Eigenvector Algorithm for Blind MA System Identification*

Dieter Boss, Björn Jelonnek, Karl-Dirk Kammeyer

University of Bremen, FB-1, Dept of Telecommunications, P.O. Box 330440, D-28334 Bremen, Germany,
Tel.: +(49)-421/218-3356, Fax: -3341, E-mail: boss@comm.uni-bremen.de, http://www.comm.uni-bremen.de

Abstract

We present a novel approach to the blind estimation of a linear time-invariant possibly mixed-phase moving average (MA) system (channel) based on second and fourth order statistics of the stationary received signal. As the algorithm incorporates the solution of an eigenvector problem, it is termed EVI standing for EIGENVECTOR APPROACH TO BLIND IDENTIFICATION. One of EVI's main features is its ability to obtain reliable estimates of the channel's MA parameters on the basis of very short records of received data samples. It is also robust with respect to an overestimation of the channel order. Furthermore, we demonstrate that, if independent additive white Gaussian noise is present, the degradation of the MA parameter estimates is minor even at low signal-to-noise ratios. By simulation results, we finally show the potential applicability of EVI to mobile radio communication channels under time-invariance conditions typically assumed in GSM receivers.

Zusammenfassung

In diesem Artikel wird ein neuartiger Ansatz zur blinden Identifikation eines linearen, zeitinvarianten, evtl. gemischtphasigen "Moving Average"-(MA)-Modelles (Übertragungskanals) vorgestellt, der auf den Statistiken zweiter und vierter Ordnung des stationären Empfangssignals basiert. Da hierzu die Lösung eines Eigenvektorproblems erforderlich ist, wird dieser Ansatz EIGENVEKTOR-ALGORITHMUS ZUR BLINDEN IDENTIFIKATION (EVI) genannt. Eines der Hauptmerkmale von EVI ist seine Fähigkeit, die MA-Parameter des Kanals auf der Basis von sehr wenigen Abtastwerten des Empfangssignals zuverlässig zu schätzen. EVI ist auch gegen Überschätzungen des Kanalgrades robust. Außerdem wird gezeigt, daß sich das Schätzergebnis selbst bei niedrigen Signal-Rausch-Verhältnissen nur unwesentlich unter dem Einfluß von unabhängigem, additivem, weißem gaußverteiltem Rauschen verschlechtert. Durch Simulationsergebnisse wird schließlich dargelegt, daß EVI unter GSM-Zeitinvarianzannahmen auch auf Mobilfunkkanäle angewendet werden kann.

Résumé

Dans cette communication nous proposons une nouvelle approche d'identification aveugle de modèles MA (canaux) à phase mixte, basée sur l'utilisation conjointe des statistiques d'ordre deux et quatre du signal réçu stationnaire. Comme cette approche incorpore la solution d'un problème de vecteurs propres, elle est appelée EIGENVECTOR APPROACH TO BLIND IDENTIFICATION (EVI). Un des intérêts principaux de EVI est sa capacité d'estimer de manière fiable les coefficients MA du canal à partir d'un très faible nombre d'échantillons. La méthode proposée est aussi robuste vis-à-vis de la surestimation de l'ordre du modèle. Par ailleurs, nous démontrons que si un bruit indépendant, additif, blanc Gaussien est présent, la dégradation des estimées du modèle est négligeable même pour un faible rapport signal à bruit. Grâce à des simulations, nous montrons finalement que EVI peut tout à fait s'appliquer aux canaux de communications mobiles sous l'hypothèse d'invariance temporelle habituellement faite dans les récepteurs GSM.

*This work is supported by the German National Science Foundation (DFG-contracts Ka 841/1 and /2).

1 Introduction

Optimum receivers in digital communication systems require the knowledge of the transmission channel's impulse response. Since this knowledge usually is not available, the problem of *channel estimation* arises. From the point of view of systems theory, channel estimation is a particular form of (linear) *system identification* which, in our case, is complicated by three main properties of the radio channel: *(i)* it consists of multiple propagation paths and is therefore frequency-selective, *(ii)* its discrete-time equivalent baseband impulse response may be mixed-phase and *(iii)* in a mobile environment, it is *time-variant*. As for the latter property, time-variance is relatively slow in many applications when compared with the symbol period so that the channel can be estimated repeatedly in periods of time where it can be assumed time-*in*variant *(piecewise* or *quasi* time-invariant).

Within such a period of time-invariance, state-of-the-art mobile communication systems transmit *training sequences* to assist the receiver in estimating the channel impulse response. For this purpose, the cross-correlation between the received (corrupted) and the stored (ideal) training sequences is calculated. However, depending on the degree of time-variance, the repeated transmission of training sequences leaves the communication system with an overhead, which, in the case of the *Global System for Mobile communications (GSM)*, amounts to 22.4% [28]. This overhead capacity could be used for other purposes such as channel coding (thus enhancing overall system performance), if the channel estimation problem was solved *blindly*.

Blind system identification

The fundamental idea of *blind* channel estimation is to derive the channel characteristics from the received signal only, i.e. *without* access to the channel input signal by means of training sequences. Depending on the different ways to extract information from the received signal, two classes of algorithms can be distinguished[1]:

- **Class HOS:** When the received signal is sampled at symbol-rate, the resulting sequence is (quasi) stationary. Since second order statistics of a stationary signal are inadequate for the identification of the complete channel characteristics (including phase information), class HOS approaches are based either explicitly or implicitly on *Higher Order Statistics*. Higher order *cumulants* contain the complete information on the channel's magnitude and phase provided that the distribution of the channel input signal is non-Gaussian (which is true for applications in digital communications). Excellent overviews on HOS and their applications can be found in [30, 27, 29, 2].
- **Class SOCS:** When the sampling period is a fraction of the symbol period (time diversity), or alternatively, the symbol-rate sampled signals received by several sensors are interleaved (antenna diversity), the resulting received sequence is (quasi) *cyclostationary* provided that some excess bandwidth is available. Generally, *Second Order Cyclostationary Statistics (SOCS)* are sufficient to retrieve the complete channel characteristics, but there are "singular" channel classes which can *not* be identified this way. They include channels with common subsystems in all polyphase subchannels (refer to [36, 42, 11] for details).

Implications of the mobile channel

In a mobile propagation environment, the channel assumes an arbitrary impulse response in any instant of time. Particularly, "singular", "criti-

[1] For the following statements, a stationary channel input sequence is assumed.

cal"[2], and mixed-phase channels can not be prevented from occurring. Furthermore, from the above quasi time-invariance assumption, it follows that the received signal can only be observed in a (short) period of time. In summary, a blind channel estimation algorithm for an application in mobile communications should satisfy the following requirements:

R1: Reliable estimates of the complex channel impulse response must be obtained from few samples of the received signal (hundreds rather than thousands of symbol periods).

R2: This should apply to arbitrary channels (whether or not they are "singular", "critical", mixed-phase etc.).

R3: As the effective channel order usually is unknown (and just *quasi* time-invariant), an overestimation must not represent a problem.

R4: The estimates should be as robust as possible with respect to stationary additive white Gaussian noise at low signal-to-noise ratios.

The algorithm we present in this paper meets the above requirements. It is a class HOS approach based on fourth (and second) order stationary statistics. SOCS-based methods are not considered because they violate requirement R2: For two algorithms proposed by Tong et al. [37] and Schell et al. [32], we have demonstrated in [7] and [3], respectively, that "singular" channels represent a severe limitation because their channel estimation performance from few samples is heavily affected even if subchannel zeros are just "close" to each other (rather than being identical).

Although our algorithm is devoted to the blind identification of finite impulse response (FIR) systems, it can also be applied to estimate the *moving average* (MA) parameters of an autoregressive moving average (ARMA) model using the *residual time series*, i.e. the AR compensated received signal [14, 4]. Together with Mendel's DOUBLE MA ALGORITHM [27], it can also be utilized to determine both the autoregressive (AR) and MA parameters of a non-causal ARMA model.

[2] Channels with zeros 'on' or 'close to' the unit circle of the complex z-plane are called "critical".

Drawbacks of existing approaches

On the particular problem of blind MA system identification, a large number of HOS-based algorithms has been proposed, e.g. [14, 35, 13, 38, 24, 39, 21, 1, 12, 40, 43, 41]. The derivation of many approaches is based on third order statistics and/or real-valued signals and systems [14, 35, 38, 39, 1, 12]. Alas, both assumptions do not apply to applications in digital communications. Extending the existing algorithms to fourth order statistics and complex signals and systems quite frequently results in a poor performance: For a given number of received data samples, they yield high values of estimation variance and/or bias, or equivalently, they require a very large block of received data samples for a satisfactory estimation of the channel's impulse response (thus violating R1). Besides, knowledge of the system order, which is not usually available, is of utmost importance [38, 1, 12], because many algorithms (e.g. [14, 35, 38, 39, 1]) are extremely sensitive to a mis-assumption of the channel order (cf. R3). However, the robust order estimation remains a difficult task, especially in noisy and/or time-variant environments (refer to [15, 44, 26], e.g.).

Purpose and organization of paper

The purpose of this paper is to derive a fast[3] algorithm for blind MA system identification from an approach to blind linear equalization. Note, however, the two principal differences between channel equalization and identification:

[3] where "fast" is meant in the sense "requiring small data blocks for a satisfactory channel estimation".

- To equalize a first order "critical" MA channel, we would require a high order (symbol-rate) FIR equalizer, e.g., whereas for system identification, a single parameter needs to be estimated. This is why many identification approaches ([10], e.g.) using a linear equalization result deliver unsatisfactory estimates of critical channels.
- Noisy case: In contrast to the channel estimate, which is supposed to be insensitive to noise, the equalizer coefficients must be adjusted differently if noise is present in order to ensure optimum equalization taking into account both intersymbol interference and noise.

These differences represent the principal obstacles when deriving an identification algorithm from an existing approach to equalization. While avoiding these obstacles, we show in this paper how to derive a sophisticated blind identification algorithm from the EIGENVECTOR ALGORITHM FOR BLIND EQUALIZATION (EVA) published recently [20].

In the following section, we state the assumptions we make in this paper. For convenience, a brief review of EVA is given, while details can be found in [22, 23, 20]. The novel algorithm is derived in section 3 from the EVA solution. It is termed EIGENVECTOR APPROACH TO BLIND IDENTIFICATION (EVI) and represents *a genuine identification approach (delivering low error channel estimates) while retaining the benefits of an equalizer algorithm (insensitivity to order overestimation).* In section 4, the performance of EVI is illustrated by simulation results. We demonstrate the excellent quality of EVI's estimates *(i)* in terms of the number of received data samples (cf. requirement R1), *(ii)* with mis-assumptions for the channel order (R3), and *(iii)* in presence of additive noise (R4). With the help of a realistic mobile radio channel example, we investigate whether or not the four requirements R1 to R4 can be satisfied *at the same time*.

2 Assumptions and review of the EigenVector Algorithm for blind equalization (EVA)

2.1 Assumptions

Fig. 1 shows an equivalent discrete-time baseband model of a digital communication system. The transmitted data $d(k)$ are an independent, identically distributed (i.i.d.) sequence of random variables with zero mean, variance σ_d^2, skewness[4] γ_3^d and kurtosis[4] γ_4^d. Each symbol period T, $d(k)$ takes a (possibly complex) value from a finite set. For this reason, the channel input random process clearly is non-Gaussian with a non-zero kurtosis ($\gamma_4^d \neq 0$), while its skewness vanishes ($\gamma_3^d = 0$) due to the even probability density function of typical digital modulation signals such as *Phase Shift Keying (PSK)*, *Quadrature Amplitude Modulation (QAM)* or *Amplitude Shift Keying (ASK)*.

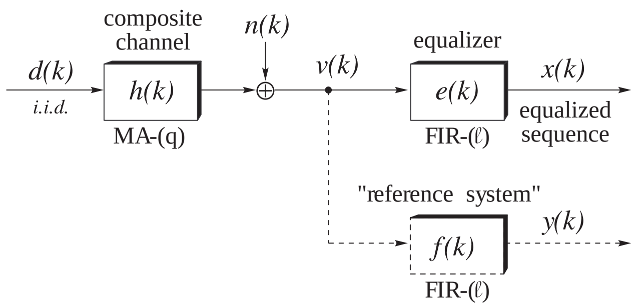


Fig. 1: Equivalent symbol-rate baseband model of a digital communication system (including a linear equalizer and a "reference system")

For the unknown composite channel, we assume the *Equivalent Discrete-Time White-Noise Filter* model [31] comprising the physical transmission channel, the transmit and receive filters, the symbol-rate sampler and the noise whitening filter. We suppose the composite channel to be (at least short time) time-invariant. It is described by the causal possibly mixed-phase finite

[4]defined as $\gamma_3^d \triangleq E\{(d(k))^3\}$ and $\gamma_4^d \triangleq E\{|d(k)|^4\} - 2\sigma_d^4 - |E\{d^2(k)\}|^2$, where $E\{\cdot\}$ denotes statistical expectation.

impulse response $h(k) = h(0), \cdots, h(q)$, where q denotes the (effective) order, and will simply be termed "channel". Apart from linear distortions, the (quasi) stationary received sequence $v(k)$ is corrupted by independent stationary zero mean additive white Gaussian noise $n(k)$.

In the receiver, an FIR-(ℓ) equalizer with impulse response $e(k) = e(0), \cdots, e(\ell)$ and an FIR filter $f(k)$ of the same order are introduced. As $f(k)$ will be used to generate an implicit sequence of training (reference) data for the subsequent iteration of the iterative approach to be explained, it is termed "reference system". For now, however, assume its coefficients $f(0), \cdots, f(\ell)$ to be fixed (arbitrarily). The output sequences of the equalizer and the reference system shall be termed $x(k)$ and $y(k)$, respectively.

All signals and systems are assumed to be complex-valued due to the equivalent baseband representation of the corresponding bandpass communication system. Notice that the assumptions mentioned so far are valid throughout the paper while the scope of the following equalization objective is limited to section 2.

2.2 Linear equalization objective

Adjust the $\ell + 1$ coefficients $e(k)$ so that the equalized sequence $x(k)$ is as close as possible to the delayed transmitted data $d(k - k_0)$ in the *MSE (mean square error)* sense

$$\begin{aligned} \mathrm{MSE}(\mathbf{e}, k_0) &\stackrel{\Delta}{=} E\{|x(k) - d(k-k_0)|^2\} \\ &\stackrel{!}{=} \min , \end{aligned} \tag{1}$$

where the vector $\mathbf{e} = [e(0), \cdots, e(\ell)]^T$ is used to simplify notation. For each order ℓ and delay k_0, the equalizer minimising (1) is called *Minimum Mean Square Error* equalizer. In brief, it is referred to as *MMSE-*(ℓ, k_0) equalizer.

2.3 Non-blind solution

If both the received sequence $v(k)$ and some transmitted data $d(k)$ *(training sequence)* are given, the MMSE-(ℓ, k_0) equalizer coefficients can be calculated using the well-known *normal equation* [17]

$$\mathbf{e}_{\mathrm{MMSE}(k_0)} = \mathbf{R}_{vv}^{-1}\,\mathbf{r}_{vd} \tag{2}$$

with

$$\begin{aligned} \mathbf{r}_{vd} &\stackrel{\Delta}{=} E\{\mathbf{v}_k \cdot d(k-k_0)\} \\ \mathbf{R}_{vv} &\stackrel{\Delta}{=} E\{\mathbf{v}_k \mathbf{v}_k^*\} , \end{aligned} \tag{3}$$

where $\mathbf{r}_{vd}$ and $\mathbf{R}_{vv}$ denote the cross-correlation vector and the non-singular $(\ell+1) \times (\ell+1)$ Hermitian Toeplitz autocorrelation matrix, respectively, and the vectors $\mathbf{v}_k$ and $\mathbf{v}_k^*$ are defined as

$$\begin{aligned} \mathbf{v}_k &\stackrel{\Delta}{=} [v^*(k), v^*(k-1), \cdots, v^*(k-\ell)]^T \\ \mathbf{v}_k^* &\stackrel{\Delta}{=} [v(k), v(k-1), \cdots, v(k-\ell)] \end{aligned} \tag{4}$$

(conjugate transpose form). The MMSE-(ℓ, k_0) equalizer

$$\mathbf{e}_{\mathrm{MMSE}(k_0)} \stackrel{\Delta}{=} [e_{\mathrm{MMSE}}(0), \cdots, e_{\mathrm{MMSE}}(\ell)]^T \tag{5}$$

according to (2) is used as a reference in this paper. In the noiseless case, it approximates the channel's inverse system (deconvolution, *zero forcing*) in order to minimize intersymbol interference (ISI). If additive noise is present, however, its coefficients are adjusted differently so as to minimize the total MSE in the equalized sequence $x(k)$ due to ISI *and noise*.

2.4 Blind "EVA solution"

With *blind* equalization, the objective is to determine the MMSE-(ℓ, k_0) equalizer coefficients *without access to the transmitted data*, i.e. from the received sequence $v(k)$ only. Similar to Shalvi and Weinstein's *maximum kurtosis* criterion [33, 34], the EVA solution to blind equalization is based on a *maximum "cross-kurtosis"* quality function. We have demonstrated in [23] that[5]

$$\begin{aligned} c_4^{xy}(0,0,0) = \\ E\{|x(k)|^2\,|y(k)|^2\} - E\{|x(k)|^2\}\,E\{|y(k)|^2\} \\ -|E\{x^*(k)\,y(k)\}|^2 - |E\{x(k)\,y(k)\}|^2 , \end{aligned} \tag{6}$$

[5]For notational simplicity, just two superscript indices are used in $c_4^{xy}(\cdot)$. They represent two occurrences of both $x(k)$ and $y(k)$.

i.e. the cross-kurtosis between $x(k)$ and $y(k)$, can be used as a measure for equalization quality:

$$\text{maximize } |c_4^{xy}(0,0,0)| \quad \begin{cases} \text{subject to} \\ r_{xx}(0) = \sigma_d^2 \end{cases} . \tag{7}$$

Rather than referring to the equalizer output $x(k)$, this scalar quality function can easily be expressed in terms of the equalizer *input* $v(k)$ by replacing $x(k)$ in equation (6) with

$$x(k) = v(k) * e(k) = \mathbf{v}_k^* \mathbf{e} , \tag{8}$$

where "$*$" denotes the convolution operator. In this way, we obtain from (7)

$$\text{maximize } |\mathbf{e}^* \mathbf{C}_4^{yv} \mathbf{e}| \quad \begin{cases} \text{subject to} \\ \mathbf{e}^* \mathbf{R}_{vv} \mathbf{e} = \sigma_d^2 \end{cases} , \tag{9}$$

where the Hermitian $(\ell+1) \times (\ell+1)$ cross-cumulant matrix

$$\begin{aligned} \mathbf{C}_4^{yv} \triangleq\ & E\{|y(k)|^2 \mathbf{v}_k \mathbf{v}_k^*\} \\ & -E\{|y(k)|^2\} E\{\mathbf{v}_k \mathbf{v}_k^*\} \\ & -E\{y(k)\mathbf{v}_k\} E\{y^*(k)\mathbf{v}_k^*\} \\ & -E\{y^*(k)\mathbf{v}_k\} E\{y(k)\mathbf{v}_k^*\} \end{aligned} \tag{10}$$

contains the fourth order cross-cumulant

$$[\mathbf{C}_4^{yv}]_{i_1,i_2} = c_4^{yv}(-i_1, 0, -i_2) \tag{11}$$

in row i_1 and column i_2 with $i_1, i_2 \in \{0, 1, \cdots, \ell\}$.

The quality function (9) is quadratic in the equalizer coefficients. Its optimization leads to a closed-form expression in guise of the generalized eigenvector problem [23]

$$\boxed{\mathbf{C}_4^{yv} \mathbf{e}_{\text{EVA}} = \lambda \mathbf{R}_{vv} \mathbf{e}_{\text{EVA}}} \quad \textbf{“EVA equation”} \tag{12}$$

which we term "EVA equation". The coefficient vector

$$\mathbf{e}_{\text{EVA}} \triangleq [e_{\text{EVA}}(0), \cdots, e_{\text{EVA}}(\ell)]^T \tag{13}$$

obtained by choosing the eigenvector of $\mathbf{R}_{vv}^{-1} \mathbf{C}_4^{yv}$ associated with the maximum magnitude eigenvalue λ is called the **"EVA-(ℓ) solution"** to the problem of blind equalization. Apart from the obvious ambiguity by a complex factor, this solution is unique if the quality function (7) has a single global maximum. In [22, 23], we have proven that this is the case if and only if the magnitude of the combined impulse response

$$w(k) \triangleq h(k) * f(k) \tag{14}$$

adopts its maximum value $w_m \triangleq \max\{|w(k)|\}$ only once, i.e.

$$\begin{array}{ll} |w(k)| = w_m & \text{if } k = k_m \\ |w(k)| < w_m & \text{otherwise} \end{array} . \tag{15}$$

2.5 EigenVector Algorithm for blind equalization (EVA)

Of course, condition (15) can not be guaranteed since the channel impulse response $h(k)$ and thus $w(k)$ are unknown. However, the effects of an "unlucky guess" of $f(k)$ resulting in a violation of (15) can be overcome by an iterative adjustment of the reference system's coefficients [22, 20]. This iterative approach makes use of the following relation between the EVA-(ℓ) solution and the desired MMSE-(ℓ, k_m) equalizer (see [20] for details and [23] for a proof):

- The EVA-(ℓ) solution $\mathbf{e}_{\text{EVA}}$ can be decomposed into a weighted sum of MMSE-(ℓ, k) equalizers $\mathbf{e}_{\text{MMSE}(k)}$ according to (2) with different delay times k around the value k_m. Each MMSE-(ℓ, k) equalizer is weighted mainly by $|w(k)|^2$. Thus, *the quicker $|w(k)|^2$ decays as its index k departs from lag k_m, the closer $\mathbf{e}_{\text{EVA}}$ will be to the MMSE-(ℓ, k_m) equalizer $\mathbf{e}_{\text{MMSE}(k_m)}$* – provided that uniqueness is ensured. For this reason, if $|w(k)|$ had *a distinct peak value*

$$|w(k_m)| \gg |w(k)| \quad \text{for} \quad k \neq k_m , \tag{16}$$

$\mathbf{e}_{\text{EVA}}$ would closely approach $\mathbf{e}_{\text{MMSE}(k_m)}$.

- It also follows from the above that if the reference system $f(k)$ was set such that[6]

$$w(k) = w(k_m) \cdot \delta(k - k_m) \qquad (17)$$

with $|w(k_m)| \neq 0$, $\mathbf{e}_{\mathrm{EVA}}$ would even be identical to $\mathbf{e}_{\mathrm{MMSE}(k_m)}$. Note that both equation (17) and (16) would obviously ensure the uniqueness condition (15).

In other words: The better the reference system deconvolves the channel (see eq. (17) and (16)), the better the resulting EVA solution will be (in the MMSE sense). For this reason, an iterative procedure to adjust the reference system's coefficients was suggested where $f(k)$ *is loaded with the equalizer impulse response calculated in the previous iteration* [22, 20]. After some iterations (based on the same block of received data samples), both the reference system and the equalizer will have the same impulse response being very close to the MMSE-(ℓ, k_m) equalizer. An improved convergence rate can be obtained by a stepwise increase of the equalizer order ℓ during the iteration procedure. The resulting overall algorithm was termed EVA standing for EIGENVECTOR ALGORITHM FOR BLIND EQUALIZATION [20].

[6] Remark: If (17) holds, $y(k)$ is proportional to $d(k-k_m)$, i.e. the matrix $\mathbf{C}_4^{yv}$ in (12) refers to a training sequence and EVA becomes a non-blind approach using $c_4^{xd}(0,0,0)$ as a quality criterion. In this case, $f(k)$ generates a sequence of reference data within the receiver. Hence its designation as "reference system".

3 Eigenvector algorithm for blind system identification (EVI)

Objective: As opposed to the previous section, the objective now is to estimate the channel impulse response $h(k)$ from the received data $v(k)$, only. Obviously, this system identification problem is closely related to equalization. A simple approach would be to take the coefficients of the inverse equalizer as an estimate for the channel impulse response. However, this method has two main problems:

P1: The required equalizer order can be very high in case of critical channels: In order to allow the FIR equalizer to perfectly approximate the inverse channel (i.e. an autoregressive model), its order would have to be infinite. In other words, even in the noiseless case, the resulting channel estimates would be biased for any finite equalizer order, especially if the effective order of the channel's inverse system was high (i.e. for critical channels). Thus, the identification of a channel as simple as $H(z) = 1 + z^{-1}$ would be quite expensive (and biased even in the noiseless case).

P2: The second disadvantage results from the influence of additive noise. As explained in section 2, EVA closely approximates the MMSE equalizer solution, which is the optimum solution for (linear symbol-rate) equalization. However, the inverse EVA solution does not correspond to the channel transfer function, if noise is present. Again, we would end up with biased channel estimates.

In this section, we present a novel method for blind MA system identification which does not reveal P1 and is quite robust with respect to P2. In section 3.1, we derive a novel equation for blind identification from the EVA equation (12). It also requires the solution of an eigenvector problem. The dimensions of the involved matrices depend on the equalizer order ℓ rather than the channel order q. Therefore, the solution reveals problem P1. In section 3.2, we demonstrate how *unbiased* channel estimates can be obtained in the noiseless case by altering the eigenvector problem. Thus, problem P1 is evaded. In order to guarantee the uniqueness of the eigenvector equation's solution and to improve estimation *variance*, we present in section 3.3 a two step approach called EVI standing for EIGENVECTOR APPROACH TO BLIND IDENTIFICATION. For an appropriate adjustment of the reference system's coefficients, EVI uses in its first step the EV$\underline{\text{A}}$ method as described in section 2. Then, the eigenvector problem mentioned above is solved in the final identification step. Notice, however, that *we do not require perfect equalization to obtain good channel estimates!* Finally, we show in section 4 that EVI has favorable properties regarding problem P2.

3.1 The fundamental EVI solution

According to section 2, the solution of the EVA equation (12) is close to the MMSE-(ℓ, k_m) solution (2)

$$\begin{aligned} \mathbf{e}_{\text{EVA}} &= \mathbf{R}_{vv}^{-1}\,\mathbf{C}_4^{yv}\mathbf{e}_{\text{EVA}}/\lambda \\ \approx\ \mathbf{e}_{\text{MMSE}(k_m)} &= \mathbf{R}_{vv}^{-1}\,\mathbf{r}_{vd}\ . \end{aligned} \tag{18}$$

From this approximation we realize that the expression $\mathbf{C}_4^{yv}\mathbf{e}_{\text{EVA}}/\lambda$ plays the same role in blind equalization as does the cross-correlation vector $\mathbf{r}_{vd}$ in non-blind equalization:

$$\frac{1}{\lambda}\,\mathbf{C}_4^{yv}\,\mathbf{e}_{\text{EVA}} \approx \mathbf{r}_{vd}\ . \tag{19}$$

For zero-mean i.i.d. channel input data $d(k)$ with variance σ_d^2, the vector $\mathbf{r}_{vd}$ is linked to the channel impulse response $h(k)$ by

$$\mathbf{r}_{vd} = \sigma_d^2 \cdot \mathbf{h}\,, \tag{20}$$

where

$$\mathbf{h} = [h^*(k_m), h^*(k_m-1), \cdots, h^*(0), 0, \cdots, 0]^T . \tag{21}$$

Choosing a sufficiently long delay value k_m, the length $\ell+1$ vector $\mathbf{h}$ contains the complete (complex conjugate) channel impulse response in reverse order. Upon insertion of equation (20) into (19), we obtain a formula to estimate the channel impulse response from the estimated equalizer coefficients

$$\sigma_d^2\,\mathbf{h}_{\text{EVI}} \triangleq \frac{1}{\lambda}\,\mathbf{C}_4^{yv}\mathbf{e}_{\text{EVA}}\ . \tag{22}$$

Multiplying the upper equation of (18) with $\mathbf{C}_4^{yv}/\lambda$ and inserting equation (22) on both sides, we obtain a direct expression for the blind identification of an MA system [25].

$$\boxed{\lambda\mathbf{h}_{\text{EVI}} = \mathbf{C}_4^{yv}\,\mathbf{R}_{vv}^{-1}\,\mathbf{h}_{\text{EVI}}} \quad \textbf{“EVI equation”} \tag{23}$$

Eq. (23) is termed “EVI equation”. Again, an eigenvector problem has to be solved. Note that it still involves the reference system (by its output $y(k)$) and the equalizer order ℓ (by the dimensions $(\ell+1)\times(\ell+1)$ of the matrices $\mathbf{C}_4^{yv}$ and $\mathbf{R}_{vv}$). The complex conjugate reverse order eigenvector associated with the maximum magnitude eigenvalue is called “EVI solution”. For the uniqueness of the EVI solution, the statements made in section 2 for the EV$\underline{\text{A}}$ solution remain valid: Condition (15) must hold for a unique channel estimate. We will deal with this requirement in the end of section 3.2.

The following remarks are in order, now, to clarify the properties of the channel estimation approach according to eq. (23) in the noiseless case: *(i)* To obtain *unbiased* channel estimates[7], the dimensions of the $(\ell+1)\times(\ell+1)$ matrices $\mathbf{C}_4^{yv}$ and $\mathbf{R}_{vv}$ must approach infinity. This evokes problem P1, which will be solved in section 3.2 by some modifications to the EVI equation. *(ii)* As for the *variance*, a deconvolving impulse response $f(k)$ leading to an approximation of (17) will result in low variance channel estimates. We will show in section 3.3 how to select $f(k)$ appropriately.

[7]Throughout the paper, we assume that the required cumulant and correlation coefficients are estimated by unbiased sample averaging.

3.2 Eigenvector problem for unbiased channel estimates (noiseless case)

Since the EVI equation (23) was derived from the EVA solution, the equalization problem is involved in so far as the dimensions of the matrix $\mathbf{C}_4^{yv}\mathbf{R}_{vv}^{-1}$ depend on the equalizer order ℓ: they are $(\ell+1)\times(\ell+1)$. As explained in the second paragraph of section 3, this leads to biased channel estimates for any given finite order ℓ. To alleviate this, ℓ must be very high for critical channels resulting in a high computational complexity. This is due to the fact that we implicitly use the inverse system for MA system identification. Consequently, the long resulting vector $\mathbf{h}_{\mathrm{EVI}}$ (length: $\ell+1$) contains just few significant elements (depending on the channel order q) whereas all the others should be zero. In the sequel, it is examined whether the matrix dimensions in the EVI equation (23) can be reduced. It is also demonstrated how unbiased channel estimates can be obtained (thus avoiding problem P1) by replacing a section of the inverse autocorrelation matrix with a modified matrix. With these modifications, the computational complexity of our approach is also reduced drastically.

3.2.1 Reducing matrix dimensions ...

As the transmitted data $d(k)$ are regarded as an i.i.d. sequence of random variables, the range of lags of non-zero cross-cumulants is restricted to the channel order q. Thus, the $(\ell+1)\times(\ell+1)$ Hermitian cumulant matrix $\mathbf{C}_4^{yv}$ is banded with bandwidth $2q+1$. The area of support of matrix $\mathbf{C}_4^{yv}$ can be illustrated by the following representation of the EVI equation (23).

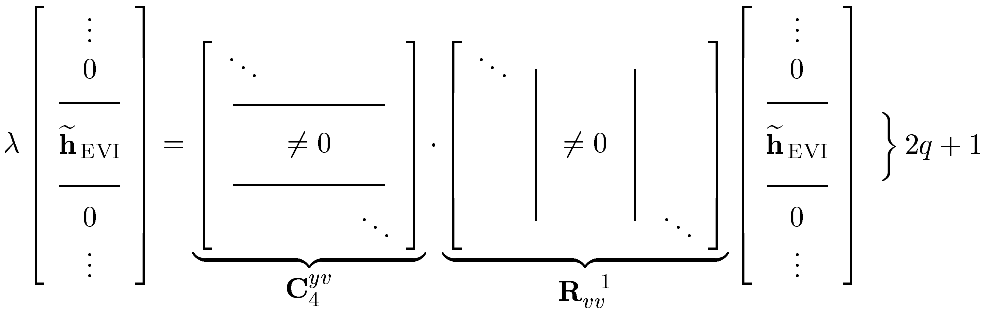
$$\lambda \cdot \mathbf{h}_{\mathrm{EVI}} = \begin{bmatrix} \ddots & & \mathbf{0} \\ & \mathbf{C}_4^{yv} & \\ \mathbf{0} & & \ddots \end{bmatrix} \mathbf{R}_{vv}^{-1}\,\mathbf{h}_{\mathrm{EVI}} \quad (24)$$

Furthermore, the reference system used with EVA determines the delay k_m caused by the combined channel/equalizer impulse response. Thus, depending on the number of minimum and maximum-phase channel zeros, the position of the channel impulse response calculated from equation (23) within the vector $\mathbf{h}_{\mathrm{EVI}}$ is fixed, too. The significant estimated channel coefficients are always located within $2q+1$ adjacent elements of $\mathbf{h}_{\mathrm{EVI}}$. Together, they form the subvector $\widetilde{\mathbf{h}}_{\mathrm{EVI}}$. The other elements of $\mathbf{h}_{\mathrm{EVI}}$ converge to zero. Therefore, just a $(2q+1)\times(\ell+1)$ submatrix of $\mathbf{C}_4^{yv}$ and a $(\ell+1)\times(2q+1)$ submatrix of $\mathbf{R}_{vv}^{-1}$ are relevant for identification. In equation (25), an illustration of the EVI equation (23) is given, where the non-zero parts are located within the horizontal and vertical stripes marked by the lines. The eigenvectors can thus be calculated from a matrix with reduced dimensions $(2q+1)\times(2q+1)$. Without loss of generality we obtain from equations (24) and (25) the regions of support shown in equation (26), where the relevant range $\mathbf{R}_{inv}$ of the inverted autocorrelation matrix $\mathbf{R}_{vv}^{-1}$ follows from the region of support of $\mathbf{C}_4^{yv}$. Just the indicated areas of support are relevant for a proper solution of the EVI equation (23). For this reason, the complexity of the eigenvector problem can be reduced by solving

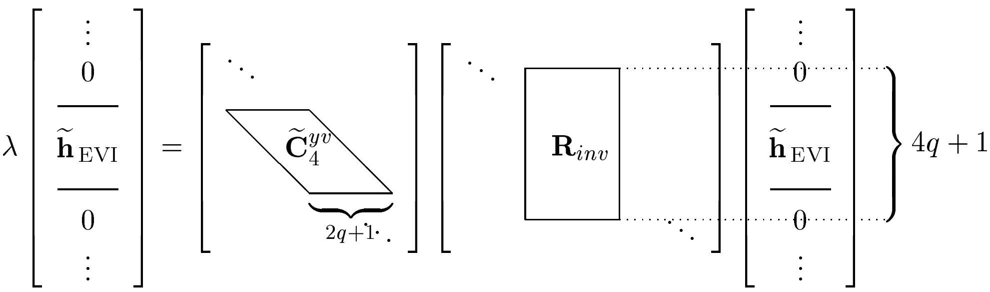
$$\lambda\widetilde{\mathbf{h}}_{\mathrm{EVI}} = \widetilde{\mathbf{C}}_4^{yv}\,\mathbf{R}_{inv}\,\widetilde{\mathbf{h}}_{\mathrm{EVI}}\,, \quad (27)$$

where the dimensions of $\widetilde{\mathbf{C}}_4^{yv}$ are $(2q+1)\times(4q+1)$ and $\mathbf{R}_{inv}$ represents the significant $(4q+1)\times(2q+1)$ submatrix of $\mathbf{R}_{vv}^{-1}$. Here, the dimensions of both matrices involved in this equation are determined by the channel order q rather than the order ℓ of the equalizer.

However, the equalizer order ℓ is still involved in equation (27), because we are required to invert the $(\ell+1)\times(\ell+1)$ autocorrelation matrix $\mathbf{R}_{vv}$ in order to select the submatrix $\mathbf{R}_{inv}$ from the inversion result $\mathbf{R}_{vv}^{-1}$. Even worse, the dimensions of $\mathbf{R}_{vv}$ need to approach infinity to deliver the optimum submatrix $\mathbf{R}_{inv}$ leading to *unbiased* channel estimates.



(25)



(26)

3.2.2 Avoiding problem P1 ...

Further considerations detailed in Appendices A to D show that unbiased channel estimates can be obtained without the inversion of an infinite dimensions matrix. Replace $\mathbf{R}_{inv}$ in equation (27) with the $(4q+1)\times(2q+1)$ Toeplitz matrix

$$\widetilde{\mathbf{R}}_{inv} \triangleq \begin{bmatrix} r_{inv}(-q) & \cdots & r_{inv}(-2q) & & 0 \\ \vdots & \ddots & & \ddots & \\ r_{inv}(q) & & \ddots & & r_{inv}(-2q) \\ \vdots & \ddots & & \ddots & \vdots \\ r_{inv}(2q) & & \ddots & & r_{inv}(-q) \\ & \ddots & & \ddots & \vdots \\ 0 & & r_{inv}(2q) & \cdots & r_{inv}(q) \end{bmatrix} \tag{28}$$

and determine its $4q+1$ different elements according to

$$\begin{bmatrix} r_{inv}(-2q) \\ \vdots \\ r_{inv}(2q) \end{bmatrix} = \widetilde{\mathbf{R}}_{vv}^{-1} \cdot [\underbrace{0,\cdots,0}_{2q},1,\underbrace{0,\cdots,0}_{2q}]^T . \tag{29}$$

Rather than inverting an infinite dimensions autocorrelation matrix, we just need to invert the modified $(4q+1)\times(4q+1)$(!) autocorrelation matrix $\widetilde{\mathbf{R}}_{vv}$. Note that this matrix is always non-singular, even if there exist channel zeros on the unit circle.

Finally, the eigenvector problem (27) can be rewritten as

$$\boxed{\lambda \widetilde{\mathbf{h}}_{\mathrm{EVI}} = \widetilde{\mathbf{C}}_4^{yv}\, \widetilde{\mathbf{R}}_{inv}\, \widetilde{\mathbf{h}}_{\mathrm{EVI}}} \quad \text{“Modified EVI equation”} \tag{30}$$

where $\widetilde{\mathbf{h}}_{\mathrm{EVI}}$ is a length $2q+1$ column vector, $\widetilde{\mathbf{C}}_4^{yv}$ represents a $(2q+1)\times(4q+1)$ submatrix of $\mathbf{C}_4^{yv}$ as indicated in eq. (26) and the $(4q+1)\times(2q+1)$ matrix $\widetilde{\mathbf{R}}_{inv}$ is defined according to equations (28) and (29). Note that it is due to the substitution of $\widetilde{\mathbf{R}}_{inv}$ for $\mathbf{R}_{inv}$ that the “modified EVI equation” (30) yields unbiased estimates (in the noiseless case) and thus avoids problem P1.

3.2.3 Properties of the modified EVI solution

Now, remember that for *unique* channel estimates, condition (15) must hold. We show with a simple example that it is sufficient for unbiased estimates that the two largest coefficients of the combined channel/reference system impulse response $w(k) = h(k) * f(k)$ have slightly different magnitudes. Consider the fourth order MA channel C1 which will be used for the simulation results given in section 4. Fig. 3a displays the zeros of $H(z) = \mathcal{Z}\{h(k)\}$ in the complex plane. The small top right subplot shows the magnitude impulse response. From these plots, we realize that C1 has two zeros relatively close to the unit circle and two identical maximum magnitude coefficients: $|h(2)| = |h(4)| = \max_k\{|h(k)|\}$. Setting the reference system to $f(k) = \delta(k-4)$, the impulse response $w(k)$ has two identical maximum magnitude coefficients, too. As we investigate the asymptotic case in this example, we use true values for the matrices $\widetilde{\mathbf{C}}_4^{yv}$ and $\widetilde{\mathbf{R}}_{inv}$ in the modified EVI equation (30). The magnitude of the resulting impulse response "estimate" $\hat{h}(k)$ is indicated by "$\times$" symbols (and dashed vertical lines) in Fig. 2. For comparison, the magnitude of the shifted true channel impulse response is depicted by circles. From Fig. 2a, we can see that the "estimate" does not correspond to the true channel impulse response. This is due to the violation of condition (15). If however, we slightly modify the channel impulse response so that there is a difference of, say, $\approx 10^{-5}$ between the magnitudes of the two largest channel coefficients, we realize from Fig. 2b that eq. (30) delivers the true channel impulse response. This demonstrates that, in practice, we will almost always obtain unbiased channel estimates.

From this example, we also realize that equation (30) is robust with respect to an overestimation of the channel order: Although the channel has 5 non-zero coefficients, 11 MA parameters are "estimated" in Fig. 2. As "knowledge of the system order is of utmost importance to many system identification algorithms" [1, 38, 12], this property must be emphasized adequately.

In summary, we state that with the modifications to the original EVI equation (23), we obtain *a genuine identification algorithm (delivering unbiased*[8] *channel estimates) while retaining the benefits of an approach to equalization (insensitivity to order overestimation)*. Furthermore, they lead to a considerable decrease in computational complexity.

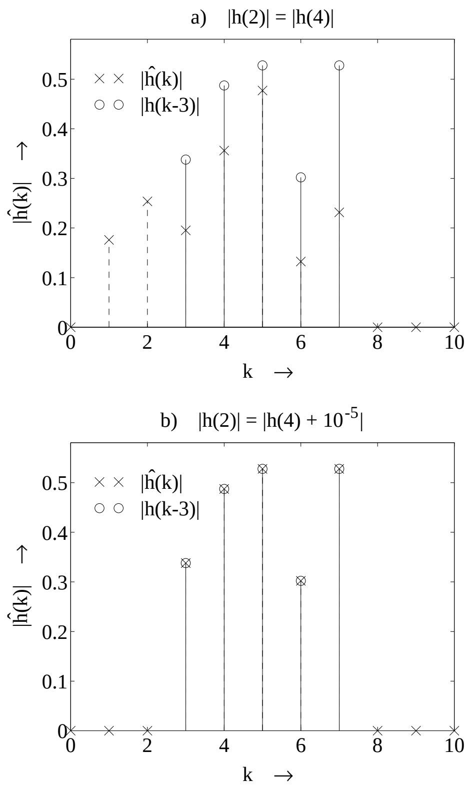


Fig. 2: EVI "estimates" obtained from eq. (30) assuming true statistics. Two largest coefficients of $w(k)$ with a) ... identical magnitudes b) ... with a difference of $\approx 10^{-5}$ in magnitudes

3.3 The two step EVI approach

Assuming the knowledge of the true values of $\widetilde{\mathbf{C}}_4^{yv}$ and $\widetilde{\mathbf{R}}_{inv}$ in equation (30), we have shown in the above example that it is very unlikely to end up with a wrong channel "estimate", because this requires $f(k)$ to be selected such that $w(k)$ has at least two largest coefficients with

[8] Strictly speaking, EVI's estimates are unbiased in the noiseless case, only. However, we show in section 4 that they are very robust with respect to additive white Gaussian noise. Although no proof is available, a justification for this robustness is given in Appendix E.

EIGENVECTOR APPROACH TO BLIND IDENTIFICATION (EVI):

S1: Execute some EVA iterations (as described in section 2 and [20]) to generate an appropriate reference system.

S2: Set up matrices $\widetilde{\mathbf{C}}_4^{yv}$ (submatrix of $\mathbf{C}_4^{yv}$ calculated in the final EVA iteration) and $\widetilde{\mathbf{R}}_{inv}$ (upon inversion of the $(4q+1) \times (4q+1)$ autocorrelation matrix $\widetilde{\mathbf{R}}_{vv}$) and solve the modified EVI equation (30). Choose the eigenvector associated with the maximum magnitude eigenvalue and take the complex conjugate reverse vector as an estimate $\hat{h}(k)$ of the channel impulse response.

Table 1: The EIGENVECTOR APPROACH TO BLIND IDENTIFICATION (EVI)

identical magnitudes. However, if these matrices are estimated from finite data blocks, the estimation *variance* will depend on the actual impulse response $w(k)$, too. To obtain minimum variance channel estimates, equation (17) must be approximated[9]. This can be done by adjusting $f(k)$ with the iterative procedure described in the end of section 2: $f(k)$ is loaded with the equalizer impulse response calculated in the previous step. This results in a two step identification algorithm, termed EIGENVECTOR APPROACH TO BLIND IDENTIFICATION (EVI), where some EV<u>A</u> iterations are executed prior to the solution of the modified EV<u>I</u> equation (30). Table 1 provides a description of EVI.

Its main parameters are:

S1:
- The order ℓ of the equalizer and the reference system.
- The number L of received data samples $v(k)$ used to estimate $\mathbf{R}_{vv}$ and $\mathbf{C}_4^{yv}$.
- The number of EVA iterations executed.

S2:
- The estimated channel order $\hat{q}$. As EVI does not suffer from choosing $\hat{q} > q$, $\hat{q}$ can also be considered the maximum expected (effective) channel order.
- The number L of received data samples $v(k)$ used to estimate $\widetilde{\mathbf{R}}_{vv}$ and $\widetilde{\mathbf{C}}_4^{yv}$.

Remark: Although EVA is executed in step S1, note that we are *not* interested in perfect equalization, here. *Even without step S1,* i.e. with the reference system set to $f(k) = \delta(k-k_0)$, EVI yields the true channel impulse response (assuming true values of the correlation and cumulant sequences as well as the noiseless case). For this to be true, we do *not* need a reference system that equalizes the channel (therefore, the equalizer order ℓ can be relatively small in step S1). We just require the combined channel/reference system impulse response $w(k)$ have a single maximum magnitude value (condition (15)). Only when considering estimation *variance* in terms of the number of data samples used to estimate the correlation and cumulant sequences, a deconvolving reference system will prove to be favorable. These are the two reasons why we execute EVA iterations in step S1 prior to the identification step S2.

Summary: Since EVI yields unbiased[8] channel estimates with low standard deviation, it represents an efficient fast approach to blind MA system identification. Remember that this is true even if the channel order q is overestimated ($\hat{q} > q$). In section 4, we show that EVI is also robust with respect to additive white Gaussian noise.

[9]Notice that this also avoids the effects of an "unlucky guess" of $f(k)$.

4 Simulation results

In this section, the estimation performance of EVI (as described in section 3.3) is illustrated by five simulation results, which are based on the channel examples C1 to C5 described in Table 2. With the synthetic time-invariant MA and ARMA channels C1 to C4, the quality of EVI's estimates is demonstrated *(i)* in terms of the number of received data samples (cf. requirement R1 in the introduction), *(ii)* with mis-assumptions for the channel order (R3), and *(iii)* in presence of additive noise (R4). With the help of channel C5, which stands for a collection of piecewise time-invariant approximations of realistic mobile radio channels under *GSM (Global System for Mobile Communications)* conditions, we investigate whether or not the four requirements R1 to R4 can be satisfied *at the same time*.

Ch. ex.	Order q	Description
C1	4	Critical MA channel with two equal magnitude coefficients and two critical zeros
C2	4	Equivalent minimum-phase system (same magnitude transfer function as C1)
C3	∞	ARMA system (first order all-pass)
C4	4	Critical MA channel with two equal magnitude coefficients and four critical zeros
C5	≤ 4	Collection of nine mobile radio sample channels

Table 2: Channel examples used for the simulations

EVI's estimation performance is also compared with the W-SLICE algorithm by Fonollosa and Vidal [12]. We will refer to this method as the WS algorithm. It computes the channel impulse response as a linear combination of 1D (auto)cumulant slices, where we apply fourth order cumulants only to avoid the problem of weighting between the sets of 4th and 2nd order cumulants (unfortunately, this problem is not addressed in [12]). To exploit a maximum amount of statistical information, we use the complete set of fourth order cumulants for WS.

Table 3 summarizes the main parameters of the simulations. Referring to Fig. 1, an i.i.d. *QPSK (Quaternary Phase Shift Keying)* or *BASK (Binary Amplitude Shift Keying)* random sequence $d(k)$ is propagated through one of the channel examples C1 to C5. The resulting steady state channel output sequence is corrupted by independent stationary zero mean additive white Gaussian noise (AWGN) $n(k)$ according to a given signal-to-noise ratio S/N. Based on a block of L received data samples $v(0), \cdots, v(L-1)$, the cumulant and correlation sequences required by the respective algorithm are estimated by unbiased sample averaging. Finally, the channel impulse response estimate $\hat{h}(k)$ is calculated according to the respective approach.
In the frame of Monte-Carlo runs, M_c different channel input sequences $d(k)$ are generated to obtain M_c estimated channel impulse responses $\hat{h}^{(\mu)}(k)$ with $\mu = 1, \cdots, M_c$. Estimation quality is assessed by the mean $\pm$ standard deviation values or the *normalized mean square error* of the estimates. However, as all blind system identification algorithms can *not* estimate one complex factor, each estimate is multiplied with the optimum constant (minimizing the estimate's Euclidean distance from the true channel impulse response) *before* estimation quality is assessed.

4.1 Synthetic channel examples C1 to C4

With the following four simulation results, QPSK signalling is applied to the synthetic channel examples. For C1 to C4, Figure 3 displays the zeros of $H(z) = \mathcal{Z}\{h(k)\}$ in the complex z-plane as well as the magnitude of the (complex) impulse responses $h(k)$.

Figure 4: Channel example C1 (see Fig. 3a) is rather unfavorable for EVI since two channel

Fig. no.	Modulation $d(k)$	Channel example	AWGN S/N	Number L of received samples	Simulation demonstrates ...
4	QPSK	C1	$=\infty$	$50, \cdots, 5000$	EVI and WS convergence rates
5	QPSK	C2	$=\infty$	$50, \cdots, 5000$	EVI and WS convergence rates
6	QPSK	C3	$=\infty$	500, 2000	EVI allpass approximation
7,11	QPSK	C4, C1	≥ 3 dB	5000	EVI performance in presence of noise
10	BASK	C5	≥ 7 dB	142 (GSM)	EVI/WS performance with short records

Table 3: Overview of the main simulation parameters

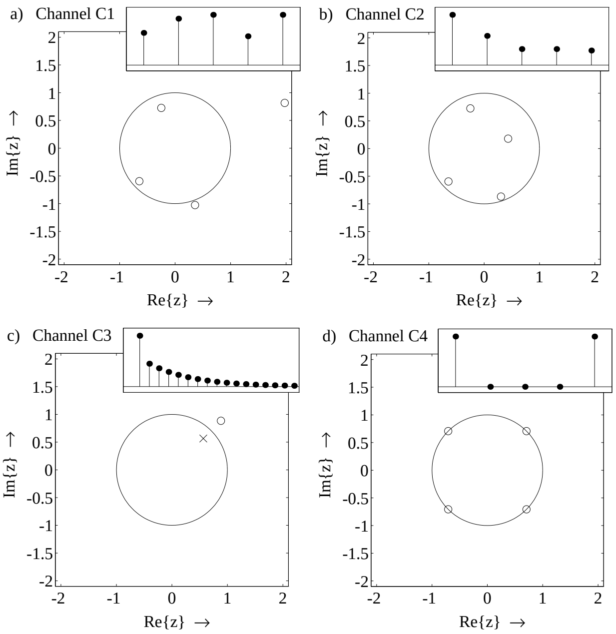


Fig. 3: Zero plots and magn. impulse responses of the synthetic channel examples C1 to C4

zeros are close to the unit circle[10] and two coefficients of the impulse response are exactly equal in magnitude. The former property would require a large equalizer order for a proper equalization (nevertheless, we select for the EVA iterations in step S1 of EVI an equalizer of order $\ell = 16$ only, which is sufficient for the generation of an adequate reference system). On account of the latter property, the optimization of the reference system requires a relatively high number of EVA iterations (eight in this example).

In Figure 4, the mean (solid lines) $\pm$ standard deviation values (dotted) of the $\hat{q}+1 = 9$ coefficients[11] of $M_c = 50$ impulse responses estimated by EVI (Fig. a) and WS (Fig. b) in the noiseless case are given in terms of the blocklength L. The magnitudes of the true channel coefficients are indicated by arrows on the right margins. This example illustrates that EVI yields unbiased estimates for $L \geq 500$ approximately, whereas the WS method delivers biased coefficients even with 5000 data samples. As for the estimation variance, EVI achieves comparable values of variance with blocklengths smaller by a factor of about 10.

Remark: For channel C1, a comparison of EVI with the GM method [14] was presented in [25]. Similar to the WS result, the GM method delivered biased coefficients with 5000 data samples (even with 500000 samples, in fact). Actually, there are quite a few algorithms (e.g. [38, 39, 1]) which are, just as the GM method, based on the same quadratic equation (see eq. (8) in [1], e.g.). As the performance levels obtained by these algorithms seem to be roughly of the same order of magnitude, we do *not* consider any representative of this class of algorithms in this paper.

Figure 5: Channel example C2 (see Fig. 3b) is the equivalent minimum-phase system of channel C1, i.e. all zeros z_0 of C1 with $|z_0| > 1$ are moved to the location $1/z_0^*$ inside the unit circle

[10] The minimum distance is 0.0797 corresponding to a spectral null which is 21.1 dB below the spectral peak. Unnormalized channel impulse response of C1: $h(k) = [0.4+0.3j,\ -0.6-0.4j,\ 0.5-0.6j,\ 0.2-0.4j,\ 0.6-0.5j]$.

[11] In this section, the "estimated channel order" $\hat{q}$ is defined as the number of estimated channel coefficients minus one.

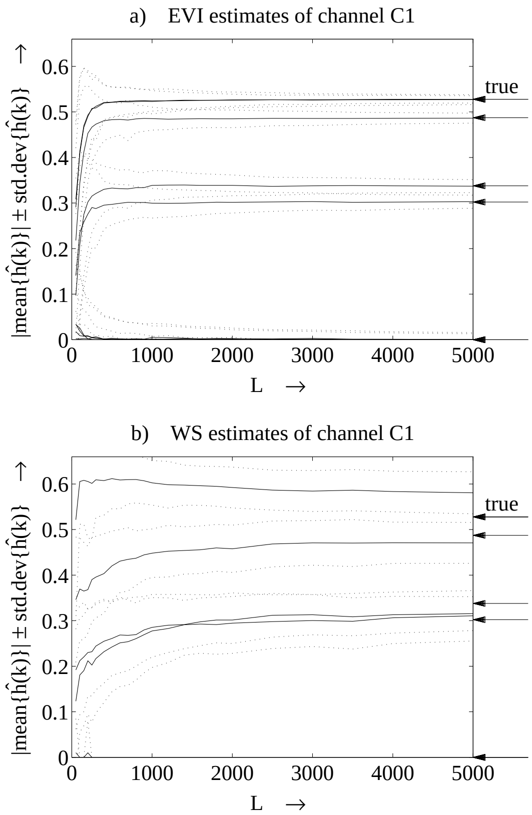


Fig. 4: Estimation of the MA channel C1 in terms of the blocklength L: a) EVI as described in section 3.3; b) W-SLICE (WS) algorithm [12].

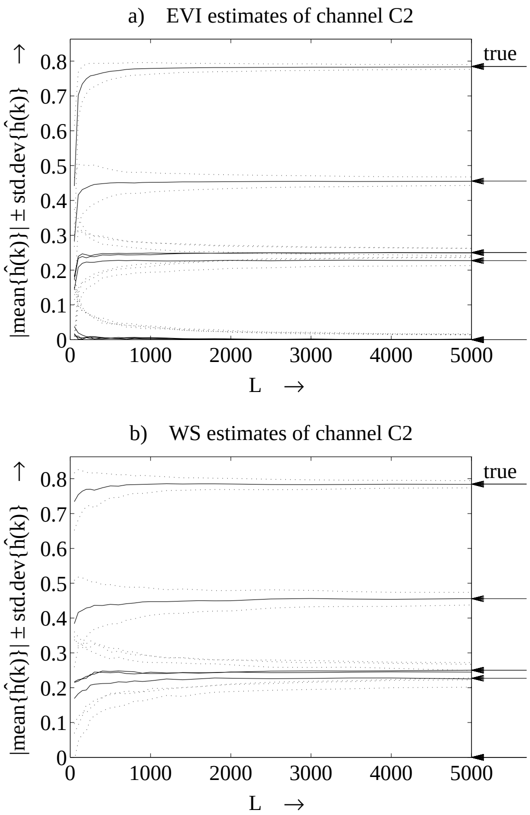


Fig. 5: Estimation of the MA channel C2 in terms of the blocklength L: a) EVI as described in section 3.3; b) W-SLICE (WS) algorithm [12].

of the complex z-plane. From Figure 5b, we realize that the WS estimation quality of minimum-phase channels such as C2 is largely superior to that of a mixed-phase channel (cf. Fig. 4b). Just as EVI's estimates of C1 in Fig. 4a, the WS estimates of C2 are unbiased from about $L = 500$ samples on. In the case of EVI, the variance of C2's estimates is also lower than that for C1, especially for small sample sizes ($L < 500$). In summary, EVI delivers a slightly better estimation performance for channel C2.

Figure 6: Channel example C3 (see Fig. 3c) is a first-order allpass. Although this is a recursive system, its impulse response can be approximated by a finite length impulse. Remember that the effective length of the allpass impulse response depends on the pole's magnitude $|z_\infty|$. In this example, we pick $z_\infty = 0.8e^{j\pi/4}$. This allows an impulse response approximation by an order 16 MA model.

Figures 6a and d display the magnitude (mean $\pm$ standard deviation) of the $M_c = 25$ impulse responses estimated on the basis of $L = 500$ and 2000 received data samples, respectively ($\hat{q} = 32$). Figures 6b,e and 6c,f show the corresponding magnitude of the frequency response and the group delay. Again, the mean $\pm$ the standard deviation values are marked (solid and dotted lines, respectively). The true values are indicated by dashed lines.

Figure 7: As opposed to fourth order cumulants, the correlation sequence is influenced by independent stationary additive Gaussian noise (AGN). Since the modified EVI equation (30) uses both the correlation and the cumulant sequences, it is obvious that the EVI solution is

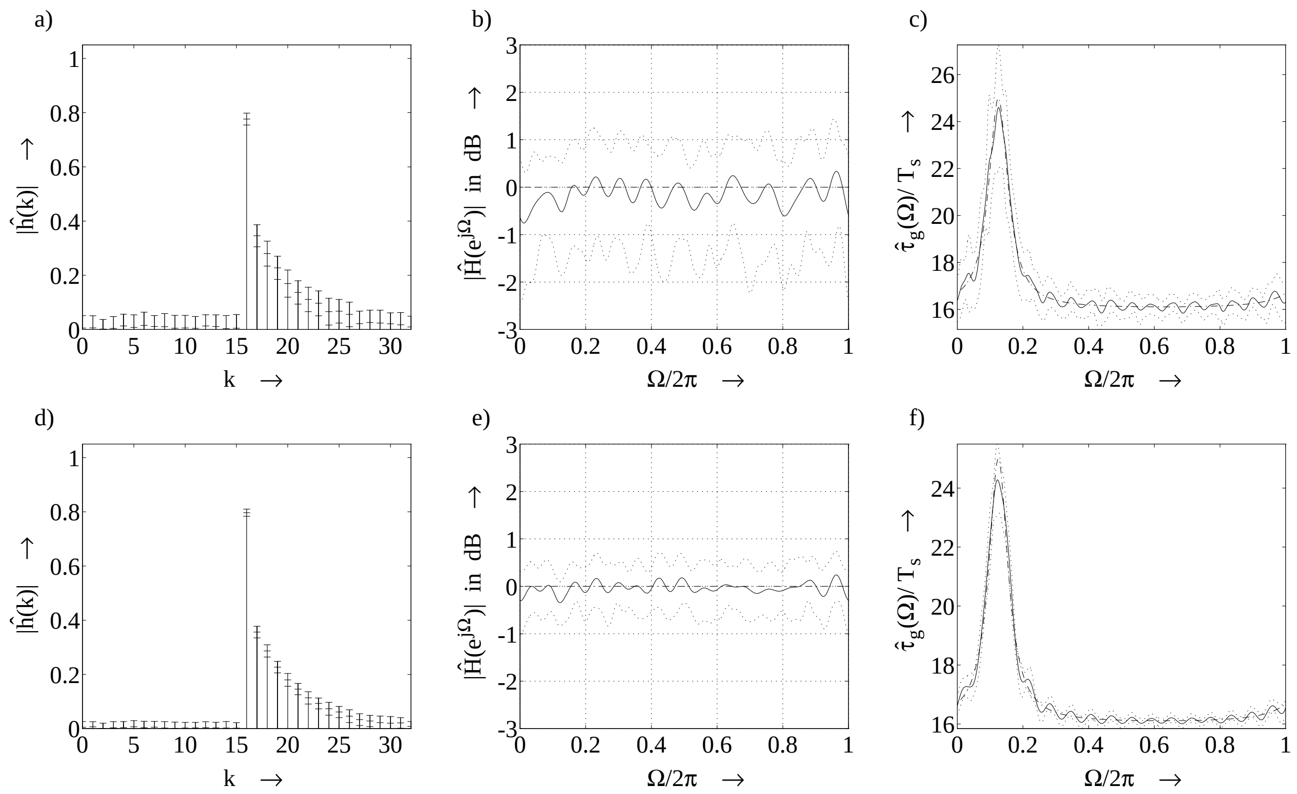


Fig. 6: EVI estimation of the ARMA channel C3 (see Figure 3c)
a-c) $L = 500$ received data samples used to estimate EVI's matrices; d-f) $L = 2000$

degraded by AGN. Some examples to illustrate this degradation due to zero mean additive *white* Gaussian noise (AWGN) are given in the sequel. In this case, the autocorrelation sequence is disturbed at lag zero only, i.e. the elements on the main diagonal of the matrix $\widetilde{\mathbf{R}}_{vv}$ are increased by the value of the noise power.

For the fourth order MA channels C4 and C1, Figure 7 demonstrates EVI's estimation performance in terms of the *noise-to-signal* (N/S) ratio of the received signal. In Figures 7a to c, channel C1 according to Fig. 3a is reconsidered, whereas in Figures 7d to f, the channel example C4 with transfer function $H(z) = 1 + 0.999\, z^{-4}$ is regarded (see Fig. 3d). As the two largest coefficients of its impulse response have nearly the same magnitudes and four zeros are extremely close to the unit circle[12], C4 is critical for a proper adjustment of the reference system coefficients in step S1 of EVI.

Figures 7a and d display the magnitudes of the estimated channel coefficients as a function of N/S, where true autocorrelation values with superimposed noise power (at lag zero) and true cumulant values are used. Thus, these results correspond to the asymptotic case, where an infinite number L of received data samples is available for the estimation of the correlation and cumulant sequences. For comparison, the true magnitude channel coefficients are indicated by arrows on the right side. First, we realize from these figures that the channels are correctly identified in the noiseless case. For channel C1 (Fig. 7a), the degradation due to AWGN is negligible up to (at least) $N/S = 0.5$ ($S/N = 3\,$dB). From Fig. 7d, we see that even with the critical channel C4, the degradation is very small, although the effect of AWGN is increased. Ignore, for the moment, the negative values of N/S.

[12] $|z_{01}| = \ldots = |z_{04}| = 0.99975$ leading to spectral nulls 66 dB below the spectral peak.

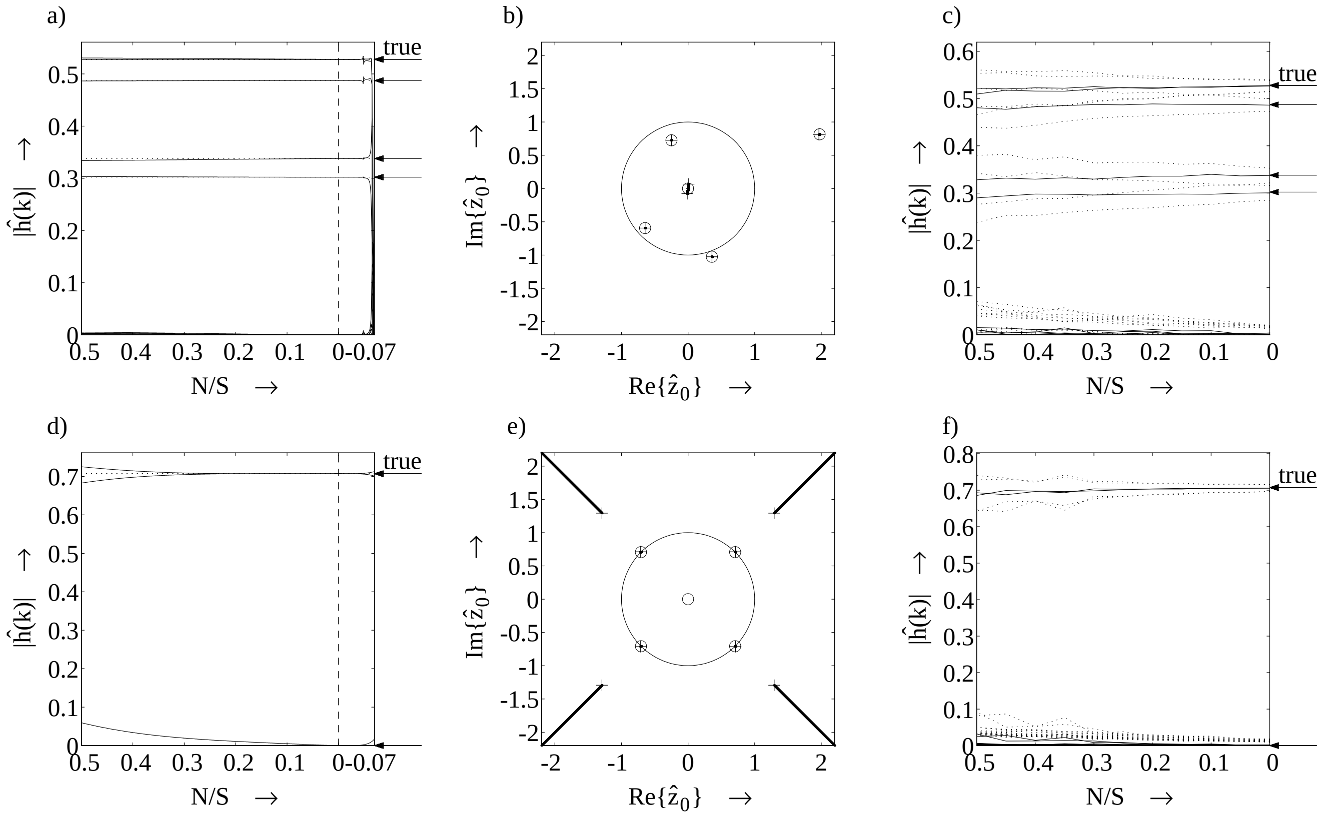


Fig. 7: EVI estimation performance in presence of additive white Gaussian noise
a-c) Channel C1 (see Figure 3a); d-f) Channel C4 (see Figure 3d)

In Figures 7b and e, the trajectories of the "estimated" zeros are drawn for N/S ratios ranging from 0 to 0.5. As we overestimate the channel order $q = 4$ in these examples ($\hat{q} = 8$), true zeros (marked with circles) are added in the origin of the complex z-plane (and in $|z| = \infty$). The trajectories emerge from the true zero locations corresponding to the noisefree case ($N/S = 0$). With the N/S ratio reaching the value of 0.5, the zeros approach their final position marked with plus symbols ("+"). We find that the channel zeros are correctly identified for the complete range of N/S ratios. Further zeros are introduced which leave their ideal positions (in $|z| = 0$ or $|z| = \infty$) as the noise power is increased. Since these additional zeros are spaced equidistantly on circles, their influence on the estimated channel impulse response is negligible. Note that a justification for EVI's robustness with respect to AWGN is given in Appendix E.

Figures 7c and f display (in terms of N/S) the mean (solid lines) $\pm$ standard deviation (dotted) values of the coefficients of $M_c = 100$ impulse responses estimated by EVI on the basis of $L = 5000$ received data samples. We can state that the mean values are quite robust to an increase in the noise power. Just the standard deviation increases as S/N degrades.

The remaining noise influence can be compensated for provided that the noise correlation sequence is known. If this is not the case, it can be estimated with an alternative method based on fourth order cumulants [24]. With this approach, the autocorrelation estimates of the received signal do not contain Gaussian noise whereas the conventional autocorrelation estimates do. The difference between both yields the noise correlation.

To compensate for the noise, EVI's matrix $\widetilde{\mathbf{R}}_{vv}$

must be rectified. In case of AWGN, its power has to be subtracted from the main diagonal only. In order to assess the influence of a *noise overcompensation*, *negative values* of the N/S ratios are also considered in the Figures 7a and d. Figure 7a reveals that an overcompensation may cause severe degradations.

4.2 Realistic mobile radio channel example C5

In this section, we attempt to estimate mobile radio communication channels on the assumptions typically made in GSM receivers *(Global System for Mobile Communications)*. According to the GSM standard (refer to [28], e.g.), information symbols are transmitted in *bursts* where each "normal" burst (see Fig. 8) contains two packets of 58 data symbols (bits) surrounding a training sequence of 26 bits. For (non-blind) channel estimation, state-of-the-art GSM receivers use the training sequence, while, for *blind* channel estimation, almost the entire burst could be used (142 symbols).

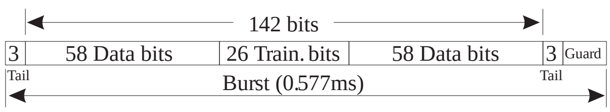


Fig. 8: GSM "normal" burst

In a mobile scenario, the physical multipath radio channel is *time-variant* with a baseband impulse response depending on the time difference τ between the observation and excitation instants as well as the (absolute) observation time t. We adopt the stochastic zero mean *Gaussian Stationary Uncorrelated Scattering (GSUS)* model [31] leading to the following impulse response of the composite channel [18]

$$h_c(\tau,t) = \frac{1}{\sqrt{N_e}} \sum_{\nu=1}^{N_e} e^{j(2\pi f_{d,\nu} t + \Theta_\nu)} \cdot g_{TR}(\tau - \tau_\nu)\,, \tag{31}$$

where N_e is the number of elementary echo paths, $g_{TR}(\tau)$ denotes the combined transmit/receive filter impulse response, and the subscript in $h_c(\cdot)$ suggests its continuous-time property. 3D sample impulse responses can easily be determined from (31) by independently drawing N_e Doppler frequencies $f_{d,\nu}$, N_e initial phases Θ_ν, and N_e echo delay times τ_ν from random variables with Jakes, uniform, and piecewise exponential probability density functions, respectively. As for the echo delay times τ_ν, we use standard COST-207[13] *Typical Urban (TU)*, *Bad Urban (BU)* and *Hilly Terrain (HT)* profiles.

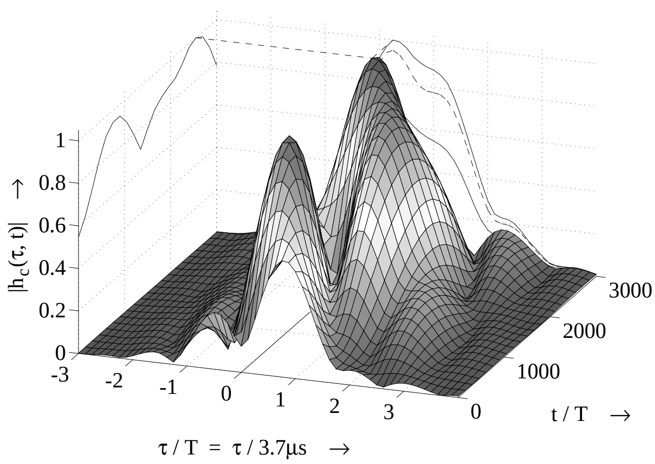


Fig. 9: Sample magnitude impulse response of a COST-207 *Bad Urban (BU)* channel with raised cosine transmit/receive filtering ($r = 0.5$), $T = 3.7\,\mu$s, $f_{d,max} = 88$ Hz

Figure 9 shows, in non-causal representation, a sample magnitude impulse response $|h_c(\tau,t)|$, obtained from equation (31) with $N_e = 100$, of a *Bad Urban* channel with raised cosine[14] transmit and receive filters ($r = 0.5$). Both time axes are normalized to the GSM symbol (bit) period $T \approx 3.7\,\mu$s. The velocity of the mobile unit is $v = 100$ km/h. Assuming a carrier frequency of 950 MHz, this leads to a maximum Doppler shift of $f_{d,max} = 88$ Hz. Equation (31) was evaluated

[13] (European) Cooperation in the field of Scientific and Technical research, project #207.

[14] The authors are well aware of the fact that GSM transmitters use non-linear GMSK modulation *(Gaussian Minimum Shift Keying)*. However, we consider linear modulation in this example in order to avoid systematic errors in the channel estimates. Simulation results taking into account transmitter non-linearity are given in [6, 9, 8].

over a t range covering one minimum Doppler period $T_{d,min} = 1/f_{d,max} = 3080T = 11.4$ ms.

Assuming *quasi time-invariance over one burst*[15], $h_c(\tau, t)$ is sampled on the t axis each 150 bits (cf. Fig. 8). This produces 21 slices within the t range of $3080T$, which can be seen in Fig. 9 as surface lines parallel to the τ axis. Furthermore, each slice is sampled at $\tau = kT$ and then constrained to the index range k where the associated sample power delay spectrum exceeds the threshold of 1% of its maximum value.

$$h(k,\xi) \triangleq h_c(\tau,t) \quad \text{for} \begin{cases} \tau = kT \\ t = \xi\, 150\, T \end{cases}, \qquad (32)$$

where $\xi = 0, \cdots, 20$.

On the above assumptions, nine different sample GSUS composite channels $h_c(\tau, t)$ were obtained from eq. (31) by combining three COST-207 propagation environments *(TU, BU, HT)* with three raised cosine transmit/receive filters: roll-off factors $r \in \{0.9, 0.5, 0.1\}$. Let *BU-(0.5)* denote the *Bad Urban* channel with roll-off factor $r = 0.5$, e.g.. According to eq. (32), each channel $h_c(\tau, t)$ was decomposed into 21 slices $h(k, \xi)$. This collection of 9 piecewise time-invariant channels will be called "channel set C5". Referring to Fig. 1 with $h(k, \xi)$ substituted for $h(k)$, $M_c = 100$ bursts of 150 i.i.d. *BASK (Binary Amplitude Shift Keying)* symbols $d(k)$ were propagated through each channel slice. Then, WS and EVI were applied to $L = 142$ samples of $v(k)$. Both algorithms were given the effective length of the sample power delay spectrum, which is equivalent to the mean length of the channel impulse response. Note that the actual effective length of a channel slice may well be shorter due to time selective fading.

Estimation quality measure: Let $\hat{h}^{(\mu)}(k, \xi)$ denote the estimate of $h(k, \xi)$ based on the μ^{th} input burst ($\mu = 1, \cdots, M_c$). For each slice index ξ, estimation quality is assessed on the basis of the averaged *Normalized Mean Square Error (NMSE)*

$$\text{NMSE}(\xi) \triangleq \frac{1}{M_c} \sum_{\mu=1}^{M_c} \frac{\sum_k |\hat{h}^{(\mu)}(k,\xi) - h(k,\xi)|^2}{\sum_k |h(k,\xi)|^2}. \qquad (33)$$

Figure 10: From the set C5 of nine sample channels described above, we have selected for Figure 10 six examples by combining the propagation environments TU, BU and HT with the roll-off factors $r = 0.5$ (Figures a to c) and 0.1 (Figures d to f). For each channel, Fig. 10 shows the NMSE(ξ)-values (in per cent) of WS' and EVI's estimates for different values of S/N, where the noiseless case is marked by "$\circ$" symbols, while "$\times$" and "+" stand for $S/N = 10$ dB and 7 dB, respectively [5]. The NMSE(ξ)-values for WS are connected by dotted lines, those for EVI by solid lines. Note the different scaling on the NMSE(ξ) axes of Figure 10.

From Figures 10a and d, we realize that EVI can estimate the TU channels very well, even at $S/N = 7$ dB ("+"): the NMSE(ξ) values are below 3% for most slices. Conversely, assuming an acceptance threshold value of 5% (dashed lines), WS can *not* identify slices $\xi = 12$ to 17 of channel *TU-(0.5)* (Fig. a) and all but the last two slices of *TU-(0.1)* (Fig. d). It turns out that these slices are mixed-phase, so that these results are in accordance with those presented in Figure 4 for the mixed-phase channel C1. Apart from the last slice ($\xi = 20$), EVI largely outperforms WS at any given S/N.

Comparable statements can be made for the BU channels in Figures 10b and e (note that the channel used for Fig. 10b was shown in Fig. 9): although there are channel slices, where WS and EVI perform almost equally well ($\xi = 6$ to 8 and 18 to 20), there is a huge performance gap for the mixed-phase channel slices ($\xi = 0$ to 2 and 10 to 14). For all slices of *BU-(0.5)* and 18 out of 21 slices of *BU-(0.1)*, EVI's NMSE(ξ)-values remain below 5%, while this is true for WS' estimates for 10 *(BU-(0.5))* and 2 *(BU-(0.1))* slices, only. Again, for all slices and S/N ratios, EVI's performance is superior to that of WS.

[15] Notice that this is also supposed in state-of-the-art (non-blind) GSM receivers.

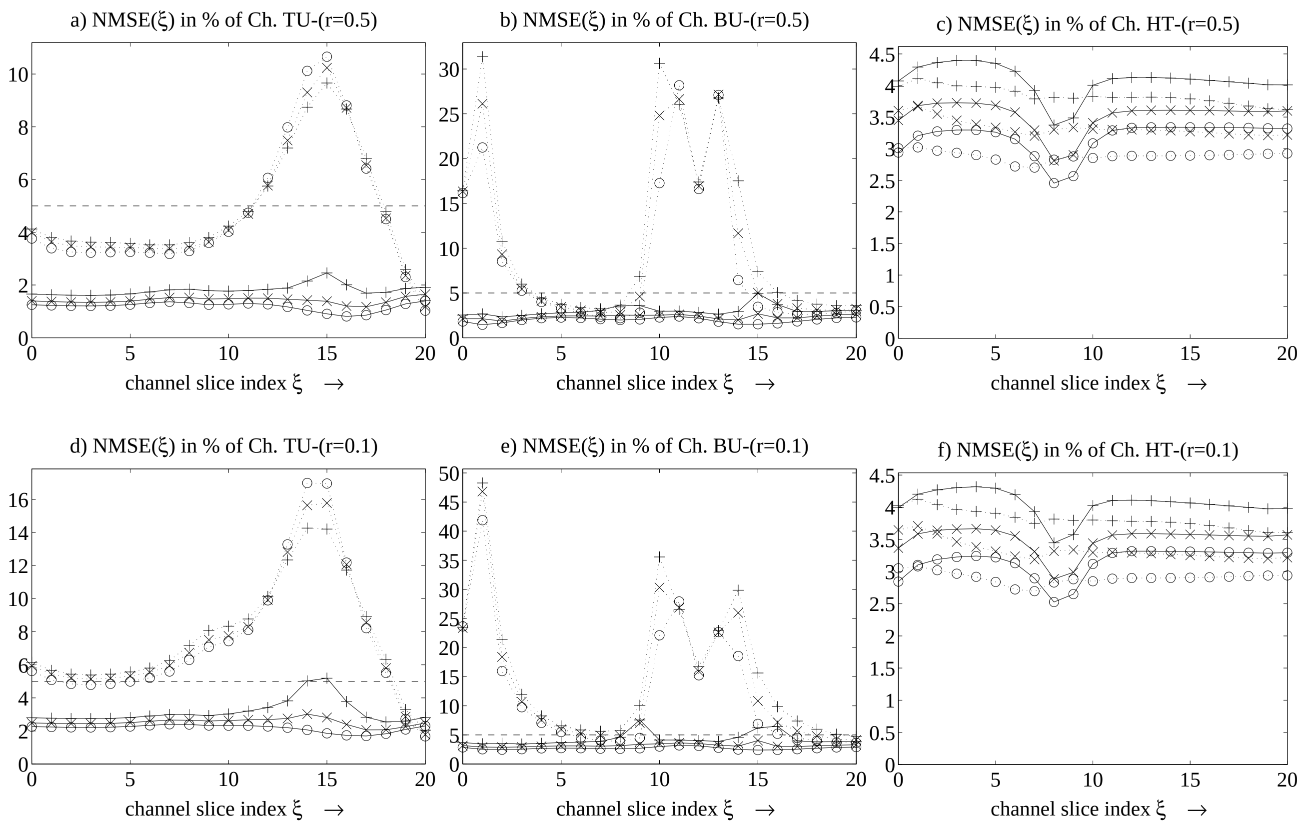


Fig. 10: EVI (solid) and WS (dotted) estimates of channels C5 from 142 samples
AWGN: "∘": $S/N = \infty$, "×": 10 dB, "+": 7 dB
a, d) *Typical Urban (TU)*, b, e): *Bad Urban (BU)*, c, f): *Hilly Terrain (HT)*
a-c): Roll-off factor $r = 0.5$, d-f): roll-off factor $r = 0.1$

In case of the *HT* channels, we can see from Figures 10c and f that both approaches successfully identify all slices: all NMSE(ξ)-values are around 3% in the noiseless case and around 4% at 7 dB. If we average NMSE(ξ) over all slice indices ξ to obtain $\overline{\text{NMSE}}$, it turns out that WS' performance is slightly better than that of EVI. It should be noted that, by mere coincidence, all channel slices were minimum-phase in this example.

Finally, Table 4 summarizes the $\overline{\text{NMSE}}$-values for the entire set C5 of sample channels, where those exceeding 5% are marked in boldface [5]. The first value in each column refers to WS, while the second applies to EVI. We realize that WS can not identify satisfactorily the *BU* channels as well as *TU-(0.1)*. On the other hand, EVI is capable of estimating all channel examples at S/N ratios down to 7 dB (10 dB) to within an outstanding $\overline{\text{NMSE}}$ bound of 4.3% (3.5%). This demonstrates that EVI can satisfy the requirements R1 to R4 *at the same time*.

It should be noted that EVI's estimates of critical channels (those with zeros very close to or on the complex plane's unit circle) can be improved by adjusting EVI's parameters: in the initial equalization step, the equalizer order and/or the number of iterations could be increased. As for WS, second order statistics could be used in addition to cumulants in order to improve performance. However, this evokes the problem of how to weight correlation estimates against cumulants. Furthermore, for low S/N ratios, the estimates will be biased.

S/N	*TU-(0.9)*	*TU-(0.5)*	*TU-(0.1)*
∞ dB	2.9 / 1.2	4.8 / 1.2	**7.5** / 2.2
10 dB	3.0 / 1.4	4.8 / 1.4	**7.6** / 2.5
7 dB	3.1 / 1.8	4.8 / 1.8	**7.7** / 3.2
S/N	***BU-(0.9)***	***BU-(0.5)***	***BU-(0.1)***
∞ dB	**7.1** / 1.2	**8.6** / 2.0	**12.2** / 2.7
10 dB	**8.3** / 1.5	**9.9** / 2.3	**14.0** / 3.2
7 dB	**9.5** / 2.0	**11.2** / 3.0	**15.6** / 4.3
S/N	***HT-(0.9)***	***HT-(0.5)***	***HT-(0.1)***
∞ dB	2.9 / 3.2	2.9 / 3.2	2.9 / 3.2
10 dB	3.3 / 3.5	3.3 / 3.5	3.3 / 3.5
7 dB	3.8 / 4.1	3.8 / 4.1	3.8 / 4.1

Table 4: $\overline{\text{NMSE}}$ [in per cent] for WS's and EVI's estimates of channels C5: *Typical Urban (TU)* (above), *Bad Urban (BU)* and *Hilly Terrain (HT)* (below)

5 Conclusions and further work

In this paper, we have presented a novel algorithm (EVI) for an efficient blind identification of possibly mixed-phase FIR systems. In the noiseless case, we have shown how to obtain unbiased channel estimates by the solution of a modified eigenvector problem derived from the closed-form EVA solution to blind equalization. The modifications to the eigenvector problem also drastically reduce the computational complexity of this approach. Furthermore, a two step procedure has been introduced in order to ensure the uniqueness of the estimate and to minimize estimation variance for a given number of received data samples used to estimate the required correlation and cumulant coefficients.

Various simulation results illustrate EVI's brilliant convergence properties. EVI's estimation performance was also investigated in presence of additive white Gaussian noise. It was demonstrated by simulation results that the true channel zeros are still identified properly. Although some additional zeros were introduced, their influence was very small even at signal-to-noise ratios as low as 3dB.

Although the HOS class of blind identification approaches is said to require excessive amounts of samples of the received signal to achieve acceptable performance levels, we have finally demonstrated in this paper that it *is* possible with EVI to blindly estimate realistic mobile radio channels from one demodulated GSM burst (142 samples, cf. requirement R1 in the introduction). At a constant signal-to-noise ratio of 7 dB (cf. R4), all channel slices (R2, R3) were identified within a normalized mean square error bound of 5 per cent. Thus, EVI can meet the requirements R1 to R4 *at the same time*. To the knowledge of the authors, no other HOS-based algorithm is capable of achieving the cited $\overline{\text{NMSE}}$ levels on the assumptions made in this paper. Furthermore, we have shown that the WS algorithm's performance level heavily depends on the actual channel impulse response.

While this paper concentrated on the quality of blind channel estimates, further work will be directed towards a comparison of the *bit error rates* attainable from blind and non-blind channel estimates. First results based on GSM were published recently [6, 9]. A comprehensive study on the feasibility of blind channel estimation in GSM systems will be provided in [8].

Remark: MATLAB programs implementing both EVA and EVI as well as compressed postscript files of preprints of related publications are readily available from our WWW server at the address `http://www.comm.uni-bremen.de`.

Acknowledgements

We would like to express our gratitude to the anonymous reviewers for having compiled comprehensive and thorough reviews leading to numerous modifications throughout the paper. We also thank the reviewers for their patience with respect to a considerable delay in the revision process due to changes of affiliation of all authors.

Appendices

A Decomposition of the cross-cumulant matrix $\widetilde{\mathbf{C}}_4^{yv}$

The cross-cumulant matrix $\widetilde{\mathbf{C}}_4^{yv}$ in equations (27) and (30) with the elements

$$[\widetilde{\mathbf{C}}_4^{yv}]_{i_1,i_2} = c_4^{yv}(-i_1 - q, 0, -i_2) \qquad (A.1)$$
$$= \gamma_4^d \sum_k h^*(k - i_1 - q)h(k - i_2)\,|w(k)|^2$$

(with $i_1 \in \{0, \cdots, 2q\}$ and $i_2 \in \{0, \cdots, 4q\}$) can be decomposed into the following triplet of matrices

$$\widetilde{\mathbf{C}}_4^{yv} = \gamma_4^d \cdot \mathbf{H}^*_{0,3q+1\times 2q+1} \cdot \mathbf{W} \cdot \mathbf{H}_{q,3q+1\times 4q+1} \quad (A.2)$$

where $\mathbf{W}$ is the $(3q+1 \times 3q+1)$ diagonal matrix

$$\mathbf{W} \triangleq \mathrm{diag}\{|w(0)|^2, \cdots, |w(3q)|^2\} \qquad (A.3)$$

and $\mathbf{H}_{0,3q+1\times 2q+1}$ and $\mathbf{H}_{q,3q+1\times 4q+1}$ represent filtering matrices. Generally, we define the channel matrix $\mathbf{H}_{a,b\times c}$ as the $(b \times c)$ Toeplitz matrix with top left element $[\mathbf{H}_{a,b\times c}]_{0,0} = h(a)$, while the following elements $h(i)$ of the first row (column) have decreasing (increasing) time lags i. Examples are

$$\mathbf{H}_{q,3q+1\times 4q+1} \triangleq \begin{bmatrix} \underline{h(q)} & \cdots & h(0) & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & h(q) & \cdots & h(0) \end{bmatrix} \qquad (A.4)$$

and

$$\mathbf{H}_{0,3q+1\times 2q+1} \triangleq \begin{bmatrix} \underline{h(0)} & & 0 \\ \vdots & \ddots & \\ h(q) & \ddots & h(0) \\ & \ddots & \vdots \\ 0 & & h(q) \end{bmatrix}. \qquad (A.5)$$

The decomposition (A.2) will be used in the sequel.

B Effect of $\mathbf{R}_{inv}$ on the eigenvector problem (27)

In Appendix C, we demonstrate that unbiased channel estimates can be obtained from eq. (27), if the $(4q+1) \times (2q+1)$ submatrix $\mathbf{R}_{inv}$ of the $(\ell+1) \times (\ell+1)$ inverse autocorrelation matrix $\mathbf{R}_{vv}^{-1}$ is replaced with a modified matrix. This can be shown by clarifying the effect of $\mathbf{R}_{inv}$ on the eigenvector problem (27). For this purpose, we insert the decomposition (A.2) of $\widetilde{\mathbf{C}}_4^{yv}$ into equation (27) to obtain

$$\lambda \widetilde{\mathbf{h}}_{\mathrm{EVI}} = \gamma_4^d\, \mathbf{H}^*_{0,3q+1\times 2q+1} \mathbf{W} \mathbf{H}_{q,3q+1\times 4q+1} \mathbf{R}_{inv} \widetilde{\mathbf{h}}_{\mathrm{EVI}}\,, \qquad (B.1)$$

where $\widetilde{\mathbf{h}}_{\mathrm{EVI}}$ would contain the $q+1$ true channel coefficients as well as q zeros, if $\mathbf{R}_{inv}$ was the submatrix of the infinite dimensions inverse autocorrelation matrix. In Appendix D, we show that, as ℓ approaches infinity, $\mathbf{R}_{vv}^{-1}$ has a submatrix $\mathbf{R}_{inv}$ approximating Toeplitz(!) structure. Therefore, $\mathbf{R}_{inv}$ can also be explained as a filtering matrix. Note that the $(4q+1) \times (2q+1)$ Toeplitz matrix $\mathbf{R}_{inv}$ contains $6q+1$ coefficients of the system correlation sequence $r_{inv}(k)$ of the infinite length inverse channel. Hence, in equation (B.1), the result of $\mathbf{H}_{q,3q+1\times 4q+1}\mathbf{R}_{inv}\widetilde{\mathbf{h}}_{\mathrm{EVI}}$ corresponds to the convolution of the channel's impulse response $h(k)$ with $6q+1$ coefficients of the inverse channel correlation sequence $r_{inv}(k)$ and with $h^*(-k)$. Thus, we have $h(k) * r_{inv}(k) * h^*(-k)\sigma_d^2 = r_{inv}(k) * r_{vv}(k) = \delta(k)$, or equivalently, in vector notation

$$\mathbf{H}_{q,3q+1\times 4q+1}\mathbf{R}_{inv}\widetilde{\mathbf{h}}_{\mathrm{EVI}}\sigma_d^2 = \underbrace{[0, \cdots, 0, 1, 0, \cdots, 0]}_{3q+1}{}^T \triangleq \mathbf{i} \qquad (B.2)$$

By the multiplication with the diagonal matrix $\mathbf{W}$ according to (B.1), the vector $\mathbf{i}/\sigma_d^2$ is weighted with the constant $|w(k_m)|^2$. Then, $|w(k_m)|^2\,\mathbf{i}/\sigma_d^2$ selects a single column from $\mathbf{H}^*_{0,3q+1\times 2q+1}$, so that equation (B.1) finally reads as shown in equation (B.3) on page 24, where $q+1$ columns contain the complete channel impulse response (as indicated by the dashed rectangles)[16]. As long as the righthand side vector does not have its non-zero element on the first (or last) q elements, this selective property

[16]Note that for $|w(k_m)|^2\,\mathbf{i}/\sigma_d^2$ to select a single column, it is also sufficient that equation (17) holds.

obviously is sufficient for a proper identification of the channel by the solution of the eigenvector problem.

C Obtaining unbiased channel estimates (noiseless case)

We show now that it is possible to replace the matrix $\mathbf{R}_{inv}$ with a computationally less expensive Toeplitz matrix $\widetilde{\mathbf{R}}_{inv}$ while maintaining the selective property mentioned above. In equation (B.2), q columns of $\mathbf{R}_{inv}$ are multiplied with the q zero coefficients of the vector $\widetilde{\mathbf{h}}_{\mathrm{EVI}}$. Thus, just $5q+1$ out of the $6q+1$ coefficients $r_{inv}(k)$ are significant in $\mathbf{R}_{inv}$. Their selection depends on the position of $h(k)$ within the vector $\widetilde{\mathbf{h}}_{\mathrm{EVI}}$. According to $r_{inv}(k) * h^*(-k) = h^*(-k) * r_{inv}(k)$ we can state $\mathbf{R}_{inv}\widetilde{\mathbf{h}}_{\mathrm{EVI}} = \mathbf{H}^*_{0,5q+1\times 4q+1}\mathbf{r}_{inv}$, so that eq. (B.2) can be transformed into equation (C.1), where the matrix has the dimensions $(3q+1)\times(5q+1)$ and the $q+1$ values of $\varepsilon \in \{0,\cdots,q\}$ correspond to the $q+1$ different positions of $h(k)$ within $\widetilde{\mathbf{h}}_{\mathrm{EVI}}$ (length: $2q+1$). The $q+1$ systems of equations according to (C.1) can be collected into the single system (C.2). These are $4q+1$ equations for the $6q+1$ unknown parameters $r_{inv}(-3q),\cdots,r_{inv}(3q)$. Thus, $2q$ parameters can be chosen arbitrarily. Letting for instance $r_{inv}(-3q),\cdots,r_{inv}(-2q-1)=0$ and $r_{inv}(2q+1),\cdots,r_{inv}(3q)=0$, the first q columns as well as the last q columns of the matrix in equation (C.2) can be omitted.

$$\widetilde{\mathbf{i}} = \widetilde{\mathbf{R}}_{vv}\widetilde{\mathbf{r}}_{inv} \triangleq \begin{bmatrix} r_{vv}(0) & \cdots & r_{vv}(-q) & & 0 \\ \vdots & \ddots & & \ddots & \\ r_{vv}(q) & & \ddots & & r_{vv}(-q) \\ & \ddots & & \ddots & \vdots \\ 0 & & r_{vv}(q) & \cdots & r_{vv}(0) \end{bmatrix} \begin{bmatrix} r_{inv}(-2q) \\ \vdots \\ r_{inv}(2q) \end{bmatrix}. \tag{C.3}$$

Thus, we have obtained a new system of equations that also guarantees the selection of a single column from the channel matrix $\mathbf{H}^*_{0,3q+1\times 2q+1}$. If we solve equation (C.3) for the $4q+1$ coefficients $r_{inv}(-2q),\cdots,r_{inv}(2q)$ and use them for the Toeplitz matrix $\widetilde{\mathbf{R}}_{inv}$ given in equation (28), we will thus achieve unbiased channel estimates (in the noiseless case) from the "Modified EVI equation" (30) *irrespective of the actual impulse response* $w(k)$.

N.B.: Remember that the uniqueness of the solution still has to be ensured. This problem is addressed by an example in the end of section 3.2 and in section 3.3. Also note that by substituting $\widetilde{\mathbf{R}}_{inv}$ for $\mathbf{R}_{inv}$, the computational effort to calculate the inverse of the $(\ell+1)\times(\ell+1)$ autocorrelation matrix $\mathbf{R}_{vv}$ is reduced to calculating the inverse of the smaller $(4q+1)\times(4q+1)$ autocorrelation matrix $\widetilde{\mathbf{R}}_{vv}$.

D Toeplitz structure of the central infinite dimensions inverse autocorrelation matrix

While in general, the inverse of a non-singular Hermitian Toeplitz matrix (such as the $(\ell+1)\times(\ell+1)$ autocorrelation matrix $\mathbf{R}_{vv}$) just is Hermitian persymmetric [16], we prove here that the inverse of $\mathbf{R}_{vv}$ has a submatrix approximating Toeplitz(!) structure as ℓ approaches infinity. Note that a MATLAB program justifying the Toeplitz assumption can be retrieved from our WWW server.

We assume an i.i.d. input sequence $d(k)$ with zero mean and variance σ_d^2 as well as a causal FIR transmission system $h(k)$ of order q. Let $v(k)$ denote its output sequence. For notational clarity, let $\mathbf{R}_{vv,\eta}$ denote the $(\eta\times\eta)$ autocorrelation matrix with $\eta=\ell+1$. It is Hermitian Toeplitz and can be represented by

$$\mathbf{R}_{vv,\eta} = \sigma_d^2\,\mathbf{H}^*_\eta\mathbf{H}_\eta\ , \tag{D.1}$$

where $\mathbf{H}_\eta$ is the $(\eta+q\times\eta)$ filtering matrix $\mathbf{H}_{0,\eta+q\times\eta}$ (as defined in Appendix A) with full column rank.

First, we build a $(\eta\times\eta)$ upper triangular matrix $\mathbf{E}_\eta$ which orthonormalizes the columns of $\mathbf{H}_\eta$. Departing from a matrix with just one column, this can be done by means of the Gram-Schmidt

$$\lambda \widetilde{\mathbf{h}}_{\text{EVI}} = \gamma_4^d \, \mathbf{H}^*_{0,3q+1\times 2q+1} \, |w(k_m)|^2 \, \frac{\mathbf{i}}{\sigma_d^2} = \begin{bmatrix} h^*(0) & \cdots & h^*(q) & & 0 & \cdots & 0 \\ & \ddots & & \ddots & & \ddots & \vdots \\ 0 & & h^*(0) & \cdots & h^*(q) & & 0 \\ \vdots & \ddots & & \ddots & & \ddots & \\ 0 & \cdots & 0 & & h^*(0) & \cdots & h^*(q) \end{bmatrix} \begin{bmatrix} \vdots \\ 0 \\ |w(k_m)|^2 \frac{\gamma_4^d}{\sigma_d^2} \\ 0 \\ \vdots \end{bmatrix} \tag{B.3}$$

$$\mathbf{i} = \sigma_d^2 \mathbf{H}_{q,3q+1\times 4q+1} \mathbf{H}^*_{0,5q+1\times 4q+1} \mathbf{r}_{inv} = \begin{bmatrix} r_{vv}(q) & \cdots & r_{vv}(-q) & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & r_{vv}(q) & \cdots & r_{vv}(-q) \end{bmatrix} \begin{bmatrix} r_{inv}(\varepsilon - 3q) \\ \vdots \\ r_{inv}(\varepsilon + 2q) \end{bmatrix} \tag{C.1}$$

$$\widetilde{\mathbf{i}} = \begin{bmatrix} r_{vv}(q) & \cdots & r_{vv}(-q) & & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & & \ddots & \vdots \\ 0 & & r_{vv}(q) & \cdots & r_{vv}(-q) & & 0 \\ \vdots & \ddots & & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & & r_{vv}(q) & \cdots & r_{vv}(-q) \end{bmatrix} \begin{bmatrix} r_{inv}(-3q) \\ \vdots \\ r_{inv}(3q) \end{bmatrix} \tag{C.2}$$

orthogonalization process. $\mathbf{E}_\eta$ is determined recursively by

$$\mathbf{E}_\eta = \left[\begin{bmatrix} \mathbf{E}_{\eta-1} \\ 0 \ldots 0 \end{bmatrix} \;\; \mathbf{e}_\eta \right] \quad \text{with} \quad \mathbf{E}_0 = [\,] \;, \tag{D.2}$$

where

$$\mathbf{e}_\eta = \frac{\widetilde{\mathbf{E}}_\eta}{\sqrt{\widetilde{\mathbf{E}}^*_\eta \mathbf{H}^*_\eta \mathbf{H}_\eta \widetilde{\mathbf{E}}_\eta}} \, ; \quad \widetilde{\mathbf{E}}_\eta = (\mathbf{H}^*_\eta \mathbf{H}_\eta)^{-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} . \tag{D.3}$$

As $\mathbf{H}_\eta \mathbf{E}_\eta$ has orthonormal columns, we have with the $(\eta \times \eta)$ identity matrix $\mathbf{I}_\eta$ and equation (D.1)

$$\left(\mathbf{E}^*_\eta \mathbf{H}^*_\eta\right) \cdot \left(\mathbf{H}_\eta \mathbf{E}_\eta\right) = \mathbf{I}_\eta \tag{D.4}$$

$$\mathbf{E}^*_\eta \left(\mathbf{R}_{vv,\eta} / \sigma_d^2\right) \mathbf{E}_\eta = \mathbf{I}_\eta \tag{D.5}$$

$$\mathbf{E}^*_\eta \mathbf{R}_{vv,\eta} \left(\mathbf{E}_\eta \mathbf{E}^*_\eta\right) / \sigma_d^2 = \mathbf{E}^*_\eta \, . \tag{D.6}$$

Due to its triangular structure with non-vanishing elements on the main diagonal, $\mathbf{E}^*_\eta$ is always non-singular and has therefore an inverse. We obtain

$$\begin{aligned} \mathbf{R}_{vv,\eta} \left(\mathbf{E}_\eta \mathbf{E}^*_\eta\right) / \sigma_d^2 &= \mathbf{I}_\eta \\ \Rightarrow \quad \mathbf{R}^{-1}_{vv,\eta} &= \left(\mathbf{E}_\eta \mathbf{E}^*_\eta\right) / \sigma_d^2 \end{aligned} \tag{D.7}$$

and thus

$$\lim_{\eta \to \infty} \mathbf{R}^{-1}_{vv,\eta} = \frac{1}{\sigma_d^2} \lim_{\eta \to \infty} \left(\mathbf{E}_\eta \mathbf{E}^*_\eta\right) \, . \tag{D.8}$$

We realize that (D.3) represents the calculation of the coefficients of the maximum-phase prediction error filter. As η approaches infinity, $\mathbf{e}_\eta$ converges to the inverse maximum-phase transmission system so that $\mathbf{E}_\eta$ approaches a Toeplitz structure (just the top left corner of $\mathbf{E}_\eta$ does not have this structure). For the inverse of the autocorrelation matrix (eq. (D.7)), the elements of any given diagonal are equal to the scalar product of two vectors, which – for $\eta \to \infty$ – differ just by a uniform shift. However, this is true in the center of this matrix, only. Thus, in its central part, all elements on the same diagonal are identical, and the inverse of the infinite dimensions autocorrelation matrix approximates a Toeplitz structure. This completes the proof. ■

E Influence of additive white Gaussian noise

We will justify the conclusion drawn from Fig. 7b,e in section 4.1 that the channel zeros "estimated" by EVI using true statistics do not change significantly under AWGN influence [19].

In Appendices B and C, we have shown that in the noisefree case

$$\mathbf{W}\mathbf{H}_{q,3q+1\times 4q+1}\widetilde{\mathbf{R}}_{inv}\widetilde{\mathbf{h}}_{\mathrm{EVI}} = |w(k_m)|^2\,\frac{\mathbf{i}}{\sigma_d^2} \qquad \text{(E.1)}$$

selects exactly one column from $\mathbf{H}^*_{0,3q+1\times 2q+1}$. However, rewriting

$$\mathbf{W}\cdot\overbrace{\mathbf{H}_{q,3q+1\times 4q+1}\widetilde{\mathbf{R}}_{inv}\widetilde{\mathbf{h}}_{\mathrm{EVI}}}^{=:\mathbf{w}_2} = \begin{bmatrix} |w(0)|^2\cdot w_2(0) \\ \vdots \\ |w(3q+1)|^2\cdot w_2(3q+1) \end{bmatrix} \overset{!}{\approx} \frac{|w(k_m)|^2}{\sigma_d^2}\cdot\begin{bmatrix} \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}, \qquad \text{(E.2)}$$

we realize that both the coefficients $w(k)$ of the combined channel/reference system and the coefficients $w_2(k)$ of a weighting vector $\mathbf{w}_2$ contribute to this selective property.

Figure 11 displays the normalized magnitudes of $w(k)$ and $w_2(k)$ as well as the product $|w(k)|^2\cdot|w_2(k)|$ for channel C1 from Figure 7b and noise-to-signal ratios of $N/S=0$ and 0.5. In the noiseless case (Fig. 11a to c), just *one* element of $\mathbf{w}_2$ does not vanish. Thus, the same property holds for the product (see Fig. 11c). Although in the noisy case quite a few coefficients $w_2(k)$ reveal non-negligible magnitudes (see Fig. e), the multiplication with $|w(k)|^2$ still ensures a distinctive peak in $|w(k)|^2\cdot|w_2(k)|$ and thus guarantees a good approximation of the selection of a single column. So, the noise robustness of EVI is due to the fact that the product $|w(k)|^2\cdot|w_2(k)|$ decays even more quickly than $|w(k)|$ and $|w_2(k)|$. As the interpretation of $\mathbf{H}^*_{q,3q+1\times 2q+1}$ as a filtering matrix (see in eq. (A.2)) nearly holds, we can write

$$h_{EVI} \propto h^*(k) * \left(|w(k)|^2\cdot w_2(k)\right), \qquad \text{(E.3)}$$

and in frequency domain

$$H_{EVI}(z) \propto H^*(z^*)\cdot G(z) \qquad \text{(E.4)}$$

with

$$G(z) \triangleq \mathcal{Z}\left\{|w(k)|^2\cdot w_2(k)\right\}. \qquad \text{(E.5)}$$

From this relation, we realize that the true channel zeros are present in EVI's estimates while excessive zeros are introduced by $G(z)$.

References

[1] A. Alshebeili, A. N. Venetsanopoulos, and A. E. Çetin. Cumulant Based Identification Approaches for Nonminimum FIR Systems. *IEEE Trans. on Signal Processing*, SP-41(4):1576–1588, April 1993.

[2] B. Boashash, E.J. Powers, and A.M. Zoubir, editors. *Higher-Order Statistical Signal Processing.* Addison Wesley Longman, NSW 2066, Australia, 1997.

[3] D. Boss, M. Boe, and K.D. Kammeyer. Exploiting Second Order Cyclostationarity or Higher Order Statistics for the Blind Identification of Mixed Phase FIR Systems? In *Proc. 15th GRETSI symposium*, volume 1, pages 57–60, Juan-les-Pins, France, September 1995.

[4] D. Boss and K.D. Kammeyer. Blind Estimation of ARMA-Systems. In *Proc. EUSIPCO-94*, volume III, pages 1105–1108, Edinburgh, Scotland, September 1994.

[5] D. Boss and K.D. Kammeyer. Blind GSM Channel Estimation. In *Proc. VTC-97*, volume 2, pages 1044–1048, Phoenix, Arizona, May 1997.

[6] D. Boss and K.D. Kammeyer. Blind GSM Channel Estimation based on Higher Order Statistics. In *Proc. ICC-97*, volume 1, pages 46–50, Montréal, Canada, June 1997.

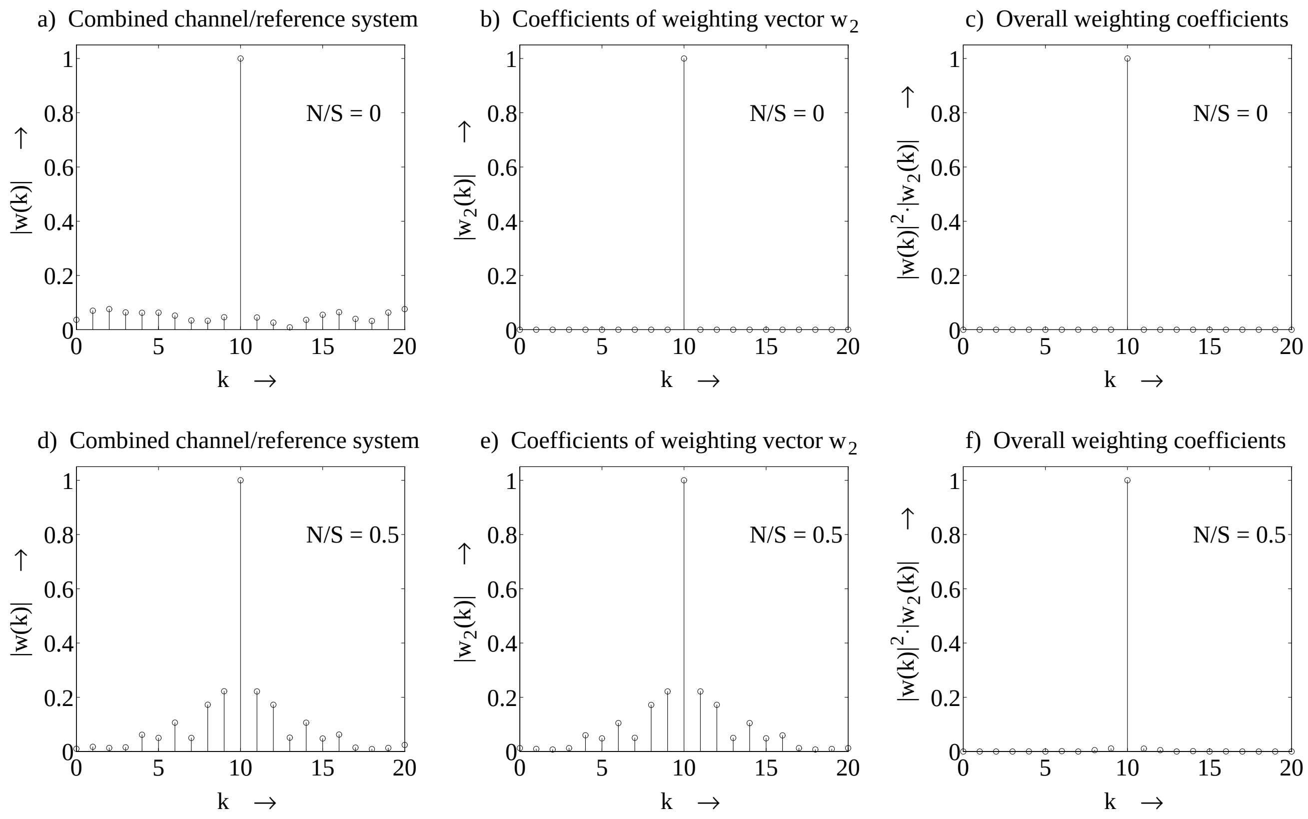


Fig. 11: Justification of EVI's robustness with respect to AWGN (Channel C1 (see Fig. 3a))
a-c) $S/N = \infty$ d-f) $S/N = 3\,\mathrm{dB}$

[7] D. Boss and K.D. Kammeyer. Blind Identification of Mixed-Phase FIR Systems with Application to Mobile Communication Channels. In *Proc. ICASSP-97*, volume 5, pages 3589–3592, Munich, Germany, April 1997.

[8] D. Boss, K.D. Kammeyer, and T. Petermann. Is Blind Channel Estimation feasible in Mobile Communication Systems? A Study based on GSM. *IEEE Journal on Selected Areas in Communications*, Q4 1998. Special issue on Signal Processing for Wireless Communications. To appear.

[9] D. Boss, T. Petermann, and K.D. Kammeyer. Impact of Blind versus Non-blind Channel Estimation on the BER Performance of GSM Receivers. In *Proc. IEEE Signal Proc. Workshop on Higher-Order Statistics*, pages 62–66, Banff, Canada, July 1997.

[10] H. H. Chiang and C. L. Nikias. Adaptive Deconvolution and Identification of Nonminimum Phase FIR Systems Based on Cumulants. *IEEE Trans. on Automatic Control*, AC-35(1):36–47, January 1990.

[11] Z. Ding. Characteristics of Band-Limited Channels Unidentifiable from Second-Order Cyclostationary Statistics. *IEEE Signal Processing Letters*, SPL-3(5):150–152, May 1996.

[12] J. A. R. Fonollosa and J. Vidal. System Identification using a Linear Combination of Cumulant Slices. *IEEE Trans. on Signal Proc.*, SP-41(7):2405–2412, July 1993.

[13] B. Friedlander and B. Porat. Asymptotically Optimal Estimation of MA and ARMA Parameters of Non-Gaussian Processes from High-Order Moments. *IEEE Trans. on Automatic Control*, AC-35(1):27–35, January 1990.

[14] G. B. Giannakis and J. M. Mendel. Identification of Nonminimum Phase Systems using Higher Order Statistics. *IEEE Trans. on Acoustics, Speech and Signal Processing*, ASSP-37(3):360–377, March 1989.

[15] G. B. Giannakis and J. M. Mendel. Cumulant-Based Order Determination of Non-Gaussian ARMA Models. *IEEE Trans. on Acoustics, Speech and Signal Processing*, ASSP-38(8):1411–1422, August 1990.

[16] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The John Hopkins University Press, London and Baltimore, MD, second edition, 1989.

[17] S. Haykin. *Adaptive Filter Theory.* Prentice-Hall, Upper Saddle River, New Jersey 07458, third edition, 1996.

[18] P. Hoeher. A Statistical Discrete-Time Model for the WSSUS Multipath Channel. *IEEE Trans. on Vehicular Technology*, VT-41(4):461–468, April 1992.

[19] B. Jelonnek. *Referenzdatenfreie Entzerrung und Kanalschätzung auf der Basis von Statistik höherer Ordnung.* PhD thesis, Dept. of Telecommunications, Hamburg University of Technology, Hamburg, Germany, March 1995.

[20] B. Jelonnek, D. Boss, and K.D. Kammeyer. Generalized Eigenvector Algorithm for Blind Equalization. *Elsevier Signal Processing*, 61(3):237–264, September 1997.

[21] B. Jelonnek and K.D. Kammeyer. Improved Methods for the Blind System Identification using Higher Order Statistics. *IEEE Trans. on Signal Processing*, SP-40(12):2947–2960, December 1992.

[22] B. Jelonnek and K.D. Kammeyer. Eigenvector Algorithm for Blind Equalization. In *Proc. IEEE Signal Proc. Workshop on Higher-Order Statistics*, pages 19–23, South Lake Tahoe, California, June 1993.

[23] B. Jelonnek and K.D. Kammeyer. A Closed-Form Solution to Blind Equalization. *Elsevier Signal Processing*, 36(3):251–259, April 1994. Special Issue on Higher Order Statistics.

[24] K.D. Kammeyer and B. Jelonnek. A Cumulant Zero-Matching Method for the Blind System Identification. In *Proc. International Signal Proc. Workshop on Higher Order Statistics*, pages 103–106, Chamrousse, France, July 1991.

[25] K.D. Kammeyer and B. Jelonnek. A New Fast Algorithm for Blind MA-System Identification based on Higher Order Cumulants. In *Proc. SPIE Advanced Signal Proc.: Algorithms, Architectures and Implementations V*, volume 2296, pages 162–173, San Diego, California, July 1994.

[26] G. Liang, M. Wilkes, and J. A. Cadzow. ARMA Model Order Estimation Based on the Eigenvalues of the Covariance Matrix. *IEEE Trans. on Signal Processing*, SP-41(10):3003–3009, October 1993.

[27] J. M. Mendel. Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications. *Proceedings of the IEEE*, 79(3):278–305, March 1991.

[28] M. Mouly and M.-B. Pautet. *The GSM System for Mobile Communications.* Published by the authors, Palaiseau, France, 1992.

[29] C. L. Nikias and J. M. Mendel. Signal Processing with Higher-Order Spectra. *IEEE Signal Processing Magazine*, 10:10–37, July 1993.

[30] C. L. Nikias and M. R. Raghuveer. Bispectrum Estimation: A Digital Signal Processing Framework. *Proceedings of the IEEE*, 75(7):869–891, July 1987.

[31] J. G. Proakis. *Digital Communications.* McGraw-Hill, New York, third edition, 1995.

[32] S. V. Schell, D. L. Smith, and W. A. Gardner. Blind Channel Identification Using 2nd-Order Cyclostationary Statistics. In *Proc. EUSIPCO-94*, volume II, pages 716–719, Edinburgh, Scotland, September 1994.

[33] O. Shalvi and E. Weinstein. New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels). *IEEE Trans. on Information Theory*, IT-36(2):312–321, March 1990.

[34] O. Shalvi and E. Weinstein. Super-Exponential Methods for Blind Deconvolution. *IEEE Trans. on Information Theory*, 39(2):504–519, March 1993.

[35] A. Swami and J. M. Mendel. Closed-Form Recursive Estimation of MA Coefficients using Autocorrelations and Third Order Cumulants. *IEEE Trans. on Acoustics, Speech and Signal Processing*, ASSP-37(11):1794–1795, November 1989.

[36] L. Tong, G. Xu, and T. Kailath. Necessary and Sufficient Conditions of Channel Identifiability Based on Second-Order Cyclostationary Statistics. In *Proc. IEEE Int. Symp. on Info. Theory*, page 188, San Antonio, Texas, January 1993.

[37] L. Tong, G. Xu, and T. Kailath. Blind Identification and Equalization Based on Second-Order Statistics: A Time Domain Approach. *IEEE Trans. on Information Theory*, IT-40(2):340–349, March 1994.

[38] J. K. Tugnait. Approaches to FIR System Identification with Noisy Data using Higher Order Statistics. *IEEE Trans. on Acoustics, Speech and Signal Processing*, ASSP-38(7):1307–1317, July 1990.

[39] J. K. Tugnait. New Results on FIR System Identification using Higher Order Statistics. *IEEE Trans. on Signal Processing*, SP-39(10):2216–2221, October 1991.

[40] J. K. Tugnait. Blind Estimation of digital Communication Channel Impulse Response. *IEEE Trans. on Communications*, COM-42(Part III):1606–1616, February-April 1994.

[41] J. K. Tugnait. Blind Estimation and Equalization of Digital Communication FIR Channels using Cumulant Matching. *IEEE Trans. on Communications*, COM-43(Part III):1240–1245, February-April 1995.

[42] J. K. Tugnait. On Blind Identifiability of Multipath Channels Using Fractional Sampling and Second-Order Cyclostationary Statistics. *IEEE Trans. on Information Theory*, IT-41(1):308–311, January 1995.

[43] J. Vidal and J. A. R. Fonollosa. Adaptive Blind Equalization using Weighted Cumulant Slices. In *Proc. ICASSP-94*, volume IV, pages 593–596, Adelaide, Australia, April 1994.

[44] X.-D. Zhang and Y.-S. Zhang. Singular Value Decomposition-Based MA Order Determination of Non-Gaussian ARMA Models. *IEEE Trans. on Signal Processing*, SP-41(8):2657–2664, August 1993.