Online Adaptive Blind Deconvolution Based on Third-Order Moments

Patrik Pääjärvi, *Student Member, IEEE,* and James P. LeBlanc, *Senior Member, IEEE*

***Abstract*—Traditional methods for online adaptive blind deconvolution using higher order statistics are often based on even-order moments, due to the fact that the systems considered commonly feature symmetric source signals (i.e., signals having a symmetric probability density function). However, asymmetric source signals facilitate blind deconvolution based on odd-order moments. In this letter, we show that third-order moments give the benefits of faster convergence of algorithms and increased robustness to additive Gaussian noise. The convergence rates for two algorithms based on third- and fourth-order moments, respectively, are compared for a simulated ultra-wideband communication channel.**

***Index Terms*—Adaptive filtering, blind equalization, third-order moments.**

I. Introduction

ADAPTIVE blind deconvolution is used for equalization or identification of unknown systems when only the output of the system can be observed. Fig. 1 shows a discrete-time signal model of a general blind deconvolution problem (the subscript n denotes a time index). The object is to find the deconvolution filter $\mathbf{f}$ that approximately inverts the system $\mathbf{c}$ with limited or no knowledge of either $\mathbf{c}$ or the source signal s_n. The system output x_n plus an additive disturbance z_n gives the observed signal u_n. The unknown, possibly time-varying system $\mathbf{c}$ may be either linear or nonlinear with minimum, maximum, or mixed phase.

A. Minimum Entropy Deconvolution

In general, due to filtration through $\mathbf{c}$, the probability distribution of the system output x_n will be closer to a Gaussian than that of s_n. This is a consequence of the central limit theorem and allows for deconvolution based on measuring the "Gaussianity" of the deconvolution filter output y_n. *Minimum entropy deconvolution* (MED) methods are based on using a *score function* $\Phi(y_n)$ to measure the Gaussianity (or entropy) of y_n. The probability distribution of y_n is then driven as far away from a Gaussian distribution as possible, by adapting the coefficients of $\mathbf{f}$. In general, $\mathbf{f}$ is a finite impulse response (FIR) filter, and all signals are assumed to have zero mean.

To measure the Gaussianity of a signal, score functions based on higher order central moments are commonly used. Such functions can typically be approximated by simple polynomial functions of y_n, making them specially suitable for online (real-time) applications, where computational efficiency is often of concern. Wiggins [1] proposed the use of the *Kurtosis* (normalized fourth-order moment) of y_n as a score function for MED. Donoho [2] generalized the theory behind MED and considered various types of score functions, including central moments of order greater than two.

Godard [3] suggested *dispersions* of y_n as score functions for blind equalization of communication channels. The dispersion of order p (p integer >0) is based on even-order moments of y_n and is defined as

$$D^{(p)} = E\{(|y_n|^p - R_p)^2\} \tag{1}$$

where R_p is a positive constant, and $E\{\cdot\}$ denotes expectation. Choosing $p = 2$ leads to the popular constant modulus algorithm (CMA) [4], which is based on fourth-order moments, similar to Wiggins' original idea.

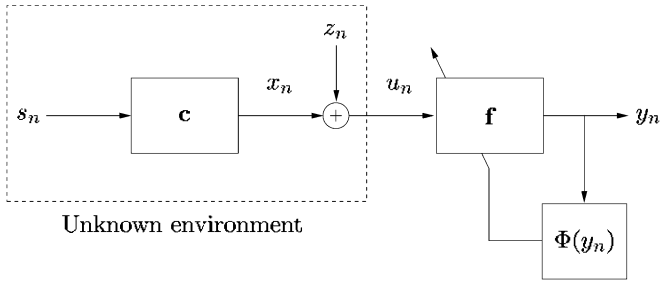

Fig. 1. Signal model of a general blind deconvolution problem.

B. Symmetric and Asymmetric Source Signals

Traditional uses of blind deconvolution include linear equalization of communication channels, deconvolution of seismic traces, and dereverberation of acoustic signals. Such applications are often assumed to feature *symmetric* source signals, i.e., zero-mean signals with a probability density function (PDF) that is symmetric around zero. Since all odd-order moments of symmetric signals are zero, most research focus in the field of blind deconvolution has hence been directed toward even-order moments. Although symmetric source signals dominate the field of applications for MED, *asymmetric* source signals, i.e., zero-mean signals with asymmetric PDFs (and thus with nonzero third central moment) occur in a wide range of acoustic, biomedical, and mechanical signals (for example, pulse oximetry signals or hammer impacts). Asymmetry is also a feature of impulse radio (IR) signals [5], a proposed signaling format for ultra-wideband (UWB) radio [5]–[7].

Manuscript received May 13, 2005; revised June 29, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Stefano Buzzi.

The authors are with the Department of Computer Science and Electrical Engineering, Luleå University of Technology, SE-97187 Luleå, Sweden (e-mail: patrik.paajarvi@ltu.se).

Digital Object Identifier 10.1109/LSP.2005.859496

In previous work [8], we noted that asymmetry in the source signal can be exploited by using a score function based on *third-order moments*, instead of the common fourth-order moments. The benefit of a lower order moment is mainly a simpler score function surface (regarding $\Phi(y_n)$ as a function of the deconvolution filter coefficients). This will, in general, give faster convergence of common gradient search algorithms.

In the work presented in this letter, we compare two simple online score functions based on third- and fourth-order moments, respectively. Since symmetric source signals have zero odd-order moments, we restrict our focus to asymmetric sources. We demonstrate that an online gradient search algorithm based on third-order moments should, in general, benefit from faster convergence and increased robustness to additive Gaussian noise, compared to algorithms based on fourth-order moments. The experimental results are obtained from simulations of an indoor UWB channel with IR signaling.

II. Notation and Model Description

Referring to the discrete-time signal model in Fig. 1, we define s_n and $\mathbf{c}$ as the unknown source signal and unknown channel, respectively. The sum of the channel output x_n and the disturbance z_n is the observed signal u_n, which is the input to the deconvolution filter $\mathbf{f}$. For simplicity, we will from here on refer to $\Phi(y_n)$ as an *objective* function of y_n, and the objective of the deconvolution problem is to find the filter $\mathbf{f}$ that maximizes $\Phi(y_n)$. In typical online situations, this is done iteratively through a gradient search algorithm.

The adaptive filter $\mathbf{f}$ is assumed to be FIR of order N. The filter after r iterations is represented by the coefficient vector

$$\mathbf{f}^{(r)} = \begin{bmatrix} f_0^{(r)} & f_1^{(r)} & \cdots & f_N^{(r)} \end{bmatrix}^{\mathrm{T}}. \tag{2}$$

Using adaption by gradient ascent, $\mathbf{f}$ is recursively updated in the direction of maximizing the objective function. The filter update rule becomes

$$\mathbf{f}^{(r+1)} = \mathbf{f}^{(r)} + \mu \nabla_{\Phi(\mathbf{f}^{(r)})} \tag{3}$$

where μ is a positive step size of adaption, and $\nabla_{\Phi(\mathbf{f}^{(r)})}$ is the gradient of Φ with respect to $\mathbf{f}^{(r)}$

$$\nabla_{\Phi(\mathbf{f}^{(r)})} = \begin{bmatrix} \dfrac{\partial \Phi}{\partial f_0^{(r)}} & \dfrac{\partial \Phi}{\partial f_1^{(r)}} & \cdots & \dfrac{\partial \Phi}{\partial f_N^{(r)}} \end{bmatrix}^{\mathrm{T}}. \tag{4}$$

Filter iteration can be performed either on a sample-by-sample basis (general applications) or on a symbol-by-symbol basis (digital communication applications). If the step size μ in (3) is small, $\mathbf{f}$ can be regarded as approximately constant in time, allowing us to drop the superscript (r). We then define the filter output at sampling instant n as

$$y_n = \mathbf{u}_n^{\mathrm{T}}\mathbf{f} = \mathbf{x}_n^{\mathrm{T}}\mathbf{f} + \mathbf{z}_n^{\mathrm{T}}\mathbf{f} = d_n + v_n \tag{5}$$

with d_n being the "filtered signal," v_n the "filtered noise," and the signal vectors defined as $\mathbf{u}_n = [u_n \; u_{n-1} \; \cdots \; u_{n-N}]^{\mathrm{T}}$, $\mathbf{x}_n = [x_n \; x_{n-1} \; \ldots \; x_{n-N}]^{\mathrm{T}}$, and $\mathbf{z}_n = [z_n \; z_{n-1} \; \ldots \; z_{n-N}]^{\mathrm{T}}$. The two objective functions we will compare are simply the third- and fourth-order central moments, respectively, of the filter output y_n

$$\Phi^{(3)}(y_n) \triangleq E\left\{y_n^3\right\} \tag{6}$$

$$\Phi^{(4)}(y_n) \triangleq E\left\{y_n^4\right\}. \tag{7}$$

The corresponding gradients with respect to $\mathbf{f}$ are

$$\nabla_{\Phi^{(3)}} \propto E\left\{y_n^2 \mathbf{u}_n\right\} \tag{8}$$

$$\nabla_{\Phi^{(4)}} \propto E\left\{y_n^3 \mathbf{u}_n\right\}. \tag{9}$$

In online applications, where computational power is often limited, it is customary to use an instantaneous estimate of the gradient in the filter update (3). This can be obtained from the two objective functions (8) and (9) by simply dropping the expectation operators.

From here on, we will make the following assumptions.

- A1) All signals are real and zero mean.
- A2) s_n is a non-Gaussian and asymmetric signal.
- A3) The disturbance z_n is a zero-mean, independent, identically distributed (i.i.d.) Gaussian noise process, independent of x_n, with variance σ_z^2.
- A4) The step-size parameter μ in (3) is small, so that the filter vector $\mathbf{f}^{(r)}$ can be regarded as approximately constant in time when compared to the signals, i.e., $\mathbf{f}^{(r)} = \mathbf{f}$.
- A5) $\mathbf{f}$ is kept at constant (unit) norm during adaption, i.e., $\|\mathbf{f}\|^2 = \sum_i f_i^2 = 1$.

Assumption A4) is customary in adaptive filtering theory and simplifies the averaging analysis in the next section. Assumption A5) is necessary since increasing the norm of any filter $\mathbf{f}$ increases both objectives (6) and (7) while leaving the Gaussianity of the filter output unchanged.

III. Comparative Performance Analysis of Third- and Fourth-Order Objective Functions

A. Objective Function Surface Topology

If the objective function Φ is regarded as a function of the deconvolution filter coefficients, adaption according to (3) can be thought of as traversing a multidimensional function surface $\Phi(\mathbf{f})$ toward any local maximum points (under the constraint of unit filter norm). The set of maximum points is a subset of the points on the function surface with zero gradient $\nabla_{\Phi(\mathbf{f})}$, with the other set members being minimum points or saddle points. Maximizing $\Phi(\mathbf{f})$ is therefore equivalent to finding a subset of solutions to

$$\nabla_{\Phi(\mathbf{f})} = \mathbf{0}. \tag{10}$$

For a filter of order N, (10) leads to a system of $N+1$ nonlinear polynomial equations in the $N+1$ unknowns $\{f_0, \ldots, f_N\}$. The highest degree of the polynomials in the equation system (10) will set an upper bound on the number of solutions, i.e., the number of stationary points on the objective surface. A large number of stationary points generally implies a large number of saddle points, which can "stall" filter adaption.

Solving (10) for the third-moment objective function (6) leads to the system of equations

$$\sum_i f_i^2 \mathcal{R}_{0}^{m-i} + \sum_{i \neq j} f_i f_j \mathcal{R}_{j-i}^{m-i} = 0 \tag{11}$$

for $m, i, j = 0, \ldots, N$, with the third moment of u_n defined as $\mathcal{R}_j^i = E\{u_n u_{n-i} u_{n-j}\}$. The highest polynomial degree of (11) is 2, which gives a Bezout [9] upper bound on the number of solutions, i.e., the number of stationary points on the function surface, equal to 2^{N+1}. The corresponding system of equations for the fourth-moment objective function (7) is

$$\sum_i f_i^3 \mathcal{R}_{m-i}^{0\,0} + 3 \sum_{i \neq j} f_i^2 f_j \mathcal{R}_{m-i}^{0\,j-i} + \sum_{i \neq j \neq k} f_i f_j f_k \mathcal{R}_{m-i}^{j-i\,k-i} = 0 \tag{12}$$

for $m, i, j, k = 0, \ldots, N$, with the fourth moment of u_n defined as $\mathcal{R}_{k}^{i\,j} = E\{u_n u_{n-i} u_{n-j} u_{n-k}\}$. The highest polynomial degree of (12) is 3, giving a Bezout upper bound on the number of solutions equal to 3^{N+1}. Note that, in general, the moments of u_n in (11) and (12) may depend on n (e.g., for a time-varying system $\mathbf{c}$), despite the notation used. This is not essential here, since a dependence on n only implies that the *shape* of the function surface changes over time. The upper bounds on the number of stationary points are still constant.

Even for moderate filter orders, the maximum number of stationary points on the third-moment function surface is considerably smaller than on the corresponding fourth-moment surface. As previously noted for offline (block-mode) algorithms in [8], lower polynomial order of score functions gives the benefit of a "simpler" objective surface, which, in general, implies fewer saddle points. Since an excessive number of saddle points can "stall" a gradient search, a simpler objective surface will therefore, in general, allow for faster adaption of such algorithms. This is of special importance in applications where the unknown system $\mathbf{c}$ is time varying, and the deconvolution filter needs to "track" changes in the system.

B. Gaussian Noise Effects on the Objective Function Surface

1) Objective Surface Analysis: In the presence of additive white Gaussian noise, as described in the model in Section II, and with the filter output decomposed into the sum of d_n and v_n, as in (5), the third-moment objective function (6) at time n becomes

$$\Phi^{(3)}(y_n) = E\left\{d_n^3\right\}. \tag{13}$$

Since all odd moments of the Gaussian disturbance are zero, (13) depends solely on the filtered signal d_n and not on the disturbance z_n. Thus, the function surface of the third-moment objective function is preserved in the presence of Gaussian noise. The corresponding expression for the fourth-moment objective function (7) in the presence of Gaussian noise is

$$\Phi^{(4)}(y_n) = E\left\{d_n^4\right\} + 3\left(\sigma_z^2\right)^2 \|\mathbf{f}\|^4 + 6\sigma_z^2 E\left\{d_n^2\right\} \|\mathbf{f}\|^2. \tag{14}$$

The Gaussian noise introduces two additional terms to the "signal" (first) term. Under Assumption A5), the second term does not depend on $\mathbf{f}$ and will therefore not change the location of the stationary points. The third term, on the other hand, which depends on $\mathbf{f}$ through d_n, will alter the location of the stationary points. Since the local maximum points have moved under the influence of noise, the ability of the algorithm to invert $\mathbf{c}$ has been reduced.

2) Gradient Analysis: With the filter output defined as in (5), the gradient of the objective function can be expressed as

$$\nabla_\Phi = \nabla_{\Phi(d)} + \nabla_{\Phi(d,v)}. \tag{15}$$

$\nabla_{\Phi(d)}$ is the "signal" component of the gradient due to the filtered source signal d_n. $\nabla_{\Phi(d,v)}$ is the perturbation of the gradient caused by the Gaussian noise. Taking the gradients of $\Phi^{(3)}$ and $\Phi^{(4)}$ with respect to $\mathbf{f}$ and separating them according to (15) yields

$$\nabla_{\Phi^{(3)}} \propto \nabla_{\Phi^{(3)}(d)} \tag{16}$$

$$\nabla_{\Phi^{(4)}} \propto \nabla_{\Phi^{(4)}(d)} + 3\left[\left(\sigma_z^2\right)^2 \|\mathbf{f}\|^2 + \sigma_z^2 E\left\{d_n^2\right\}\right]\mathbf{f} + 3\sigma_z^2 \|\mathbf{f}\|^2 E\{d_n \mathbf{x}_n\}. \tag{17}$$

At "true" local maximum points, the signal gradients $\nabla_{\Phi^{(3)}(d)}$ and $\nabla_{\Phi^{(4)}(d)}$ are zero. As indicated by (13), the function surface of $\Phi^{(3)}$ is not affected by noise. Therefore, an instantaneous estimate of $\nabla_{\Phi^{(3)}}$ [obtained by dropping the expectation operator in (8)] will be unbiased in the presence of Gaussian noise. For $\Phi^{(4)}$, the noise causes a perturbation of the gradient in the direction of $E\{d_n \mathbf{x}_n\}$, causing a corresponding instantaneous estimate of $\nabla_{\Phi^{(4)}}$ to become biased. This adversely affects the algorithms ability to invert the unknown system, as indicated by (14). Although the perturbation in the direction of $\mathbf{f}$ does not introduce a bias under unit-norm constraints, a large noise variance may have a negative effect on the convergence rate of the algorithm on finite-precision machines.

IV. Experimental Results From a Simulated UWB Radio Channel

Wireless communication over UWB radio channels has attracted interest in recent years. One of the proposed signaling formats for UWB communication is IR [5], which consists of pulse-position modulated pulses of extremely short duration, typically on the order of a nanosecond, transmitted without the use of a sinusoidal carrier. The short pulses used give IR signals a bandwidth from near dc to several gigahertz, giving them good material-penetrating abilities and resolvable multipath delays down to about 30 cm. To allow for multiple user access, an additional pseudo-random time-hopping modulation scheme is used. This reduces the risk of catastrophic collisions with other IR transmitters and also avoids interference with coexisting narrow-band signals by "spreading" the spectrum of the signal [5].

Although the large bandwidth of IR signals makes them robust to fading, the large multipath spread of a typical indoor UWB channel is likely to cause intersymbol interference (ISI) at higher data rates [7], [10]. The asymmetry of typical IR signal

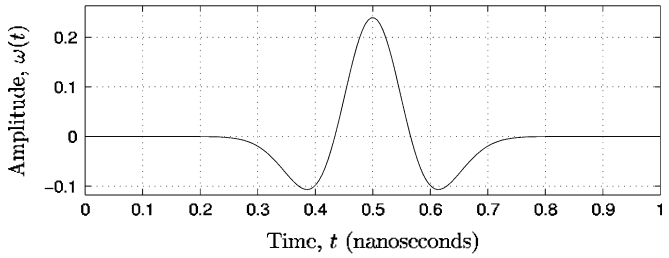

Fig. 2. Signal pulse shape $\omega(t)$ used in the experiment. $\omega(t) = [1 - 4\pi(t/\tau_m)^2]\exp[-2\pi(t/\tau_m)^2]$, with $\tau_m = 0.2333$.

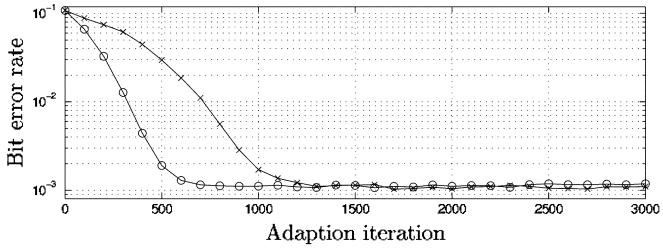

Fig. 3. Bit-error rate versus adaption iteration for third-moment (∘) and fourth-moment (×) objective functions, averaged over 13 runs. SNR per bit = 11 dB.

pulses motivates the use of a blind adaptive linear equalizer based on third-order moment maximization to mitigate ISI.

A numerical experiment was conducted in which the two objective functions (6) and (7) were used to implement two fractionally spaced, adaptive linear equalizers for a UWB channel. IR signals with i.i.d. symbols were simulated based on the model described in [5], using a pulse shape

$$\omega(t) = [1 - 4\pi(t/\tau_m)^2]\exp[-2\pi(t/\tau_m)^2] \qquad (18)$$

with $\tau_m = 0.2333$, giving a pulse duration of about 1 ns. The pulse shape is shown in Fig. 2. The sampling interval was chosen to give each pulse a support of 15 samples, based on results from [11]. The IR signals used binary orthogonal modulation at a bit rate of 10 Mb/s. A UWB channel impulse response with a rich multipath spread up to approximately 200 ns was synthesized with the aid of a recipe from [12]. Although only a single transmitter was simulated, the interference from a large number of adjacent transmitters can in many situations be modeled as a Gaussian random process [5].

The receiver structure consisted of a filter matched to (18) followed by the linear equalizers. The two FIR equalizers of order $N = 400$ were implemented with adaption using third-order moment and fourth-order moment maximization, respectively. The equalizers were recursively updated at the symbol instants, using instantaneous estimates of (8) and (9), respectively, starting from the customary "center-tap" initialization. The individual step sizes of adaption, $\mu^{(3)} = 9.5\cdot 10^{-3}$ for the third-moment algorithm and $\mu^{(4)} = 6\cdot 10^{-3}$ for the fourth-moment algorithm, were chosen so that both algorithms gave equal bit-error rate performance at convergence. Fig. 3 shows the bit-error rate versus adaption iteration for the third- and fourth-order-moment-based objective functions. The curves show the average results from 13 runs for a signal-to-noise ratio per bit of 11 dB.

As seen in Fig. 3, the algorithm that uses the third-order moment objective function converges approximately twice as fast as the corresponding fourth-order moment version. This confirms the results in Section III-A, namely, that the lower order polynomial structure of the third-order moment results in a "simpler" function surface. In general, this should imply faster convergence of filter adaption, which is important for the algorithm's ability to track a time-varying channel. Since typical indoor UWB channels are indeed time-varying, third-order-moment-based blind deconvolution, with its ability to exploit the source asymmetry, seems to be a suitable option for UWB channel equalization.

V. Conclusion

We have compared the performance of two objective functions for adaptive blind deconvolution based on third-order moments and fourth-order moments, respectively. Asymmetric source signals offer opportunities to use objective functions based on third-order moments, as an alternative to the commonly used fourth-order moments. Both the analytical and the experimental results indicate that a lower order objective function results in fewer stationary points on the objective function surface, which, in general, allows for faster convergence of online blind adaptive algorithms. The analysis of gradient estimation in the presence of Gaussian noise further highlights the advantages of using third-order moments. The faster convergence and increased robustness to additive Gaussian noise makes third-order-moment-based methods interesting candidates for blind adaptive equalization in UWB communication.

References

[1] R. A. Wiggins, "Minimum entropy deconvolution," *Geoexploration*, no. 16, pp. 21–35, 1978.
[2] D. L. Donoho, "On minimum entropy deconvolution," in *Applied Time Series Analysis*, D. F. Findley, Ed. New York: Academic, 1981.
[3] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COM-28, no. 11, pp. 1867–1875, Nov. 1980.
[4] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust. Speech, Signal, Process.*, vol. ASSP-31, no. 2, pp. 459–472, Apr. 1983.
[5] M. Z. Win and R. A. Scholtz, "Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple access communications," *IEEE Trans. Commun.*, vol. 48, no. 4, pp. 679–691, Apr. 2000.
[6] R. A. Scholtz, R. Weaver, E. Homier, J. Lee, P. Hilmes, A. Taha, and R. Wilson, "UWB radio deployment challenges," in *Proc. IEEE Int. Symp. Personal, Indoor, Mobile Radio Communications*, vol. 1, London, U.K., Sep. 2000, pp. 620–625.
[7] D. Porcino and W. Hirt, "Ultra-wideband radio technology: Potential and challenges ahead," *IEEE Commun. Mag.*, vol. 41, no. 7, pp. 66–74, Jul. 2003.
[8] P. Pääjärvi and J. P. LeBlanc, "Skewness maximization for impulsive sources in blind deconvolution," in *Proc. IEEE Nordic Signal Processing Symp.*, Espoo, Finland, Jun. 2004, pp. 304–307.
[9] E. W. Weisstein, *CRC Concise Encyclopedia of Mathematics*. London, U.K.: Chapman & Hall/CRC, 1999.
[10] A. G. Klein and C. R. Johnson, Jr., "MMSE decision feedback equalization of pulse position modulated signals," in *Proc. IEEE Int. Conf. Communications*, vol. 5, Paris, France, Jun. 2004, pp. 2648–2652.
[11] V. Lottici, A. D'Andrea, and U. Mengali, "Channel estimation for ultra-wideband communications," in *IEEE J. Sel. Areas Commun.*, vol. 20, Dec. 2002, pp. 1638–1645.
[12] D. Cassioli, M. Z. Win, and A. F. Molisch, "The ultra-wide bandwidth indoor channel: From statistical model to simulations," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1247–1257, Aug. 2002.