θ space coincide ($(dF_1(\Theta))/d\Theta = 0 \Leftrightarrow (d\bar{F}_2(\Theta))/d\Theta = 0$). Moreover, from Eqs. (8) and (9) together, the classification of each stationary point is preserved. ('Classification' here refers to whether a given stationary point is a local minimum, global minimum, saddle point, local maximum, or global maximum).

2. *Complex equalizer output.* If the output process $\{\bar{y}(\cdot)\}$ is circular for all choices of the equalizer parameters (see [6] for general conditions guaranteeing circularity), then $E(\bar{y}_n^2) = E[(\bar{y}_n^*)^2] = 0$. In this case, Eq. (6) reduces to

$$\bar{F}_2(\Theta) = \gamma^2 \left(1 - \frac{1}{2 + F_1(\Theta)} \right).$$

The arguments showing that $F_1(\Theta)$ and $\bar{F}_2(\Theta)$ are related by a monotonic deformation, and all conclusions following from this, proceed in direct parallel to the real case above.

If, on the other hand, circularity of the equalizer output \bar{y}_n does not hold, as may happen for certain nonlinear processes, then the Godard and Shalvi–Weinstein criteria can have different convergent points. We may remark that the utility of the Shalvi–Weinstein criteria appears, based on the original article [10], strongly conditioned on various source and channel hypotheses for the complex case. Similar remarks may apply to the Godard algorithm, and the superiority of one or the other for noncircular complex processes remains to be investigated.

4. A leptokurtic Godard algorithm

The above generic equivalence, either for real of complex circular processes, suggests a simple modification to the Godard algorithm to render it applicable to leptokurtic sources. We begin by specializing our model to a common linear model used in communication contexts, viz.

$$\mathbf{u}_n = \sum_{l=0}^{\infty} \mathbf{h}_l s_{n-l} + \mathbf{b}_n,$$

in which $\{s_n\}$ is a scalar source sequence, $\{\mathbf{h}_l\}$ is the single-input/ k -output channel impulse response (possibly infinite length), and \mathbf{b}_n is the channel noise.

We let $\{f_i\}$ denote the combined (channel-equalizer) impulse response:

$$f_i = \sum_{l=0}^{L-1} \mathbf{g}_l^T \mathbf{h}_{i-l}, \quad i = 0, 1, 2, \dots$$

The Shalvi–Weinstein approach was first proposed for the setting in which $\{s_n\}$ is an i.i.d. source, and the noise is absent. Criterion (1) is then known to assume the form

$$F_1(\Theta) = \frac{c_4(\bar{y}_n)}{(E[\bar{y}_n^2])^2} = \frac{c_4(s_n)}{(E[s_n^2])^2} \left(\frac{\sum_{i=0}^{\infty} f_i^4}{(\sum_{i=0}^{\infty} f_i^2)^2} \right),$$

in which the parenthetic term involving the combined response lies between zero and one. The upper limit of unity is attained if and only if the combined response reduces to a sole nonzero term [10,12]. Thus, if the source is platykurtic [$c_4(s_n) < 0$], the objective function $F_1(\Theta)$ should be

minimized (i.e., pushed to an extremal negative value), with clearly the opposite strategy for leptokurtic sources [$c_4(s_n) > 0$].

The Godard algorithm, by contrast, appears ill-suited to leptokurtic sources, since minimizing the cost function (3) actually worsens intersymbol interference for that case (see, e.g. [5] for a nice analysis). From our equivalence, it is clear that the correct strategy to adopt for leptokurtic sources is to maximize the reduced Godard cost function $\bar{F}_2(\Theta)$, as this is equivalent to maximizing the Shalvi–Weinstein criterion.

As the reduced error surface $\bar{F}_2(\Theta)$ arises from minimizing with respect to the radial parameter r , such a ‘min–max’ procedure would appear difficult for a conventional transversal parametrization of the equalizer. The polar parametrization studied here, though, renders such an approach straightforward. A (stochastic approximation to a) gradient descent procedure applied to Eq. (3) appears as

$$\Theta_{n+1} = \Theta_n \pm \mu \operatorname{Re} \left(y_n^* (\gamma - |y_n|^2) \frac{dy_n}{d\Theta} \Big|_{\Theta=\Theta_n} \right) \quad \mu > 0,$$

in which the ‘ \pm ’ sign is chosen opposite to the sign of the source kurtosis. The gradient vector $dy_n/d\Theta$ may be computed in a straightforward fashion (see, e.g. [7] or [1] for the real case, with straightforward extension to the complex case). The scale factor r may be adapted either by a gradient descent procedure, or a least-squares-type algorithm, among other possibilities. Further algorithm development, and performance comparisons between the various possibilities, is a worthwhile topic for the interested reader.

We should emphasize that the Godard and Shalvi–Weinstein schemes may still exhibit some practical differences, some of which are worth noting:

- Because their respective objective functions are related by a deformation, the eigenvalue spread of the Hessian matrix about a minimum point may differ between the two algorithms, indicating that local convergence properties may not show the same speeds.
- Noise sources, whether they be channel noise, gradient estimation noise, or roundoff noise in the computations, can manifest themselves

differently in the two algorithms, indicating that misadjustment effects need not be equivalent. It remains true that *mean* performance degradation induced by channel noise will remain equivalent, provided the misadjustment effects in either case are negligible compared to the mean degradation. Ref. [9], though, shows a noteworthy example where misadjustment effects should *not* be neglected.

Further comparative studies between the two algorithm classes should thus focus on these and other conditioning aspects.

5. Concluding remarks

We have presented an elementary proof of equivalence between the Godard and Shalvi–Weinstein criteria for real signals which assumes little more than stationarity to fourth order of the equalizer input, in contrast to previous comparisons which tend to invoke further hypotheses (e.g., linear channel, i.i.d. source, etc.). The complex case also carries through, but subject to more stringent circularity conditions. Various performance measures, such as existence of and susceptibility to local minima, the ability to open the eye, or mean performance degradation due to channel noise and/or source correlation properties, are thus common to the two schemes, even when applied to nonlinear channels (for real signals). We have noted that conditioning aspects of the two algorithms, such as misadjustment severity or convergence speed, can nonetheless differ due to the deformation involved. Further studies should thus focus on these questions, as well as on any possible superiority of one over the other when applied to non-circular complex processes.

References

- [1] N. Delfosse, P. Loubaton, Adaptive blind separation of independent sources: A deflation approach, *Signal Processing* 45 (July 1995) 59–83.
- [2] A. Ding, R.A. Kennedy, On the whereabouts of local minima for blind adaptive equalizers, *IEEE Trans. Circuits Systems II* 39 (February 1992) 119–123.
- [3] D.N. Godard, Self-recovering equalization and carrier tracking in two-dimensional data communication systems, *IEEE Trans. Commun.* 28 (November 1980) 1867–1875.
- [4] C.R. Johnson Jr., B.D.O. Anderson, Godard blind equalizer error surface characteristics: White, zero-mean, binary source case, *Internat. J. Adaptive Control Signal Process.* 9 (July–August 1995) 301–324.
- [5] J.P. LeBlanc, Effects of source distributions and correlation on fractionally spaced blind constant modulus equalizers, Ph.D. Dissertation, Cornell University, August 1995.
- [6] B. Picinbono, P. Bondon, Second-order statistics of complex signals, *IEEE Trans. Signal Process.* 45 (February 1997) 411–420.
- [7] P.A. Regalia, An adaptive unit-norm filter with applications to signal analysis and Karhunen–Loève transformations, *IEEE Trans. Circuits Systems* 37 (May 1990) 646–649.
- [8] P.A. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*, Marcel Dekker, New York, 1995.
- [9] M. Reuter, J.R. Zeidler, Non-Wiener effects in LMS-implemented adaptive equalizers, *Proc. ICASSP, Munich*, April 1997, Vol. 3, pp. 2509–2512.
- [10] O. Shalvi, E. Weinstein, New criteria for blind deconvolution of nonminimum phase systems (channels), *IEEE Trans. Inform. Theory* 36 (March 1990) 312–321.
- [11] J.R. Treichler, B.G. Agee, A new approach to multipath correction of constant modulus signals, *IEEE Trans. Acoust. Speech Signal Process.* 31 (April 1983) 459–472.
- [12] J.K. Tugnait, Comments on ‘New criteria for blind deconvolution of nonminimum phase systems (channels)’, *IEEE Trans. Inform. Theory* 38 (January 1992) 210–213.