# Bussgang Blind Deconvolution for Impulsive Signals

Heinz Mathis and Scott C. Douglas, *Senior Member, IEEE*

***Abstract*—Many blind deconvolution algorithms have been designed to extract digital communications signals corrupted by intersymbol interference (ISI). Such algorithms generally fail when applied to signals with impulsive characteristics, such as acoustic signals. While it is possible to stabilize such procedures in many cases by imposing unit-norm constraints on the adaptive equalizer coefficient vector, these modifications require costly divide and square-root operations. In this paper, we provide a theoretical analysis and explanation as to why unconstrained Bussgang-type algorithms are generally unsuitable for deconvolving impulsive signals. We then propose a novel modification of one such algorithm (the Sato algorithm) to enable it to deconvolve such signals. Our approach maintains the algorithmic simplicity of the Sato algorithm, requiring only additional multiplies and adds to implement. Sufficient conditions on the source signal distribution to guarantee local stability of the modified Sato algorithm about a deconvolving solution are derived. Computer simulations show the efficiency of the proposed approach as compared with various constrained and unconstrained blind deconvolution algorithms when deconvolving impulsive signals.**

***Index Terms*—Blind deconvolution, blind equalization, constant-modulus algorithm, impulsive signals, Sato's algorithm, super-Gaussian signals.**

## I. Introduction

BLIND deconvolution is a signal processing task that is important for several applications. In broadcast communications, blind deconvolution can be used to remove intersymbol interference (ISI) without devoting channel capacity to embedded training sequences [1]. In teleconferencing, blind deconvolution can be used to remove reverberation effects from speech signals for which no training signals are available [2]. Fig. 1 shows the general blind deconvolution task in which the source signal $s_t$ is filtered by an unknown transmission channel, which is denoted by $h(z)$ with unknown impulse response $h_t$, $0 \leq t < \infty$.

The received signal $x_t$, which is given by

$$x_t = \sum_{k=0}^{\infty} h_k s_{t-k} + n_t \tag{1}$$

contains mixtures of the source signal elements at numerous lags plus additive noise, which we will neglect for the moment. The goal is to "undo" the mixing process by a linear filtering operation, which is denoted by $w(z)$ in the figure. For simplicity,

Manuscript received November 27, 2001; revised January 17, 2003. The associate editor coordinating the review of this paper and approving it for publication was Prof. Nicholas D. Sidiropoulos.

H. Mathis is with the University of Applied Sciences, Rapperswil, Switzerland (e-mail: h.mathis@ieee.org).

S. C. Douglas is with the Department of Electrical Engineering, School of Engineering, Southern Methodist University, Dallas, TX 75275-0338 USA (e-mail: douglas@engr.smu.edu).

Digital Object Identifier 10.1109/TSP.2003.812836

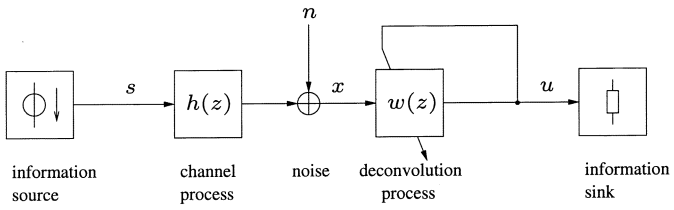



Fig. 1. Blind deconvolution system.

we will approximate the deconvolution filter by the FIR model

$$u_t = \sum_{k=0}^{L} w_k x_{t-k}. \tag{2}$$

The blind deconvolution task is solved if

$$u_t = \epsilon s_{t-d} \tag{3}$$

where $\epsilon$ is an unknown gain, and $d$ is an unknown integer delay. In such situations, the original source signal is recovered in the deconvolution system output, except for perhaps an arbitrary gain and delay. To solve the blind deconvolution task, an iterative algorithm is applied to the coefficients $\{w_k\}$ to adjust them to the vicinity of a deconvolving solution. In the sequel, we will employ a vector notation to represent certain filtering operations, such that $\boldsymbol{w}_t \triangleq [w_{0t}\ w_{1t}\ \cdots\ w_{Lt}]^T$ is the time-varying coefficient vector, $\boldsymbol{x}_t \triangleq [x_t\ x_{t-1}\ \cdots\ x_{t-L}]^T$ is the input signal vector, and

$$u_t = \boldsymbol{w}_t^T \boldsymbol{x}_t \tag{4}$$

is the deconvolved output signal at time $t$.

The Sato algorithm was one of the first blind deconvolution procedures to appear in the signal processing literature [3]. This algorithm remains quite popular today, having become known as the decision-directed algorithm for blind equalization of BPSK signals [4]. The coefficient updates for the Sato algorithm are

$$\boldsymbol{w}_{t+1} = \boldsymbol{w}_t + \mu(\gamma \text{sign}(u_t) - u_t)\boldsymbol{x}_t \tag{5}$$

where $\text{sign}(u_t)$ is one if $u_t$ is positive, zero if $u_t$ is zero, and $(-1)$ if $u_t$ is negative. The coefficient $\gamma$, which is known as the dispersion constant, controls the output scaling of the extracted signal and is normally computed as

$$\gamma \triangleq \frac{E\{s^2\}}{E\{|s|\}}. \tag{6}$$

The parameter $\mu$ controls the convergence performance of the algorithm and is normally chosen to be positive. The Sato algorithm can be viewed as a stochastic-gradient procedure for minimizing the cost function

$$\begin{aligned} J(\boldsymbol{w}_t) &= E\left\{(\gamma \text{sign}(u_t) - u_t)^2\right\} \\ &= E\left\{(|u_t| - \gamma)^2\right\}. \end{aligned} \tag{7}$$

The Sato algorithm can be seen as one version of the Bussgang family of blind deconvolution algorithms that also includes the

Godard and constant modulus algorithms (CMA) [5], [6]. Each of these algorithms can be written in the general form

$$\boldsymbol{w}_{t+1} = \boldsymbol{w}_t + \mu(g_1(u_t) - g_2(u_t))\boldsymbol{x}_t \tag{8}$$

where $g_1(u)$ and $g_2(u)$ are nonlinearities with specific, usually monomial, forms. It should be noted that the global convergence behavior of the Sato algorithm is largely unknown, whereas much is known about the constant modulus algorithm and its properties. In particular, CMA is known to be globally convergent to a valid equalizing solution for a doubly infinite equalizer filter for negative-kurtosis source sequences [16].

It is well-known that many Bussgang-type adaptive algorithms perform well when deconvolving communication signals. The ability of any Bussgang algorithm to perform deconvolution depends on the amplitude distribution of the source sequence and the choice of nonlinearities $g_1(u)$ and $g_2(u)$ within the coefficient updates. For impulsive signals, such as many common acoustic signals in seismic and audio signal processing, such methods often diverge, making them inappropriate for more general tasks such as acoustic dereverberation and geophysical exploration [2], [15]. The work in [15] is noteworthy in that Bussgang methods in batch form are described with specific application to impulsive signals. One possible remedy for sequential or online versions of this algorithm class has been suggested by Benveniste *et al.* [7]. Their approach employs a prewhitening stage followed by a unit-norm-constrained coefficient vector adaptation procedure. Such a solution complicates an otherwise simple adaptive procedure through additional processing stages. In [8], Shalvi and Weinstein study a constrained kurtosis maximization method for blind deconvolution and describe two different gradient adaptive algorithms for this task. The first method imposes unit-norm constraints on the coefficient vector and requires a divide and square root at each iteration to implement. The second method employs a modified cost function that combines second- and fourth-order criteria whose successful design depends on the sign of the source kurtosis. Both methods are globally convergent to an equalizing solution in the absence of noise; however, their performance, when applied to impulsive signals, is often poor due to the fourth-order moments used within the deconvolution criterion.

The use of unit-norm coefficient vector constraints for Bussgang blind deconvolution of impulsive signals has been proposed in [1] without an explicit analysis of why such constraints are useful for the task. This work also presents Bussgang algorithms employing optimum nonlinearities that have the general form

$$\boldsymbol{w}_{t+1} = \boldsymbol{w}_t + \mu\left(u_t + \frac{p_s'(u_t)}{p_s(u_t)}\right)\boldsymbol{x}_t \tag{9}$$

where $p_s(s)$ is the probability density function (pdf) of the source sequence $s_t$, and $p_s'(s)$ is its corresponding derivative. Such algorithms can be shown to be statistically efficient in their deconvolution capabilities [1] as long as they are locally stable about a deconvolving solution. To our knowledge, no simple algorithms for general-form (i.e., nonkurtosis-based) unconstrained Bussgang blind deconvolution of impulsive signals have been described in the scientific literature.

In this paper, we provide a theoretical justification as to why a class of unconstrained Bussgang blind deconvolution algorithms fail to deconvolve impulsive signals from their filtered measurements. Our results are quite general in that arbitrary monomial forms of $g_1(u)$ and $g_2(u)$ in (8) are considered so that both the Sato and constant modulus algorithms are included within this algorithm class. We then provide a novel modification of the Sato algorithm to enable it to deconvolve impulsive signals. Our modified version is computationally simple; it avoids signal prewhitening, divides, and square roots and increases the complexity of the already-simple Sato algorithm by a fixed number of multiplies and adds, independent of the equalizer length. We determine sufficient statistical conditions to guarantee local stability of the algorithm about a deconvolving solution. Computer simulations indicate that the modified Sato algorithm is more efficient in deconvolving Laplacian-distributed source signals than the unconstrained Shalvi–Weinstein algorithm in moderate SNR, and it provides nearly identical performance to a unit-norm-constrained Bussgang method without divides or square roots.

## II. Stability of Bussgang Algorithms

We will study the local stability properties of Bussgang algorithms to determine their capabilities in extracting signals with various distributions assuming a noise-free channel. Local convergence is often enough for tracking purposes of an adaptive algorithm. The global stability properties of general-form Bussgang algorithms is not considered and remains a challenging research topic. The procedure used for this analysis is the linearized ordinary differential equation (ODE) approach. For additional details on this method, see [9].

To analyze the local stability properties of (8), we express the deviations of the deconvolution coefficients from their optimal values as

$$w_k = w_{k,\mathrm{opt}} + \Delta_k. \tag{10}$$

The output of the deconvolution block

$$u_t = \sigma_u s_{t-d} + \sum_k \Delta_k x_{t-k} \tag{11}$$

is a delayed and scaled version of the original signal sequence perturbed by convolutive noise, where $\sigma_u$ allows for a different output energy. We will assume a two-sided infinitely long deconvolution impulse response in order to circumvent problems with filter truncation. The summation limits of the last section are therefore changed to $\pm\infty$. The perturbation vector $\boldsymbol{\Delta} \triangleq [\Delta_{-\infty}, \ldots, \Delta_0, \ldots, \Delta_\infty]^T$ now propagates similarly to (8) as

$$\begin{aligned}\boldsymbol{\Delta}_{t+1} &= \boldsymbol{\Delta}_t + \mu\left(g_1(u_t) - g_2(u_t)\right)\boldsymbol{x}_t \\ &= \boldsymbol{\Delta}_t + \mu\left(g_1\left(\sigma_u s_{t-d} + \sum_k \Delta_k x_{t-k}\right)\right. \\ &\quad \left. - g_2\left(\sigma_u s_{t-d} + \sum_k \Delta_k x_{t-k}\right)\right)\boldsymbol{x}_t \end{aligned} \tag{12}$$

where we have used (11). We now approximate the nonlinearities $g_1(\cdot)$ and $g_2(\cdot)$ by their respective first-order Taylor series about $\sigma_u s_{t-d}$:

$$g_{\{1,2\}}\left(\sigma_u s_{t-d} + \sum_k \Delta_k x_{t-k}\right) \approx g_{\{1,2\}}(\sigma_u s_{t-d}) + g'_{\{1,2\}}(\sigma_u s_{t-d}) \cdot \sum_k \Delta_k x_{t-k}. \quad (13)$$

Using (13) in (12) and taking the expectation on both sides yields

$$E\{\boldsymbol{\Delta}_{t+1}\} = E\{\boldsymbol{\Delta}_t\} + \mu E\{(g_1(\sigma_u s_{t-d}) - g_2(\sigma_u s_{t-d}))\boldsymbol{x}_t\} + \mu E\left\{(g'_1(\sigma_u s_{t-d}) - g'_2(\sigma_u s_{t-d}))\sum_k \Delta_k x_{t-k}\boldsymbol{x}_t\right\}. \quad (14)$$

At a stationary point of the update equation given by (8), we have

$$E\{(g_1(\sigma_u s_{t-d}) - g_2(\sigma_u s_{t-d}))\boldsymbol{x}_t\} = \mathbf{0}. \quad (15)$$

This condition also determines the output power so that we are left with the last term in (14), which is the difference of the derivatives, as the driving term. Employing the ODE method on the above update, we obtain

$$\frac{d\boldsymbol{\Delta}}{dt} = \mu \boldsymbol{A}\boldsymbol{\Delta}. \quad (16)$$

The elements of the transition matrix $\boldsymbol{A}$ are obtained by looking at

$$\frac{d\Delta_i}{dt} = \mu E\left\{(g'_1(\sigma_u s_{t-d}) - g'_2(\sigma_u s_{t-d}))\sum_k \Delta_k x_{t-k}x_{t-i}\right\} \quad (17)$$

$$= \mu E\left\{(g'_1(\sigma_u s_{t-d}) - g'_2(\sigma_u s_{t-d})) \cdot \sum_k \Delta_k \sum_j h_j s_{t-k-j}x_{t-i}\right\} \quad (18)$$

$$= \mu E\left\{(g'_1(\sigma_u s_{t-d}) - g'_2(\sigma_u s_{t-d})) \cdot \sum_k \Delta_k \sum_j h_j s_{t-k-j}\sum_l h_l s_{t-i-l}\right\}. \quad (19)$$

Thus, the matrix $\boldsymbol{A} \triangleq \boldsymbol{A}_1$ may be expressed as (20), shown at the bottom of the page, where the nonlinear moments are given by

$$\zeta_0 \triangleq E\left\{(g'_1(\sigma_u s) - g'_2(\sigma_u s))s^2\right\} \quad (21)$$

$$\zeta_1 \triangleq E\left\{g'_1(\sigma_u s) - g'_2(\sigma_u s)\right\}E\left\{|s|^2\right\} = E\left\{g'_1(\sigma_u s) - g'_2(\sigma_u s)\right\}. \quad (22)$$

The last equality in (22) stems from the unit-variance assumption. Alternatively, $\boldsymbol{A}_1$ may be written as

$$[A_1]_{ik} = \zeta_1 \sum_j h_{j+k-i}h_j + (\zeta_0 - \zeta_1)h_{d-i}h_{d-k} \quad i,k = -\infty, \ldots, 0, \ldots, \infty. \quad (23)$$

The matrix $\boldsymbol{A}$ is both infinite dimensional and symmetric. Symmetric matrices have real spectra. Thus, the corresponding ODE is locally stable if the spectrum of $\mu\boldsymbol{A}$ is always negative. By defining the channel matrix

$$\boldsymbol{H} \triangleq \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \iddots \\ \cdots & h_d & h_{d+1} & h_{d+2} & \cdots \\ \cdots & h_{d-1} & h_d & h_{d+1} & \cdots \\ \cdots & h_{d-2} & h_{d-1} & h_d & \cdots \\ \iddots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (24)$$

and the channel vector $\overline{\boldsymbol{h}} \triangleq [\ldots, h_{d+1}, h_d, h_{d-1}, \ldots]^T$, respectively, we can formulate $\boldsymbol{A}$ as a sum of an autocorrelation matrix and an outer product

$$\boldsymbol{A} = \zeta_1 \boldsymbol{H}^T\boldsymbol{H} + (\zeta_0 - \zeta_1)\overline{\boldsymbol{h}}\,\overline{\boldsymbol{h}}^T. \quad (25)$$

The bar in $\overline{\boldsymbol{h}}$ indicates the reverse order of its elements compared with the usual definition of a vector. As an infinite-dimensional matrix, $\boldsymbol{H}$ is Toeplitz with all its corresponding properties, and the eigenfunctions of $\boldsymbol{H}$ are given by the basis functions of the form $e^{-j\omega k}$ for each row of $\boldsymbol{H}$ [10]. Hence, if the channel transfer function has no spectral nulls, $\boldsymbol{H}$ is nonsingular, and thus, $\boldsymbol{H}^T\boldsymbol{H}$ is positive definite. The matrix $\overline{\boldsymbol{h}}\,\overline{\boldsymbol{h}}^T$ is positive semidefinite with a rank of one. Sufficient conditions for $\mu\boldsymbol{A}$ to be negative definite are therefore

$$\mu\zeta_1 < 0 \quad (26)$$

$$\mu\zeta_0 \leq \mu\zeta_1. \quad (27)$$

$$\boldsymbol{A}_1 = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \iddots \\ \cdots & \zeta_0 h_{d+1}^2 + \zeta_1 \sum_{j\neq d+1} h_j^2 & \zeta_0 h_d h_{d+1} + \zeta_1 \sum_{j\neq d} h_j h_{j+1} & \zeta_0 h_{d-1} h_{d+1} + \zeta_1 \sum_{j\neq d-1} h_j h_{j+2} & \cdots \\ \cdots & \zeta_0 h_d h_{d+1} + \zeta_1 \sum_{j\neq d+1} h_{j-1} h_j & \zeta_0 h_d^2 + \zeta_1 \sum_{j\neq d} h_j^2 & \zeta_0 h_{d-1} h_d + \zeta_1 \sum_{j\neq d-1} h_j h_{j+1} & \cdots \\ \cdots & \zeta_0 h_{d-1} h_{d+1} + \zeta_1 \sum_{j\neq d+1} h_{j-2} h_j & \zeta_0 h_{d-1} h_d + \zeta_1 \sum_{j\neq d} h_{j-1} h_j & \zeta_0 h_{d-1}^2 + \zeta_1 \sum_{j\neq d-1} h_j^2 & \cdots \\ \iddots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (20)$$

Many Bussgang-type algorithms can be written as a difference of two polynomials

$$\boldsymbol{w}_{t+1} = \boldsymbol{w}_t + \mu(a\text{sign}(u_t)|u_t|^p - \text{sign}(u_t)|u_t|^q)\boldsymbol{x}_t \quad (28)$$

which is simply a special form of (8), with

$$g_1(u) \triangleq a\text{sign}(u)|u|^p \quad (29)$$

$$g_2(u) \triangleq \text{sign}(u)|u|^q. \quad (30)$$

Note that the scaling factor $a$ allows an adjustment of the output power at the stationary point. The respective nonlinear moments for some known Bussgang equalizers are summarized in Tables I (general output power) and II (normalized output power), respectively. The output power $\sigma_u^2$ can be obtained from the evaluation of the stationary point, which is given by the third column of Table I. The resulting conditions for stability are given in Table III.

In the following, we are going to show that for Bussgang-type algorithms of the form given by (28), sufficient stability conditions can be decoupled with respect to $\zeta_0$ and $\zeta_1$. To this end, we need the following result from Loève [11]:

**LEMMA 1**

*For the higher-order moments of a random variable $x$ we note that*

$$E\{|x|^{a_1}\}E\{|x|^{b-a_1}\} \geq E\{|x|^{a_2}\}E\{|x|^{b-a_2}\}$$

*for all* $0 < a_1 < a_2 \leq b/2$. (31)

A proof of this can be found in [12, p. 343]. Using Lemma 1 allows us to derive the following result:

**COROLLARY 1**

*The ratio* $\dfrac{E\{|x|^{m+1}\}}{E\{|x|^{m-1}\}}$ *is a nondecreasing function in $m$ for any distribution* $p_x(.)$.

*Proof:* By setting $b \triangleq 2m+\delta$, $a_1 \triangleq m-1$, $a_2 \triangleq m+\delta-1$, with $\delta$ being a positive real number, $0 < \delta \leq 2$, and rearranging (31), we get

$$\frac{E\{|x|^{m+\delta+1}\}}{E\{|x|^{m+\delta-1}\}} \geq \frac{E\{|x|^{m+1}\}}{E\{|x|^{m-1}\}} \quad (32)$$

from which we conclude that $E\{|x|^{m+1}\}/E\{|x|^{m-1}\}$ is a nondecreasing function in $m$. □

We are now ready to show the following:

**LEMMA 2**

*For a Bussgang-type equalizer in the form of a difference of two monomials according to (28) sufficient decoupled stability conditions on the nonlinear moments are*

$$\mu\zeta_1 < 0 \quad (33)$$

$$\mu\zeta_0 \leq 0. \quad (34)$$

*Proof:* For the nonlinear moments of (28), we get, by using (21) and (22) and still assuming unit-variance source signals

$$\zeta_0 = ap\sigma_u^{p-1}E\{|s|^{p+1}\} - q\sigma_u^{q-1}E\{|s|^{q+1}\} \quad (35)$$

$$\zeta_1 = ap\sigma_u^{p-1}E\{|s|^{p-1}\} - q\sigma_u^{q-1}E\{|s|^{q-1}\}. \quad (36)$$

From the evaluation of the stationary point of (28), we get

$$\sigma_u = \left(a\frac{E\{|s|^{p+1}\}}{E\{|s|^{q+1}\}}\right)^{1/(q-p)}. \quad (37)$$

Thus, using (37) in (35) and (36), we have

$$\zeta_0 = a^{(q-1)/(q-p)}\frac{E\{|s|^{p+1}\}^{(q-1)/(q-p)}}{E\{|s|^{q+1}\}^{(p-1)/(q-p)}}(p-q) \quad (38)$$

$$\zeta_1 = a^{(q-1)/(q-p)}\frac{E\{|s|^{p+1}\}^{(q-1)/(q-p)}}{E\{|s|^{q+1}\}^{(p-1)/(q-p)}} \cdot \left(p\frac{E\{|s|^{p-1}\}}{E\{|s|^{p+1}\}} - q\frac{E\{|s|^{q-1}\}}{E\{|s|^{q+1}\}}\right) \quad (39)$$

as already listed in Table I. The scaling factor $a$ must be positive; otherwise, no equilibrium can be reached. Assume for the time being that $\mu > 0$. In order for $\zeta_0$ to be smaller than zero, we must choose $p < q$. Inside the bracket term of (39), $p$ gets multiplied by an equal or higher value than $q$ due to Corollary 1, which implies that $E\{|s|^{m-1}\}/E\{|s|^{m+1}\}$ is a nonincreasing function in $m$ (reciprocal). Thus, we have

$$\mu\zeta_0 \leq \mu\zeta_1. \quad (40)$$

A similar argument can be made for negative $\mu$; we end up with the same inequality (40). □

## III. GENERALIZED GAUSSIAN DISTRIBUTIONS

Now, let us have a closer look at the algorithms provided in Table I. The first three algorithms are known to be stable update equations for sub-Gaussian signals. Sub-Gaussian signals are a class of distributions whose pdf is of the form

$$p_s(s) = \frac{\alpha}{2\beta\Gamma\left(\frac{1}{\alpha}\right)}e^{-(|s|/\beta)^\alpha} \quad (41)$$

with $\alpha > 2$ and $\beta$ some constant. On the other hand, signal distributions given by (41) with $0 < \alpha < 2$ are called super-Gaussian. We now prove that for such distributions, e.g., Laplacian distributed signals ($\alpha = 1$), Bussgang-type algorithms fail to provide local stability. Our proof uses the following intermediate result:

**LEMMA 3**

*For $0 < \alpha < 2$ and $p < q$ the ratio of Gamma functions behaves as*

$$\frac{\Gamma\left(\frac{p+2}{\alpha}\right)}{p\Gamma\left(\frac{p}{\alpha}\right)} < \frac{\Gamma\left(\frac{q+2}{\alpha}\right)}{q\Gamma\left(\frac{q}{\alpha}\right)}. \quad (42)$$

TABLE I
NONLINEAR MOMENTS FOR DIFFERENT BUSSGANG-TYPE DECONVOLUTION ALGORITHMS. ASSUMPTION: $E\{|s|^2\} = 1$

| alg. | update term | at stationary point | $\zeta_0$ | $\zeta_1$ |
|---|---|---|---|---|
| Sato | $+\mu\,(\gamma\,\mathrm{sign}(u) - u)\,x$ | $\gamma E\{|s|\} - \sigma_u = 0$<br>$\sigma_u = \gamma E\{|s|\}$ | $-1$ | $2\frac{p_s(0)}{E\{|s|\}} - 1$ |
| CMA | $+\mu\,(R_2 u - u^3)\,x$ | $R_2\sigma_u - \sigma_u^3 E\{|s|^4\} = 0$<br>$\sigma_u = \sqrt{\frac{R_2}{E\{|s|^4\}}}$ | $-2R_2$ | $R_2\left(1 - \frac{3}{E\{|s|^4\}}\right)$ |
| Godard | $+\mu\,(R_p u^{p-1} - u^{2p-1})\,x$ | $R_p\sigma_u^{p-1}E\{|s|^p\}$<br>$-\sigma_u^{2p-1}E\{|s|^{2p}\} = 0$<br>$\sigma_u = \sqrt[p]{R_p\frac{E\{|s|^p\}}{E\{|s|^{2p}\}}}$ | $-pR_p^{\frac{2p-2}{p}}\cdot$<br>$\cdot\frac{E\{|s|^p\}^{\frac{2p-2}{p}}}{E\{|s|^{2p}\}^{\frac{p-2}{p}}}$ | $R_p^{\frac{2p-2}{p}}\frac{E\{|s|^p\}^{\frac{2p-2}{p}}}{E\{|s|^{2p}\}^{\frac{p-2}{p}}}\cdot$<br>$\cdot\left((p-1)\frac{E\{|s|^{p-2}\}}{E\{|s|^p\}}\right.$<br>$\left.-(2p-1)\frac{E\{|s|^{2p-2}\}}{E\{|s|^{2p}\}}\right)$ |
| general | $+\mu(a\,\mathrm{sign}(u)\lvert u\rvert^p$<br>$-\mathrm{sign}(u)\lvert u\rvert^q)x$ | $a\sigma_u^p E\{|s|^{p+1}\}$<br>$-\sigma_u^q E\{|s|^{q+1}\} = 0$<br>$\sigma_u = \left(a\frac{E\{|s|^{p+1}\}}{E\{|s|^{q+1}\}}\right)^{\frac{1}{q-p}}$ | $(p-q)a^{\frac{q-1}{q-p}}\cdot$<br>$\cdot\frac{E\{|s|^{p+1}\}^{\frac{q-1}{q-p}}}{E\{|s|^{q+1}\}^{\frac{p-1}{q-p}}}$ | $a^{\frac{q-1}{q-p}}\frac{E\{|s|^{p+1}\}^{\frac{q-1}{q-p}}}{E\{|s|^{q+1}\}^{\frac{p-1}{q-p}}}\cdot$<br>$\cdot\left(p\frac{E\{|s|^{p-1}\}}{E\{|s|^{p+1}\}} - q\frac{E\{|s|^{q-1}\}}{E\{|s|^{q+1}\}}\right)$ |

TABLE II
NONLINEAR MOMENTS FOR DIFFERENT BUSSGANG-TYPE DECONVOLUTION ALGORITHMS. ASSUMPTION: $E\{|s|^2\} = 1$ AND PROPER CHOICE OF SCALING FACTOR RESULTING IN $\sigma_u^2 = 1$

| algorithm | scaling constant | $\zeta_0$ | $\zeta_1$ |
|---|---|---|---|
| Sato | $\gamma = \frac{1}{E\{|s|\}}$ | $-1$ | $2\frac{p_s(0)}{E\{|s|\}} - 1$ |
| CMA | $R_2 = E\{|s|^4\}$ | $-2E\{|s|^4\}$ | $E\{|s|^4\} - 3$ |
| Godard | $R_p = \frac{E\{|s|^{2p}\}}{E\{|s|^p\}}$ | $-pE\{|s|^{2p}\}$ | $(p-1)\frac{E\{|s|^{2p}\}}{E\{|s|^p\}}E\{|s|^{p-2}\} - (2p-1)E\{|s|^{2p-2}\}$ |
| general | $a = \frac{E\{|s|^{q+1}\}}{E\{|s|^{p+1}\}}$ | $(p-q)E\{|s|^{q+1}\}$ | $p\frac{E\{|s|^{q+1}\}}{E\{|s|^{p+1}\}}E\{|s|^{p-1}\} - qE\{|s|^{q-1}\}$ |

TABLE III
STABILITY CONDITIONS FOR BUSSGANG-TYPE DECONVOLUTION ALGORITHMS

| algorithm | stability conditions |
|---|---|
| Sato | $\mu > 0, \quad 2\gamma p_u(0) < 1$ |
| CMA | $\mu > 0, \quad \kappa_4 < 0$ |
| Godard | $\mu > 0, \quad \zeta_1 < 0$ |
| monomials | $\mu\zeta_0 < 0, \quad \mu\zeta_1 < 0$ |

*Proof:* We have to show that for

$$f(m) \triangleq \frac{\Gamma\left(\frac{m+2}{\alpha}\right)}{m\Gamma\left(\frac{m}{\alpha}\right)} \tag{43}$$

we have

$$f'(m) \triangleq \frac{df(m)}{dm} > 0, \qquad 0 < \alpha < 2. \tag{44}$$

We define

$$\delta \triangleq \frac{2-\alpha}{\alpha} > 0, \qquad 0 < \alpha < 2 \tag{45}$$

and make use of the recursive property of the Gamma function to note that

$$\Gamma\left(\frac{m+2}{\alpha}\right) = \Gamma\left(\frac{m}{\alpha} + \delta + 1\right) = \left(\frac{m}{\alpha} + \delta\right)\Gamma\left(\frac{m}{\alpha} + \delta\right). \tag{46}$$

Using (46) in (43) and taking the derivative with respect to $m$ leads to (47), shown at the bottom of the next page. Using (cf. [13])

$$\Gamma'(x) = \xi(x)\Gamma(x) \tag{48}$$

with

$$\xi(x) \triangleq -C - \frac{1}{x} + \sum_{n=1}^{\infty} \frac{x}{n(x+n)} \tag{49}$$

where $C$ is Euler's constant, we can write

$$f'(m) = \frac{\Gamma\left(\frac{m}{\alpha}+\delta\right)}{m^2\Gamma\left(\frac{m}{\alpha}\right)} \cdot \left(\frac{m}{\alpha}\left(\frac{m}{\alpha}+\delta\right)\left(\xi\left(\frac{m}{\alpha}+\delta\right)-\xi\left(\frac{m}{\alpha}\right)\right)-\delta\right) \tag{50}$$

which leads after some calculation and using (49) to

$$f'(m) = \frac{\Gamma\left(\frac{m}{\alpha}+\delta\right)}{m^2\Gamma\left(\frac{m}{\alpha}\right)} \cdot \frac{m^2+\delta\alpha m}{\alpha^2} \cdot \left(\sum_{n=1}^{\infty} \frac{\frac{m}{\alpha}+\delta}{n\left(\frac{m}{\alpha}+\delta+n\right)} - \sum_{n=1}^{\infty} \frac{\frac{m}{\alpha}}{n\left(\frac{m}{\alpha}+n\right)}\right). \tag{51}$$

The first two factors of (51) are positive for $0 < \alpha < 2$, and the difference of the two sums is

$$\sum_{n=1}^{\infty} \frac{\frac{m}{\alpha}+\delta}{n\left(\frac{m}{\alpha}+\delta+n\right)} - \sum_{n=1}^{\infty} \frac{\frac{m}{\alpha}}{n\left(\frac{m}{\alpha}+n\right)} = \sum_{n=1}^{\infty} \frac{\delta}{\left(\frac{m}{\alpha}+\delta+n\right)\left(\frac{m}{\alpha}+n\right)} > 0 \tag{52}$$

making $f'(m)$ positive for all $m > 0$. □

We can now show that Bussgang-type algorithms generally do not work for super-Gaussian signals.

**LEMMA 4**

*No blind Bussgang-type algorithm of the general form*

$$\boldsymbol{w}_{t+1} = \boldsymbol{w}_t + \mu(a\text{sign}(u_t)|u_t|^p - \text{sign}(u_t)|u_t|^q)\boldsymbol{x}_t \tag{53}$$

*can be stabilized by any choice of p and q and any super-Gaussian distribution.*

*Proof:* Due to Lemma 2, the nonlinear moments $\zeta_0$ and $\zeta_1$ need to have the same sign. Then, stability can be controlled by proper choice of the sign of the step size $\mu$. Without loss of generality, we assume that $p < q$, which makes $\zeta_0 < 0$, as can be seen from Table I, but due to Lemma 3, we note that

$$\frac{p\Gamma\left(\frac{p}{\alpha}\right)}{\Gamma\left(\frac{p+2}{\alpha}\right)} - \frac{q\Gamma\left(\frac{q}{\alpha}\right)}{\Gamma\left(\frac{q+2}{\alpha}\right)} > 0 \tag{54}$$

making $\zeta_1 > 0$ (see again Table I), thus enforcing a different sign of $\zeta_1$. □

## IV. ALGORITHM MODIFICATIONS

Given that Bussgang-type algorithms fail to deconvolve impulsive signals, a simple modification to such procedures that enables their use would be highly desirable. Our goal is to modify the coefficient updates to drive the length of the coefficient vector $\boldsymbol{w}_t$ to a constant value, where this constant can be potentially source or channel dependent. We propose to modify the Bussgang update in (8) by inserting terms proportional to the squared $L_2$-norm of the coefficient vector within the updates

$$\boldsymbol{w}_{t+1} = \boldsymbol{w}_t + \mu(||\boldsymbol{w}_t||^{\tilde{p}} g_1(u_t) - ||\boldsymbol{w}_t||^{\tilde{q}} g_2(u_t))\boldsymbol{x}_t \tag{55}$$

where $||\boldsymbol{w}||^2 = \boldsymbol{w}^T\boldsymbol{w}$ and $\tilde{p}$ and $\tilde{q}$ are even-valued integers. Note that such modifications do not significantly increase the complexity of the update. In fact, it is straightforward to compute a scalar-only update to $||\boldsymbol{w}_t||^2$ as

$$||\boldsymbol{w}_{t+1}||^2 = ||\boldsymbol{w}_t||^2 + 2\mu(||\boldsymbol{w}_t||^{\tilde{p}} g_1(u_t) - ||\boldsymbol{w}_t||^{\tilde{q}} g_2(u_t))u_t + \mu^2(||\boldsymbol{w}_t||^{\tilde{p}} g_1(u_t) - ||\boldsymbol{w}_t||^{\tilde{q}} g_2(u_t))^2 r_t \tag{56}$$

where $r_t$ can be updated as

$$r_t = r_{t-1} + x_t^2 - x_{t-L-1}^2. \tag{57}$$

Hence, complicated procedures such as prewhitening and divides are avoided. The design goal is to choose $\tilde{p}$ and $\tilde{q}$ so that the corresponding algorithm is locally stable. Performing this design requires us to repeat our analysis procedure for (55).

By approximating

$$||\boldsymbol{w}+\boldsymbol{\Delta}||^{\tilde{p}} = \left(\sum_k (w_k+\Delta_k)^2\right)^{\tilde{p}/2} \approx ||\boldsymbol{w}||^{\tilde{p}} + \tilde{p}||\boldsymbol{w}||^{\tilde{p}-2}\sum_k w_k\Delta_k \tag{58}$$

the first-order Taylor series approximation of the scalar part of the update term becomes

$$\begin{aligned} &||\boldsymbol{w}||^{\tilde{p}} g_1\left(\sigma_u s_{t-d} + \sum_i \Delta_i x_{t-i}\right) - ||\boldsymbol{w}||^{\tilde{q}} g_2\left(\sigma_u s_{t-d} + \sum_i \Delta_i x_{t-i}\right) \\ &\approx ||\boldsymbol{w}||^{\tilde{p}} g_1(\sigma_u s_{t-d}) + ||\boldsymbol{w}||^{\tilde{p}} g_1'(\sigma_u s_{t-d})\sum_i \Delta_i x_{t-i} \\ &\quad + \tilde{p}||\boldsymbol{w}||^{\tilde{p}-2}\sum_i \Delta_i w_i g_1(\sigma_u s_{t-d}) \\ &\quad - ||\boldsymbol{w}||^{\tilde{q}} g_2(\sigma_u s_{t-d}) - ||\boldsymbol{w}||^{\tilde{q}} g_2'(\sigma_u s_{t-d})\sum_i \Delta_i x_{t-i} \\ &\quad - \tilde{q}||\boldsymbol{w}||^{\tilde{q}-2}\sum_i \Delta_i w_i g_2(\sigma_u s_{t-d}). \end{aligned} \tag{59}$$

---

$$f'(m) = \left(\frac{\left(\frac{m}{\alpha}+\delta\right)\Gamma\left(\frac{m}{\alpha}+\delta\right)}{m\Gamma\left(\frac{m}{\alpha}\right)}\right)' = \frac{\left(\Gamma\left(\frac{m}{\alpha}+\delta\right)+\left(\frac{m}{\alpha}+\delta\right)\Gamma'\left(\frac{m}{\alpha}+\delta\right)\right)\frac{m}{\alpha}\Gamma\left(\frac{m}{\alpha}\right) - \left(\Gamma\left(\frac{m}{\alpha}\right)+\frac{m}{\alpha}\Gamma'\left(\frac{m}{\alpha}\right)\right)\left(\frac{m}{\alpha}+\delta\right)\Gamma\left(\frac{m}{\alpha}+\delta\right)}{m^2\Gamma^2\left(\frac{m}{\alpha}\right)} \tag{47}$$

The additional factors $||\boldsymbol{w}||^{\tilde{p}}$ and $||\boldsymbol{w}||^{\tilde{q}}$ in (55) shift the stationary point of this update such that the output power $\sigma_u^2$ depends on $\tilde{p}$ and $\tilde{q}$. At the stationary point, we get

$$E\left\{\left(||\boldsymbol{w}||^{\tilde{p}} g_1(\sigma_u s_{t-d}) - ||\boldsymbol{w}||^{\tilde{q}} g_2(\sigma_u s_{t-d})\right)\boldsymbol{x}_t\right\} = \boldsymbol{0} \quad (60)$$

where the power of the output signal is given by the combined transfer function of the channel, and the equalizer as

$$\sigma_u^2 = ||w * h||^2 \sigma_s^2 = ||w * h||^2. \quad (61)$$

Equations (60) and (61) define a system of equations for the two unknown $\sigma_u$ and $||\boldsymbol{w}||^2$. Going through a similar analysis as before, we obtain the ODE

$$\frac{d\boldsymbol{\Delta}}{dt} = \mu \boldsymbol{A}\boldsymbol{\Delta} = \mu(\boldsymbol{A}_1 + \boldsymbol{A}_2)\boldsymbol{\Delta} \quad (62)$$

where $\boldsymbol{A}_1$ is given by (20), and

$$\boldsymbol{A}_2 = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \iddots \\ \cdots & \tilde{\zeta}_1 h_{d+1} w_{-1} & \tilde{\zeta}_1 h_{d+1} w_0 & \tilde{\zeta}_1 h_{d+1} w_1 & \cdots \\ \cdots & \tilde{\zeta}_1 h_d w_{-1} & \tilde{\zeta}_1 h_d w_0 & \tilde{\zeta}_1 h_d w_1 & \cdots \\ \cdots & \tilde{\zeta}_1 h_{d-1} w_{-1} & \tilde{\zeta}_1 h_{d-1} w_0 & \tilde{\zeta}_1 h_{d-1} w_1 & \cdots \\ \iddots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (63)$$

The nonlinear moments in $\boldsymbol{A}_1$ are now different from the ones given by (21) and (22)

$$\zeta_0 \triangleq ||\boldsymbol{w}||^{\tilde{p}} E\left\{g_1'(\sigma_u s)s^2\right\} - ||\boldsymbol{w}||^{\tilde{q}} E\left\{g_2'(\sigma_u s)s^2\right\} \quad (64)$$

$$\zeta_1 \triangleq ||\boldsymbol{w}||^{\tilde{p}} E\left\{g_1'(\sigma_u s)\right\} - ||\boldsymbol{w}||^{\tilde{q}} E\left\{g_2'(\sigma_u s)\right\}. \quad (65)$$

Furthermore, we have introduced

$$\tilde{\zeta}_1 \triangleq \tilde{p}||\boldsymbol{w}||^{\tilde{p}-2} E\left\{g_1(\sigma_u s)s\right\} - \tilde{q}||\boldsymbol{w}||^{\tilde{q}-2} E\left\{g_2(\sigma_u s)s\right\}. \quad (66)$$

The short-hand notation leads to

$$[A]_{ik} = \zeta_1 \left[\sum_j h_{j+k-i} h_j\right]_{ik} + (\zeta_0 - \zeta_1)\left[h_{d-i} h_{d-k}\right]_{ik} + \tilde{\zeta}_1 \left[h_{d-i} w_k\right]_{ik}, \qquad i, k = -\infty, \ldots, 0, \ldots, \infty. \quad (67)$$

Although $\boldsymbol{A}_1$ is still a symmetric matrix, $\boldsymbol{A}$ is no longer symmetric. Assigning the equalizer vector $\boldsymbol{w} \triangleq [\ldots, w_{-1}, w_0, w_1, \ldots]^T$, we can write $\boldsymbol{A}$ as

$$\boldsymbol{A} = \zeta_1 \boldsymbol{H}^T \boldsymbol{H} + (\zeta_0 - \zeta_1)\overline{\boldsymbol{h}}\,\overline{\boldsymbol{h}}^T + \tilde{\zeta}_1 \overline{\boldsymbol{h}}\,\boldsymbol{w}^T \quad (68)$$

$$= \zeta_1 \boldsymbol{H}^T \boldsymbol{H} + \overline{\boldsymbol{h}}\,\tilde{\boldsymbol{h}}^T \quad (69)$$

where we have defined

$$\tilde{\boldsymbol{h}} \triangleq (\zeta_0 - \zeta_1)\overline{\boldsymbol{h}} + \tilde{\zeta}_1 \boldsymbol{w}. \quad (70)$$

Although the two last terms of (68) contain two outer products, the sum has still rank one and, thus, has only one nonzero spectra owing to the outer products sharing the same column vector. Hence, $\overline{\boldsymbol{h}}\tilde{\boldsymbol{h}}^T$ is either positive semidefinite or negative semidefinite with an impulsive spectrum. Its definiteness depends on

$$\lambda = \mathrm{tr}\left(\overline{\boldsymbol{h}}\,\tilde{\boldsymbol{h}}^T\right) = \overline{\boldsymbol{h}}^T \tilde{\boldsymbol{h}}. \quad (71)$$

The first part of (62) requires

$$\mu\zeta_1 < 0. \quad (72)$$

For a negative semidefinite expression for the second part of (62), we must ensure that

$$\mu\overline{\boldsymbol{h}}^T\tilde{\boldsymbol{h}} = \mu\overline{\boldsymbol{h}}^T((\zeta_0 - \zeta_1)\overline{\boldsymbol{h}} + \tilde{\zeta}_1\boldsymbol{w} \leq 0. \quad (73)$$

Without loss of generality, we can restrict the channel response to unit energy, and thus, $\overline{\boldsymbol{h}}^T\overline{\boldsymbol{h}} = 1$. We designate $\boldsymbol{w}_{\mathrm{opt}}$ as the solution of the equalizer coefficients that perfectly inverts the channel $\overline{\boldsymbol{h}}$, i.e., $\overline{\boldsymbol{h}}^T\boldsymbol{w}_{\mathrm{opt}} = 1$, so that the output power is $\sigma_u^2 = ||h * w_{\mathrm{opt}}||^2 = 1$. Because we have added a norm factor in the update equation, we can no longer expect the output power $\sigma_u^2$ to be controllable purely by adjusting the gain parameter $a$. Moreover, $\sigma_u^2$ depends on the channel impulse response or, more accurately, on the norm of the equalizer impulse response that inverts this channel. The equalizer taps are adjusted to a scaled version of $\boldsymbol{w}_{\mathrm{opt}}$

$$\boldsymbol{w} = \sigma_u \boldsymbol{w}_{\mathrm{opt}}. \quad (74)$$

Subsequently, we also have

$$||\boldsymbol{w}||^m = \sigma_u^m ||\boldsymbol{w}_{\mathrm{opt}}||^m. \quad (75)$$

Therefore, the convolution of the channel impulse response and the equalizer impulse response yields

$$\overline{\boldsymbol{h}}^T\boldsymbol{w} = \sigma_u. \quad (76)$$

Hence, the second condition for negative definiteness of $\mu\boldsymbol{A}$ is

$$\mu(\zeta_0 - \zeta_1 + \tilde{\zeta}_1\sigma_u) \leq 0. \quad (77)$$

For a modified version of an algorithm according to (55), (72) and (77) are sufficient conditions for local stability. In summary, the first stability condition, which is given by (72), stays the same for the algorithm modification, although the nonlinear moment $\zeta_1$ changes. We can satisfy (72) by a proper choice of the sign of the step size $\mu$. If this choice results in the failure of the previous second stability condition as given by (34), we have now an additional term $\tilde{\zeta}_1$ in (77) to stabilize the algorithm.

## V. Example: Modified Sato Algorithm

In the following, we look at a specific design example for a modified Sato algorithm in more detail. More examples and modifications for general Bussgang-type algorithms are given in [14].

Suppose that the first stability condition of the Sato algorithm in (26), i.e., $\mu(2p_s(0)/E\{|s|\} - 1) \geq 0$, is not satisfied for a particular signal distribution and any positive value of $\mu$. Clearly, we can satisfy this condition by choosing a negative $\mu$. Such a choice causes the norm $||\boldsymbol{w}_t||$ to fail to converge since the second condition, i.e., (27), is no longer satisfied. We can modify Sato's algorithm by appending the norm factor $||\boldsymbol{w}||^{\tilde{p}}$ to the update equation to get

$$\boldsymbol{w}_{t+1} = \boldsymbol{w}_t + \mu(||\boldsymbol{w}||^{\tilde{p}}\gamma\,\mathrm{sign}(u_t) - u_t)\boldsymbol{x}_t. \quad (78)$$

By determining the stationary point of (78) and using (75), we obtain

$$\sigma_u = \|\boldsymbol{w}_{\text{opt}}\|^{\tilde{p}/(1-\tilde{p})} \left(\gamma E\left\{|s|\right\}\right)^{1/(1-\tilde{p})} \tag{79}$$

and the norm relation

$$\|\boldsymbol{w}\|^m = \|\boldsymbol{w}_{\text{opt}}\|^{m/(1-\tilde{p})} \left(\gamma E\left\{|s|\right\}\right)^{m/(1-\tilde{p})}. \tag{80}$$

In order to test this modified algorithm for local stability, we need to compute the nonlinear moments. Using the definitions in the previous sections and the fact that the mode of a pdf is inversely proportional to its standard deviation, we get

$$\zeta_0 = -1 \tag{81}$$

$$\begin{aligned}\zeta_1 &= \|\boldsymbol{w}\|^{\tilde{p}} 2\gamma p_u(0) - 1 \\ &= \|\boldsymbol{w}\|^{\tilde{p}} 2\gamma \frac{p_s(0)}{\sigma_u} - 1 \\ &= \|\boldsymbol{w}_{\text{opt}}\|^{\tilde{p}/(1-\tilde{p})} \left(\gamma E\left\{|s|\right\}\right)^{\tilde{p}/(1-\tilde{p})} \\ &\quad \cdot 2\gamma \frac{p_s(0)}{\|\boldsymbol{w}_{\text{opt}}\|^{\tilde{p}/(1-\tilde{p})} \left(\gamma E\left\{|s|\right\}\right)^{1/(1-\tilde{p})}} - 1 \\ &= \frac{2p_s(0)}{E\left\{|s|\right\}} - 1\end{aligned} \tag{82}$$

$$\begin{aligned}\tilde{\zeta}_1 \sigma_u &= \tilde{p}\|\boldsymbol{w}\|^{\tilde{p}-2} \gamma E\left\{|s|\right\} \sigma_u \\ &= \tilde{p}\|\boldsymbol{w}_{\text{opt}}\|^{(\tilde{p}-2)/(1-\tilde{p})} \left(\gamma E\left\{|s|\right\}\right)^{(\tilde{p}-2)/(1-\tilde{p})} \\ &\quad \cdot \gamma E\left\{|s|\right\} \|\boldsymbol{w}_{\text{opt}}\|^{\tilde{p}/(1-\tilde{p})} \left(\gamma E\left\{|s|\right\}\right)^{1/(1-\tilde{p})} \\ &= \tilde{p}\|\boldsymbol{w}_{\text{opt}}\|^{-2}.\end{aligned} \tag{83}$$

For negative $\mu$, the stability condition in (72) now means

$$\frac{2p_s(0)}{E\left\{|s|\right\}} > 1 \tag{84}$$

which is exactly the contrary of the stability condition of the original Sato algorithm, whereas the stability condition in (77) can be written as

$$\|\boldsymbol{w}_{\text{opt}}\|^2 \leq \tilde{p} \frac{E\left\{|s|\right\}}{2p_s(0)}. \tag{85}$$

These conditions allow for the deconvolution of super-Gaussian signals as long as the optimum coefficient vector is not overly long. It should be noted that these stability conditions are sufficient but not necessary due to our use of overly stringent conditions on the components of matrix $\boldsymbol{A}$. Moreover, the results indicate that some knowledge of the source sequence distribution is required to design the proper algorithm. This situation is similar in spirit, but not exactly identical, to the cost function design method used in [8] to obtain a globally convergent blind deconvolution method without using unit-norm equalizer filter length constraints.

## VI. Computer Simulations

### A. Channel Model

The channel model used in this section for the computer simulation of a modified Bussgang-type algorithm for a Laplacian distributed source signal is a seven-tap FIR filter with the impulse response $h(z) = 0.4 + z^{-1} - 0.7z^{-2} + 0.6z^{-3} + 0.3z^{-4} - 0.4z^{-5} + 0.1z^{-6}$. Its properties are depicted in Fig. 2.

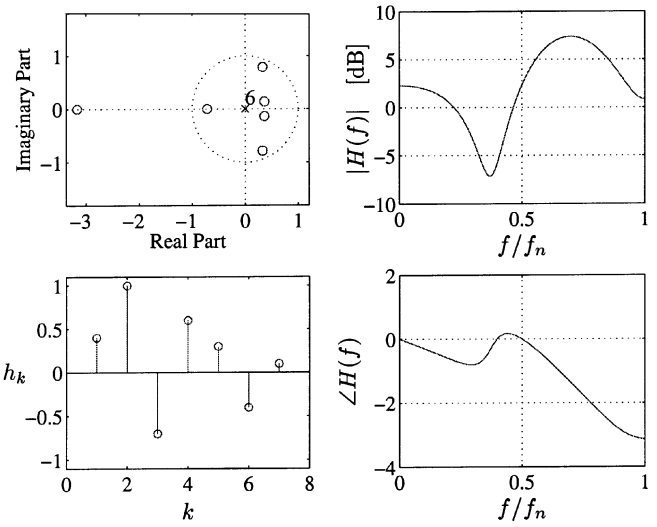



Fig. 2. Channel model for simulation of extended Sato algorithm. (Top left) Zero-pole plot. (Top right) Frequency response (magnitude). (Bottom left) Impulse response. (Bottom right) Frequency response (phase).

### B. Simulation Results

Computer simulations have been carried out to verify the stable behavior of the modified Sato algorithm for deconvolving super-Gaussian, in this case Laplacian, signals. It is important to note that in this case, the modified Sato algorithm resembles the Bussgang algorithm with optimum nonlinearity given by (9), where $-p_s'(u)/p_s(u)$ is proportional to $\text{sgn}(u)$ for the Laplacian distribution. For this reason, we expect our modified Sato algorithm to outperform methods that are based on other criteria, such as kurtosis-maximization methods, in terms of statistical efficiency. Four different simulation types have been performed, as we now describe.

In the first simulation, we compared the original Sato algorithm with the sign of update term reversed, the modified Sato algorithm proposed in this paper, and the unconstrained Shalvi–Weinstein algorithm (SWA) for one typical run when moderate channel noise is present, resulting in a received signal-to-noise-ratio of SNR $=$ 30 dB. The step size was chosen as $\mu = 0.0006$ for both Sato algorithms, and the step size for the SWA algorithm was chosen as $\mu = 0.000\,17$. A 24-tap FIR filter was chosen as the blind adaptive equalizer filter. The exponents of the norms in the modified Sato algorithm were chosen as $\tilde{p} = 2$, $\tilde{q} = 0$. The ISI measurement, which is defined as

$$J_{\text{ISI}}(c(z)) = \frac{\sum_{k=0}^{3L} |c_k|^2}{\max_k |c_k|^2} - 1 \tag{86}$$

where $c_k = w_k * h_k$ is the combined channel-plus-equalizer impulse response, was then computed for all algorithms at each iteration. The results are shown in the top plot of Fig. 4. The original Sato algorithm, whose performance is not shown on this figure, fails to converge, as would be expected for this type of

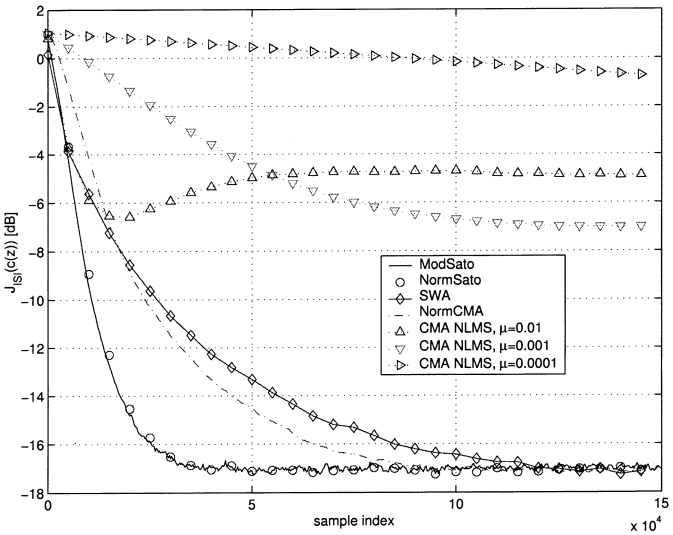



Fig. 3. Convergence comparison of the five blind deconvolution algorithms for a noise-free channel.

distribution, whereas the modified Sato algorithm converges as predicted by the theory. The SWA algorithm also converges, but its convergence speed is slower than that of our proposed Sato algorithm for a similar residual ISI. Additional simulations in this section provide more detailed comparisons in this regard. The coefficient norm of the modified algorithm converge toward a value different from that of the optimal inverting filter. The stable norm value depends on the choice of the norm exponents and the channel.

The second set of simulations were designed to verify the abilities of the modified Sato and SWA blind deconvolution methods to obtain an unbiased estimate of the optimum equalizer filter in the absence of noise. To this end, ten sequences of 200 000 samples of the received signal $x_t$ were generated from Laplacian-distributed source sequences and used to compute a single least-squares estimate of the channel inverse, where $L = 20$ and the delay value of $d = 6$ were chosen. The ISI measurement was then computed for the least-squares coefficient estimate, resulting in a value of $J_{\text{ISI}}(c(z)) = -27.3$ dB. This equalizer estimate was then used as an initial coefficient vector for both the modified Sato algorithm and the unconstrained Shalvi–Weinstein algorithm (SWA) [8], which were allowed to adapt on these ten sequences in the following way. For the first 100 000 samples, both equalizers were given larger step sizes of $\mu = 0.0005$ and $\mu = 0.000\,04$ to aid convergence to their stationary points. These step sizes were then reduced to one fifth of their starting values for the remaining 100 000 samples to obtain more-accurate stationary point estimates. Finally, the last 40 000 coefficient vectors of each simulation were used to compute an average coefficient vector for each algorithm run, and these coefficient vectors were then averaged over the ten simulation runs to compute a final average coefficient vector $\boldsymbol{w}_{ave}$ for each algorithm. The normalized squared distance of each averaged coefficient vector from the least-squares solution was then computed as

$$\mathcal{D}(\boldsymbol{w}_{ave}, \boldsymbol{w}_{opt}) = \left\| \frac{\boldsymbol{w}_{ave}}{\|\boldsymbol{w}_{ave}\|} - \frac{\boldsymbol{w}_{opt}}{\|\boldsymbol{w}_{opt}\|} \right\|^2 \tag{87}$$

and a corresponding ISI level for the averaged coefficient vector of each algorithm was also computed. The resulting normalized squared distance values for the modified Sato and SWA algorithms were $3.68 \times 10^{-5}$ and $1.12 \times 10^{-4}$, respectively. The corresponding ISI values for these stationary point solutions were $-27.2$ and $-27.0$ dB, respectively. The small values of $\mathcal{D}(\boldsymbol{w}_{ave}, \boldsymbol{w}_{opt})$ and the nearly identical ISI levels produced by the three coefficient solutions indicate that the two blind methods provide essentially unbiased estimates of the equalizer solution. For these reasons, the local convergence performance of these algorithms can be compared fairly based on their adaptation speeds and measured total ISI values in steady state.

We then compared the convergence performances of the modified Sato algorithm and the SWA algorithm for the chosen channel under noise-free conditions, where $L = 20$ has been chosen. In this case, we have used step sizes of 0.0007 and 0.0002 for these algorithms, respectively, so that both methods produce an average total ISI of $-17$ dB in steady state. In addition, we have simulated the performance of three additional unit-norm-constrained algorithms for comparison:

- a unit-norm-constrained Sato algorithm (Sato-NC, $\mu = 0.0007$);
- a unit-norm-constrained constant-modulus algorithm (CMA-NC, $\mu = 0.000\,22$);
- a unit-norm-constrained normalized constant-modulus algorithm as modified from [17] (NCMA-NC, $\mu = 0.01$, $\mu = 0.001$, and $\mu = 0.0001$).

In each unit-norm-constrained method, the equalizer coefficient vector length was normalized to unity after every update. Note that these latter three algorithms are more computationally complex than the former two unconstrained approaches, as they require divides and square roots. This is particularly true for real-time applications, whereas for batch algorithms, this advantage does not appear to be important. Their convergence behaviors, as computed from averages of 100 simulation runs, are shown in Fig. 3, where it is seen that the modified Sato and unit-norm-constrained Sato algorithms outperform the other three methods. As the modified Sato algorithm requires about two multiply/adds per adaptive filter coefficient, it is to be preferred. Moreover, the normalized CMA method is noncompetitive due to the excessive noise enhancement that the normalization procedure imposes for every choice of the step size $\mu$. Interestingly, the two Sato algorithms provide nearly identical equalization performance, despite imposing different constraint values on the lengths of the coefficient vectors ($\|\boldsymbol{w}_t\| \approx 0.8$ for the modified Sato algorithm).

Last, we performed a noise analysis. To this end, we compared the performances of the modified Sato and SWA algorithm for a 20-tap equalizer and a noisy received signal with different signal-to-noise ratios (SNRs). In such cases, neither equalizer filter is guaranteed to converge to the Wiener solution for this task, although both do reduce ISI to some extent. Step sizes for each algorithm have been adjusted so that a steady-state ISI of $-15$ dB is achieved at convergence except for very low SNR, where a steady-state ISI of $-12$ dB and $-10$ dB has been aimed at by proper choice of the respective step sizes of the individual algorithms. Fifty different data realizations have been

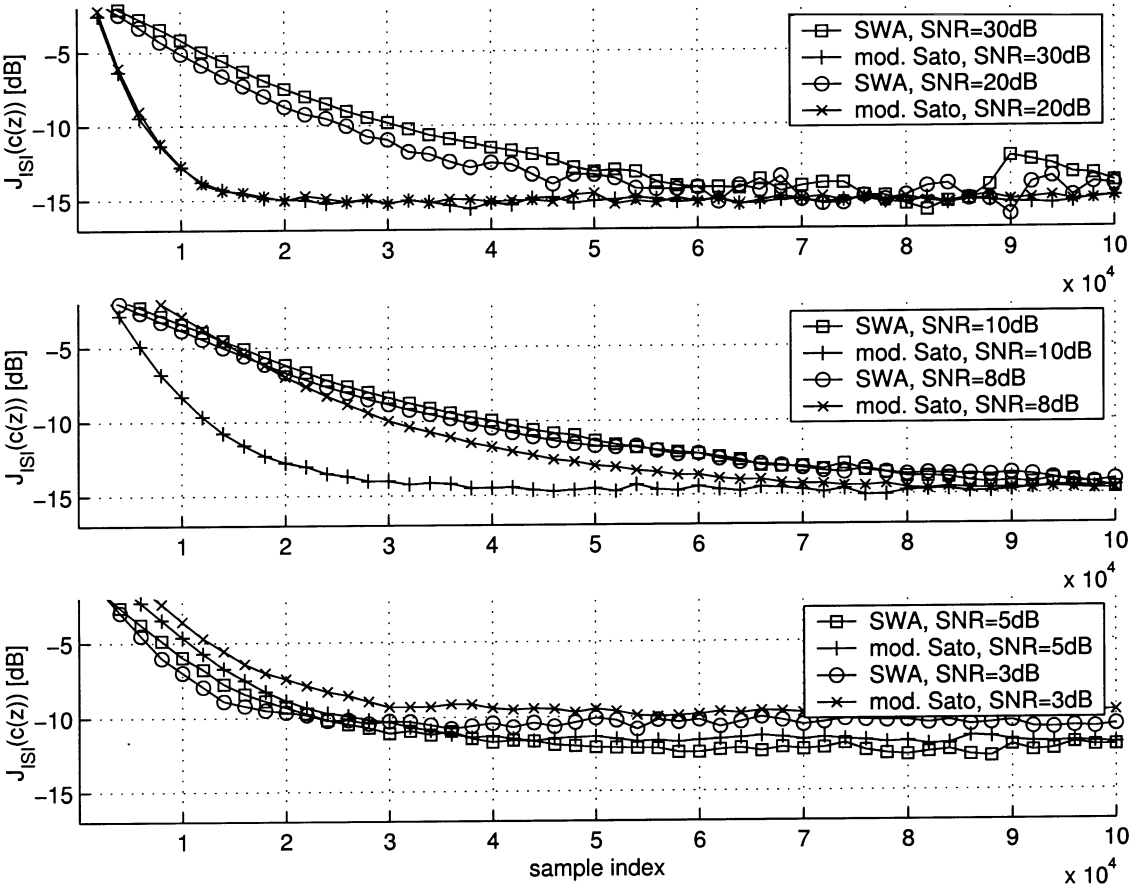



Fig. 4. Convergence comparison of the Shalvi–Weinstein algorithm (SWA) and the modified Sato algorithm for a noisy channel.

used to smooth the convergence curves. Fig. 4 shows the favorable behavior of the modified Sato algorithm compared with the SWA in this simulation, except for very low SNR. The slow convergence of SWA for high SNR is indicative of the higher order-statistics-based procedure used, whereas the Sato algorithm employs much more efficient estimation. For very low SNR, the performance of the Sato algorithm starts to degrade until it is inferior to the performance of the SWA at around SNR of 5 dB and less. For moderate SNR such as 10 dB and even 8 dB, the convergence time of the modified Sato algorithm is still shorter than that of the SWA.

## VII. Conclusion

It is well known that unconstrained Bussgang-type algorithms fail to deconvolve signals with impulsive distributions. In this paper, we have proven why these algorithms cannot work for such distributions using local stability analysis. Our analysis provides a way of designing modified Bussgang algorithms to obtain locally stable, but not necessarily globally convergent, behavior for various source signal distributions. Simulations of a modified Sato algorithm confirm the efficacy of the modifications. Our analysis method allows one to predict the suitability of a particular blind deconvolution algorithm to a chosen source distribution, and its extension to unit-norm-constrained blind deconvolution approaches is a potential future research topic.

## Acknowledgment

The first author appreciates the discussions held with M. Joho and P. Vontobel on the issue of eigenvalue distributions and spectra of circulant and Toeplitz matrices, respectively.

## References

[1] S. Bellini, "Bussgang techniques for blind deconvolution and equalization," in *Blind Deconvolution*, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1994, pp. 8–59.

[2] B. W. Gillespie, H. Malvar, and D. A. F. Florencio, "Speech dereverberation via maximum-kurtosis subband adaptive filtering," in *Proc. ICASSP*, vol. 6, Salt Lake City, UT, May 7–11, 2001, pp. 3701–3704.

[3] Y. Sato, "A method of self-recovering equalization for multilevel amplitude-modulation systems," *IEEE Trans. Comput.*, pp. 679–682, June 1975.

[4] R. W. Lucky, "Techniques for adaptive equalization of digital communication systems," *Bell Syst. Tech. J.*, vol. 45, pp. 255–286, Feb. 1966.

[5] 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.

[6] J. R. Treichler and B. G. Agee, "A new approach to the multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 331–344, Apr. 1983.

[7] A. Benveniste, M. Goursat, and G. Ruget, "Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communications," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 385–399, June 1980.

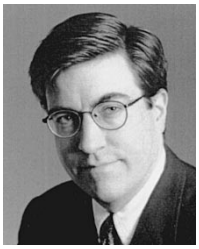
[8] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–321, Mar. 1990.

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

[10] R. M. Gray, "Toeplitz and circulant matrices: A review," Stanford Electron. Lab., Stanford, CA, Tech. Rep. 6502-1, June 1971.
[11] M. Loève, *Probability Theory*. New York: D. van Nostrand, 1960.
[12] J. E. Peĉarić, F. Froschan, and Y. L. Tong, *Convex Functions, Partial Orderings, and Statistical Applications*. New York: Academic, 1992.
[13] E. Artin, *The Gamma Function*. New York: Holt, Rinehart, and Winston, 1964.
[14] H. Mathis, "Nonlinear functions for blind separation and equalization," Ph.D. dissertation, ETH Zürich, Zurich, Switzerland, Nov. 2001.
[15] R. Godfrey and F. Rocca, "Zero-memory nonlinear deconvolution," *Geophys. Prospecting*, vol. 29, pp. 189–228, 1981.
[16] G. J. Foschini, "Equalizing without altering or detecting data," *ATT Tech. J.*, vol. 64, no. 8, pp. 1885–1911, Oct. 1985.
[17] C. B. Papadias and D. T. M. Slock, "Normalized sliding window constant modulus and decision-directed algorithms: A link between blind equalization and classical adaptive filtering," *IEEE Trans. Signal Processing*, vol. 45, pp. 231–235, Jan. 1997.

**Heinz Mathis** was born in Zurich, Switzerland, in 1968. He received the Diploma in electrical engineering from the Swiss Federal Institute of Technology (ETH), Zurich, in 1993 and the Ph.D. degree in electrical engineering from the same university in 2001, respectively.

From 1993 to 1997, he was a DSP and RF engineer with different companies in Switzerland and England. From 1997 to 2001, he was a Research Assistant at the Signal and Information Processing Laboratory, ETH Zurich. He is currently a Professor of mobile communications at the University of Applied Sciences, Rapperswil, Switzerland. His research interests include signal processing for wireless communication systems.



**Scott C. Douglas** (SM'98) received the B.S. (with distinction), M.S., and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1988, 1989, and 1992, respectively.

From 1992 to 1998, he was an Assistant Professor with the Department of Electrical Engineering, University of Utah, Salt Lake City. Since August 1998, he has been with the Department of Electrical Engineering, School of Engineering and Applied Science, Southern Methodist University, Dallas, TX, as an Associate Professor. His research activities include adaptive filtering, active noise control, blind deconvolution and source separation, and VLSI/hardware implementations of digital signal processing systems.