Toni Huovinen

Independent Component Analysis in DS-CDMA Multiuser Detection and Interference Cancellation

Preliminary Examination Edition

Thesis for the degree of Doctor of Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TBXXX, at Tampere University of Technology, on the XX-th of xxxxxx 2008, at 12 noon.
Abstract

Multiuser detection (MUD) and interference cancellation (IC) have emerged as important theoretical and practical problems among spread spectrum multiple access researchers during the last decades. Conventionally, MUD and IC methods are based on second order statistics. They assume low statistical correlation between desired signal components and interfering ones. One relatively new idea is to use also higher order statistics (HOS) to improve the system performance. Especially, HOS based blind source separation (BSS) techniques are attractive, since they are able to separate signals from a mixture of original source signals in a completely blind manner, i.e., without explicit knowledge of signals waveforms. Consequently, many types of interference sources, for instance, internal interferences due to multiple access and non-Gaussian external interference signals, can be suppressed out. However, possible performance gains due to BSS or HOS processing, in general, compared to conventional second order methods are not studied extensively in the literature – until this PhD dissertation.

In this dissertation, the main emphasis is on BSS assisted MUD and IC methods and, especially, on those which employ an independent component analysis (ICA). Advanced MUD/IC strategies are considered and applied in direct sequence code division multiple access (DS-CDMA) uplink receivers. In particular, a new family of BSS/ICA assisted successive interference cancellation (SIC) schemes are introduced and performance of such a receivers are set against to conventional methods. The main emphasis is on the overcomplete data which situation arises in challenging highly loaded systems. The new receiver structures combine the main benefits of HOS signal processing and conventional non-linear MUD methods, that is, (i) inherent mitigation of various types of interference sources by BSS/ICA, (ii) robustness against parameter estimation errors due to BSS/ICA, (iii) greatly improved interference suppression capability due to novel combination of SIC ideology and BSS/ICA.
# Contents

List of Publications vii
List of Figures xi
List of Symbols and Abbreviations xiii

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Symbols</td>
<td>xiii</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xv</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background and Motivation</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Scope of the Dissertation</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Outline and Organization</td>
<td>4</td>
</tr>
<tr>
<td>2 Direct-Sequence CDMA</td>
<td>5</td>
</tr>
<tr>
<td>2.1 Signal Model</td>
<td>7</td>
</tr>
<tr>
<td>2.2 Conventional Single-User Detection</td>
<td>9</td>
</tr>
<tr>
<td>2.3 On Code Synchronization</td>
<td>12</td>
</tr>
<tr>
<td>3 Independent Component Analysis</td>
<td>15</td>
</tr>
<tr>
<td>3.1 Definition of ICA</td>
<td>16</td>
</tr>
<tr>
<td>3.1.1 Basic Noise-Free ICA</td>
<td>16</td>
</tr>
<tr>
<td>3.1.2 Noisy ICA</td>
<td>18</td>
</tr>
<tr>
<td>3.1.3 ICA with Overcomplete Basis</td>
<td>18</td>
</tr>
<tr>
<td>3.2 ICA Model in DS-CDMA Reception</td>
<td>19</td>
</tr>
</tbody>
</table>
### CONTENTS

3.3 ICA Algorithms .................................................. 21  
3.3.1 Pre-whitening the data ......................................... 21  
3.3.2 Basic ICA algorithms ........................................... 22  
3.3.3 Noisy ICA algorithms ........................................... 25  
3.3.4 Algorithms for ICA with overcomplete basis ................. 26  
3.4 On Performance of ICA Algorithms ............................. 26  

4 Multiuser Detection .................................................. 31  
4.1 Optimum MUD ....................................................... 33  
4.2 Linear suboptimum MUD ............................................ 34  
4.2.1 Decorrelating detection ......................................... 35  
4.2.2 Linear MMSE detection ......................................... 37  
4.3 Nonlinear suboptimum MUD ......................................... 38  
4.3.1 Successive interference cancellation (SIC) ................... 38  
4.3.2 Parallel interference cancellation (PIC) ..................... 39  
4.3.3 Comparison of SIC and PIC ..................................... 43  
4.4 Blind MUD ........................................................... 44  
4.4.1 Subspace decorrelating and LMMSE detection ............... 45  
4.4.2 ICA assisted MUD ............................................... 46  
4.4.3 Discussion on complexity of ICA assisted receivers ........ 51  

5 Rejection of External Interference .................................. 53  
5.1 External Interference in Signal Model ........................... 53  
5.2 Conventional Narrowband Interference Rejection ............... 55  
5.3 External Interference Sources and ICA .......................... 58  

6 Summary of Publications .............................................. 61  
6.1 Overview of Contents .............................................. 61  
6.2 Author’s Contributions ............................................. 65  

7 Conclusions ............................................................ 67  

A Wise FastICA Initialization ......................................... 71  

B On Complexity .......................................................... 75  

Bibliography .............................................................. 77  

II Original Publications ................................................ 89
List of Publications

This dissertation consists of one journal article, nine conference articles and one research report listed in the following:


LIST OF PUBLICATIONS


Other Publications of the Author

The author of this thesis has also published (or submitted) the following three articles which are not included in this dissertation:


LIST OF PUBLICATIONS
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Time–frequency occupancy of the multiple access schemes</td>
<td>5</td>
</tr>
<tr>
<td>2.2</td>
<td>DS-modulation of BPSK symbols</td>
<td>6</td>
</tr>
<tr>
<td>4.1</td>
<td>Conventional multiuser detector</td>
<td>32</td>
</tr>
<tr>
<td>4.2</td>
<td>Decorrelating detector</td>
<td>36</td>
</tr>
<tr>
<td>4.3</td>
<td>Successive interference cancellation, SIC</td>
<td>39</td>
</tr>
<tr>
<td>4.4</td>
<td>Multistage PIC detector</td>
<td>41</td>
</tr>
<tr>
<td>4.5</td>
<td>Needed symbol estimates in the 2-stage PIC</td>
<td>42</td>
</tr>
<tr>
<td>4.6</td>
<td>Successive ICA assisted receiver</td>
<td>50</td>
</tr>
<tr>
<td>5.1</td>
<td>NBI cancellation: estimator–subtractor</td>
<td>56</td>
</tr>
<tr>
<td>5.2</td>
<td>Frequency domain NBI cancellation</td>
<td>57</td>
</tr>
<tr>
<td>5.3</td>
<td>Cutting off spectrum peaks</td>
<td>57</td>
</tr>
<tr>
<td>5.4</td>
<td>ICA based external interference cancellation</td>
<td>59</td>
</tr>
<tr>
<td>A.1</td>
<td>Effect of Wise Initialization on num. of ICA iterations: SNR=10 dB</td>
<td>73</td>
</tr>
<tr>
<td>A.2</td>
<td>Effect of Wise Initialization on BER: SNR=10 dB</td>
<td>74</td>
</tr>
<tr>
<td>A.3</td>
<td>Effect of Wise Initialization on BER: Variable SNR</td>
<td>74</td>
</tr>
</tbody>
</table>
List of Symbols and Abbreviations

Symbols

:= Definition (left hand side is defined by right one)
(·)† Pseudo inverse of a matrix
(·)* Complex conjugate
(·)T, (·)H Matrix transpose and Hermitian transpose
| · | Absolute value
∥ · ∥ Euclidean norm
0, 0n Origin (n-dimensional vector of zeros)
β = [b₁, ..., bₖ]T Vector of user symbols
δk, δkl Delay Estimation Error
Δ = diag(a₁, ..., aₖ) Diagonal matrix of channel coefficients
Γ Code cross-correlation matrix
η(t) Additive noise signal
η Additive noise vector
η̃, η̃ Noise in MF output(s)
θ(k, k′) Correlation for user identification
Λ Eigenvalue matrix of ICA observations or received data
ξₖ(·) Spreading code sequence of the k-th user
ρₖ₁ Cross-correlation between k-th and l-th users
\overline{q}_{k,l}, \rho_{k,l}, \overline{p}_{k,l} Correlations between asynchronous users’ spreading codes
σ² Noise variance
Σ Noise covariance matrix
τk, τkl, \tilde{τ}_{kl} Delay of the k-th user’s l-th path and its estimate
Υk, \tilde{Υ}_k Multiaccess interference component and its estimate
\begin{itemize}
\item \( \Upsilon_{\text{ext}}(t), \ \Upsilon_{\text{ext}}[m] \) \quad External interference signal
\item \( \Psi(v) \) \quad Update term of EASI algorithm
\item \( \Omega_K(\cdot) \) \quad Contrast function of jointly optimum MUD
\item \( \imath \) \quad Imaginary unit
\item \( a_k, a_{kl} \) \quad Path coefficient of \( k \)-th user’s \( l \)-th path
\item \( A \) \quad ICA mixing matrix
\item \( b_k, b_k[m] \) \quad \((m\text{-th}) \) symbol of \( k \)-th user
\item \( b \) \quad Vector containing “early”, “middle” and “late” symbols of all users
\item \( B, \breve{B} \) \quad De-mixing matrix and its estimate
\item \( (e_i)_{i=1}^\infty \) \quad Binary chip sequence
\item \( C \) \quad Number of chips per symbol
\item \( C_x \) \quad Covariance matrix of ICA observations
\item \( C_r \) \quad Covariance matrix of received data vector
\item \( \mathbb{C} \) \quad Set of complex numbers
\item \( e_n \) \quad M-GEF vector
\item \( \mathbf{E}s \) \quad Principal eigenvector matrix of \( C \)
\item \( \mathbb{E}(\cdot) \) \quad Expected value
\item \( f : \mathbb{R} \to \mathbb{R} \) \quad ICA nonlinearity (derivative of \( F \))
\item \( F : \mathbb{R} \to \mathbb{R} \) \quad ICA contrast function
\item \( g_k, \breve{g}_k \) \quad One column of matrix \( \mathbf{G} \) and its estimate
\item \( \mathbf{G}, \breve{\mathbf{G}} \) \quad Code matrix after propagation channel and its estimate
\item \( h_n \) \quad Column of a mixing matrix
\item \( I \) \quad Identity matrix
\item \( \Im\{\cdot\} \) \quad Imaginary part of complex number
\item \( J \) \quad Number of ICA observations
\item \( K \) \quad Number of users
\item \( L \) \quad Number of paths
\item \( M \) \quad Number of symbol in one frame or block
\item \( N \) \quad Number of ICA sources
\item \( m_k, \breve{m}_k \) \quad LMMSE transformation of \( k \)-th (user) component and its estimate
\item \( \mathbf{M}_0 \) \quad LMMSE matrix
\item \( O(\cdot) \) \quad Order of
\item \( p_T(t) \) \quad Chip wave form
\item \( P \) \quad Number of PIC stages
\item \( \mathbb{P}(\cdot) \) \quad Probability measure
\end{itemize}
$Q(\cdot)$: $Q$-function, i.e., a complementary cumulative distribution function (CDF) of the unit variance normal random variable

$r(t)$: Received signal
$r$: Received (stochastic) data vector of length $2C$
$r[m]$: $m$-th sample $r$ (contains contributions $m$-th symbols)
$R_n$: Observation covariance of desired component
$R'_n$: Observation covariance of interference and noise
$\mathbb{R}$: Set of real numbers
$\mathbb{R}\{\cdot\}$: Real part of complex number
$s = [s_1, s_2, \ldots, s_N]^T$: ICA source component vector
$s_n$: ICA source component
$\text{sign}(\cdot)$: Sign-function
$T$: Symbol duration
$T_c = T/C$: Chip duration
$U$: Arbitrary unitary matrix
$V$: Whitening transformation
$z$: Whitened ICA observation vector
$w_i$: Whitened ICA basis vector
$W$: Whitened ICA mixing matrix
$x_i$: ICA observation
$x = [x_1, x_2, \ldots, x_M]^T$: ICA observation vector
$y_k, y_k[m]$: Matched filter output
$y = [y_1, y_2, \ldots, y_K]$: Vector of matched filter outputs
$y_{\text{DD}}$: Soft output of decorrelating detector
$y_{\text{LMMSE}}$: Soft output of LMMSE-detector

**Abbreviations**

- **AWGN**: Additive White Gaussian Noise
- **BEP**: Bit Error Probability
- **BER**: Bit Error Rate
- **BPSK**: Binary Phase Shift Keying
- **BSS**: Blind Source Separation
- **CDF**: Cumulative Distribution Function
- **CDMA**: Code Division Multiple Access
- **DD**: Decorrelating Detector
- **DS**: Direct Sequence
<table>
<thead>
<tr>
<th>Symbol/Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS-CDMA</td>
<td>Direct Sequence Code Division Multiple Access</td>
</tr>
<tr>
<td>DSP</td>
<td>Digital Signal Processing</td>
</tr>
<tr>
<td>EASI</td>
<td>Equivariant Adaptive Source Identification</td>
</tr>
<tr>
<td>FH</td>
<td>Frequency Hopping</td>
</tr>
<tr>
<td>FT</td>
<td>Fourier-Transformation</td>
</tr>
<tr>
<td>M-GEF</td>
<td>(SINR) Maximizing Generalized EigenFilter</td>
</tr>
<tr>
<td>GS</td>
<td>Gram–Schmidt orthogonalization algorithm</td>
</tr>
<tr>
<td>HOS</td>
<td>Higher Order Statistics</td>
</tr>
<tr>
<td>I/Q</td>
<td>Inphase and Quadrature (parts of signal)</td>
</tr>
<tr>
<td>IC</td>
<td>Interference Cancellation</td>
</tr>
<tr>
<td>ICA</td>
<td>Independent Component Analysis</td>
</tr>
<tr>
<td>IFT</td>
<td>Inverse Fourier-Transformation</td>
</tr>
<tr>
<td>ISR</td>
<td>Interference-to-Signal-Ratio</td>
</tr>
<tr>
<td>JADE</td>
<td>Joint Approximation Diagonalization Estimation</td>
</tr>
<tr>
<td>LMMSE</td>
<td>Linear Minimum Mean Square Error</td>
</tr>
<tr>
<td>MAI</td>
<td>MultiAccess Interference</td>
</tr>
<tr>
<td>MF</td>
<td>Matched Filter</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple Input Multiple Output</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum Likelihood</td>
</tr>
<tr>
<td>MMSE</td>
<td>Minimum Mean Square Error</td>
</tr>
<tr>
<td>MOE</td>
<td>Minimum Output Energy</td>
</tr>
<tr>
<td>MRC</td>
<td>Maximal Ratio Combining</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Square Error</td>
</tr>
<tr>
<td>MUD</td>
<td>MultiUser Detection</td>
</tr>
<tr>
<td>NBI</td>
<td>NarrowBand Interference</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PIC</td>
<td>Parallel Interference Cancellation</td>
</tr>
<tr>
<td>QPSK</td>
<td>Quadrature Phase Shift Keying</td>
</tr>
<tr>
<td>RAKE</td>
<td>Conventional multipath receiver</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>SIC</td>
<td>Successive Interference Cancellation</td>
</tr>
<tr>
<td>SINR</td>
<td>Signal-to-Interference-and-Noise-Ratio</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise-Ratio</td>
</tr>
<tr>
<td>SO</td>
<td>Second order (e.g., statistics)</td>
</tr>
<tr>
<td>SS</td>
<td>Spread Spectrum</td>
</tr>
<tr>
<td>SUD</td>
<td>Single-User Detection</td>
</tr>
<tr>
<td>WCDMA</td>
<td>Wideband Code Division Multiple Access</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Background and Motivation

Code division multiple access (CDMA) is a multiplexing technique, which enables multiple users to access simultaneously and asynchronously to a common radio frequency (RF) channel by modulation and spreading of the information-bearing signals with a user-specific spreading sequences or spreading codes. For example, third generation’s mobile communication systems as well as many other latest radio systems are based on the CDMA technology. In addition, CDMA or some of its variants are good candidates for multiplexing technique also in future commercial systems, since they fit well to increasing demand for higher bit rates. Different ways to spread the signal over a wide and common frequency band results in different multiple access schemes. Direct sequence (DS) and frequency hopping (FH) (and hybrids of them) are examples of single-carrier modulation schemes, and have been under extensive research [89, 113]. The WCDMA standard, for instance, is based on DS-modulation [42]. Also in this dissertation, we study CDMA systems where spreading is carried out by DS-modulation.

An inherent capability to resist interference is the most important feature of spread spectrum communication – and, thus, also of CDMA. As a matter of fact, the concept of spread spectrum communication was originally introduced particularly for military communication and especially for antijamming purposes [93]. Later, the spread spectrum techniques were harnessed to the
civilian use. For example, a possibility to use the spread spectrum system on top of traditional narrowband systems without causing insuperable interference increment is very fascinating in the commercial communication where RF band has become scarce resource [92]. Nevertheless, also CDMA systems – like any radio communication systems – are interference limited. Furthermore, in practice, also an internal multiple access interference (MAI) due to the non-ideal cross-correlations between users’ spreading sequences limits the performance of a CDMA system. Conventionally, this interference is considered as an additional background noise which increases a total noise floor and, eventually, makes conventional detectors inadequate in highly loaded systems. Conventional detection is also quite sensitive to external interference sources like adjacent channel interference or jamming which has lead to development of numerous interference rejection techniques. An effective interference suppression or cancellation (IC) is a natural way to enhance system performance.

In CDMA literature (see, e.g., [89, 112, 113]), MAI reduction is often called as multiuser detection (MUD). In it, the existence of the interfering users is taken into account when detecting the desired user(s), unlike in conventional single-user detection. Conventional MUD/IC methods are typically based on second order statistical properties of data, hence, being founded on low statistical correlation assumption between desired signal components and interfering ones. One relative new idea is to use higher order statistics (HOS), which makes the methods very robust against problems related to incomplete cross-correlation properties and also against a near-far problem (i.e., differences in received power) which is an another limiting factor for conventional detection and IC. Especially, HOS based blind source separation (BSS) techniques [29, 53] are attractive, since they are able to separate signals from a mixture of original source signals in a completely blind manner, i.e., without explicit knowledge of waveform structure (modulation) or mixing coefficients. Consequently, many types of interference sources, for instance, internal interferences due to multiple access and out-of-cell interferences (intentional and unintentional) due to cellular network, can be mitigated in DS-CDMA systems. Here, our main emphasis is particularly in BSS assisted MUD. In addition, we also touch the external interference suppression briefly.

1.2 Scope of the Dissertation

In this dissertation work, the goal is to study possible benefits of using higher order statistics in DS-CDMA multiuser detection and interference cancellation.
This is carried out study by developing new non-linear MUD/IC method which employ HOS based BSS and comparing their performance to existing MUD methods. The main emphasis is on the asynchronous uplink (mobile-to-base) reception and, especially, on highly loaded systems which means in DS-CDMA that the number of active connections or users is high compared to number of chips per bit being used, that is, to a processing gain of the system. From BSS’s point of view, the highly loaded system model appears to correspond to so called “more sources than sensors” -problem [53], which is considered a challenging problem in literature.

Novel successive interference cancellation (SIC) schemes obeying the fundamental procedures of conventional SIC [43, 112], are proposed. That is to say, (i) the estimation of interference, (ii) re-generation of interference and finally (iii) subtraction of the re-generated interference from the original data. What makes the schemes advanced compared to conventional SIC is the way of estimating the interference using HOS based BSS by means of independent component analysis (ICA) [22, 53] and a sensitive selection of the interference sources to be subtracted. The proposed receiver structures combine the main benefits of BSS/ICA and conventional SIC methods:

- inherent mitigation of various types of interference sources by BSS/ICA,
- robustness against parameter estimation errors due to BSS/ICA,
- greatly improved interference suppression capability in highly loaded systems due to novel combination of SIC ideology and HOS signal processing of BSS/ICA.

Best of all, all the benefits are achieved using single receiver antenna only. Notice also that the last item in the list above can also be seen as a way to somewhat circumvent the “more sources than sensors” -problem of BSS/ICA.

Noisy models have been another challenge in BSS/ICA area [53]. In most of practical applications, nevertheless, the presence of an additive background noise can not be neglected. Especially in telecommunications, some level of Gaussian noise is always interfering the reception. For this reason, this dissertation study considers also the noisy ICA problem. It is shown that conceptually many basic ICA algorithms designed for noise-free ICA models, can actually provide the best linear source separation solution possible in terms of input-output signal-to-interference-and-noise (SINR) gain. This seems not to be well-understood in the literature earlier. From CDMA reception perspective, this founding means that the performance of BSS/ICA assisted reception is indeed exceedingly competitive against any suboptimum MUD detector.
One should also recognize the fact that in interference subtractive receivers (in SIC, for example), a very precise channel state information is needed in the interference subtraction to prevent interference enhancement. This is why accurate code acquisition and tracking is of great importance. However, an important property of BSS/ICA is its inherent capability to cope with erroneous timing estimates. This feature of BSS/ICA is exploited to refine tentative timing estimates after the ICA processing, but before interference re-generation and substraction, such that also the timing refinement procedure takes advantage of interference suppression due to blind separation. Consequently, the requirement of extremely precise code tracking can be relaxed in the proposed ICA-SIC schemes. This kind of relaxation is not inherently possible in conventional interference subtractive receivers.

1.3 Outline and Organization

This PhD dissertation consists of two parts. The first part gives an introduction and background knowledge of the research area to which the topic of the dissertation work belongs. Also the earlier literature on the area is reviewed. The second part is a compilation of eleven original research publications reporting the novel scientific contributions of this dissertation work in detailed manner. In the first part, these publications are referred as [P1–P11].

Rest of the first part is organized as follows. In Chapter 2 we describe basics of DS-CDMA systems starting with definitions of different system models and continuing with description of conventional single-user detection. In Chapter 3, we discuss the central signal processing method of this dissertation, independent component analysis (ICA). We give definition of ICA, reveal the connection between ICA and DS-CDMA models and review the most well known ICA algorithms. In the end of the chapter, we discuss on performance of the ICA algorithms. In Chapter 4, we deal with multiuser detection (MUD). We summarize the basics of optimum and suboptimum MUD as well as blind MUD. The former includes also the ICA assisted MUD. In Chapter 5, we discuss the rejection of external interference. Finally, in Chapters 6 and 7, we give summary of the publications and draw a conclusions of the work, respectively.
Chapter 2

Direct-Sequence CDMA

Frequency division multiple access (FDMA) and time division multiple access (TDMA) enable multiple users to share a common RF medium via disjoint use of frequency band and time, respectively. In a signal space language, this means that users are mutually orthogonal. Orthogonal signaling is possible also with signals overlapping both in time and frequency by spreading users’ narrowband data streams to considerably wider frequency band using mutually orthogonal code sequences. This type of multiplexing of users data carrying signals is called as code division multiple access (CDMA). Fig. 2.1 illustrates the time–frequency occupancy of the different multiple access techniques.

![Figure 2.1: Time–frequency occupancy of the FDMA, TDMA and CDMA schemes. The figure depicts three users marked with different filling patterns.](image-url)
In Direct sequence (DS-) CDMA the spreading is carried out by direct modulation of the information bearing signals with a user-specific wideband signature sequences or codes. The principle is shown for BPSK symbols (that is, symbols having values in \{-1, 1\}) in Fig. 2.2. The main benefit of the DS-CDMA is that the strict orthogonality between the signature sequences, and hence, also between users’ data bearing signals, can be relaxed. E.g., the performance of the simple de-modulator is degraded in the presence of non-orthogonal interfering users, but the degradation can be kept in tolerable levels provided that cross-correlations between users are moderate. Relaxing the orthogonality requirement makes DS-CDMA even more attractive as a multiple access technique. For instance, users can access the system asynchronously and unlike in strictly orthogonal multiple access schemes (such as TDMA and FDMA), it is possible to trade off reception quality for increased capacity. However, due to multiaccess interference, additional interference mitigation or multiuser detection (MUD) techniques are usually needed to improve system performance. These techniques are the main theme of this dissertation.

Rest of this chapter introduces the principle of DS-CDMA more in detail. First, in Section 2.1, we present the short-code DS-CDMA system model which
we then study in the rest of the dissertation. In Section 2.2, we introduce a
principle of conventional single-user detection (SUD) which is based on spread-
ing code matched filtering. In addition, we define a multiple access interference
experienced by conventional detector. Finally, in Section 2.3, we discuss shortly
on code synchronization.

2.1 Signal Model

Next, we present the complex-valued baseband model of the carrier modulated
short-code DS-CDMA signal. This model is commonly used in literature and
also assumed in [P1–P8]. Let us assume that system has $K$ active users all
sending a symbol, $b_k$, drawn randomly from some predetermined, unit-energy,
complex-valued symbol constellation, over a radio-frequency channel with ad-
\-ditive white Gaussian noise (AWGN). Then we write the received multiuser
DS-CDMA signal, in its’ simplest form, as [112]

$$r(t) = \sum_{k=1}^{K} a_k b_k \xi_k(t) + \eta(t), \quad t \in [0, T) \quad (2.1)$$

where $T$ is a symbol duration, $a_k \in \mathbb{C}$ is a complex path gain of the $k$-th
user and $\xi_k(\cdot)$ denotes the pseudo-random spreading code wave associated with
the $k$-th user. We adopt here the typical assumption made in the theoretical
literature that the code waveforms are supported by the time interval $[0, T)$. Under this assumption the system characterized by the model (2.1) do not have
intersymbol interference. We also assume that the code waves have unit energy,
\-i.e., $\int_0^T \xi_k(t)^2 dt = 1$. In addition, $\eta(t)$ denotes an additive white Gaussian noise
signal.

A proper selection of spreading codes has an important role in the elabora-
tion of DS-CDMA systems. Properties of the codes affect directly on system’s
capability to prevent active users from interfering one another too much, as
well as, to cope with multipath propagation. Basically, cross-correlations or an
inner products between pairs of code waves, that is,

$$\rho_{kl} := \langle \xi_k, \xi_l \rangle = \int_0^T \xi_k(t)\xi_l(t) dt, \quad (2.2)$$

measure system’s inherent multiaccess interference resistance. Ideally, $\rho_{kl} \approx 0$
for $k \neq l$. Correspondingly, an autocorrelation function of a code wave character-
izes the ability to cope with multipaths. If the correlation between any two

© 2008 TUT / Toni Huovinen

Preliminary Examination Edition
different phases of any code is small (close to zero), the system is able to resist multipath propagation inherently. [95, 112]

Usually the code wave is formed by modulating a so called chip wave, \( p_{T_c} \), antipodally with a binary code sequence \((c_i)_{i=1}^{\infty}\), i.e., [113]

\[
\xi(t) = \sum_{i=1}^{C} (-1)^{c_i} p_{T_c}(t - (i - 1)T_c) \tag{2.3}
\]

where \( C \) is the spreading factor and \( T_c = T/C \) is the chip “duration”. In practice, the chip waves do not have to be strictly time-limited, but they are required to satisfy the condition

\[
\int_{-\infty}^{\infty} p_{T_c}(t)p_{T_c}(t - nT_c)dt = 0, \tag{2.4}
\]

for all \( n \in \{1, 2, \ldots\} \). Two theoretical examples of such waveforms are a rectangular pulse,

\[
p_{T_c}(t) = \begin{cases} 
1, & \text{if } 0 \leq t < T_c \\
0, & \text{otherwise}
\end{cases}, \tag{2.5}
\]

and sinc-pulse,

\[
p_{T_c}(t) = \text{sinc}\left(\frac{2t}{T_c} - 1\right). \tag{2.6}
\]

These are examples of time-limited and bandwidth-limited waves, respectively. In practice, some compromise of these two are selected due to the achieved spectral efficiency. For example, raised cosine with parameter \( 0 \leq \beta \leq 1 \),

\[
p_\beta(t) = \text{sinc}\left(\frac{t}{T_c}\right) \cos\left(\frac{\beta\pi t}{T_c}\right) \frac{1}{1 - (2\beta t/T_c)^2} \tag{2.7}
\]

has a band-width \( (1 + \beta)/(2T_c) \) [95].

The simple DS-CDMA model (2.1) is so called (symbol) synchronous model. It means that wide-band symbols of each user are assumed to arrive to the receiver exactly at the same time. Consequently, it is sufficient to consider only the reception of one symbol (per user) at a time. This approach is valid mainly in the downlink communications where the base station transmits the signals of all users in the system. Thus, the synchronization of transmissions is easy to implement. In the uplink, in turn, each active user sends its data independently and, in addition, they all have own, more or less random, propagation delays. For this reason, we need also an asynchronous signal model. The model
must take into account a block or a frame of, say, $M$ symbols, since due to asynchronism the demodulation of a particular users symbol is now affected by more than one symbol of the other users. Hence, we write the asynchronous DS-CDMA model as [112]

$$r(t) = \sum_{k=1}^{K} \sum_{m=1}^{M} a_k b_k[m] \xi_k(t - mT - \tau_k) + \eta(t). \quad (2.8)$$

Here, $\tau_k \in [0, T)$ defines a (relative) propagation delay and $b_k[m]$ the $m$-th symbol for $k$-th user. We assume that the users are in ascending order with respect to their delays, i.e., that $\tau_1 \leq \tau_2 \leq \ldots \leq \tau_K$. This assumption simplifies some following notations and has no effect on generality. Further, we actually assume in (2.8) that the reception of each user’s data block is beginning (as well as ending) within one symbol duration $T$.

In practice, the receiver observes several reflections of the transmitted signal. They all have traveled different radio paths and, hence, have different path gains and, especially, propagation delays [95,113]. In the case of $L$ propagation paths, we write the asynchronous (multipath) DS-CDMA model as [113]

$$r(t) = \sum_{k=1}^{K} \sum_{m=1}^{M} \sum_{l=1}^{L} a_{kl} \xi_k(t - mT - \tau_{kl}) + \eta(t), \quad (2.9)$$

where $a_{kl} \in \mathbb{C}$ and $\tau_{kl}$ are the complex path gain coefficient and the propagation delay for the $l$-th path of the $k$-th user, respectively. The delays of each path is assumed to be within one symbol duration $T$. The model 2.9 is often referred as “fixed multipath” model.

### 2.2 Conventional Single-User Detection

Let us consider, first, a so called single-user channel model, [112,113]

$$r(t) = ab\xi(t) + \eta(t), \quad t \in [0, T), \quad (2.10)$$

that is, the model (2.1) with one user ($K = 1$). We define the (coherent) matched filter (MF) output as

$$y = a^* \langle \xi, r \rangle = a^* \int_0^T \xi(t)r(t) \, dt = |a|^2 b + \tilde{\eta}, \quad (2.11)$$
in which
\[ \tilde{\eta} := \sigma \int_0^T \xi(t)\eta(t) \, dt \] (2.12)
is Gaussian random variable with zero mean and variance \( \sigma^2 \). The MF output, \( y \), is sufficient statistics to achieve the minimum probability of error when detecting the transmitted symbol \( b \) [95,112]. More precisely, the minimum distance decision after MF results in the minimum bit error probability (BEP) among all detectors given that noise realizations are not known. The error probability of such detector is given as [112]
\[ \text{BEP} = Q \left( \frac{|a|^2}{\sigma} \right), \] (2.13)
where \( Q(\cdot) \) is a complementary CDF of the unit variance and zero mean Gaussian random variable. This sets the unified upper bound for the system performance of any DS-CDMA receiver, and is called as the single-user bound. The MF detector is also called as the conventional single-user detector (SUD).

Let us turn our attention to the synchronous \( K \)-user channel model (2.1). The output of the \( k \)-th user matched filter is now, [112, 113]
\[ y_k = a_k^* \langle \xi_k, r \rangle = a_k^* \int_0^T \xi_k(t)r(t) \, dt \] (2.14)
\[ = |a_k|^2 b_k + a_k^* \sum_{j \neq k} a_j b_j \rho_{jk} + \tilde{\eta}_k \] (2.15)
where \( \tilde{\eta}_k \) is the filtered Gaussian noise term defined analogously to (2.12), and \( \rho_{jk} \) is the cross-correlation between the signature waveforms. The sum term in (2.15) is multiaccess interference (MAI) affecting on detection of \( k \)-th user. If signature waves, \( \xi_k, \ k = 1, \ldots, K \), are orthogonal, i.e., \( \rho_{jk} = 0 \ \forall j \neq k \), the MAI term vanishes, and \( y_k \) equals to single-user MF output for all users. Consequently, the MF detector is optimum (in the sense that it minimizes BEP) also in the orthogonal multiuser system. In general, the MAI term naturally deteriorates the detection compared to single-user bound, and the MF output, \( y_k \), is not any more the sufficient statistics to minimize the bit error probability [112].

Multiaccess interference affecting on the conventional single-user reception gets more complicated in asynchronous multiuser channel (2.1). Assuming we know the propagation delay, \( \tau_k \), we define the MF output for \( k \)-user’s \( m \)-th
2.2. CONVENTIONAL SINGLE-USER DETECTION

symbol as [112, 113]

\[ y_k[m] = a_k^* \int_{mT + \tau_k}^{(m+1)T + \tau_k} \xi_k(t - mT - \tau_k)r(t) \, dt \]

(2.16)

where \( Y_k[m] \) is a MAI term. Before characterizing this term rigorously, we define so called asynchronous cross-correlations between the signature waveforms as follows:

\[ \overline{\varrho}_{kl} := \int_{-\infty}^{\infty} \xi_l(t + T - \tau_l)\xi_k(t - \tau_k) \, dt, \]

(2.17)

\[ \varrho_{kl} := \int_{-\infty}^{\infty} \xi_l(t - \tau_l)\xi_k(t - \tau_k) \, dt \]

and

(2.18)

\[ \underline{\varrho}_{kl} := \int_{-\infty}^{\infty} \xi_l(t - T - \tau_l)\xi_k(t - \tau_k) \, dt. \]

(2.19)

These correlations in a sense determines how the current (e.g., \( m \)-th) symbol of the \( k \)-th user is affected, respectively, by the preceding ((\( m - 1 \))-th), the current (\( m \)-th) and the succeeding ((\( m + 1 \))-th) symbols of the \( l \)-th user. Notice, that

\[ \begin{cases} 
\overline{\varrho}_{kl} = 0 & \text{when } \tau_l < \tau_k \ \\
\underline{\varrho}_{kl} = 0 & \text{when } \tau_l > \tau_k. 
\end{cases} \]

(2.20)

Now, assuming \( \tau_1 \leq \tau_2 \ldots \leq \tau_K \) we can write the multiaccess interference term in (2.16) as

\[ Y_k[m] = \sum_{l<k} \left( a_l b_l[m] \varrho_{kl} + a_l b_l[m + 1] \overline{\varrho}_{kl} \right) + \sum_{l>k} \left( a_l b_l[m - 1] \overline{\varrho}_{kl} + a_l b_l[m] \underline{\varrho}_{kl} \right). \]

(2.21)

Not only the correlations \( \overline{\varrho}_{kl} \), \( \varrho_{kl} \) and \( \underline{\varrho}_{kl} \) but also the complex path gain coefficients \( a_k, k = 1, 2, \ldots, K \), affect on the multiaccess interference, \( Y_k[m] \), in (2.21). If the coefficients have clearly different moduli, the users with small moduli suffer from the effect of the multiaccess interference more than the users with relatively large moduli [112]. This situation appears especially in DS-CDMA uplink due to the spatial diversity of transmitting users. If all users transmit their signals with equal powers, then for example the users that are far
from the base station are received with weaker powers than the users close to the base station. This negative effect of the path gain coefficients with different moduli, or in other words the users with different received powers, is called as near-far problem.

Finally, we consider the conventional SUD in multipath channel (2.9). In many radio techniques multipath phenomenon and especially a fading phenomenon caused by it deteriorates the performance of the system [95]. However, as already discussed in the previous section, good autocorrelation properties of users’ spreading codes makes the DS-CDMA systems to resist distortion due to multipath propagation. Basically, this means that low correlations between code phases pushes down the contribution due to interfering multipath components in MF output. Naturally, code autocorrelation can never be fully ideal (i.e., zero autocorrelation between any miss-aligned code phases) in practice and, hence, existence of multipath interference increases the interference floor also in single MF output. A RAKE receiver [113] is an extension of the conventional detector which can take diversity advantage of the multipath environment. The RAKE receiver of $k$-th user consists of several MF detectors that are matched to code phases of different paths. These MF branches are often called as fingers. RAKE receiver combines finger outputs, say $y_{kl}[m]$, $l = 1, \ldots, L$, e.g., according to principle of maximal ratio combining (MRC) [95], that is,

$$y_k'[m] = \sum_{l=1}^{L} a_{kl}^* y_{kl}[m].$$

(2.22)

This improves a signal-to-noise(or interference)-ratio and, by this way, also the system performance.

2.3 On Code Synchronization

Conventional matched filter detection (and also most of the existing more an advanced detection methods) requires accurate knowledge of users’ propagation delays or, more precisely, (temporal) phase of the spreading sequences of decided users. We denoted this quantity earlier by $\tau_k$ (or $\tau_{kl}$ in the multipath context). Typically DS-CDMA receivers consist of two code synchronization or delay estimation stages: acquisition stage and delay tracking stage [113]. The purpose of acquisition is to obtain a coarse pre-estimate of the delays in precision of roughly a half chip duration. Code tracking then refines this pre-estimate and also adapts to a possible slow fluctuation in delay after the acquisition stage.
Describing details of these two code synchronization stages is out of scope in this dissertation. However, we assume in the following that the acquisition is performed, i.e., that receivers know coarse delay estimates for all desired users (and propagation paths), and denote the estimates by $\hat{\tau}_k$ (or $\hat{\tau}_{kl}$), $k = 1 \ldots K$. In addition, we denote the delay estimation errors by $\delta_k := \hat{\tau}_k - \tau_k$ (or $\delta_{kl} := \hat{\tau}_{kl} - \tau_{kl}$) and assume that $|\delta_k|$ (or $|\delta_{kl}|$) < 1/2 for all $k \in \{1 \ldots K\}$ (and $l \in \{1 \ldots L\}$).

Further, we can assume that acquisition is performed in discrete time (digital) domain and that the received continuous time signal is, first, sampled by chip matched filter. Consequently, resolution of code acquisition is one chip duration and, thus, $\hat{\tau}_k = d_k T_C$ (and $\hat{\tau}_{kl} = d_{kl} T_C$) for some integer(s) $d_k$ (and $d_{kl}$) $\in \{1, \ldots, C\}$. 
Before moving to more advanced DS-CDMA reception principles, let us take a look at the central signal processing method of this thesis, namely at independent component analysis (ICA) [29, 52, 53]. ICA is a statistical method for searching independent source random variables or signals from a set of observed linear combinations of them. (In this dissertation we consider only linear ICA.) One of the main applications of ICA is blind source separation (BSS), which has become an attractive field of research in statistical signal processing and neural network communities. ICA performs purely in a blind manner, i.e., without any explicit knowledge of original source variables or a mixing transformation. It relies on assumption of statistical independency of sources. Although independency is a very strong assumption from a theoretical point of view, it is often quite a realistic assumption in practice. Consequently, ICA has drawn a lot of attention in various application fields lately.

The idea of ICA was first introduced in a neurophysiological setting in the early 1980s by researchers (J. Hérault, C. Jutten and B. Ans) who needed a blind method to separate the neural impulses coming from different parts of the human body [5, 40, 41]. Later on, ICA has been studied and applied in several very different signal processing contexts like in audio and biomedical signal processing, feature extraction, finance, seismology, etc. Telecommunications related applications of ICA have been found earlier, e.g., in MIMO systems [62, 101, 117, 118], I/Q processing receivers [107], DS-CDMA blind multi-user detection [98] and DS-CDMA out-of-cell interference cancellation [9, 96, 97]. More inclusive review
of history of ICA and applications using ICA is given, e.g., in [53,56,57].

In this chapter, we first give a definition of linear ICA (Section 3.1). Next, in Section 3.2, we show that DS-CDMA model actually equals readily to ICA model. Then we continue by reviewing different algorithms briefly (Section 3.3) and, finally, discussing on performance issues of ICA algorithms (Section 3.4).

### 3.1 Definition

In this dissertation, we cover only, so called, linear independent component analysis, which we define in this section. Also other, extended, forms of ICA is known in literature (see, e.g., definitions of convolutive ICA and non-linear ICA in [53]). In the following, we give definitions for the basic noise-free (linear) ICA, noisy (linear) ICA and (linear) ICA with overcomplete basis in Sections 3.1.1, 3.1.2 and 3.1.3, respectively.

#### 3.1.1 Basic Noise-Free ICA

Prior to giving the definition of basic independent component analysis, we need to introduce the generative ICA model. In literature, the ICA model was first given for real-valued data and mixtures [22] and later on generalized straightforwardly for the complex-valued case. Here, we present the complex-valued model since it coincides with DS-CDMA data model defined earlier as will be seen in Section 3.2. Let \( \mathbf{s} = [s_1, s_2, \ldots, s_N]^T \) be a random vector, such that components \( s_1, s_2, \ldots \) and \( s_N \) are mutually independent, non-Gaussian, complex-valued random variables – to be precise, one and only one of them is allowed to be Gaussian. Further, let \( \mathbf{A} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_N] \in \mathbb{C}^{J \times N} \) be a constant complex-valued mixing matrix which we assume to have a full column rank. The column vectors, \( \mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_N \), are called as ICA basis vectors. Now, the basic noise-free ICA model is defined simply as linear model,

\[
\mathbf{x} = \mathbf{A}\mathbf{s},
\]

in which \( \mathbf{x} = [x_1, x_2, \ldots, x_J]^T \) is the observed \( J \)-dimensional random vector. The components of the vector \( \mathbf{s} \) and \( \mathbf{x} \) are called as independent components and ICA observations, respectively. Notice, that we assume above that the number of linear combinations \( x_i \) is greater or equal than the number of independent components \( s_i \) (\( J \geq N \)). In other words, it means that there is at least as many observations as sources.
3.1. **DEFINITION OF ICA**

Now we define independent component analysis as estimating independent components $s_1, s_2, \ldots, s_N$ and mixing matrix $A$ from model (3.1) while only the vector $x$ is known. In other words, the goal is essentially to invert the model (3.1) blindly, that is, to find a de-mixing matrix $B \in \mathbb{C}^{N \times J}$ such that $BA$ is as close to identity as possible by using only the observations $x$.

Because of the blindness of the problem, the original sources $s_n$ can not be recovered fully uniquely. To see this, recall first that the source components can only be determined up to some multiplicative constant, since for all complex constants $\lambda_n \neq 0$

$$x_j(t) = \left( \sum_{n=1}^{N} a_{mn}s_n \right)$$

$$= \left( \sum_{n=1}^{N} \left( \frac{a_{mn}}{\lambda_n} \right)(\lambda_n s_n) \right),$$

(3.2)

where $a_{mn}$ stand for $(j, n)$-th element of the mixing matrix. In other words, the sources and/or basis vectors can be estimated at most up to arbitrary scale.

Another trivial indeterminacy of ICA is the order of the estimated independent components. Since a sum is independent from the order of summation, the source components $s_n$ can be directly estimated only up to some permutation. Expect for these two indeterminacies, the ICA model (3.1) is unique and identifiable [22, 32, 33]. Thus, with ICA it is possible to find a random vector

$$s' = [\lambda_1 s_{p(1)}, \lambda_2 s_{p(2)}, \ldots, \lambda_N s_{p(N)}]^T,$$

(3.3)

in which $\lambda_k, k = 1, 2, \ldots, N$, are some arbitrary non-zero complex numbers, $p : \{1, 2, \ldots, N\} \rightarrow \{1, 2, \ldots, N\}$ is some permutation and $s = [s_1, s_2, \ldots, s_N]^T$ is an original source vector. In other words, de-mixing matrix, $B$, can be estimated unique only up to left multiplication by an arbitrary permutation and diagonal matrices.

Due to the scale indeterminacy of ICA, we can assume without loss of generality, that the source components, $s_1, s_2, \ldots$ and $s_N$, have unit variance. In addition, the sources can be assumed to be zero mean, since otherwise the sources can be made zero mean by subtracting the observation mean from observation ($\hat{x} = x - \mathbb{E}(x)$). This preprocessing task does not change the mixing matrix ($\hat{x} = A(s - \mathbb{E}(s))$). Together with the independence assumption, these assumptions imply that $\mathbb{E}(ss^H) = I$.

In the literature of complex valued ICA (see, e.g., [11]), the complex valued sources are usually assumed also to be circularly symmetric at least up to
second-order. A complex valued random variable, say \( z \), is said to be second-order circularly symmetric, if \( \mathbb{E} (z^2) = 0 \) \[91\]. Circular symmetry of the source components is assumed also here. This assumption yields (again together with independency and zero mean assumptions) that \( \mathbb{E} (s s^T) = 0 \).

### 3.1.2 Noisy ICA

Needless to say, the noise-free model (3.1) is unrealistic in most of the applications, especially, in telecommunications. For this reason, also ICA models with noise has been considered in literature \[53\]. Here, we define a noisy ICA model simply by adding an additive Gaussian noise term, \( \eta \), to the basic model, that is,

\[
x = A s + \eta.
\] \hfill (3.4)

In addition, we assume that the noise term is statistically independent from the sources \( s \). The identifiability of this model is treated in \[23\]. The main conclusion of that paper is the mixing matrix \( A \) can be still determined blindly up to the same indeterminacies as discussed in case of the noise free model. Nevertheless, the sources cannot fully be separated from noise in blind manner due to power or variance ambiguity between them.

### 3.1.3 ICA with Overcomplete Basis

A major drawback for standard ICA methods is the case where the number of source signals is greater than the number of observations made \((N > J)\), in which case standard ICA model does not hold anymore. This problem is commonly known as a “more sources than sensors”-problem, or ICA with overcomplete basis. Here, we say also that the data is over-saturated when the data have an overcomplete basis. This terminology is also used in \[P4–P8\].

Separating source signals from over-saturated data is more a difficult task, since mixing model is not invertible anymore. In other words, estimation of the mixing matrix is not alone enough to recover the source components \[53\]. In this respect, ICA with overcomplete basis is similar to noisy ICA. Consequently, typical ICA methods for over-saturated systems are computationally more expensive compared to standard ICA. Identifiability and separability of ICA model with overcomplete basis is discussed in \[32,33\]. The main conclusion is that also this model is identifiable, i.e., the mixing matrix can be uniquely determined up to trivial ambiguities, if none of the source signals are Gaussian. Nevertheless, complete separation of the sources is not possible anymore.
3.2 ICA Model in DS-CDMA Reception

In order to utilize ICA in DS-CDMA multiuser reception, we first need to reveal ICA model in DS-CDMA context. One way to do this, is to use antenna arrays. In that case, signals transmitted by independent transmitters acts as independent components and inputs of different antennas in antenna array as observations. This is due to fact, that each antenna receives differently weighted combination of transmitted signals (spatial diversity). Hence, signal by independent transmitters can be separated using ICA. In [97], the antenna array driven ICA is used for external interference cancellation in DS-CDMA downlink. Although ICA model can be generated as explained above basically in any telecommunication system, the disadvantage of this approach is the need for more complex and also more expensive physical system layer. In addition, antenna array based ICA does not take into account possibly beneficial characteristics of the system in which it is applied.

Other – DS-CDMA specific – way to reveal ICA model assumes single antenna only and is based on linearity of the data model (2.9). Let us assume first, that continuous-time signal is sampled by chip-matched filter. Since we have assumed that path delays of the users are within one symbol duration, using processing window length of two symbols, that is, 2

\[ C \]

chips, ensures that all contribution of (at least) one symbol of all users falls in that window. That is to say, we can collect the sampled data into vectors of length 2

\[ C \]

as follows: [98], [10]

\[
\mathbf{r}[m] := \sum_{k=1}^{K} \sum_{l=1}^{L} a_{kl} (b_k[m-1] \mathbf{g}_{kl} + b_k[m] \mathbf{g}_{kl} + b_k[m+1] \mathbf{g}_{kl}) + \mathbf{n}[m]
\]  

(3.5)

Here \( \mathbf{n}[m] \) denotes noise vector and code vectors of length 2

\[ C \] are defined as

\[
\mathbf{g}_{kl} := (1 - |\delta_{kl}|) \mathbf{c}_{kl}(d_{kl}) + |\delta_{kl}| \mathbf{c}_{kl}(d_{kl} + \text{sign}(\delta_{kl}))
\]

\[
\mathbf{g}_{kl} := (1 - |\delta_{kl}|) \mathbf{c}_{kl}(d_{kl}) + |\delta_{kl}| \mathbf{c}_{kl}(d_{kl} + \text{sign}(\delta_{kl}))
\]

(3.6)

\[
\mathbf{g}_{kl} := (1 - |\delta_{kl}|) \mathbf{c}_{kl}(d_{kl}) + |\delta_{kl}| \mathbf{c}_{kl}(d_{kl} + \text{sign}(\delta_{kl}))
\]

where \( d_{kl} = \tilde{\tau}_{kl}/T_C \in \{0, 1, \ldots, C\} \). Recall (from Section 2.3) that \( \tilde{\tau}_{kl} \) is a delay estimate of k-th user’s l-th path which we assume to be an integer multiple of chip duration and \( \delta_{kl} = \tilde{\tau}_{kl} - \tau_{kl} \) is a delay estimation error. In addition,

\[
\mathbf{c}_{kl}(d_{kl}) := [\xi_k[C - d_{kl} + 1] \ldots \xi_k[C] \mathbf{0}_{2C - d_{kl}}^T]^T
\]

\[
\mathbf{c}_{kl}(d_{kl}) := [\mathbf{0}_{d_{kl}}^T, \xi_k[1] \ldots \xi_k[C] \mathbf{0}_{C - d_{kl}}^T]^T
\]

(3.7)

\[
\mathbf{c}_{kl}(d_{kl}) := [\mathbf{0}_{C - d_{kl}}^T, \xi_k[1] \ldots \xi_k[C - d_{kl}]]^T.
\]
In (3.6), we have modeled an effect of delay estimation error after chip matched filtering basically as linear interpolation between two subsequent phases of code sequences. This model assumes the use of rectangular chip waveform. In general, the interpolation is non-linear.

With a simple manipulation, we can get a compact representation for the data,

\[ \mathbf{r}[m] = \mathbf{G}\mathbf{b}[m] + \mathbf{\eta}[m]. \]  

(3.8)

The \( 2C \times 3K \) dimensional code matrix \( \mathbf{G} \) contains the code vectors and path strengths, while the \( 3K \)-vector \( \mathbf{b}_m \) contains the symbols:

\[ \mathbf{G} := \left[ \cdots \sum_{l=1}^{L} a_{kl} \mathbf{g}_{kl} \sum_{l=1}^{L} a_{kl} \mathbf{g}_{kl} \sum_{l=1}^{L} a_{kl} \mathbf{g}_{kl} \cdots \right] \]  

(3.9)

and

\[ \mathbf{B}_m := [\cdots, b_k[m-1], b_k[m], b_k[m+1], \cdots]^T. \]  

(3.10)

Notice that \( m \)-th symbols of each user are included in the three successive vectors \( \mathbf{b}[m-1], \mathbf{B}_m \) and \( \mathbf{b}[m+1] \). Hence, we say that these vectors are “early”, “middle” and “late” parts of the \( m \)-th symbols, respectively.

Next, we interpret \( \mathbf{r}[m], \mathbf{B}_m \) and \( \mathbf{\eta}[m], m = 1, \ldots, M \), as ergodic samples of some underlying random vectors \( \mathbf{r}, \mathbf{b} \) and \( \mathbf{\eta} \) that also follows the affine relationship,

\[ \mathbf{r} = \mathbf{G}\mathbf{b} + \mathbf{\eta}. \]  

(3.11)

It is natural to consider \( \mathbf{\eta} \) as a jointly Gaussian (noise) vector with zero mean and covariance \( \mathbf{\Sigma} = \sigma^2 \mathbf{I} \), since the original noise process \( \eta(t) \) is white Gaussian process. In addition, it is reasonable to assume that users’ symbols are mutually and temporally independent, and consequently, to consider random vector \( \mathbf{b} \) to have independent components.

The stochastic, matrix algebraic representation (3.11) of DS-CDMA signal model is readily a noisy ICA model (3.4) with \( 2C \) observations of \( 3K \) source components provided that matrix \( \mathbf{G} \) has full rank. The full rank requirement of mixing matrix, that is the matrix \( \mathbf{G} \) here, can be clearly met only if \( 2C \geq 3K \) or, equivalently, \( K \leq \frac{2}{3}C \). In highly loaded system (i.e., \( K \approx C \)) the model (3.11) equals the noisy ICA model with overcomplete basis. We assume the model (3.11) in ICA assisted receiver structures proposed in this dissertation.
3.3 Algorithms

Basically, ICA algorithms can be characterized, in a few words, as optimization algorithms that search for extremum points of some suitable non-linear real-valued function depending on observed data [52,53]. These suitable functions – often called as contrast functions – are designed such that their extreme points equal to the ICA basis. In some of the ICA algorithms, the objective is to find the de-mixing matrix, $\mathbf{B}$. In that case, contrast functions are defined on $\mathbb{C}^{N \times J}$ and, ideally, $\mathbf{B}$ is their (global) extremum point. We refer these algorithms as multi-unit algorithms. Multi-unit contrast functions are based, e.g., on stochastic concepts of likelihood, entropy, mutual information, higher-order non-linear moments, etc. An extensive survey on different multi-unit contrast function can be found in [53]. One-unit algorithms, in turn, are meant to search for the basis vectors or independent components one by one. Their contrast functions are defined on $\mathbb{C}^{J}$. Basically, these contrast functions measure the non-Gaussianity of the inner product $\mathbf{w}^H \mathbf{x}$ for $\mathbf{w} \in \mathbb{C}^{J}$. Loosely speaking, the inner product that equals to some independent source component, has locally the most non-Gaussian distribution thanks to the central limit theorem. One-unit contrast functions can be based, e.g., on negentropy, that is, difference between the differential entropies of a given random variable and a Gaussian random variable with same variance. One popular class of one-unit contrast functions are based on higher-order cumulants, like a kurtosis of random variable, or some non-linear generalizations of them. Again, [53] includes good review on these functions.

In the following, we discuss briefly on algorithms for solving the optimization of ICA contrast functions. First, in Section 3.3.1, we introduce a pre-whitening which is often performed prior to actual ICA optimization. Then we continue with discussion of ICA algorithms for the basic, noisy and over-saturated model in Sections 3.3.2, 3.3.3 and 3.3.4, respectively.

3.3.1 Pre-whitening the data

A whitening or a sphering is a common preprocessing task in ICA algorithms, since it simplifies the remaining separation procedure. It is a linear transform $\mathbf{V}$ which de-correlates the observed mixtures, and normalize the observations components to have unit variances. Thus, the whitened data, $\mathbf{z} := \mathbf{Vx}$, satisfies

$$\mathbb{E}(\mathbf{zz}^H) = \mathbf{I}. \quad (3.12)$$

Such $\mathbf{V}$ exists always, but is not a unique transformation [53]. One way to find such a whitening transformation is to use a principal component analysis.
(PCA) [24], which gives the whitening transform as:

\[ V = \Lambda^{-1/2}E^H, \]  

(3.13)

in which (unitary) matrix \( E \) have the principal eigenvectors of the observation covariance matrix \( C_x := \mathbb{E}(xx^H) \) as columns, and the diagonal matrix \( \Lambda \) contains the corresponding eigenvalues on its diagonal. Clearly, all matrixes of form \( UV \), for any unitary matrix \( U \) and whitening matrix \( V \), whiten the data also. Hence, we get a notionally effective whitening matrix by multiplying (3.13) from left by \( E \). The resulting matrix is, actually, the inverse square root of the covariance \( C_x \), and it is denoted, in short, by \( C_x^{-1/2} \) [53].

Assuming the basic ICA model (3.1), the whitening (3.12) implies directly that the new (whitened) mixing matrix, \( W := VA \), is unitary, since

\[ I = \mathbb{E}(zz^H) = VA\mathbb{E}(ss^H)\Lambda^H V^H = WW^H. \]  

(3.14)

(We have assumed that \( \mathbb{E}(ss^H) = I \)). In other words, the new (whitened) ICA basis vectors, i.e., the columns of the matrix \( W \) are orthogonal vectors lying in the unit sphere. For this reason, many ICA algorithms assume the observed data to be first whitened and then constraint the search for ICA basis vectors (or de-mixing vectors) to the unit sphere and assume them to be mutually orthogonal.

### 3.3.2 Basic ICA algorithms

The basic (noise-free) ICA model has catch the most attention from researchers during the past, partly because of it was the first model considered and, partly, because it is the simplest (yet adequate to many applications) model. Consequently, numerous algorithms based on different criteria have been developed. Here, we list the most important families of algorithms. More a comprehensive survey on algorithms can be found, e.g., in [16, 52, 53, 57]. In addition, we describe two algorithms – FastICA [11, 46] and Equivariant Adaptive Source Identifications (EASI) [17] – more in detail since they are used in the publications compiled to this dissertation.

The Jutten–Hérault algorithm [58] is considered as the ground-breaking ICA algorithm. It is based on non-linear decorrelation, that is, canceling out certain non-linear (higher-order) correlations. Further, more an advanced, algorithms of this kind has proposed in [17, 19, 21]. An another family of algorithms worth to mention is based on maximization of network entropy, (also known as infomax principle) [1, 2, 6, 7]. Closely related algorithms [36, 90] use the the maximum
3.3. ICA ALGORITHMS

likelihood (ML), which is actually equal to infomax principle under suitable conditions [15, 80]. Higher-order cumulant tensor based algorithms are proposed, e.g., in [12–14, 18, 66]. The algorithm in [18] is well known as JADE algorithm (Joint Approximation Diagonalization Estimation). The algorithm in [66] is based merely on HOS, that is to say, it do not use second order statistics by any means. Also neural computation principles has adopted in ICA field. For instance, a non-linear neural PCA algorithm [83] is capable for ICA separation [59, 81, 82]. Other neural network based ICA algorithms are proposed, e.g., in [35, 37, 54, 55, 60, 74].

Most of the ICA algorithms consider the observed time signals as samples of some underlying random variables. They do not pay attention to possible temporal structure of data. However, there is also algorithms, for example the ones given in [8, 76, 106], that are based, in particular, the temporal structure or time correlations in observations and source signals.

FastICA algorithm [11, 45, 46, 48] is a fixed-point algorithm which operates on a block of observed data samples. Basically, FastICA algorithm searches for extreme points of $\mathbb{E} (F(|w_i^H z|^2))$ in unit sphere. Here, $F : \mathbb{R} \rightarrow \mathbb{R}$ stand for an one-unit contrast function discussed above. In effect, the extreme points points are, thus, the ICA basis vectors. The FastICA algorithm converges faster than typical stochastic gradient decent algorithms [53], hence the name. This is also why FastICA has become popular in application field. A one-unit version of the algorithm recovers one ICA basis vector in an iterative or recursive manner. Assuming the pre-whitened observations (3.12), the recursion step is given as [11, 53]

$$w_{i+1} = \mathbb{E} \left( z (w_i^H z)^* f(|w_i^H z|^2) \right) - \mathbb{E} \left( f(|w_i^H z|^2) \right) + |w_i^H z|^2 f'(|w_i^H z|^2) w_i, \tag{3.15}$$

in which $f : \mathbb{R} \rightarrow \mathbb{R}$ is derivative of a given contrast function $F$. In addition, $w_{i+1}$ must be normalized to have unit norm before proceeding to the next step. Assuming a simple kurtosis based contrast function, $F(y) = \frac{1}{2} y^2$, we can simplify (3.15) to [53, 98]

$$w_{i+1} = \mathbb{E} \left( z (w_i^H z)^* |w_i^H z|^2 \right) - \gamma w_i, \tag{3.16}$$

where the scalar coefficient $\gamma = 3$ for real valued data and $\gamma = 2$ for complex valued data. The iterations steps (3.15) or (3.16) are computed until a convergence, i.e., until $|w_{i+1}^H w_i|$ is close enough to one. After the convergence, the corresponding independent component is $w_j^H z$, where $j$ is the last index of
CHAPTER 3. INDEPENDENT COMPONENT ANALYSIS

Recursion. Several independent components are recovered by repeating the one-unit algorithm successively starting the iterations from different, e.g., randomly selected initial points ($w_0$). However, without control, recursions can converge to the same basis vector for several times. This is prevented by making the vector $w_i$ orthogonal to all basis vectors already estimated, between each iteration. Recall, that the basis vectors are orthonormal after the pre-whitening. The orthogonalization can be accomplished, for instance, with Gram-Schmidt algorithm [38]. Repeating the one-unit algorithm successively is often referred as deflation. Another way to estimate several ICA basis vectors is to use, so-called, symmetric FastICA [11,53]. This algorithm runs several, say $K$, one-unit iterations in parallel manner and, between each iteration step, orthogonalizes the vectors $w_i(k), k=1\ldots K$, using symmetric orthogonalization. The matrix $W_i = [w_i(1) \ldots w_i(K)]$ can be orthogonalized symmetrically, e.g., as [53]

$$\hat{W}_i = W_i (W_i^H W_i)^{-\frac{1}{2}}. \quad (3.17)$$

EASI algorithm [17], on turn, is one of the algorithms based on non-linear decorrelation mentioned above. To be more precise, it is a recursive online algorithm which operates on individual samples of observed data. The explicit pre-whitening is not assumed in this algorithm. One recursion step of the EASI algorithm, i.e., of searching the de-mixing matrix $B \in \mathbb{C}^{N \times J}$, is given as

$$B_{m+1} = B_m - \mu \Psi(v_m)B_m, \quad (3.18)$$

in which $\mu$ is a scalar step size, $v_m = B_m x[m]$ and the update matrix, $\Psi(v_m) \in \mathbb{C}^{N \times N}$, is defined as

$$\Psi(v_m) = v_m v_m^H - I + f(v_m) v_m^H - v_m f(v_m)^H. \quad (3.19)$$

Here, $f: \mathbb{C}^N \rightarrow \mathbb{C}^N$ is an arbitrary non-linear function. On right-hand side of (3.19), two first terms tend to whiten the output $v_m$, thus, the algorithm uses implicitly the orthogonality constraint discussed in the previous section. Since only the current sample is used in each step of the algorithm, the update matrix (3.19) does not vanish asymptotically. Instead, a stationary point of the algorithm is defined stochastically as follows: an $N \times J$ matrix $B'$ is a stationary point of the EASI algorithm if the expected value of the update term, $\Psi(v)$, is zero, i.e.,

$$E(\left( v v^H - I + f(v) v^H - v f(v)^H \right)) = 0, \quad (3.20)$$

for $v = B' x$ [17]. The convergence of the EASI algorithm is given as stability of the stationary point [17]. In other words, at least local convergence is guaranteed in theory.
The selection which algorithm to use depends, in practice, largely on the applications at hand. In theory, all the algorithms are, of course, equivalent at least with respect to the outcome. They all recover the mixing transformations and independent components. The practical validity of each algorithm may vary a lot depending on system or model parameters as dimensionality of the data, available sample size, etc. In Publications [P1–P11], we selected the FastICA and EASI algorithm to represent ICA due to their popularity and relatively easy implementations. In conceptual level, the validity of the proposed methods and shown results, should not be depend on the choice of the algorithm.

3.3.3 Noisy ICA algorithms

Formally the noise-free (3.1) and noisy (3.4) ICA models do not seem to differ that much. In practise, however, the noisy ICA has turn out to be a challenging task. Consequently, the number of ICA algorithms capable to recapture the mixing matrix under the noisy model is comparatively small. Recall also, that estimating the mixing (or de-mixing) matrix is not sufficient to separate sources from noise. The topic of noisy ICA is considered, e.g., in [20,26,47,66,77].

Basically, FastICA algorithm can be used also for the noisy ICA if the covariance of noise, $\Sigma := \mathbb{E}(\eta^H \eta)$, is known [50,51,53]. The only modification to the basic algorithm is in pre-whitening stage, which is replaced with, so called, quasiwhitening. A quasiwhitening transformation or matrix is defined as

$$\tilde{V} := \left( C_x - \Sigma \right)^{-1/2}, \quad (3.21)$$

where $C_x$ is the observation covariance. This matrix actually equals to the covariance of noise-free part of the observation $(x - \eta)$ and, in effect, whitens that part of the observation. The whole observation is not white after quasiwhitening. After the quasiwhitening, FastICA basically converges to the same solution than in corresponding noise-free model provided that the contrast function is not affected by additive Gaussian noise. Such contrast functions are discussed in [51].

Since the basic noise-free ICA algorithms are clearly better known than the noisy algorithms in literature, applications of ICA often assume a noisy linear model, but exploit one of the basic ICA algorithms anyway. A presence of reasonable level of additive noise is thought to cause “only” some feasible distortion due to the model mismatch. Surprisingly, one finding of this dissertation [P9,P11] is that applying noise-free ICA algorithms to noisy data can, actually, result in the best possible linear source recovery – clearly better what
can obtained by the conventional de-mixing transformation, that is, by the in-
verse transformation of the mixing transformation. We discuss more on this
finding in Section 3.4.

3.3.4 Algorithms for ICA with overcomplete basis

Also the ICA model with overcomplete basis has received some algorithm devel-
opment. For instance, [3,70,71,84,85] cover this issue. From the computational
complexity point of view, a modification of a FastICA algorithm [49] is one of
the most attractive ICA algorithm designed to operate in over-saturated sys-
tems. The modification is based on concept of quasiorthogonality, which roughly
speaking means that we now consider a set of nearly orthogonal basis for the
representation of the data, thus enabling the increase in dimensionality com-
pared to strictly orthogonal basis. Quasiorthogonalization of set of $N$ basis
vectors can be carried out by an iterative orthogonalization algorithm:

$$W_{t+1} = \frac{3}{2}W_t - \frac{1}{2}W_tW_t^HW_t$$

(3.22)

with $W_0 = W \in \mathbb{C}^{J \times N}$. Here, $W$ stands for the matrix of the estimated
basis vectors. If $W$ has a full rank (which, of course, is not the case when
$N > J$), this iterative algorithm converges to an orthogonal matrix. If $N > J$,
the algorithm anyway makes the basis vectors closer to orthogonal in each
iteration step although the convergence to orthogonal matrix is impossible [53].
Numerical experiments in [P4–P8] and, e.g., in [49] indicates that only a few
(or even one) iteration steps are enough for the FastICA algorithm. Since only
change to the standard FastICA algorithm is in orthogonalization, the modified
algorithm have the same computational complexity than the standard one.

3.4 Discussion on Performance Study of ICA

Algorithms

In principle, all the ICA algorithms are statistical estimation algorithms whose
purpose is to estimate elements of the mixing matrix or, equally, de-mixing ma-
trix. Typically performance of ICA algorithms is studied by measuring asymp-
totic difference of the original mixing matrix and its estimate. The difference
can be measured, for instance, as an asymptotic efficiency of an ICA estimator. For
example, [64] introduces and analyzes the versions of FastICA algorithm that
produce an unbiased estimator of de-mixing matrix which attains the Cramer-Rao lower bound.

Also different performance or rejection indexes are used to measure the difference in the literature. They indicate how well the matrix product between an estimated de-mixing matrix and the original mixing matrix (or vice versa) resembles the identity or permutation matrix which are the ideal cases. For example, [8] and [2] use indexes

\[ I_1 = \sum_{i \neq j} \mathbb{E} |\kappa_{ij}|^2 \]  

(3.23)

and

\[ I_2 = \sum_{i=1}^{N} \left( \sum_{j=1}^{N} \frac{|\kappa_{ij}|}{\max_k |\kappa_{ik}|} - 1 \right) + \sum_{j=1}^{N} \left( \sum_{i=1}^{N} \frac{|\kappa_{ij}|}{\max_k |\kappa_{kj}|} - 1 \right), \]  

(3.24)

respectively. Here, \( \kappa_{ij} := [\bar{\mathbf{B}} \mathbf{A}]_{ij} \) where \( \mathbf{A} \) is the original mixing matrix and \( \bar{\mathbf{B}} \) is an estimate of de-mixing matrix. The index (3.23) is a stochastic index which measures the mean disparity between the product \( \bar{\mathbf{B}} \mathbf{A} \) and the unity matrix. Recall, that the product is basically a random matrix since \( \bar{\mathbf{B}} \) can be taken as a random estimator of the de-mixing matrix. Hence, this type of index can be used in analytical performance studies. Typically, analytical evaluation of this type of stochastic index is, nevertheless, difficult or impossible for higher order ICA methods and, for this reason, (3.23) is applicable mainly for second order BSS/ICA algorithms. The index (3.24), in turn, measures the difference of the product \( \bar{\mathbf{B}} \mathbf{A} \) from arbitrary permutation matrix. It does not assume the product \( \bar{\mathbf{B}} \mathbf{A} \) to be random, but rather takes \( \bar{\mathbf{B}} \) as output matrix of ICA algorithm. Consequently, this index is suitable for numerical performance analysis.

Assuming the noise free model (3.1), measuring the quality of the mixing or de-mixing matrix estimate is essentially equal to measuring the quality of corresponding independent component estimates, since the noise free ICA model is invertible. However, this is not the case in the noisy model (3.4), since transforming the observation vector by the de-mixing matrix can, naturally, cause uncontrolled noise amplification. Basically, the performance indexes (3.23) and (3.24) defines merely an output interference-to-signal-ratio (ISR) (or, strictly speaking, the sum of ISR’s wrt. to all independent components) given as the average power ratio between the contribution of one (desired) independent component and the total contribution of the other (interfering) components in the
de-mixed output that corresponds to the desired component. In many applications, like in the ICA assisted CDMA detectors considered in this dissertation, the ultimate purpose is to recover the original independent components from noisy mixtures. For this reason, better performance indicator is an output interference-and-noise-to-signal ratio or, equivalently, an output signal-to-interference-and-noise ratio (SINR) which takes into account also the additive noise.

To give a definition of SINR, let \( w \in \mathbb{C}^J \setminus \{0\} \) be an arbitrary linear filter and \( y_w = w^H x \) the corresponding filtered output. In terms of ICA, \( w^H \) can be taken as the row of an estimated de-mixing matrix which corresponds to, say, \( n \)-th source component. SINR wrt. this source component, at the output \( y_w \) is then defined as

\[
\text{SINR}_n(w) := \frac{w^H R_n w}{w^H R'_n w},
\]

(3.25)

in which

\[
R_n := \mathbb{E} \left( h_n s_n s_n^* h_n^H \right) = h_n h_n^H
\]

(3.26)

and

\[
R'_n := \mathbb{E} \left( \sum_{k \neq n} h_k s_k + \eta \right) \left( \sum_{k \neq n} h_k s_k + \eta \right)^H = \sum_{k \neq n} h_k h_k^H + \sigma^2 I.
\]

(3.27)

Since all eigenvalues of the Hermitian matrix \( R'_n \) are greater than or equal to \( \sigma^2 \), the Hermitian form in the denominator of (3.25) is positive definite, i.e., strictly positive for all \( w \in \mathbb{C}^J \setminus \{0\} \), provided that \( \sigma^2 > 0 \). Consequently, (3.25) is well-defined for all \( w \in \mathbb{C}^J \setminus \{0\} \).

Now, as seen in (3.25), maximizing SINR among all linear transformations of observed data, i.e., maximizing \( \text{SINR}_n(w) \), equals to solving the generalized eigenvalue problem [104] associated with matrix pair \((R_n, R'_n)\). Hence,

\[
\max_{w \in \mathbb{C}^J \setminus \{0\}} \text{SINR}_n(w) = \lambda_n
\]

(3.28)

and

\[
\arg \max_{w \in \mathbb{C}^J \setminus \{0\}} \text{SINR}_n(w) = e_n,
\]

(3.29)

in which \( \lambda_n \) stands for the greatest eigenvalue of the Hermitian matrix \((R'_n)^{-1} R_n\) and \( e_n \) for the corresponding eigenvector. To be specific, since \( \text{SINR}_n(w) \) is scale
invariant, $e_n$ can be any vector in one-dimensional eigensubspace corresponding to the eigenvalue $\lambda_n$.

Also the linear minimum mean square error (LMMSE) estimator of a source can be shown to yield the maximum SINR among linear transformations. This is basically stated in [116] and in references therein. The LMMSE transformation for $n$-th source in the noisy ICA model (3.4) is given as [61]

$$m_n = C_x^{-1}h_n,$$

in which $C_x = E(xx^H)$ is the observation covariance and $h_n$ stands for the $n$-th column of the mixing matrix in the model (3.4). This linear transformation, thus, gives an explicit solution to the generalized eigenvalue problem above, i.e., $e_n = m_n$. Nevertheless, the solution assumes the knowledge of the mixing coefficients and noise variance and, for this reason, is not a blind method as such, but rather an appropriate reference method for ICA algorithms under the noisy ICA model.

Naturally, a transformation that approximately maximizes the linear SINR can be constructed blindly as the estimated LMMSE transformation after the identification of the mixing matrix, $A$, given that the observation covariance, $C_x$, is also estimated. However, results in Publications [P9, P11] (and also numerical results in [63, 65]) suggest that some ICA algorithms developed for noise-free models are able to provide directly input-output SINR gains very close to the best linear gain possible, in particular, the gains clearly better than with inverse transform of $A$. This can be explained by pre-whitening the observed signal $x$ to whitened signal $z$, which actually also transforms the original mixing matrix to the LMMSE matrix of $z$ [P11]. Hence, after whitening, ICA algorithms using an optimization criterion that is invariant to additive Gaussian noise [53] actually estimate the LMMSE transformations directly. This seems not to be well-understood earlier in literature. For instance, authors of [63, 65] do not give mathematical explanation why their ICA methods reach so close to performance of LMMSE in the papers. Above reasoning, however, explains also good performance of their algorithms.

Only problem in the most of the well known ICA algorithms is their assumption of orthogonality (or unitarity in in the complex valued case) of the mixing matrix after whitening, which is valid, in general, only if the linear model is noise-free. In the noisy model (3.4), this assumption is not true, or in other words, the whitened mixing matrix is not orthogonal (unitary) in general. For this reason, the algorithms using the orthogonality constraint can not attain the M-GEF (or LMMSE) solution exactly in theory. In Appendix A, we propose an
simply DS-CDMA specific modification of the FastICA algorithm that enables relaxation of the original orthogonality constraint and, consequently, provides variant of the FastICA algorithm that attains the LMMSE solution exactly in theory.
Chapter 4

Multiuser Detection

Conventional way to detect all $K$ users of the DS-CDMA system, e.g., in base station, is to use $K$ parallel single-user receivers, each receiving one user independently [112]. Such a receiver structure is shown in Fig. 4.1. If spreading codes of users is chosen wisely such that cross-correlations between users is low for all relative delays, acceptable performance can be maintained using the bank of conventional single-user detectors. Nevertheless, designing the adequate signature codes is very demanding when the number of users is high and, especially, in asynchronous systems (2.8) and (2.9). Moreover, the conventional reception is very sensitive to the near-far problem even under low cross-correlation. This has led to developing of more sophisticated multiuser detectors (MUD) that take into account the existence of interfering users and, consequently, provide performance improvements compared to the conventional single-user detection [27, 109, 111].

While the multiuser detection improves the system performance under multiaccess interference (MAI), they also have some drawbacks. First, the receiver should know the spreading codes of each user in the system in order to detect any of the users. In the conventional receivers only the code of desired user is needed. Second, the multiuser detectors are computationally more complex than the conventional receivers. The first drawback, however, is circumvented by blind MUD methods that take the existence of MAI into account implicitly or, in other words, learn the information needed to suppress the effect of MAI by observing the data.
Figure 4.1: Conventional multiuser detector. Bank of matched filters has one MF for each user. Multiaccess interference is not taken into account.
4.1. OPTIMUM MUD

In this chapter, we review the most important MUD methods. We start with optimum MUD in Section 4.1 and continue with linear and non-linear suboptimum MUD in Sections 4.2 and 4.3, respectively. Finally, in Section 4.4, we consider different blind MUD receivers including those that are based on the independent component analysis.

4.1 Optimum MUD

In [109], Verdú develops and analyzes an optimum multiuser detector, that is, the detector that minimizes the probability of error among all possible detectors. In other words, its objective is to maximize a posteriori probability of transmitted symbol for all users, i.e., to find the value $\hat{b}_k[m]$ from the symbol alphabet that maximizes the conditional probability

$$P( b_k[m] \mid r(t), t \in [0, \infty) ),$$

(4.1)

where $b_k[m]$ is the transmitted symbol (interpreted as a random variable) and $r(t)$ is the received signal. This detection strategy is called the individually optimum detection. In the synchronous system (2.1), observation of $m$-th symbol interval ($t \in [(m-1)T, mT]$) is sufficient for optimum detection. In the asynchronous system (2.8), in turn, reception of the entire frame of transmitted symbols ($t \in [0, MT]$) is needed.

An alternative detection strategy (also introduced by Verdú), jointly optimum detection, maximizes the joint a posteriori probability which, for the synchronous model, is given as

$$P( \beta[m] \mid r(t), t \in [(m-1)T, mT) ),$$

(4.2)

where $\beta[m] = [b_1[m], b_2[m], \ldots, b_K[m]]^T$ is the vector of $m$-th transmitted symbols [112]. For asynchronous model, the jointly optimum detection is defined as maximization of the joint conditional probability of symbol vector $\beta_{\text{all}}$ containing a whole frame of symbols for each user given corresponding frame of the received signal, that is,

$$P( \beta_{\text{all}} \mid r(t), t \in [0, M) ).$$

(4.3)

The jointly optimum detection is not exactly equal to the individual one. However, unless the SNR is very low, error probabilities of the two strategies are close to each other. More precisely, the error probability of the jointly optimum strategy converges to error probability of individually optimum one as noise goes
down [112]. Consequently, the joint detection is favored due to its lower complexity. Assuming equiprobable and independent symbols, we can write [112] the jointly optimum detection more explicitly as

$$
\hat{\beta} = \arg \max_{\beta} \Omega_K(\beta),
$$

where maximization is performed over all possible combinations of $m$-th symbols (synchronous case) or of symbols in whole frame (asynchronous case). In the synchronous case, $\Omega_K$ is defined for $m$-th symbols as

$$
\Omega_K(\beta[m]) = \sum_{k=1}^{K} A_k b_k[m] y_k[m] - \sum_{i,j=1}^{K} A_i A_j b_i[m] b_j[m] \rho_{ij}.
$$

and in the asynchronous case, for entire frame of symbols as

$$
\Omega_K(\beta_{all}) = \sum_{k=1}^{K} \sum_{m=-M}^{M} A_k b_k[m] y_k[m]
$$

$$
- \frac{1}{2} \sum_{k=1}^{K} \sum_{m=-M}^{M} \left[ \sum_{j<k} A_k A_j (b_k[m] b_j[m] \rho_{jk} + b_k[m] b_j[m+1] \rho_{kj}) + \sum_{j>k} A_k A_j (b_k[m] b_j[m-1] \rho_{jk} + b_k[m] b_j[m] \rho_{kj}) \right].
$$

Above $y_k[m]$ is the matched filter output of the $k$-th user’s $m$-th symbol defined in Section 2.2.

The optimum multiuser detection gives a great improvement in performance and is also more robust to the near-far problem compared with the conventional single-user detector. However, the major drawback of the optimum detection is its’ exponential complexity in the number of users due to the combinatorial nature of the involved optimization task [110]. To get rid of this disadvantage suboptimum multiuser detectors, which still outperforms the conventional multiuser detection but with a moderate level of increase in computational complexity, have been developed. In the next sections, we survey the most important such detection methods.

### 4.2 Linear suboptimum MUD

As the name suggests, suboptimum multiuser detection (MUD) refers to methods that take the existence of multiaccess interference into account – explicitly or
implicitly – to outperform the conventional single-user detection, but are relaxed from the computationally heavy requirement of optimal performance. In this section, we introduce the two most well known linear suboptimum MUD principles: decorrelating reception and linear minimum mean square error (LMMSE) reception. We adopt, here, the stochastic data model,

\[ r = Gb + \eta, \]  

that we introduced in Section 3.2 (Eq. (3.11)).

### 4.2.1 Decorrelating detection

Let us assume that we know spreading code vectors, channel coefficients and propagation delays for all active users in the system, then we have all the building blocks to construct the matrix \( G \) in (4.7). Now, if \( G \) has full rank it has pseudo-inverse \( G^\dagger \), and we have that

\[ y_{DD} := G^\dagger r = b + \hat{\eta}, \]

where \( \hat{\eta} := G^\dagger \eta \) is a transformed Gaussian noise vector. Hence, the components of the vector \( y_{DD} \) are entirely free of multiaccess interference. Nevertheless, since this transformation does not take into the account existence of the additive noise, it tends to enhance the noise floor.

In literature, the transformation by pseudo-inverse of \( G \) followed by the element-wise minimum distance decision is known as decorrelating detection (DD) [112]. Given the synchronous multiuser data model (2.1), DD is often presented using vector of MF outputs, \( y = [y_1, y_2, \ldots, y_K] \) and a cross-correlation matrix, \( \Gamma \), of spreading waveforms whose elements are defined as

\[ [\Gamma]_{kl} := \langle \xi_k, \xi_l \rangle = \rho_{kl}, \quad k, l = 1, \ldots, K. \]  

Namely, these two quantities, \( y \) and \( \Gamma \), are related as

\[ y = \Gamma \Delta b + \tilde{\eta}, \]

in which \( \Delta = \text{diag}(a_1, \ldots, a_K) \), \( b = [b_1, \ldots, b_K]^T \) and \( \tilde{\eta} \) is the vector of matched filtered noise components. Provided that the square matrix \( \Gamma \) is invertible, we have analogously to (4.8) that

\[ y'_{DD} := (\Gamma \Delta)^{-1} y = \beta + \tilde{\eta}. \]
in which \( \tilde{\eta} \) denotes again a transformed noise vector. Recall, that the right hand sides of (4.8) and (4.11) are essentially equal – they both include current symbols of all users.

From (4.11) we see that the knowledge of channel coefficients is employed only to scale and rotate independently the complex symbols, \( \beta \), to have their original moduli and phases prior to the minimum distance decision. In other words, dropping \( \Delta \) from (4.11) does not change component wise output signal-to-noise ratio (SNR). Particularly, that is to say, that DD ignores the received (i.e., input) SNR of the users’ data, which is seen as gratuitous noise enhancement.

A linear algebraic interpretation of DD is that the decorrelating linear filter for \( k \)-th user component, i.e., \( k \)-th row of \( (\Gamma \Delta)^{-1} \) (or, correspondingly, \( G^\dagger \)), is orthogonal to the subspace spanned by the interfering users as shown in Fig. 4.2.1. This also supports the the fact that DD does not use the information of the amplitudes of interfering user components, but just their ”directions in signal space”.

Figure 4.2: Decorrelating detector (DD) is orthogonal to signal subspace spanned by interfering users (the dashed arrows). The solid unlabeled arrow depicts the user of interest.
4.2. LINEAR SUBOPTIMUM MUD

4.2.2 Linear MMSE detection

As discussed in the previous section, the decorrelating detector does not actually exploit the knowledge of users’ received amplitudes and noise level and, for this reason, suffers from performance degradation especially in low SNR scenarios. In this section, we explore how the performance of linear multiuser detection can be improved if we use that knowledge. The improvement must exist, since even conventional single-user detection outperforms decorrelating detector when SNR is sufficiently low [112].

A key to the performance improvement is the linear minimum mean square error (LMMSE) detection that maximizes the MSE function,

\[
\text{MSE}(\mathbf{M}) := \mathbb{E} \left( \| \mathbf{b} - \mathbf{M}^H \mathbf{r} \|_2^2 \right)
\]

with respect to the matrix variable \( \mathbf{M} \in \mathbb{C}^{2C \times 3K} \) [112]. The matrix that maximizes (4.12) is [61]

\[
\mathbf{M}_0 = \mathbf{C}_r^{-1} \mathbf{G},
\]

in which \( \mathbf{C}_r = \mathbb{E}(\mathbf{r}\mathbf{r}^H) \) is covariance matrix of the received data (cf. LMMSE vector (3.30) in Section 3.4). Decision statistics of LMMSE detection is

\[
\mathbf{y}_{\text{LMMSE}} := \mathbf{M}_0^H \mathbf{r}.
\]

An important property of the LMMSE transformation (4.13), which we already discussed also in Section 3.4, is that it maximizes the linear signal-to-interference-and-noise ratio (SINR). Or to be more precise, the \( k \)-th column of \( \mathbf{M}_0, \mathbf{m}_0 = \mathbf{C}_r^{-1} \mathbf{g}_k \), maximizes the linear SINR, i.e., SINR wrt. the symbol \( b_k \) at the output \( y_w = \mathbf{w}^H \mathbf{r} \) among all linear filters \( \mathbf{w} \in \mathbb{C}^{2C} \). Recall that we gave the rigorous definition of SINR in (3.25) in Section 3.4. In other words, the components of the LMMSE output, \( \mathbf{y}_{\text{LMMSE}} \), has maximum linear SINR wrt. the corresponding components of \( \mathbf{b} \). Of course, SINR is scale-invariant and, consequently, scaling columns of the matrix \( \mathbf{G} \) in (4.13) does not change component-wise SINR. However, the covariance matrix \( \mathbf{C}_r \) in (4.13) includes implicitly the information of users’ received amplitudes and noise level making the LMMSE detection more sophisticated compared to decorrelating one. We can think that LMMSE detection is optimal linear trade-off between the complete mitigation of multiaccess interference and the best noise reduction. This is strictly true only in the SINR sense. In the bit-error-probability sense, the optimality (among linear receivers) is not exactly attained by maximizing SINR unless the component-wise total interference after linear transformation
is Gaussian [112]. In practise, however, the interference is approximately Gaussian thanks to the well known central limit theorem provided that number of active users is relatively high.

4.3 Nonlinear suboptimum MUD

The decorrelating detection and LMMSE detection described in the previous section are the most important examples of linear MUD methods. In this section, we consider non-linear MUD methods, that is, the methods that can not be presented as a linear transformation of received data (followed by minimum-distance decision). Typically such methods have the interference subtractive nature meaning that the multiaccess interference or at least some portion of it is first estimated and then subtracted from received data prior to demodulation of the user’s symbols.

4.3.1 Successive interference cancellation (SIC)

Successive interference cancellation (SIC) is a non-linear suboptimum multiuser detection method. The basic idea in SIC is to detect the symbols of interfering users one by one. After detecting a symbol of one user the wave form carrying the symbol is recreated and subtracted from the remaining signal before detecting the next user’s symbol [112]. Recreation of the wave form (as well as detection of symbols) is possible only if the spreading codes of all interfering users are known by receiver. In addition, to be able to recreate the wave forms of the users, the delays and path gains of each user should have been estimated [43].

The main principle of SIC detector is shown in Fig. 4.3. The signal \( \tilde{r}_k(t) \), \( k = 1, 2, \ldots, K \), which is used as received signal for detection of the \( k \)-th user can now be formulated as [43,112]

\[
\tilde{r}_k(t) = \begin{cases} 
  r(t) & \text{, when } k = 1 \\
  r(t) - \sum_{l=1}^{k-1} \sum_{m=1}^{M} a_l \hat{b}_l[m] \xi_l(t - mT - \tau_l) & \text{, otherwise,}
\end{cases}
\]

where \( r(t) \) is the original received signal and \( \xi_l(\cdot) \), \( a_l \), \( \tau_l \) and \( \hat{b}_l[m] \) are the spreading code, path gain coefficient, delay and estimate of the \( m \)-th symbol of the \( l \)-th user, respectively. Note that the symbols from the first to the \( (k-1) \)-th user have already been estimated when the signal \( \tilde{r}_k(t) \) is about to be generated.
Typically, the conventional single-user matched filter detector is used to obtain the symbol estimates [112].

If the decision of user’s symbol is correct, the multiaccess interference (and of course total interference also) decreases during every subtraction. If the detection is incorrect, subtraction actually enhances the interference. Consequently, the order in which users are demodulated and subtracted affects the reliability of SIC. Thus, the user which is the most likely to be detected correctly or, in other words, has the best received SINR, should be processed first and so on [43, 112]. One way to proceed is use the decreasing order of received power estimates [43]. This method is popular since it is simple to implement. However, it does not take into account cross-correlations between users spreading codes. A better way to order the user is to use, e.g., MF output of the user [112].

The performance of SIC is analyzed in [87, 112]. SIC was found to outperform the conventional detectors clearly. SIC increases both the system performance in a multipath fading environment and the robustness against near-far effect. Thus, the capacity of the system can also be increased with a SIC based receiver [87]. Compared to optimum MUD, SIC does not give quite as good performance improvement [112]. On the other hand, the number of cancellations made in SIC is (at most) $K - 1$, where $K$ is the number of users, and every cancellation has a constant computational complexity in the number of user. So, the complexity of SIC is (almost) linear in the number of users [112] (the only non-linear part in SIC is ordering the users) whereas it is exponential for the optimum MUD.

### 4.3.2 Parallel interference cancellation (PIC)

Parallel interference cancellation (PIC) is another non-linear suboptimum multuser receiver structure. Like the successive interference cancellation, the PIC is also based on the idea of estimating and subtracting the multiaccess interfer-
ence (MAI) from the received signal before processing it. Though, interference cancellations are accomplished in parallel manner. We consider here only the data detection methods using PIC, but the PIC ideology is also used for channel estimation and delay tracking (see, e.g., [67, 69]).

In the basic PIC [39, 69], the received multiuser signal is first passed through the bank of $K$ matched filter, where each filter is matched to spreading code of one user. Then the MAI affecting the demodulation of each user is estimated and subtracted from the matched filter outputs. In order to estimate the MAI, the channel parameters should naturally be first estimated. Finally, the minimum distance decision is used to obtain the hard estimates of users’ symbols.

The multistage algorithm by Varanasi and Aazhang is one of the PIC based methods for data detection [108]. The detection of users’ symbols is done in $P$ stages (see Fig. 4.4). Before making the $p$-th, $p = 1, 2, \ldots, P$, stage decision of user’s symbol the MAI is estimated and subtracted from the received signal. The MAI estimation in the $p$-th stage is made using the symbol estimates of the $(p-1)$-th stage. Hence, the $p$-th stage decision statistics of the $k$-th user’s the $m$-th symbol can be written as [108]

$$y_k^p[m] = y_k[m] - \tilde{\Upsilon}_k^{p-1}[m]$$ (4.16)

where $y_k[m]$ is the MF output and $\tilde{\Upsilon}_k^{p-1}[m]$ is the $(p-1)$-th state’s estimate of MAI affecting on the $m$-th symbol of the $k$-th user. The $p$-th stage’s MAI estimate is defined more exactly as

$$\tilde{\Upsilon}_k^p[m] := \sum_{l=k+1}^{K} a_l \hat{b}_l^p[m-1] \overline{\rho}_{kl} + \sum_{l \neq k} a_l \hat{b}_l^p[m] \rho_{kl} + \sum_{l=1}^{K-1} a_l \hat{b}_l^p[m+1] \rho_{kl},$$ (4.17)

where $a_l$ denotes a channel coefficient of the $l$-th user and $\hat{b}_l^p[m]$ an $p$-th stage’s symbol estimate of $l$-th user’s $m$-th symbol. $\overline{\rho}_{kl}$, $\rho_{kl}$ and $\rho_{kl}$ are the asynchronous correlations between users’ spreading codes defined rigorously in Section 2.2. Recall, that the MAI affecting on $y_k[m]$ is characterized as

$$\Upsilon_k[m] = \sum_{l=k+1}^{K} a_l b_l[m-1] \overline{\rho}_{kl} + \sum_{l \neq k} a_l b_l[m] \rho_{kl} + \sum_{l=1}^{K-1} a_l b_l[m+1] \rho_{kl}.$$ (4.18)

Thus, the MAI is totally cancelled in the $p$-th stage if the symbol estimates of the $(p-1)$-th stage are correct (assuming ideal channel and timing estimations). In the first stage, the MAI estimation can be done using the symbol estimates.
Figure 4.4: The multistage PIC detector. A bank of $K$ matched filters which is followed by $K$ $P$-stage processors. Each processor estimates the MAI affecting on the symbols of the corresponding user in each stage. The MAI estimated on the $(p - 1)$-th stage is subtracted from the MF output before the estimation of the $p$-th stage.
obtained directly from matched filter outputs. The symbol estimates obtained in the last (P-th) stage is considered as the final decisions.

As seen in (4.17), the three successive symbol estimates from previous stages for each user is needed to obtain one, say m-th, symbol estimates for each user in current stage. Recursively, three symbol estimates from the stage (P-1), five symbol estimates from the stage (P-2), . . . and finally 1 + 2P symbol estimates from the initial stage is needed to obtain the final (the P-th stage’s) estimates of the m-th symbols [67, 108]. Figure 4.5 shows an example of the needed symbol estimates in case of the 2-stage detector. Consequently, all needed symbol estimates should be stored until they are not needed anymore [108] and, hence, implementations should contain relatively large amount of memory. In addition, roughly half of the previous stages’ symbol estimates needed in the current stage, are estimates of the subsequent symbols with respect to the symbol being processed. As a result, an additional processing delay is involved in multistage algorithm [108].

The performance of the multistage PIC detector is experimented numerically in [108]. A significant performance improvement was found compared to the conventional single-user detector even with small number of cancellations.
4.3. **NONLINEAR SUBOPTIMUM MUD**

stages. Moreover, the improvement was getting even more outstanding when the number of stages was getting larger. Also the computational complexity of the multistage algorithm is dealt with in [108] and a conclusion is that the algorithm is linear with respect to the number of users, i.e., the algorithm has the complexity of $O(K)$. This is valid if only one user is detected. Hence, for detection of all active users (which is usually the case in the uplink receivers), the PIC algorithm has the computational complexity of $O(KK) = O(K^2)$.

Some modifications to the basic multistage algorithm have also been done. The improved version of the multistage algorithm is derived by in [25]. In the original multistage algorithm, the whole estimated MAI is subtracted in every stage. In the improved version, on the contrast, only a fraction of the estimated MAI is subtracted in the early stages such that the amount of MAI being subtracted is getting larger as the algorithm proceeds. Since the reliability of tentative decisions is relatively poor in the beginning of the algorithm and gets better towards the end, the use of gradually increasing weights for the MAI estimates reduces the influence of erroneous tentative decisions. This improves the performance of the multistage algorithm.

### 4.3.3 Comparison of SIC and PIC

The performance of the SIC and basic PIC based detectors is compared in [88]. The main findings are that, first, if all users are received with equal strength, then the PIC based receiver outperforms the SIC based one. Secondly, in the case of inequal received powers SIC gives better performance than PIC. Recall that in the SIC method the users are first ordered such that in every successive cancellation the user cancelled is the strongest one of the remaining users. However, if all received users have nearly equal strength, this ordering cannot be done properly. This explains why SIC can provide relatively better performance if received powers are not equal. Thus, the SIC method was found to resist the near-far problem more than PIC.

Nevertheless, only 1- and 2-stage PIC was considered in the experiments in [88]. The multistage PIC also with more than two cancellation stages is considered as a reference method in [P1–P8]. An interesting finding is that the performance of the PIC receiver with more than three cancellation stages seem to be dramatically different from the performance of the PIC receiver with only one or two cancellation stages in the system with a severe near-far problem. In our experiments, the performance of the PIC receivers with more than three cancellation stages was actually getting better while near-far problem was getting worse. Moreover, the PIC receivers with more than three
cancellation stages outperform the SIC receiver clearly also in the cases of a bad near-far problem.

We shall also compare the complexity of SIC and PIC shortly. Like seen in Sections 4.3.1 and 4.3.2 the SIC method has linear and the multistage PIC method quadratic computational complexity in the number of users. However, the processing delay is also linear in the number of users in the successive cancellation [112]. Whereas, the multistage PIC can be implemented such that it has a constant processing delay in the number of users and linear in the number of cancellation stages [108]. Thus, if the number of stages used in the PIC is relatively small and the number of users is relatively large, the PIC receiver can provide shorter processing delay than the SIC receivers.

4.4 Blind MUD

As “blindness” has been a vogue concept in the signal processing society during the past years, the word ‘blind’ has been attached to names of numerous and quite different signal processing methods and algorithms. Accordingly, a meaning of the word varies from a context to another. Also this thesis treats with two distinct concepts that are referred to blind in the literature, but where the notion has different meaning. In Chapter 3, we already discussed the blind source separation (BSS), i.e., the signal processing concept aiming to recover source variables or signals from their linear mixtures. One could say that BSS methods are blind in the strict sense, since they are not allowed to use basically any explicit knowledge of the sources (except maybe the number of them) and the mixing transformation. These methods rely only on some very general identifiability assumptions like bijectivity of the mixing transformation and, for example, statistical independence of the sources.

The other concept referred as blind and considered here, is blind multiuser detection [111, 112, 114, 115]. When detecting a DS-CDMA user, blind MUD receivers assume the same knowledge as the conventional single-user receiver, that is, the spreading code and channel state information (including timing) of the user of interest only. In addition, they implicitly assume an existence of MAI from other users of the system and try to learn or mitigate it in the blind manner. Thus in this case, the notion of blindness is somewhat looser than in the former case – from BSS’s point of view, blind MUD receivers use partial knowledge of mixing transformations explicitly. The following subsection (Subsection 4.4.1) consider blind counterparts of the suboptimum MUD methods discussed in the previous sections. Also BSS/ICA assisted MUD receiver structures (Subsection
4.4. BLIND MUD

4.4.2 can be seen as blind MUD methods, although some of them (in particular, receiver structures proposed in [P1–P8]) consist of two separate stages: a fully blind separation stage and an non-blind identification stage following after the separation. This type of structure is often referred as ‘semi-blind’.

In rest of this section, we use again the stochastic data model,

\[ r = Gb + \eta, \]

that we introduced in Section 3.2 (Eq. (3.11)). Recall also, that we have assumed the additive noise term to have the covariance matrix \( \mathbb{E} (\eta\eta^H) = \sigma^2 I \) with \( \sigma^2 \) being the noise variance.

### 4.4.1 Subspace decorrelating and LMMSE detection

The purpose of subspace detection methods [112,115] is to achieve a blind MUD receiver in the sense described above. Thus, we assume here that receiver knows only one column of the matrix \( G \) in (4.19), the column corresponding to the (“middle” part) of the user of interest. That column vector – let us assume it is the \( k \)-th column and denote it by \( g_k \) – consist, essentially, of the spreading code vector of the user with correct timing. The other columns of \( G \), we assume as unknown. In addition, we assume the noise variance \( \sigma^2 \) to be known by the receiver.

When deriving a blind variant of decorrelating detector (DD), the objective is to find the \( k \)-th row vector of the pseudo inverse \( G^\dagger \) by observing received data \( r \) and using the vector \( g_k \). This is equal to finding a vector orthogonal to the signal subspace spanned by multiaccess interferers, that is, the other users than the one of interest. The row vector in question is found (blindly) by inspecting the the covariance matrix of received data [112]. We can write the covariance matrix as

\[ C_r := E[rr^H] = GG^H + \sigma^2 I. \]  

(4.20)

Now, since the covariance matrix is hermitian and positive definite, it has an eigendecomposition of the form \( C_r = \mathbf{E} \Lambda \mathbf{E}^H \) with an unitary matrix \( \mathbf{E} \) and a positive definite diagonal matrix \( \Lambda \) [104]. Consequently, we can write the covariance matrix also as

\[ C_r = \mathbf{E} (\Lambda - \sigma^2 I) \mathbf{E}^H + \sigma^2 I. \]  

(4.21)

Comparing (4.20) and (4.21), we conclude that

\[ GG^H = \mathbf{E} (\Lambda - \sigma^2 I) \mathbf{E}^H \]  

(4.22)
and, hence,
\[
G^\dagger := G^H (GG^H)^{-1} = G^H E(\Lambda - \sigma^2 I)^{-1} E^H.
\] (4.23)

Finally, we obtain the \(k\)-th row vector of \(G^\dagger\) from the most right hand side of (4.23) as
\[
g_k^\dagger := g_k^H E(\Lambda - \sigma^2 I)^{-1} E^H. \tag{4.24}
\]

Since the covariance \(C_r\) can be learnt by observing received data, \(g_k\) is sufficient information to construct the DD receiver of the desired user in (4.24) assuming that noise power is known or estimated. Especially, explicit information of MAI is not needed, thus, this variant of DD is a blind MUD method in the sense described above. In practise, to be able to learn the covariance \(C_r\) or its eigendecomposition adequately, receiver must observe received data from longer time period than just over the one symbol, which is enough in non-blind DD reception. In other words, the receiver is able to learn the signal subspace in which MAI lies by observing received data over a longer period period of time.

Similarly, we can construct subspace variant of LMMSE receiver straightforwardly from (4.13) as
\[
m_k = E \Lambda^{-1} E^H g_k. \tag{4.25}
\]

This is, thus, the \(k\)-th column of LMMSE matrix represented in terms of signal subspace formalism. Also other blind adaptive MUD algorithms with theoretical performance equal to LMMSE receiver, are proposed in the literature [34, 72, 73, 112]. These are based on minimizing output energy of “canonical” matched filter output, that is essentially, minimizing
\[
\text{MOE}(v) := E \left( (g_k + v)^H r \right)^2, \tag{4.26}
\]
with respect to \(v \in \mathbb{C}\) and subject to orthogonality constraint \(v^H g_k = 0\). It can be shown that, the canonical form \(g_k + v'\) maximizes LMMSE for minimum point of (4.26) denoted by \(v'\) [73]. This minimum output energy (MOE) formalism enables the adaptive gradient descent algorithms that need only the same information than the conventional single-user detector.

Also a concept of blind multistage interference cancellation and blind successive cancellation are proposed in [99] and [100], respectively.

### 4.4.2 ICA assisted MUD

Recall from Section 3.2, that DS-CDMA model (4.19) is actually the noisy ICA model. Therefore, extracting source components from received data, \(r\), with ICA, is basically what is required of multiuser detector. Nevertheless,
4.4. BLIND MUD

due to ICA’s phase and permutation unambiguity, separation alone is not quite enough for detection of users’ data – source components need to be identified after separation or, alternatively, the unambiguities must be removed before it. This can be done if the receiver knows at least the spreading code(s) of the user(s) of interest. In principle, ICA assisted receivers considered here use the same information as needed in blind MUD methods, i.e., the column(s) of the matrix $G$ corresponding to the desired user(s).

At this point, it is good time to set the question: “Why in earth wants someone use some considerably weighty blind statistical ICA algorithm if we anyway must know essentially all the information the algorithm can give us?” Fortunately enough, the answer is simply: in practice, the receiver does not know the needed columns of $G$ exactly. It only knows estimates of them that are affected by channel and timing estimation errors. However, ICA algorithms, as being fully blind, are not affected by these errors. The ambiguity removal task after the actual ICA separation, in turn, is not very sensitive to these errors [P1–P8]. Thus, the effect of erroneous knowledge of the matrix $G$ can be compensated using ICA. In the following, we assume that the receiver knows only erroneous estimate of $G$ or its $k$-th column and denote them by $\tilde{G}$ and $\tilde{g}_k$, respectively.

From performance point of view, ICA assisted MUD receivers are, conceptually, closely related to decorrelating and LMMSE receivers. Namely the objective of ICA algorithms is, essentially, to find (pseudo-)inverse of the matrix $G$ or at least one row of it (cf. one-unit algorithms). In other words, the algorithms try to find the decorrelating detectors of the active users. Under noisy model, this is strictly true only if, so-called, noisy ICA algorithms (see Section 3.1.2) is used. However, as described in Section 3.4, the basic (noise-free) ICA algorithms that include pre-whitening stage, e.g., the FastICA algorithm, tend to estimated the LMMSE matrix related to the model in hand. Hence, using these algorithms resembles basically LMMSE reception.

One of the first BSS/ICA assisted DS-CDMA MUD receiver structure was proposed in [98]. Starting from the model (4.19), the proposed receiver simply uses the one-unit FastICA algorithm to separate the desired user from multiaccess interference. However, the one-unit FastICA applied such as, can converge to any user component in the system due to ICA’s permutation indeterminacy. To circumvent this problem, that is, to make FastICA to converge to the desired user, the receiver uses the the vector $\hat{g}_k$, or strictly speaking, the corresponding LMMSE vector $\hat{m}_k := E \Lambda^{-1} E^H \hat{g}_k$ (cf. (4.25)), to initialize FastICA. The idea is to start FastICA algorithm from initial point close enough to the desired (and unknown) target point and by this way guarantee the converge to the right ba-
sis vector. Nevertheless, sometimes the FastICA algorithm may converge to a wrong basis vector although the initial value is selected wisely. In these cases, the conventional detection provides likely a better performance than ICA. For this reason the the receiver includes also “a branch switching” between the conventional and ICA assisted detection methods as a final step of the reception. The branch switching is performed by measuring the difference between initial and final point(s) (or vector(s)) of FastICA. If the difference is small enough, the ICA assisted branch is selected. Otherwise, the FastICA algorithm has the most likely run away from neighborhood of the desired basis vector and converged to another one, thus, output of conventional receiver is used.

In Publications [P1,P2], we derived and analyzed a semi-blind receiver structure for the DS-CDMA uplink. In uplink receivers (e.g., in base stations of cellular network), all the active users need to be detected. Hence, the receiver uses the symmetric FastICA algorithm to separate all the active users from one another. The receiver, nevertheless, is not totally blind, since users are identified using the pre-knowledge of the users, that is, the pre-estimated matrix \( \tilde{G} \), after the separation. When identifying the \( k \)-th user, the receiver evaluate correlations between \( \tilde{g}_k \), and ICA’s basis vectors, say \( w_{k'} \), that are estimated by FastICA. That is to say,

\[
\theta(k, k') = \frac{\tilde{g}_k w_{k'}^H}{|\tilde{g}_k||w_{k'}|} \quad (4.27)
\]

is computed for each \( k' \). Finally, the source which corresponds to the best correlation is chosen to be that user’s data. Furthermore, the phase of complex number \( \theta(k, k') \) in (4.27) equals to arbitrary and unknown phase shift produced by ICA. This makes it possible to correct ICA’s phase ambiguity. This simplest idea of using ICA to separate DS-CDMA users from each other has been reinvented several times in literature [28,30,31,102,103] since the author introduced it first time in his M.Sc. thesis [44].

Sometimes FastICA fails to recover some of the data components satisfyingly due to noise and other imperfections in the model. Poorly estimated data components impair the performance of reception, especially when the number of users is high. However, if an estimated source signal is of low quality, the corresponding column of estimated mixing matrix is usually defective, also. This makes it possible to recognize poorly estimated users. To elude this problem, we introduced an idea of adapting the concept of successive interference cancellation (SIC) to the ICA assisted receiver in Publication [P3]. The principle of this successive ICA receiver is rather straightforward. First, a correlation
threshold, say $\theta_1$, is set to indicate proper detection. This is needed because subtraction of erroneously detected signal will actually enhance interference. Thus, after the symmetric FastICA separation, only the signal corresponding to the user $k$ with $\theta(k, k') > \theta_1$ for some $k'$ is subtracted from the receiver signal. After subtraction of all such users’ signals, a new ICA separation is performed for the interference-subtracted signal. The order of ICA model decreases in subtraction, which help ICA to recover the remaining users. The procedure is repeated successively until all users are detected. The successive ICA receiver is illustrated in Fig. 4.6.

Another recent study has also found blind source separation techniques to be beneficial in interference subtractive receivers [105]. In that study an iterative least squares technique for digital signal separation was enhanced by an interference cancellation stages, a procedure similar to conventional successive or parallel interference cancellation schemes. Namely, the method takes advantage of an alternating estimation of the mixing matrix and user symbol streams in until converged, while employing an interference cancelling stages during the symbol streams’ estimation.

Recall from Chapter 3, that the major drawback for standard ICA methods is the oversaturated model where the number of source signals is greater than the number of observations made in which case standard ICA model does not hold anymore. And further recall, that, from the single antenna DS-CDMA reception point of view, the data model become oversaturated if the number of users is greater than two third of processing gain ($K > \frac{2}{3}C$), assuming processing window size of two symbols. Hence, the number of users in the system must be rather low to make standard ICA applicable. In Publications [P4–P6], we showed that this “more sources than sensors”-problem can be somewhat circumvented in the successive-type ICA receiver, in particular, if received powers of the users are inequal. This is mainly based on the PCA pre-whitening task of FastICA, which is originally intended to make the remaining separation procedure simpler. However, since relative magnitudes of the eigenvalues are, in practice, directly proportional to the strengths of independent components the whitening by PCA tends to prefer the strong components, and some weak components might be lost [78]. In general, this is an undesirable phenomenon, since ICA algorithms should not be sensitive to the differences in component strengths. Nevertheless, the successive ICA receiver structure benefits of this phenomenon. Namely, as PCA processing prefers relatively strong users, these users are detected more likely after ICA separation. On the other hand, interference subtractions are performed for original, unwhitened data, which is not affected by PCA. Hence, originally relatively weak users are intact after
Figure 4.6: Successive ICA assisted receiver structure. $r^i(t)$ and $r^i$ are a received data signal and a received data vector after the $(i-1)$-th IC round, respectively ($r^1(t) = r(t)$ and $r^1 = r$). Vector $y_{ICA}^i$ is source components by FastICA separation. Number of components in that vector decreases in each IC round. Soft and hard decisions of users’ symbols are denoted by $\tilde{b}^i$ and $\hat{b}^i$, respectively. $\Upsilon_i(t)$ denotes interference due to users detected in $i$-th round.
the subtractions. How these users can be detected after following ICA separations depends on their relative strengths and number. This all is to say, that in a sense, PCA preprocessing hides the weak users in early separation stages whereas subtractions reveals them again in the later ones. Therefore, the successive ICA receiver structure provides the best performance gain when number of users having (nearly) equal received powers is small enough.

One should recognize the fact that, conventionally, in interference subtractive receivers, like in SIC, a very precise channel state information is needed in the interference subtraction to prevent interference enhancement. This is why also accurate code timing acquisition and tracking is of great importance. However, an important property of ICA assisted receivers is its inherent capability to cope with erroneous parameter estimates. Roughly speaking, the estimate of interference by ICA implicitly includes also the timing information of the user signals, thus giving a possibility to refine preliminary delay estimates before re-generation of interfering signals and its substraction. This kind of delay estimate refinement is not inherently possible in conventional interference subtractive receivers, but additional delay tracking circuitry is needed. We proposed in Publications [P6, P8] a modification of successive ICA receiver which is, in practice, immune to delay estimation errors. The modification constitutes of a two-step delay tracker within ICA processing where a coarse pre-tracking helps the identification procedure during ICA and a finer post-tracking improves the interference re-generation, and hence, enables more effective interference cancellation. It is of primary importance to stress out, that all timing refinements are done after the ICA processing. Hence, also the delay tracking procedure takes advantage of interference suppression due to ICA separation.

The above mentioned ICA assisted MUD-receivers are based on stochastic single-antenna DS-CDMA model directly. Also receiver structures using ICA that are based on multi-antenna model is proposed in literature. For example, [86] introduces such a structure. More in detail, the proposed method maximizes the multiuser kurtosis among linear transformations of received data. An other example of multi-antenna methods is proposed recently in [79]. It uses a tensorial ICA method.

### 4.4.3 Discussion on complexity of ICA assisted receivers

Considering practical implementation of any receiver structure, the computational complexity of the structure in hand is, naturally, one of the first attributes taken under inspection. However, the most of ICA algorithms – including FastICA – are iterative statistical optimization algorithm, where a needed number
of iterations before convergence is random in its nature. For this reason, deriving overall complexity of ICA algorithms is difficult. In [44], the author concludes that the complexity of the simple (non-successive) ICA receiver \([P1, P2]\), measured as needed “multiply and add”-operations, is \(O(K^2M) + O(CK^2)\) in which \(C, K, M\) \((M >> C, K)\) denote the length of spreading code, number of users and number of symbols (per user) in one received data block, respectively. The derivation of this result is also included for a easy reference in Appendix B. However, this simple analysis does not take into account how a needed number of FastICA iterations depends on parameters \(C, K\) and \(M\), but rather considers the number of iterations only as a constant.

The most natural way to affect the convergence speed of an ICA algorithm or, in other words, keep the needed number of iterations in moderate level, is to use the known and pre-estimated system parameters when initializing the algorithm. Considering the FastICA algorithm, for instance, and assuming that we initialize it with the “whitened” LMMSE matrix related to \(\tilde{G}\), we actually start the iterations from some close neighborhood of the actual ICA solution provided that \(\tilde{G} \approx G\). Hence, we can expect to obtain a convergence after a very small number of iterations in that case. This is also demonstrated by new results in Appendix A. Moreover, using information included in \(\tilde{G}\) to wise initialization is, in general, a sensible and simple way to embed the approximative pre-knowledge of the system in ICA separation stage.

The subtractive ICA assisted structures \([P3–P8]\), in turn, consists of (a few) successive repetitions of the symmetric FastICA separation stage. In addition, the fist separation is computationally the most demanding, since a dimension of ICA model (actually, the parameter \(K\)) decreases after each stage. Consequently, the complexity the successive structures is basically of the same order than the complexity of the simple ICA receiver.
Rejection of External Interference

In the previous chapter we discussed several methods of mitigating an internal MAI due to non-ideal cross-correlations between spreading sequences in a DS-CDMA system. Naturally, also some external sources can interfere the DS-CDMA reception. Here, we mean by an external interference signal, basically, any non-Gaussian signal component other than MAI. These signals can be, e.g., some spurious narrowband signals originated from nearby RF channel or wideband spread spectrum signals from adjacent cell of a cellular system. Also an intentional jamming is possible, especially, in military systems. In this chapter, we discuss the rejection of external interference shortly. First, we define the external interference signal rigorously in Section 5.1. Then, in Section 5.2 we give brief overview of conventional ideas to fight against narrowband interference. In the end, in Section 5.3, we continue by discussion on an inherent capability of the ICA assisted receivers to resist external interference.

5.1 External Interference in Signal Model

We model the external interference signal, say $\Upsilon_{\text{ext}}(t)$, as a continuous wave (possibly pulsed at the symbol level) with an amplitude $A_T \in (0, \infty)$, frequency $f_T \in \mathbb{R}$ and phase $\phi_T \in [0, 2\pi)$:

$$\Upsilon_{\text{ext}}(t) = \chi_p(t) A_T e^{i(2\pi f_T t + \phi_T)}, \quad (5.1)$$
where a stochastic process

\[ \chi_p(t) = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1 - p \end{cases} \]

during a symbol duration, i.e., \( \chi_p(t) \equiv \chi_p[m] \in \{0, 1\} \) for \( t \in [(m - 1)T, mT) \).

Hence, the interference corresponds to a continuous wave when \( p = 1 \) and pulsed wave at the symbol level otherwise.

We obtain a special case of the (5.1) assuming that the frequency, \( f_\Upsilon \), is locked to the carrier frequency used in the system. In this case, the external interference signal remains constant during one symbol duration also after the down conversion (assuming that the phase of the signal is constant). Hence, in this case, we can write the (base-band) interference signal as

\[ \Upsilon_{\text{ext}}(t) = \Upsilon_{\text{ext}}[m], \quad \text{for } t \in [(m - 1)T, mT), \]

where \( \Upsilon_{\text{ext}}[m] := \chi_p[m]A_\Upsilon e^{i\phi_\Upsilon[m]} \). In addition, we assume that the phase \( \phi_\Upsilon[m] \) is a sample of an uniformly distributed random variable having values in \([0, \pi)\). This presentation is assumed also in Publications \([P1–P3, P6–P8]\) and, e.g., in \([96, 97]\).

The received signal has now the form

\[ r(t) = r_{\text{If}}(t) + \Upsilon_{\text{ext}}(t + \tau_\Upsilon), \]

where \( r_{\text{If}}(t) \) is the original external interference-free signal and \( \tau_\Upsilon \in [0, T) \) is a formal temporal phase of the interference signal. For sake of simplicity, we assume that \( \tau_\Upsilon \) is an integer multiple of chip duration \( T_C \).

We can embed the external interference also into the stochastic, matrix algebraic representation (3.11) defined in Section 3.2. Namely, after chip-matched filtering, the \( m \)-th received data vector of length \( 2T_C \) becomes

\[ r[m] = r_{\text{If}}[m] + \Upsilon_{\text{ext}}[m - 1]\mathbf{1} + \Upsilon_{\text{ext}}[m]\mathbf{1} + \Upsilon_{\text{ext}}[m + 1]T\mathbf{1}, \]

where \( r_{\text{If}}[m] \) the interference-free data vector and \( 2T_C \)-vectors \( \mathbf{1}, \mathbf{1} \) and \( \mathbf{1} \) are defined as

\[ \mathbf{1} := [1 \ldots 1 0_{2T_C - \tau_\Upsilon}]^T, \]
\[ \mathbf{1} := [0_{T_\Upsilon}^T 1 \ldots 1 0_{T_C - \tau_\Upsilon}]^T \quad \text{and} \]
\[ \mathbf{1} := [0_{T_C + \tau_\Upsilon}^T 1 \ldots 1]^T. \]
5.2. CONVENTIONAL NARROWBAND INTERFERENCE REJECTION

Hence, if we extend the original code matrix, $G$, with three column vectors $\mathbf{1}$, $\mathbf{1}$ and $\mathbf{1}$ and the original symbol vector, $B_m$, with three interference symbols $\Upsilon_{\text{ext}}[m-1]$, $\Upsilon_{\text{ext}}[m]$ and $\Upsilon_{\text{ext}}[m-1]$, we get the same representation,

$$r[m] = GB_m + \eta[m],$$

(5.6)

for the data as in interference-free case. Again, we interpret $r[m]$, $B_m$ and $\eta[m], m = 1, \ldots, M$, as ergodic samples of some underlying random vectors $r$, $b$ and $\eta$ that also follows the affine relationship,

$$r = Gb + \eta.$$  

(5.7)

Thus, the data model with external interference still corresponds to ICA model with three additional source components.

5.2 Conventional Narrowband Interference Rejection

Several techniques to improve the performance of the spread spectrum system in the presence of narrowband interference (NBI) have been proposed during the past. Recall, that the external interference signal defined in (5.1) and (5.2) corresponds to NBI signal if pulse probability $p = 1$. An inclusive review on traditional NBI cancellation methods can be found in [4], [75], [94] and references therein. Here we go through the most typical NBI methods briefly in conceptual level.

In contrast to an energy of the spread spectrum signal, all energy of narrowband interference is concentrated to some slender frequency band near a central frequency. Consequently, if a spread spectrum signal is interfered by a narrowband signal and total energies of the signals are somewhat in the same level, the power density spectrum has a clean peak on the band of the narrowband signal. The typical proposed NBI cancellation methods exploit this spectral feature in one way or the other.

The simplest NBI cancellators use some (adaptive) notch filter whose stop band equals approximately to the band of the interferer. This approach surely removes the interference signal almost totally, but in pursuance, causes some damage to the desired data signal. However, if the attenuated band is very narrow with respect to the total band width of the spread spectrum signal, the damage caused might be more tolerable compared to the interfering effects of the narrowband signal. More sophisticated NBI cancellation approach proposed,
is depicted in Fig. 5.1. The main idea in these cancellators is to estimate the narrowband signal and then subtract the estimate from the received signal before the detection of users. If the estimate subtracted equals exactly to the original interference signal, the interference is totally cancelled and no damage is caused to the spread spectrum signal (cf. subtractive MUD methods). Naturally, the perfect interference estimate cannot be obtained, but some estimation error is always made. The estimation of the interference is based on the fact that the narrowband signal is well predictable compared to the spread signal which is almost random in time domain. The NBI estimator can be implemented, for instance, by linear or non-linear filtering.

Another typical NBI cancellation scheme, which can be considered more sophisticated than simple notch filtering, is frequency domain NBI suppression. The functionality of such a cancellator structure is shown in Fig. 5.2. A received signal is first Fourier-transformed. This reveals the spectral power peak due to narrowband interference. Now, the peak can be attenuated by cutting off the power density spectrum from some predetermined power level (see Fig. 5.3). This threshold level should be determined such that the power density of the desired spread spectrum signal stays only just under the level in whole frequency band. After frequency domain handling, the modified signal is inverse-transformed to time domain before the detection. In principle, the frequency domain cancellation does not harm the spread spectrum signal at all. On the other hand, the interfering signal is not removed totally. Some portion of the interference energy always remains after this kind of cancellation procedure.
5.2. CONVENTIONAL NARROWBAND INTERFERENCE REJECTION

**Figure 5.2:** Frequency domain NBI cancellator. The received signal (which consists of the spread spectrum signal, $r(t)$, and narrowband interference, $\Upsilon_{\text{ext}}(t)$) is first Fourier-transformed (FT). Then the NBI cancellation is performed in frequency domain. Finally before the detection of users in the system, the modified signal is inverse-transformed (IFT).

**Figure 5.3:** NBI cancellation in frequency domain. The received signal is processed in frequency domain by cutting off the parts of the power density spectrum which exceed some predefined power level. This attenuates the narrowband interference signal that has its all energy concentrated near one frequency.
5.3 External Interference Sources and ICA

If some external interference signals are present in the system, they can be naturally assumed to be statistically independent from desired signal components. Consequently, ICA considers them as additional source components and inherently tends to separate them as well. Moreover, since ICA is a time-domain method, it does not make difference between narrow- and wideband interferers. Understandably, possible interference components can not be identified after separation, since known code sequences are not associated with them. Nevertheless, identification of undesired interference components is not necessary. Matter of substance is that the desired components, i.e., the components due to users’ data, are not affected by external interference in ICA assisted receiver structures. Our results in Publications [P1–P3, P6–P8] show that all the ICA assisted receiver structures proposed in this dissertation cope with external interference as such.

Two rather similar receiver structures that are particularly designed for external interference suppression in DS-CDMA downlink is proposed in [9, 97]. These structures assume an use of antenna array to obtain ICA model and consider a transmitted multiuser signal as one source component and the external interference signal as other. Notice, that an input of the $j$-th antenna, say $u_j$, is basically a noisy linear combination of downlink multiuser signal, say $r_{\text{dl}}(t)$, and the interference signal, that is,

$$ u_j(t) = \alpha_j r_{\text{dl}}(t) + \beta_j \Upsilon_{\text{ext}}(t) + \eta(t) \quad (5.8) $$

with $\alpha_j, \beta_j \in \mathbb{C}$. The receivers then use ICA to separate the interference signal and the multiuser signal before detection of the desired user’s symbols. Thus, in these receivers ICA is not applied to separate the CDMA users from each other, i.e., ICA does not reduce multiaccess interference. The detection itself is proposed to be performed using some standard DS-CDMA detector. A selection of the correct source signal prior to detections is performed either with help of pilot information assumed to be embedded in the original multiuser signal or using known information of spreading codes of the desired users. The main principle of such receiver is illustrated in Fig. 5.4.
5.3. **EXTERNAL INTERFERENCE SOURCES AND ICA**

Figure 5.4: ICA based interference mitigation structure which use an antenna array of $J$ antennas. $r_p(t)$ stands for a pilot signal which is used to select the desired source component after separation. Here, the conventional single-user detector (or RAKE) is used for the final DS-CDMA detection. Basically, RAKE can be replaced with any detection method.
Chapter 6

Summary of Publications

The second part of this dissertation is a compilation of eleven original research publications reporting the novel scientific contributions of this dissertation work in detailed manner. We have referred to these publication as [P1–P11] throughout the dissertation. In this chapter, we summarizes the content of the publications briefly (Section 6.1) and describe the author’s contribution in them (Section 6.2).

6.1 Overview of Contents

Publication [P1] introduces the idea of employing BSS/ICA in DS-CDMA multiuser detection. It evaluates the effectiveness of the simple BSS/ICA assisted linear receiver against conventional non-linear interference cancellation schemes. The systems with and without additional external interference are considered. Numerical experiments indicate a clear performance gain in comparison to SIC and roughly the equal performance as PIC with few cancellation stages when no jamming was present. However, even with a moderate level of additional jamming, BSS/ICA out-performs both the non-linear schemes. The results thus show how the blindness of BSS/ICA, even when used in the simple linear receiver, is valuable in making the uplink reception robust against intentional/unintentional out-of-cell interference.

Publication [P2] extends the performance evaluation of the simple linear BSS/ICA receiver in DS-CDMA uplink. Especially, the effects of multipaths,
related channel estimation errors and additional external interference are examined and compared to conventional SIC and $P$-stage PIC schemes. The paper demonstrates by numerical experiments that benefits of BSS are emphasized when a state of propagation channel is not known perfectly. Moreover, the results show how robust BSS/ICA methods are in a fixed multipath scenario. As a matter of fact, the performance of BSS/ICA-receiver is greatly improved when the number of paths grows.

Publication [P3] adapts an ideology of successive interference cancellation to the BSS/ICA assisted uplink reception. The effectiveness of the new method is evaluated against the existing linear BSS/ICA assisted reception and conventional non-linear interference cancellation schemes (SIC and PIC). The systems with and without additional external interference are considered also in this paper. Numerical experiments indicate a clear performance gain in comparison to reference methods. Above all, the successive BSS/ICA method outperforms the conventional schemes clearly also when the basic linear BSS/ICA method barely outdo them. All this is achieved even though ICA processing was terminated with a very mild condition, resulting in low computational load.

Publication [P4] proposes a modification of the BSS/ICA-SIC–type receiver, which, unlike standard BSS/ICA, is able to operate in highly loaded systems. The paper shows the new receiver’s ability to drastically enhance the number of users a system can support, given a certain bit-error-rate requirement, compared to that of conventional PIC and SIC schemes and linear BSS/ICA-alone receiver. The paper also shows how the proposed scheme relaxes the inherent weakness of standard BSS/ICA. Namely, the inadequateness in highly loaded systems due to the “more sources than observations”-problem, which is somewhat circumvented in the proposed successive ICA receiver.

In Publication [P5], the BSS/ICA-SIC–type receiver structure where the users’ delays are simultaneously tracked is introduced and evaluated via numerical examples. The main finding was that SIC-ideology combined with ICA for oversaturated systems is notably beneficial given that certain key parameters like users’ delays are tracked simultaneously. This is due to fact that the delay tracking is performed after the blind separation, which basically outputs interference mitigated signals. Consequently, the delay estimates and, especially, the multiaccess interference estimates to be subtracted, are more precise. The example cases assumed $m = 2C = 62$ observations from the mixture of $n = 3K = 66 \rightarrow 90$ sources. Hence, all examples have more sources than observations.

the paper deals with the advanced BSS/ICA assisted interference cancellation strategies applied in DS-CDMA uplink receivers. The considered receiver structures combine the main benefits of pure BSS/ICA and SIC methods: (i) inherent mitigation of various types of interference sources by BSS/ICA, (ii) robustness against parameter estimation errors due to BSS/ICA, (iii) greatly improved interference suppression capability due to novel combination of SIC ideology and ICA, especially, in the challenging case of highly loaded systems. Numerical experiments with DS-CDMA uplink data are given to illustrate the achieved gains in user capacity and bit-error-rate performance compared to conventional SIC and PIC schemes as well as more an advanced LMMSE-PIC receiver and variants of BSS/ICA techniques alone.

Publications [P7] gives a brief survey of ICA/BSS assisted DS-CDMA methods in the both uplink and downlink reception. In the uplink, the paper present a single antenna interference cancellation scheme introduced basically in Publications [P3–P5]. For the downlink case, the paper look at an antenna array scheme [96,97], where ICA acts as a pre-filter to conventional detection, separating the information component from the interference component and thus providing an interference-free signal to conventional receivers.

Publication [P8] considers the BSS/ICA-SIC receivers in a DS-CDMA reception with multi-path propagation. Thus this paper is an extension to Publications [P5, P6]. Numerical experiments reveals the robustness of the proposed scheme against the delay estimation errors which are known to be one of the expected imperfections at the DS-CDMA receivers.

Publications [P9] and [P11] considers the noisy ICA problem generally, that is, they do not address DS-CDMA detection directly. However, the results can be applied when analyzing the performance of the ICA assisted receiver structures as discussed in Section 3.4. The former publication, [P9], illustrates that basic ICA designed for noise-free linear models is able to provide essentially the best possible output SINR among all linear transformations of received data, in the challenging case of having both additive noise and interference disturbing the desired signal observation in a multi-antenna receiver context. Thus in effect, the ICA is able to do joint diversity reception and interference cancellation in a blind manner, such that the output SINR is maximized. In particular, the experiments indicates that one of the most widely applied ICA algorithms, EASI algorithm, is, in practice, identical with SINR maximizing generalized eigenfilter (M-GEF) in terms of SINR. In theory, EASI can not attain exactly the M-GEF bound when both noise and interference are present, but difference was negligible (< 0.1 dB-unit) in all of the experiments. The paper also shows that, in an important special case of single-source system (i.e., one user and additive
noise), the EASI algorithm provides precisely the greatest linear diversity gain blindly, i.e., performs as a blind maximal ratio combiner (MRC).

Publication [P10] gives an example of taking advantage of the findings in [P9, P11]. In the paper, a novel DSP method for mitigating the dynamic offset interference due to RF blocker self-mixing in direct conversion diversity receivers is proposed. The proposed technique is stemming from modeling the dynamic offset interference as a linear signal mixture model, and then applying ICA on the set of observed signal vectors. Computer simulations are used to assess the achievable compensation performance, and based on the obtained results, the offset interference can be efficiently mitigated using the ICA based approach, while also simultaneously obtaining diversity gain against ordinary channel noise. Based on the simulated interference tolerance, the paper also discusses through an RF dimensioning example, how the ICA based compensator can relax the RF filtering requirements for out-of-band blocker signals in practical receiver design, when compared to ordinary MRC processing. The other main advantage of ICA based method over the other existing techniques is that the proposed method in this paper does not need the channel state information, and thus channel estimation can be basically avoided.

Publication [P11] completes the theme of [P9] taking more an analytical approach. The paper shows that conceptually the ICA designed for noise-free linear models is able to solve, blindly and directly, the generalized eigenvalue problem exactly, i.e., to provide the best possible output SINR among all linear transformations of observed data. However, ICA algorithms constraining the estimated de-mixing matrix to be orthogonal (or unitary) can not exactly attain the optimal solution in general, but in a sense they produce an orthogonalized version of the solution. In addition, the paper gives the necessary and sufficient condition under which the stationary point of the EASI algorithm maximizes the linear output SINR, and also proves that, in the special case of interference-free (that is, noise only) system, the EASI algorithm can attain exactly the greatest diversity gain blindly, i.e., performs as a blind maximal ratio combiner (MRC). Further numerical results are also given to show that the performance of the EASI algorithm is remarkably close to the optimal (i.e., the maximal output SINR among all linear transforms of observed data) also in cases in which the above mentioned theoretical optimality condition is not met.
6.2 Author’s Contributions

The author conducted the scientific research work for this thesis mainly at the Department (former Institute) of Communications Engineering, Tampere University of Technology. That part of the work is reported in [P3–P11]. The studies reported in [P1, P2] date back to the time when the author worked at Department of Mathematical Information Technology, University of Jyväskylä. All the work for the thesis has been done under supervision of Prof. Tapani Ristaniemi.

The author was the main contributor for the content of [P1–P6, P8, P9, P11]. In all of these publications, the author was the primary developer of the novel ideas, carried out the theoretical analysis and simulated proposed methods numerically. The author also wrote the bulk of the script for [P1–P4,P6,P9,P11]. The scripts of [P5, P8] were mainly written by Tapani Ristaniemi. (Recall that the author was the main contributor also in these publications in the sense described above.)

In [P7], the author contributed the part dealing with ICA assisted uplink (single-antenna) reception. The rest of the paper, that is, the downlink (multi-antenna) reception related part, was contributed by Karthikesh Raju. In [P10], the author had considerably smaller role. In that paper the author more or less only guided the main contributor, Ali Shahed haghadam, in the ICA related issues.
Chapter 7

Conclusions

In this dissertation, our target was to study possible benefits of using higher order statistics (HOS) in DS-CDMA multiuser detection (MUD) and interference cancellation (IC). We carried out this study by developing new non-linear MUD/IC methods which employ HOS based blind source separation (BSS) and, in particular, independent component analysis (ICA) and comparing their performance to existing methods. The main emphasis was on the asynchronous uplink reception and, especially, on highly loaded systems. From ICA’s point of view, the highly loaded system model equals to challenging model with more sources than observation.

The dissertation consists of two parts. The first (current) part gave an introduction and background knowledge of the research area to which the topic of the dissertation work belongs and surveyed the existing literature on the area. The second part (which follows after Appendixes and Bibliography) is a compilation of Publications [P1–P11] that reports the novel scientific contributions of this dissertation work in detailed manner. In this concluding chapter, we summarize the contribution of the dissertation.

To begin with, we proposed novel ICA assisted successive interference cancellation (SIC) schemes that employ the fundamental procedures of conventional SIC. That is to say, (i) the estimation of interference, (ii) re-generation of interference and finally (iii) subtraction of the re-generated interference from the original data. What makes the schemes advanced compared to conventional SIC is the way of estimating the interference using ICA and a sensitive selec-
CHAPTER 7. CONCLUSIONS

tion of the interference sources to be subtracted. The proposed receiver structures combine the main benefits of BSS/ICA and conventional SIC methods. First, they are capable to mitigate various types of interference sources thanks to ICA. Consequently, they do not cancel out only multiaccess interference as traditional MUD methods but also external interference sources. Second, the proposed structures are very robust against channel parameter estimation errors due to HOS signal processing of BSS/ICA. One should recall that, traditionally, a very precise channel state information is needed in the interference subtractive receivers (like SIC and PIC) to prevent interference enhancement. This is why accurate code acquisition and tracking has been of great importance in these receivers. However, an important property of BSS/ICA is its inherent capability to cope with erroneous timing estimates. We exploited this feature in the proposed receiver structures, to refine tentative timing estimates after the ICA processing before interference re-generation and subtraction such that also the timing refinement procedure takes advantage of interference suppression due to blind ICA separation. Consequently, we were able to relax the requirement of extremely precise code tracking. This kind of relaxation is not inherently possible in traditional interference subtractive receivers. The third benefit of the proposed receiver structures is their good performance in highly loaded system due to novel combination of SIC ideology and ICA. Notice, that this can also be seen as a way to somewhat circumvent the “more sources than sensors”-problem of BSS/ICA.

Noisy models have been another challenge in BSS/ICA area. In most of practical applications, nevertheless, the presence of an additive background noise can not be neglected. Especially in telecommunications, some level of Gaussian noise is always interfering the reception. For this reason, we considered also the noisy ICA problem. We showed that conceptually many basic ICA algorithms designed for noise-free ICA models, can actually provide the best linear source separation solution possible in terms of input-output SINR gain. This seems not to be well-understood in the literature earlier. From CDMA reception perspective, this founding means that the performance of BSS/ICA assisted reception is indeed exceedingly competitive against any suboptimum MUD detector.

Finally, we conclude that HOS signal processing makes it possible to exploit “hidden” statistical features of information carrying signals in a way which is not possible with traditional second-order (SO) methods. For example, in spread spectrum communications, natural statistical independency between narrowband data streams transmitted by different users is lost in wideband domain due to nonideal auto- and cross-correlation properties of spreading sequences.
This results in unavoidable performance lost in conventional SO de-spreading and detection. HOS assisted methods can, however, employ the original independence and avoid performance degradation due to multiple access and, as well, external interference.
Appendix A

Wise FastICA Initialization in Successive ICA-Receiver

In the following “last-minute” numerical experiments, the successive ICA receiver (ICA-SIC) with the wise FastICA initialization is evaluated numerically against an advanced LMMSE and LMMSE-PIC receivers [68]. We initialize FastICA with estimated whitened LMMSE matrix, \( \tilde{M}_0 = C_z^{-1} \tilde{V} \tilde{G} = \tilde{V} \tilde{G} \). Here, \( C_z = I \) is the covariance matrix of whitened data, \( V \) is a whitening matrix and \( \tilde{G} \) is an estimated code matrix which differ from the exact code matrix due to erroneous delay or timing estimation. With this initialization, we actually start the FastICA iterations from some close neighborhood of the actual ICA solution, that is, the exact whitened LMMSE matrix (see Section 3.4), provided that \( \tilde{G} \approx G \). Hence, we can expect to obtain a convergence after a very small number of iterations. Further, as discussed in [P11] and Section 3.4, any ICA algorithm that performs orthogonalization of the basis vectors can not converge exactly to LMMSE solution under noisy model. The original purpose of this orthogonalization is to prevent multiple ICA units to converge to the same ICA basis vector. We presume that, after the wise initialization of FastICA algorithm, the starting point of iterations is such close to point of convergence that the orthogonality constraint is not needed any more. Thus, we modify the original FastICA algorithm slightly by simply removing the orthogonalization in order to provide a possibility of the exact convergence to LMMSE solution. Consequently, we expect clear performance gains compared to results in our
earlier ICA-SIC related publications [P3–P8].

The simulated system model is asynchronous DS-CDMA model (see Eq. (2.8) in Section 2.1) with $K = 30$ users. The users are assumed to be in two equal sized service classes, which is modeled as a power difference of 10 dB between the two user groups. Inside both groups all the users are assigned the same power. Each user is supposed to sent data blocks of $M = 5000$ QPSK symbols. Symbols are spread using Gold codes of length $C = 31$. In ICA-SIC, value $\rho = 0.9$ is used for the identification threshold. All results are based on average bit error rates (BER) over 1000 independent repetitions. BER values are computed with respect to all $K$ users. For estimated LMMSE methods, we have used also the estimated LMMSE matrix instead of the exact one. In addition, we have plotted performance of exact LMMSE methods in some figures below.

Fig. A.1 shows performance of the ICA-SIC receiver as a function of number of iterations in FastICA. The results indicate relatively short convergence time. The best input-output performance gain of the ICA-SIC is obtained, essentially, after 10–25 FastICA iterations depending on SNR level. Comparing this to our earlier results in which we initialized FastICA randomly, the required convergence time has degraded even to one tenth. E.g., in [P6], FastICA converged to adequate performance level only after 110 iterations in extremely highly loaded system, i.e., with $K = 30$ and $C = 31$.

Figs. A.2 and A.3 compare the ICA-SIC receiver to LMMSE and LMMSE-PIC receivers under erroneous timing estimation. The figures shows clear performance gain of ICA-SIC. Most interestingly, ICA-SIC provides better performance also with perfect timing knowledge (that is, with zero error variance) in Fig. A.2. In our earlier results, both LMMSE and LMMSE-PIC outperformed ICA-SIC in that case. This performance improvement on the earlier results can be explained by relaxation of orthogonalization-step of the FastICA algorithm. Hence, the new results support the theoretical founding of [P9, P11] (see also discussion in Section 3.4) that ICA is able to provide the best input-output SINR gain in noisy model.
Figure A.1: BERs as a function of number of FastICA iterations. FastICA is initialized with estimated LMMSE matrix. Users’ estimated delays are assumed to have an error $\delta_k$ with variance $\sigma_\delta^2$ (unit of error corresponds to one chip duration). SNRs are fixed to 10 dB (top) and 15 dB (bottom).
APPENDIX A. WISE FASTICA INITIALIZATION

Figure A.2: BERs as a function of number of variance of delay estimation error. FastICA is initialized with estimated LMMSE matrix. SNR is fixed to 10 dB.

Figure A.3: BERs as a function of number of received SNR. FastICA is initialized with estimated LMMSE matrix. Users’ estimated delays are assumed to have an error $\delta_k$ with variance $\sigma_\delta^2 = 0.025$ (unit of error corresponds to one chip duration).
Appendix B

On Complexity of FastICA Assisted Detection

This appendix is substantially direct citation of M.Sc. thesis of the author [44, Section 6.3.2]. The remark on oversaturated data at an end of the appendix, is added to this PhD dissertation.

Next, we derive the computational complexity of a receiver structure which uses the symmetric FastICA algorithm. Particularly, the needed number of “multiply and add”-operations is considered. The following derivation does not take needed number of iterations into account. The complexity of the receiver structure using the one-unit FastICA algorithm is discussed in [98]. Here, we use conclusions given in that paper to the extent that the symmetric and one-unit algorithm conjoin.

As previously, $C$, $K$ and $M$ denote the length of spreading codes, number of users and number of symbols (per user) in one received data block, respectively. Note, that now $M >> C > K$. The both FastICA algorithms consist partly of preprocessing tasks, namely the dimension reduction and whitening. Typical eigenvalue based methods for dimension reduction have the complexity of $O(C^3)$, which reduces to $O(C^2K)$ if only $K$ eigenvalue is needed like in this application [98]. The whitening, for one, can be performed with the complexity of $O(CK)$, when it is assumed that $K < C$ [98].

For the complexity analysis of symmetric FastICA iterations themselves, we cannot directly apply the complexity obtained in [98]. Anyway, the complexity
in the one-unit case is \(O(KM)\) per iteration. To estimate all components simultaneously, the one-unit iteration is, basically, repeated for all components. This raises the complexity to \(O(K^2M)\). In addition, the tentative estimate vectors should be made orthogonal between each iteration. Applying the Gram–Schmidt orthogonalization algorithm (GS), for instance, has the well-known complexity of \(O(CK^2)\) \[38\]. Thus with GS, the complexity of one FastICA iteration is \(O(K^2M) + O(CK^2)\).

The selection of the right source components should also taken into account here. Recall, that the selection is based on correlations

\[
\theta_{ki} = \frac{\langle g_k|\tilde{g}_i \rangle}{\|g_k\| \|\tilde{g}_i\|} \in \mathbb{C},
\]  

where \(g_k\) and \(\tilde{g}_i\) are essentially \(2C\)-vectors. Hence, computing one correlation has the complexity of \(O(C)\). Further, \(3(K + 1)\) correlations is computed for each \(K\) users. Thus on the whole, the complexity of computing correlations is \(O(K^2C)\).

Altogether, given that \(M >> C > K\), the complexity of the Pure ICA receiver is \(O(CK^2) + O(K^2M)\).

Remark. If we assume oversaturated data \((K > C)\), which is more practical in DS-CDMA systems as seen in this PhD dissertation, only the complexity of the whitening process changes from \(O(CK)\) to \(O(C^2)\) in above analysis. Further, since \(C^2 < CK^2\) (in this case), the final conclusion of the complexity stays same.
Bibliography


77


interference. IEEE Transaction on Vehicular Technology, 33(3):144–155, 
1984.

[117] V. Zarzoso and A.K. Nandi. Exploiting non-gaussianity in blind identifi-
cation and equalisation of mimo fir channels. IEE Proceedings – Vision, 

tion and equalization of mimo fir channels based on second-order statistics 
Part II

Original Publications
COMPARISON OF NONLINEAR INTERFERENCE CANCELLATION AND BLIND SOURCE SEPARATION TECHNIQUES IN THE DS-CDMA UPLINK

by

Toni Huovinen and Tapani Ristaniemi

Comparison of Nonlinear Interference Cancellation and Blind Source Separation Techniques in the DS-CDMA Uplink

Toni Huovinen and Tapani Ristaniemi
Department of Mathematical Information Technology, University of Jyväskylä, P.O.Box 35 (Agora), FIN-40041, University of Jyväskylä, Finland

Abstract—In this paper we compare the performance of conventional nonlinear interference cancellation and blind source separation (BSS) techniques. Namely, serial interference cancellation (SIC), M-stage parallel interference cancellation (PIC) and BSS based on independent component analysis (ICA) are employed in the uplink of Direct-Sequence Code Division Multiple Access (DS-CDMA) system. Existing studies on BSS/ICA in DS-CDMA have mainly concerned blind interference cancellation, which have been argued by the ability of BSS to deal with unknown (interfering) signals while demodulating a specific user. The performance of BSS/ICA techniques compared to conventional single user detection and blind (linear) multiuser detectors are quite well understood.

In this paper we high-light the effectiveness of BSS/ICA against conventional nonlinear interference cancellation schemes. The systems with and without additional external interference, like adjacent channel interference or jamming, are considered. Numerical experiments are give to evaluate the performance in different settings. They indicate clear performance gains in comparison to SIC and PIC with just a few stages when no jamming is present. However, even with a moderate level of additional jamming, BSS/ICA can outperform both the nonlinear schemes. The results thus show how the blindness of BSS/ICA makes it a robust method in the presence of intentional/unintentional out-of-cell interference.

I. INTRODUCTION

Interference cancellation is an important practical problem in direct sequence code division multiple access (DS-CDMA) communication systems, and have been studied extensively in the past [1]. When the spreading codes of all the active users are known, like in the uplink (e.g. mobile to base) direction, successive and parallel interference cancellation schemes (abbreviated SIC and PIC, respectively) [2]–[5] are traditional ways to mitigate the effects of multiple access in the modulation of a single user. In these schemes, the contribution of interfering users are subtracted iteratively from the received signal either in serial or parallel manner, which results in a signal with lower level of interference and thus more reliable demodulation.

Since conventional SIC and PIC relies on matched filters or RAKE receivers, they are expected to have a performance floor due to the residual interference after successive/parallel cancellation stages. This performance floor depends mainly on the power distribution of the users. Moreover, if intentional or unintentional interference (like adjacent channel interference or intentional jamming) is present, conventional detection may fail completely resulting in the failure of PIC/SIC, too.

In this paper we evaluate the potential of blind source separation [6] compared to traditional PIC/SIC schemes. Blind source separation (BSS) has drawn a lot of attention in statistical signal processing and neural network communities. The goal in BSS is to separate signals from the received mixture of signals in completely blind manner. Existing studies on BSS in DS-CDMA have mainly concerned blind interference cancellation (see e.g. [6], [7] and references therein), which have been argued by the ability of BSS to deal with unknown users while demodulating a specific user. In the uplink direction the interfering users within the cell are known in the base station. Hence, it can be questioned whether BSS still can compete with the schemes that utilize the interference explicitly. In the following we illustrate by numerical simulation that remarkable benefits of BSS still exists. Namely, the results thus show how the blindness of BSS makes it a robust method in the presence of intentional/unintentional out-of-cell interference.

II. BLIND SOURCE SEPARATION BY INDEPENDENT COMPONENT ANALYSIS

Independent component analysis (ICA) [6] is a fairly new statistical technique by which BSS can be performed. In ICA a set of observed signals or random variables are tried to express as linear combinations of statistically independent components, which are often called sources or source signals. The ICA problem is blind, because not only the source signals but also the mixing coefficients are unknown.

In standard linear ICA, the \( m \) observed signals \( x_1(t), \ldots, x_m(t) \) at the time instant \( t \) are assumed to be linear combinations of \( n \) unknown but statistically independent source signals \( s_1(t), \ldots, s_n(t) \) at the time \( t \). By introducing the data vector \( x(t) = [x_1(t), \ldots, x_m(t)]^T \) for the observed signals, and the source vector \( s(t) = [s_1(t), \ldots, s_n(t)]^T \) for the source signals, the instantaneous noisy linear ICA mixture model is given by

\[
x(t) = As(t) + n(t)
\]  

(1)

Here the \( m \times n \) unknown but constant mixing matrix \( A \) contains the mixing coefficients, and \( n(t) \) denotes the additive

0-7803-7822-9/03/$17.00 © 2003 IEEE.
noise vector at time $t$. The ICA model (1) is readily a DS-CDMA signal model as shown in [7] for the downlink case.

In ICA, the source signals $s(t)$ are estimated using only the observations $x(t)$ by finding an $n \times m$ unmixing matrix $W$. This matrix should be such that the $n$-vector $Wx(t)$ recovers the set of original sources as well as possible. Because of the blindness of the problem, only the waveforms of the sources can be estimated. For estimating the unmixing (separating) matrix $W$, many different methods have been proposed [6]. Most of these are ICA methods exploiting the statistical independence of the sources, but there exist other approaches which utilize temporal correlations or nonstationarity of the sources. The mutual performance of these methods depends largely on the validity of the assumptions made on them in the problem at hand.

In numerical experiments, we applied the so-called FastICA algorithm for complex mixtures [6, 7]. FastICA is a fast method for performing linear ICA, and its basic form relies on the sample fourth-order statistics kurtonian. However, other forms of the algorithm employing more robust lower-order statistics have been developed [6]. Instead of FastICA, other ICA algorithms developed for complex-valued mixtures could be used.

### III. Signal Model

A standard asynchronous spread spectrum system [8] with direct sequence spreading is assumed with additional out-of-cell interference. This interference, $j_p(t)$, is here modeled as a continuous wave (possibly pulsed at the symbol level) with a frequency, phase offsets and power equal to $\Delta f$, $\theta$ and $J$, respectively:

$$j_p(t) = \delta_p \sqrt{J_0 e^{j(2\pi \Delta f t + \theta)}}$$

Here $\delta_p(t) = 1$ with a probability $p$ during a symbol. Hence, the interference corresponds to a continuous wave when $p = 1$ and pulsed wave at the symbol level otherwise. The above two cases are used as examples of narrowband and wideband interference.

Therefore, the data describing the received signal at the $m$th symbol interval (assuming AWGN channel and uplink direction), have the form

$$r_m(t) = \sum_{k=1}^{K} b_{k,m} a_k s_k(t - mT - d_k T_c) + n(t) + j_p(t),$$

in which the $m$th symbol $b_{k,m}$ is sent by user $k$. The complex coefficient of the $k$th user’s channel is denoted by $a_k$, which is assumed to remain the same during the data block of, say, $M$ symbols. $s_k(t)$ is $k$th user’s binary chip sequence, supported by $[0, T_c]$, where $T$ is the symbol duration. $T_c$ is the chip duration. For notational simplicity, the user delays are assumed to be discretized, and hence $d_k \in \{0, \ldots, C-1\}$, where $C$ is number of chips in spreading code. The delays are assumed to remain constant during the block of $M$ data symbols. $n(t)$ denotes noise.

By chip-matched filtering, and using processing window size of two symbols, we get the sampled data

$$r_m = \sum_{k=1}^{K} a_k (b_{k,m-1} g_k + b_{k,m} g_k + b_{k,m+1} g_k) + n_m$$

$$+ j_p((m-1)T) \mathbf{1} + j_p(mT) \mathbf{1} + j_p((m+1)T) \mathbf{I}$$

(4)

Without loss of generality, it is here assumed that the interference remains constant during a symbol. Here $n_m$ denotes noise vector and the code vectors of length $2C$ are defined as

$$g_k = g_k(d_k) e(kC - d_k C + 1) \ldots S_k C 0^{C-d_k C}$$

$$g_k = g_k(d_k) = \begin{bmatrix} g_k[d_k] \ldots g_k[C] 0^{C-d_k C} \end{bmatrix}^T$$

$$1 = \begin{bmatrix} 1 \ldots 0 \end{bmatrix}^T$$

$$\mathbf{I} = \begin{bmatrix} 0^{C-1} \end{bmatrix}^T$$

With a simple manipulation, we can get a compact representation for the data.

$$r_m \stackrel{def}{=} G b_m + n_m.$$  

The $2C \times 3(K + 1)$ dimensional code matrix $G$ contains the code vectors and path strengths, while the $3(K + 1)$-vector $b_m$ contains the symbols and sampled external interference:

$$G \stackrel{def}{=} \begin{bmatrix} \leq G(d_k) g_k (d_k) g_k (d_k) \ldots \leq \mathbf{1} \end{bmatrix}^T$$

$$b_m \stackrel{def}{=} \begin{bmatrix} \leq b_{k(m-1)} b_{k(m+1)} \ldots \leq j_p((m-1)T) j_p(mT) j_p((m+1)T) \end{bmatrix}^T.$$  

(8)

### IV. Numerical Experiments

The performance of ICA based receiver was tested and compared with the performances of SIC and PIC receivers numerically. Path delays and phases of complex path coefficients were chosen randomly. The moduli of path coefficients were chosen either such that all users had equal received power or such that half of users had 10 dB stronger received power than other half. In latter case signal-to-noise-ratio (SNR) and signal-to-jammer-ratio (SJR) were proportioned to the weaker users. Further, the path delays and coefficients were assumed to be known for all users by the receiver. Number of users in system was $K = 8$ (except in those experiments where $K$ was varied), and each user was spread with Gold Codes [8, 9] of length $C = 31$. One data block of each user contained $M = 5000$ randomly chosen QPSK symbols. Simulations were run with different values of jamming probability $p$. In some simulations value $p = 0$ was used. That way also situations without jammer was concerned.

Symbols of each user were detected using ICA, SIC and PIC methods in the receiver end, and after detection bit-error-rates (BER) for each method were computed with respect to all users. All simulations were repeated independently 5000 times and finally BER values from each repeat were averaged to obtain final values of BERs.
Fig. 1. Bit-error-rates as a function of SNR. The system was a DS-CDMA uplink channel with AWGN and there were $K = 8$ users with equal received powers.

In the ICA receiver number of independent components in model was assumed to be $3(K + 1)$ (three components per each user and, likewise, three components per jammer). Same number of components was used also in those simulations, where the jammer was absent ($p = 0$). Hence, one does not have to know if the jammer really is present or not. After separating the components with the FastICA the right components were selected using knowledge of users’ spreading codes as well as the channel parameters. First, each user’s spreading code was translated by the user’s path delay and multiplied by the user’s path coefficient. Then the correlations between this modified spreading code and each column of mixing matrix got from the FastICA, were computed for each user. Finally, the data component which corresponded to the best correlation was chosen to be the user’s data.

In the SIC receiver the users were first ordered with respect to their received powers, and then number of cancellations performed was $K - 1$ (the maximum number of cancellations). The PIC receiver used the multistage algorithm proposed in [4]. Number of cancellation stages in PIC was ranged from 2 to 5.

Several simulation sets with different variables were run. Used variables were: SNR, SJR, $K$ and near-far-level (power ratio between two user groups). Next each simulation set is dealt with one by one.

A. Varying SNR

Effect of noise was examined in the system without jammer ($p = 0$). SNR was varied from 0 to 30 dB. The results in cases of equal power users and two power groups are shown in Figs. 1 and 2, respectively. The figures show that the ICA receiver outperformed the SIC receiver clearly. On the other hand, the performance of the PIC receivers with more than two stages seems to have been better than the performance of the ICA receiver. The ICA receiver and the PIC receiver with two stages had almost equal performance, especially in the case of two power groups. Further, changing between the two power cases had quite a little effect on the ICA receiver, whereas the effect was greater on the SIC receiver and the PIC receivers with small number of cancellation stages.

B. Varying SJR

The systems with jammer was examined in another set of simulations. In these simulations SNR was fixed to 20 dB while SJR was varied from -30 to 30 dB. The big values of SJR (i.e. cases where the spread signal is stronger than the jammer) was included to correspond the situation where some other jammer suppression method is first used but some part
Probability of bit being jammed was $p = 0.25$.

The results of simulations in cases of equal power users with $p = 0.75$ and two power groups with $p = 0.25$ are shown in Figs. 3 and 4, respectively. (Figures of results with $p = 0.25$ in the former case and with $p = 0.75$ in the latter case are excluded, because they were very similar to the shown figures.)

The figures show that the ICA receiver is superior to the SIC and PIC receivers in the SJR range from -30 to 5 dB in the equal power case and in the range from -20 to 15 dB in the case of two power groups. This is, naturally, understandable, since the SIC and PIC methods do not take account the presence of jammer signal. The performance of ICA receiver was also very stable in these SJR ranges and beyond them as well.

However, the performance of ICA receiver is getting worse when the jammer is strong enough in proportion to spread signals. This can be seen especially from Fig. 4. The phenomenon can be explained with the whitening process performed before the FastICA algorithm in the ICA receiver. The whitening is performed with the principal component analysis (PCA) which, among others, reduce the dimension of received data. Now, if the components in system (i.e. the users and the jammer) have very different powers, PCA may lose some weak components during this dimension reduction [6], [10]. The distribution of correct bits in one block of 5000 symbols (10000 bits) is shown in figure 5. The system behind the figure is same as in Fig. 4 but now the SJR is fixed to be -30 dB. The distributions in the cases of ICA and SIC receivers are depicted. The figure discovers the duality of distribution in case of ICA receiver. Some minority of blocks had only roughly a half of their bits detected correctly, whereas majority of blocks (practically all the other blocks) had almost all bits detected correctly. On the other hand, the distribution in the case of SIC receiver has not this kind of clear duality although the performances of both receivers were at the same level. The duality of distribution supports the theory that some components have been lost by PCA when SJR is negative enough. Several algorithms have been proposed to improve the performance of PCA in case of components with unequal strengths [10]. Replacing PCA with some of them in the ICA receiver will be one subject of our further research.

C. Varying number of users

Number of users in system ($K$) was also varied in one set of simulations. Jammer was absent in these simulations ($p = 0$). SNR was again fixed to 20 dB and $K$ was varied between 4 and 18. Again both cases, the one with equal power users and the one with two power groups, was simulated. Figs. 6 and 7 show the results, respectively. The figures show that the ICA receiver outperformed the SIC receiver in the both cases. For instance, at the BER level of $10^{-3}$ the ICA receiver offered about 8 users better capacity than the SIC receiver in the system with the equal power users. In the system with two power groups the capacity improvement was, yet, about 6 users. The PIC receivers, for one, outperformed the ICA.
The results are shown in Fig. 8. The figure shows that the performance of ICA receiver was keeping between the performances of PIC receivers and the performance of SIC receiver also with other values of near-far-level than 10 dB in the system without jammer. However, the ICA receiver was more robust against changes of near-far-level than e.g. the SIC receiver or the PIC receiver with two stages. The decreasing behaviour of BER in the case of SIC receiver and the increasing behaviour of BER in the case of PIC receiver with two stages equals to the simulation results by Patel and Holtzman in [11]. Nevertheless, one interesting point is that the performance of PIC receiver was actually getting better as near-far-level was getting greater when the number of cancellation stages was more than three.

V. CONCLUSIONS

In this paper we evaluated the effectiveness of BSS/ICA against conventional nonlinear interference cancellation schemes. The systems with and without additional external interference were considered. Numerical experiments indicated a clear performance gain in comparison to SIC and PIC with few cancellation stages when no jamming was present. However, even with a moderate level of additional jamming, BSS/ICA outperformed both the nonlinear schemes. The results thus showed how the blindness of BSS/ICA is valuable in making the uplink receiver robust against intentional/unintentional out-of-cell interference.

REFERENCES

EFFECT OF CHANNEL ESTIMATION AND MULTIPATH ON INTERFERENCE CANCELLATION EMPLOYING BLIND SOURCE SEPARATION IN THE DS-CDMA UPLINK

by

Toni Huovinen and Tapani Ristaniemi

Effect of Channel Estimation and Multipath on Interference Cancellation Employing Blind Source Separation in the DS-CDMA Uplink

Toni Huovinen
Department of Mathematical Information Technology
University of Jyväskylä
P.O.Box 35 (Agora), FIN-40041, Univ. of Jyväskylä, Finland
Email: toni@cc.jyu.fi

Tapani Ristaniemi
Institute of Communications Engineering
Tampere University of Technology
P.O.Box 553, FIN-33101, Tampere, Finland
Email: Tapani.Ristaniemi@tut.fi

Abstract—In this paper we consider DS-CDMA uplink interference cancellation based on blind source separation (BSS) technique. While BSS assisted interference suppression compared to linear receivers is quite well reported, a comparison of BSS with nonlinear receivers is overlooked in the literature. [11] was the first one to make this comparison in AWGN channel with perfect channel information. In this paper we make further evaluation. Especially, the effects of multipaths, related channel estimation errors and additional external interference (like adjacent channel interference or intentional jamming) are examined and compared to conventional successive interference cancellation (SIC) and M-stage parallel interference cancellation (PIC) schemes. We demonstrate by numerical experiments that benefits of using BSS exists. Moreover, results indicate that the performance gain compared to SIC and PIC is even greater in multipath scenario and in the presence of channel estimation imperfections.

I. INTRODUCTION

Concept of interference cancellation in Direct-Sequence Code Division Multiple Access (DS-CDMA) systems is studied a lot in the past [1]. One relatively new idea is to employ blind source separation (BSS) techniques [2]. What makes BSS techniques attractive is their ability to separate signals from a mixture of original source signals in a blind manner i.e. without explicit knowledge of signals waveforms. Consequently, both internal interference due to multiple access and out-of-cell interferences (intentional and unintentional) can be mitigated in DS-CDMA systems. In addition, typical BSS methods like independent component analysis (ICA) relies solely on higher-order statistical properties of data, which makes the methods robust against the problems due to near-far scenario and incomplete cross-correlation properties of the users. Applications of BSS have been found e.g. in MIMO systems [3], I/Q processing receivers [4], DS-CDMA blind multi-user detection [5] and DS-CDMA out-of-cell interference cancellation [6].

Successive and parallel interference cancellation schemes (abbreviated SIC and PIC, respectively) [7]–[9] are traditional ways to mitigate the effects of multiple access in the uplink reception. These suboptimal multiuser detection (MUD) solutions are often considered as reasonable alternatives to optimal MUD [10], since they have essentially lower computational complexity. In SIC and PIC, the contribution of interfering users are subtracted iteratively from the received signal either in serial or parallel manner, which results in a signal with lower level of interference and thus more reliable demodulation.

Conventional SIC and PIC are expected to have a performance floor due to the residual interference after successive/parallel cancellation stages, since they relies on matched filters (MF) or RAKE receivers. The performance floor depends mainly on the power distribution of the users. In addition, the state of channel should be known in order to re-build interfering signals properly for the subtractions. Possible errors in channel estimates even strengthen the residual interference and, thus, impair a performance of the receivers.

Also typical BSS assisted interference cancellation methods need channel estimation. Though, channel estimates (path delays and coefficients) are not used in a blind source separation task itself, but in either pre- or post-processing of the separation. For instance, BSS assisted single user detector in [5] uses some conventional detection method like RAKE or MMSE receiver to obtain initial values for BSS block. In BSS based uplink receiver in [11], channel estimates are needed after blind source separation in user identification block. Hence, it is important to learn how different kind of imperfections in channel estimation affects on the performance of BSS based methods.

Benefits of BSS assisted interference suppression techniques compared to conventional single user detection and blind (linear) multiuser detectors are quite well understood (see e.g. [5] and references therein). In addition, recent studies had also indicated that BSS techniques can provide clear performance gains compared SIC and M-stage PIC schemes in an AWGN channel with perfect knowledge of the state of the channel [11]. We show here that noticeable gains still exists in a multipath channel, and that BSS techniques have better tolerance against imperfect channel estimation than SIC/PIC schemes. In addition, remarkable gains are achieved in the presence of additional external interference.
II. SIGNAL MODEL

Studies channel is DS-CDMA uplink with fixed multipaths and additive white gaussian noise (AWGN) [13]. In addition, channel is assumed to contain an external interference signal \( j_p(t) \), which is either pulsed at the symbol level with probability \( p \) or a continuous wave \( (p = 1) \). Hence, the received signal has the form

\[
r(t) = \sum_{k=1}^{K} \sum_{m=1}^{M} b_{km} \sum_{l=1}^{L} a_{kl}s_k(t - mT - d_{kl}) + n(t) + j_p(t),
\]

in which \( b_{km} \) is the \( m \)th symbol sent by user \( k \in \{1, \ldots, K\} \). Each symbol travels through \( L \) paths. The complex coefficients and delays of the \( k \)th user’s \( l \)th path is denoted by \( a_{kl} \) and \( d_{kl} \in [0, T] \), respectively. They are assumed to remain the same during the data block of \( M \) symbols. \( s_k(\cdot) \) is \( k \)th user’s binary chip sequence, supported by \([0, T]\), where \( T \) is the symbol duration. \( n(t) \) denotes noise.

The sampled data, which is obtained by chip-matched filtering, and using processing window size of two symbols, can be written as [5], [14]

\[
r_m(t) = \sum_{k=1}^{K} \sum_{l=1}^{L} a_{kl}(b_{km+1}g_{kl} + b_{km}g_{kl} + b_{km+1}g_{kl}) + n_m + j_p((m - 1)T)I + j_p(mT)I + j_p((m + 1)T)I
\]

Without loss of generality, it is assumed that the interference remains constant during a symbol. Here \( n_m \) denotes noise vector and code vectors of length \( 2C \) are defined as

\[
\begin{align*}
\mathbf{g}_{kl} & \equiv (1 - |\delta_{kl}|) \mathbf{e}_{kl}(\hat{d}_{kl}) + |\delta_{kl}| \mathbf{e}_{kl}(\hat{d}_{kl} + \text{sign}(\delta_{kl})) \\
\mathbf{e}_{kl} & \equiv (1 - |\delta_{kl}|) \mathbf{c}_{kl}(\hat{d}_{kl}) + |\delta_{kl}| \mathbf{c}_{kl}(\hat{d}_{kl} + \text{sign}(\delta_{kl})) \\
\mathbf{\mathcal{G}}_{kl} & \equiv (1 - |\delta_{kl}|) \mathbf{c}_{kl}(\hat{d}_{kl}) + |\delta_{kl}| \mathbf{c}_{kl}(\hat{d}_{kl} + \text{sign}(\delta_{kl}))
\end{align*}
\]

\[
\begin{align*}
\mathbf{I} & = [1 \ldots 1 \ 0_{2C - d_k}^{T}]^{T} \\
\mathbf{T} & = [0_{T_{d_k}}^{T} \ldots 0_{T_{C - d_k}}^{T}]^{T}
\end{align*}
\]

where \( \hat{d}_{kl} \in \{0, 1, \ldots, C\} \) is a delay estimate of \( k \)th user’s \( l \)th path, \( \hat{d}_{kl} = d_{kl} - d_{kl} \) is an estimation error in delay and

\[
\begin{align*}
\mathbf{e}_{kl}(d) & \equiv [s_k[C - d + 1] \ldots s_k[C] \ 0_{C - d}^{T}]^{T} \\
\mathbf{c}_{kl}(d) & \equiv [0_{C}^{T} \ s_k[1] \ldots s_k[C] \ 0_{C - d}^{T}]^{T} \\
\mathbf{\mathcal{G}}_{kl}(d) & \equiv [0_{C}^{T} \ s_k[1] \ldots s_k[C - d]]^{T}
\end{align*}
\]

In addition, the delay estimation error is assumed to be smaller than one chip duration.

With a simple manipulation, we can get a compact representation for the data,

\[
r_m = \mathbf{G}b_m + n_m.
\]

The \( 2C \times 3(K+1) \) dimensional code matrix \( \mathbf{G} \) contains the code vectors and path strengths, while the \( 3(K+1) \)-vector \( \mathbf{b}_m \) contains the symbols and sampled external interference:

\[
\begin{align*}
\mathbf{G} & \equiv \begin{bmatrix}
\cdots \sum_{l=1}^{L} a_{kl}g_{kl} & \sum_{l=1}^{L} a_{kl}g_{kl} & \sum_{l=1}^{L} a_{kl}g_{kl} \cdots 11^T
\end{bmatrix} \quad (6) \\
\mathbf{b}_m & \equiv \begin{bmatrix}
\cdots b_{k(m-1)}b_{km}b_{k(m+1)} \cdots \\
\cdots j_p((m - 1)T)j_p(mT)j_p((m + 1)T)^T
\end{bmatrix} \quad (7)
\end{align*}
\]

Notice that \( m \)th symbols are included in three successive vectors \( \mathbf{b}_{m-1}, \mathbf{b}_m \) and \( \mathbf{b}_{m+1} \). Hence, we say that these vectors are “early”, “middle” and “late” parts of \( m \)th symbols, respectively.

III. INDEPENDENT COMPONENT ANALYSIS

Independent component analysis (ICA) [2] is a statistical BSS-method which have received lot of attention recently. In ICA a goal is to express a set of observed random variables or signals as linear combinations of statistically independent components, which are often called sources or source signals. The ICA problem is blind, because not only the source signals but also the mixing coefficients are unknown.

In standard linear ICA, the \( m \) observed signals \( x_1(t), \ldots, x_m(t) \) at the time instant \( t \) are assumed to be linear combinations of \( n \) \( (n \leq m) \) unknown but statistically independent source signals \( s_1(t), \ldots, s_n(t) \) at the time \( t \). By introducing the data vector \( \mathbf{x}(t) = [x_1(t), \ldots, x_m(t)]^{T} \) for the observed signals, and the source vector \( \mathbf{s}(t) = [s_1(t), \ldots, s_n(t)]^{T} \) for the source signals, the instantaneous noisy linear ICA mixture model is given by

\[
\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)
\]

Here the \( m \times n \) unknown but constant mixing matrix \( \mathbf{A} \) contains the mixing coefficients, and \( \mathbf{n}(t) \) denotes the additive noise vector at time \( t \). The DS-CDMA signal model (5) is readily a ICA model with \( m = 2C \) observations of \( n = 3(K+1) \) source components as shown in [5] for the downlink case.

In ICA, the source signals \( \mathbf{s}(t) \) are estimated using only the observations \( \mathbf{x}(t) \) by finding an \( m \times n \) unmixing matrix \( \mathbf{W} \). This matrix should be such that the \( n \)-vector \( \mathbf{W}\mathbf{x}(t) \) recovers the set of original sources as well as possible. Because of the blindness of the problem, only the waveforms of the sources can be estimated. For estimating the unmixing (separating) matrix \( \mathbf{W} \), many different methods have been proposed [2]. Most of these are ICA methods exploiting the statistical independence of the sources, but there exist other approaches which utilize temporal correlations or nonstationarity of the sources. The mutual performance of these methods depends largely on the validity of the assumptions made on them in the problem at hand.

IV. BSS/ICA BASED RECEPTION

A BSS/ICA-assisted uplink receiver was recently considered in [11]. The receiver separates all the users simultaneously using FastICA algorithm for complex mixtures [12]. FastICA is a fast method for performing linear ICA, and its basic form relies on the sample fourth-order statistics kurtosis. However,
other forms of the algorithm employing more robust lower-order statistics have been developed [2].

The iteration algorithm of symmetric FastICA for complex valued data is the following:

1) Center the original data (i.e. make its mean zero).
2) Whiten the centered data as follows: \( \mathbf{y}_m = \mathbf{T} \mathbf{r}_m \), such that \( E[\mathbf{y}_m \mathbf{y}_m^H] \) equals the identity matrix. Such \( \mathbf{T} \) can be simply found in many ways [2].
3) Take a random \( 2C \times 3(K + 1) \) matrix \( \mathbf{W} = [\mathbf{w}_1 \ldots \mathbf{w}_3(K + 1)] \).
4) Orthogonalize matrix \( \mathbf{W} \) as in step 5 and normalize its columns.
5) FastICA: For all \( i = 1, \ldots, 3(K + 1) \), let
   \[
   \mathbf{w}_i \leftarrow E[\mathbf{y}_m (\mathbf{w}_i^H \mathbf{y}_m)^* | \mathbf{w}_i^H \mathbf{y}_m |^2] - 2 \mathbf{w}_i
   \]
6) Orthogonalization: \( \mathbf{W} \leftarrow \mathbf{W} (\mathbf{W}^H \mathbf{W})^{-\frac{1}{2}} \)
7) Normalize columns of matrix \( \mathbf{W} \).
8) Go to FastICA-step until converged.

Above \( z^* \) denotes a complex conjugate of complex number \( z \) and \( \mathbf{M}^H \) stands for a Hermitian (i.e. an element wise complex conjugate and transpose) of matrix \( \mathbf{M} \).

After FastICA-separation, users are identified by their spreading codes. Hence, the BSS/ICA-assisted receiver is semi-blind. Strictly speaking, each user’s spreading code is translated by the user’s path delay and multiplied by the user’s path coefficient. Then the correlations between this re-built spreading code, say \( \tilde{g}_k \), and a column of the mixing matrix got from the ICA, say \( \mathbf{w}_k \), are computed for each user. That is to say, \( \rho(k, k') = ||\tilde{g}_k \mathbf{w}_k^H|| / ||\tilde{g}_k|| ||\mathbf{w}_k|| \) is computed for each \( k, k' \). Finally, the source which corresponds to the best correlation is chosen to be that user’s data.

Notice, that only user identification block needs a knowledge of the state of propagation channel in BSS/ICA-assisted receiver. Thus, possible imperfections in channel estimation do not affect on quality of separated components themselves, but only on capability to choose components that corresponds to users. On the contrary, conventional receivers, like MF- or RAKE-based receivers, suffer from channel estimation errors directly. Further, the effect of errors is expected to be emphasized in subtractive SIC and PIC receivers, since the residual interference after successive/parallel cancellation stages strengthens due to estimations errors.

V. NUMERICAL EXPERIMENTS

Effect of errors in channel estimation and multipath propagation model on the performance of BSS/ICA based receiver was studied with numerical experiments. A system with \( K = 8 \) users and Gold codes of length \( C = 31 \) were considered. Each user was supposed to send a data block of \( M = 5000 \) QPSK symbols. All results are based on average bit error rates (BER) over 5000 independent repetitions. BER values are computed with respect to all \( K \) users.

Path delays and phases were chosen randomly. The modulus of the path coefficients were chosen such that either all the users had equal received power or half of the users had 10 dB stronger received power than other half. In the latter case, signal-to-noise-ratio (SNR) and signal-to-jammer-ratio (SJR) were proportioned to a single user in the weaker users’ group. Further, the path delays and coefficients were assumed to be estimated for all users and propagation paths at the receiver’s end. Channel estimates included a normally distributed error with a fixed variance. The probability of external jamming was either \( p = 0.5 \) or \( p = 0 \) on the symbol level.

A. Effect of channel estimation error

The results show that ICA-based receiver is more immune to channel estimation errors than SIC and PIC receivers, even when out-of-cell interference sources are absent. The effect of error in delay estimation in the case of one path and two power groups is depicted in Fig. 1. The figure shows that ICA-based receiver outperforms SIC and PIC receivers clearly when the variance of error in delay estimation is beyond 0.05 (note that the duration of a chip is normalized to 1). For instance, if variance is order of 0.1, only ICA-based receiver is able to provide bit-error-rate lower than \( 10^{-2} \).

The experiments show also that ICA-based receiver resists estimation errors in channel coefficients better than SIC and PIC. This is illustrated in Fig. 2 and 3. In addition, the latter figure shows how multistage-PIC receivers totally fail to improve the performance of conventional MF (RAKE) receiver when estimation error in modulus increases.

B. Effect of external jamming

Adding an out-of-cell interference signal into the system reveals remarkable benefits of BSS/ICA. ICA-based receiver does not suffer from external interference at all, while performances of SIC and PIC receivers collapse even with a moderate level of interference. Fig. 4 shows an example of a one path system with out-of-cell interference. The variance
of error in delay estimates is 0.1. The figure shows how ICA-based receiver is superior to SIC and PIC receivers in whole SJR range from -10 to 25 dB. Whereas same experiment with ideal channel estimation showed that PIC receivers outperformed ICA-based receiver when SJR exceeded 15 dB [11].

Stability of ICA-based receiver against external interference can be also seen in Fig. 1. Compared to Fig. 5, ICA-based receiver behaves practically similarly in system with and without external jammer when variance of delay error is varied. Reference methods, for one, fail in the whole variance range. Similar results were obtained also when other channel estimation errors were considered.

C. Effect of multipaths

We also found out, that number of propagation paths does not impair the performance of ICA-based receiver. In fact, the performance gets better when number of paths increases. This can be explained by increasing number of different path delays. Namely, ICA assumes that rank of mixing equals to number of independent components i.e. that mixing matrix has a full rank. This means here that for code matrix $G$ (see Eq. 7), $\text{rank}(G) = 3(K + 1)$. Nevertheless, DS-CDMA model itself does not ensure that the rank assumption holds literally. For example, if two users happens to have path delay one in one-path system, they both have “late” code vector $g$ (see Eq. 3) such that only the last element of it is non-zero. It immediately follows that two columns of $G$ are mutually linearly independent i.e. that $G$ can not have full rank. However, when the number of paths increases, matrix $G$, loosely speaking, becomes more random due to different delays between paths and summations of code vectors ($g_{kl}$, $g_{kl}^1$ or $g_{kl}^2$). Consequently, the rank condition holds more likely resulting in more reliable symbol estimation.

In case of ideal channel estimation, ICA based receiver outperforms SIC receiver and PIC receiver with few cancellation stages, when the number of paths ranges from 1 to 9. PIC receivers with more cancellation stages, for one, have better performance than ICA receiver in this case. However, ICA receiver seems to resist channel estimation errors better than SIC and PIC receiver also in multipath case. This can be seen from Fig. 6 in which BER as function of number of paths is shown. The variance of error is 0.05 in all the parameters (path delay, moduli and phase of the path coefficient).
VI. CONCLUSION

In this paper we evaluated the performance of BSS/ICA-assisted receiver in DS-CDMA uplink. Systems with and without external interference sources (like adjacent channel interference or intentional jamming) were considered. Especially, the effects of multipaths, related channel estimation errors and additional external interference were examined and compared to conventional serial interference cancellation (SIC) and M-stage parallel interference cancellation (PIC) schemes. We highlighted by numerical experiments that benefits of BSS are emphasized when a state of propagation channel is not known perfectly. Moreover, the results showed how robust BSS/ICA methods are in a fixed multipath scenario. As a matter of fact, the performance of BSS/ICA-receiver is greatly improved when the number of paths grows.

REFERENCES

BLIND SOURCE SEPARATION BASED SUCCESSIVE INTERFERENCE CANCELLATION IN THE DS-CDMA UPLINK

by

Toni Huovinen and Tapani Ristaniemi

Recent studies have found interference cancellation (IC) schemes employing blind source separation (BSS) methods to be promising competitors to conventional IC schemes in DS-CDMA reception, since they are able to mitigate both internal multiaccess interference and external interference. In this paper we introduce a new successive BSS based IC scheme for DS-CDMA uplink. The new scheme combines the benefits of existing BSS based schemes and conventional successive interference cancellation (SIC). Numerical experiments are given to illustrate the superiority of the successive BSS scheme compared to existing BSS schemes and conventional successive and parallel interference cancellation (PIC) schemes. A system with and without additional external interference, like adjacent channel interference or jamming, are considered.

1. INTRODUCTION

Interference cancellation in Direct-Sequence Code Division Multiple Access (DS-CDMA) system has been studied a lot in past [1]. Successive and parallel interference cancellation (abbreviated SIC and PIC, respectively) [2]-[5] are conventional methods to mitigate effects of multiple access in uplink (e.g. mobile to base), where spreading codes are known for all the active users. In these nonlinear methods, the contribution of interfering users are subtracted iteratively from the received signal either in serial or parallel manner, which results in a signal with lower level of interference and thus more reliable demodulation. However, conventional SIC and PIC are quite sensitive to external interference sources (like adjacent channel interference or intentional jamming), since they both are based on matched filters or RAKE receivers.

One relatively new idea is to employ blind source separation (BSS) techniques [6] on interference cancellation. What makes BSS techniques attractive is their ability to separate signals from a mixture of original source signals in a blind manner i.e. without explicit knowledge of signals’ wave forms. Consequently, both internal interference due to multiple access and external interference sources can be mitigated in DS-CDMA systems. Benefits of BSS-assisted interference suppression techniques compared to conventional single user detection and blind (linear) multuser detectors are quite well understood (see e.g. [7] and references therein). In addition, recent studies had also indicated that BSS techniques can provide clear performance gains compared to SIC and PIC in DS-CDMA uplink, especially in the presence of external interference [8], [9].

In this paper a new interference cancellation strategy is introduced, which combines the concepts of blind source separation and successive interference cancellation. Namely, each user is first detected by the means of BSS after which the contribution of that user is subtracted from the received signal. In addition, we utilize a novel strategy during BSS, which is motivated by the fact that BSS is able to detect all the consecutive symbols within the processing window independently. Considering each users’ signals, this means that BSS estimates also the delayed symbols within a processing window. Therefore, a maximal ratio combining of them (with appropriate delaying) is proposed to get better SNR for the hard decision. The potential of the new receiver structure is evaluated with respect to existing BSS receivers, SIC and PIC in DS-CDMA uplink.

2. SIGNAL MODEL

A DS-CDMA uplink channel with additive white gaussian noise (AWGN) [10] is studied. In addition, channel is assumed to contain an external interference signal \( j_p(t) \), which is modeled as a continuous wave (possibly pulsed at the symbol level) with a frequency, phase offsets and power equal to \( \Delta f, \theta \) and \( J \) respectively:

\[
j_p(t) = \delta_p \sqrt{J} e^{j(2\pi \Delta f t + \theta)}
\]

Here \( \delta_p(t) = 1 \) with a probability \( p \) during a symbol. Hence, the interference corresponds to a continuous wave when \( p = 1 \) and pulsed wave at the symbol level otherwise. The above two cases are used as examples of narrowband and wideband interference.

Hence, the received one-path signal has the form

\[
r(t) = \sum_{m=1}^{M} \sum_{k=1}^{K} b_{mk} d_k s_k(t - mT - d_k T_p) + n(t) + j_p(t),
\]

where

\[
0-7803-8379-6/04/$20.00 ©2004 IEEE.

775
in which \( b_{kn} \) is the \( n \)th symbol sent by \( k \)th user. The complex coefficient of the \( k \)th user’s channel is denoted by \( a_k \), which is assumed to remain the same during the data block of, say, \( M \) symbols. \( s_k(t) \) is \( k \)th user’s binary chip sequence, supported by \([0, T]\), where \( T \) is the symbol duration. \( T_c \) is the chip duration. For notational simplicity, the user delays are assumed to be discretized, and hence \( b_k \in \{0, \ldots, C - 1\} \), where \( C \) is number of chips in the spreading code. The delays are assumed to remain constant during the block of \( M \) data symbols. \( n(t) \) denotes noise.

The received signal is assumed to be sampled by chip-matched filtering, and using processing window size of two symbols in BSS based receivers. The sampled data can be written as [7], [11]

\[
r_m = \sum_{k=1}^{K} a_k b_{k,m-1} s_k + b_{km} g_k + b_{k,m+1} g_k + n_m + j_p((m-1)T)1 + j_p(mT)1 + j_p((m+1)T)1
\]  
(3)

Without loss of generality, it is here assumed that the interference remains constant during a symbol. Here \( r_m \) denotes noise vector and the code vectors of length \( 2C \) are defined as

\[
g_k = s_k [C - b_k + 1] \ldots s_k [C - b_k + 1] \begin{bmatrix} 0 \cdots 0 \\ 1 \cdots 1 \end{bmatrix}^T
\]

\[
g_k = s_k [b_k] \begin{bmatrix} 0 \cdots 0 \\ 1 \cdots 1 \end{bmatrix}^T
\]

\[
g_k = s_k [b_k] \begin{bmatrix} 0 \cdots 0 \\ 1 \cdots 1 \end{bmatrix}^T
\]

\[1 = \begin{bmatrix} 1 \cdots 1 \end{bmatrix}^T
\]

\[T = \begin{bmatrix} 1 \cdots 1 \end{bmatrix}^T
\]

With a simple manipulation, we get a compact representation for the data,

\[
r_m \overset{\text{def}}{=} Gb_m + n_m.
\]  
(5)

The \( 2C \times 3(K + 1) \) dimensional code matrix \( G \) contains the code vectors and path strengths, while the \( 3(K + 1) \)-vector \( b_m \) contains the symbols and sampled external interference:

\[
G \overset{\text{def}}{=} \begin{bmatrix} s_k [C - b_k + 1] \ldots s_k [C - b_k + 1] \\ \vdots \\ s_k [b_k] \begin{bmatrix} 0 \cdots 0 \\ 1 \cdots 1 \end{bmatrix}^T \\ \vdots \\ s_k [b_k] \begin{bmatrix} 0 \cdots 0 \\ 1 \cdots 1 \end{bmatrix}^T \\ \begin{bmatrix} 1 \cdots 1 \end{bmatrix}^T \\ \begin{bmatrix} 1 \cdots 1 \end{bmatrix}^T \end{bmatrix}
\]

\[
b_m \overset{\text{def}}{=} \begin{bmatrix} b_{km} s_k [b_k] \begin{bmatrix} 0 \cdots 0 \\ 1 \cdots 1 \end{bmatrix}^T \\ \vdots \\ b_{km} s_k [b_k] \begin{bmatrix} 0 \cdots 0 \\ 1 \cdots 1 \end{bmatrix}^T \\ \begin{bmatrix} 1 \cdots 1 \end{bmatrix}^T \\ \begin{bmatrix} 1 \cdots 1 \end{bmatrix}^T \end{bmatrix}
\]

Notice that \( n \)th symbols are included in three successive vector \( b_{m-1} \), \( b_m \) and \( b_{m+1} \). Hence, we say that these vectors are “early”, “middle” and “late” parts of \( n \)th symbols, respectively.

3. BLIND SOURCE SEPARATION BY INDEPENDENT COMPONENT ANALYSIS

Independent component analysis (ICA) [6] is a fairly new statistical technique by which BSS can be performed. In ICA a set of observed signals or random variables are tried to express as linear combinations of statistically independent components, which are often called sources or source signals. The ICA problem is blind, because not only the source signals but also the mixing coefficients are unknown.

In standard linear ICA, the \( n \) observed signals are assumed to be linear combinations of \( n \) unknown but statistically independent source signals \((m \geq n)\). By introducing the data vector \( x(t) = [x_1(t), \ldots, x_n(t)]^T \) for the observed signals, and the source vector \( s(t) = [s_1(t), \ldots, s_n(t)]^T \) for the source signals, the instantaneous noisy linear ICA mixture model is given by

\[
x(t) = As(t) + n(t)
\]  
(8)

Here the \( m \times n \) unknown but constant mixing matrix \( A \) contains the mixing coefficients, and \( n(t) \) denotes the additive noise vector at time \( t \). The sampled representation of DS-CDMA signal model (5) equals readily to the ICA model (8) as shown in [7] for the downlink case.

In numerical experiments, we applied the so-called FastICA algorithm for complex mixtures [6], [7]. FastICA is a fast method for performing linear ICA, and its basic form relies on the sample fourth-order statistics kurtosis. However, other forms of the algorithm employing more robust lower-order statistics have been developed [6]. Instead of FastICA, other ICA algorithms developed for complex-valued mixtures could be used, too.

4. BSS/ICA BASED SUCCESSIVE INTERFERENCE CANCELLATION

The BSS/ICA-assisted uplink receiver considered recently in [8] and [9] separates all the users simultaneously. After separation, users were identified by their spreading codes. Strictly speaking, each user’s spreading code was translated by the user’s path delay and multiplied by the user’s path coefficient. Then the correlations between this re-built spreading code, say \( c_k \), and a column of the mixing matrix got from the ICA, say \( w_{kj} \), were computed for each user. That is to say, \( \rho(k, k') = \frac{c_k w_{kj}}{|c_k| |w_{kj}|} \) is computed for each \( k, k' \). Finally, the source which corresponds to the best correlation is chosen to be that user’s data.

The principle of successive interference cancellation is adapted to BSS/ICA-assisted receiver to have all the users detected properly. First of all, a threshold is set to indicate proper detection. This is needed because subtraction of erroneously detected signal will actually enhance interference. Thus, only the signal corresponding to the user \( k \) with \( \rho(k, k') > \rho_k \) for some \( k' \) is subtracted from the receiver signal. After subtraction of all such users’ signals, the next ICA separation is performed for the interference-subtracted signal. The order of ICA model decreases in subtraction, which help ICA to recover the remaining users. The procedure is repeated successively until all users are detected. The successive ICA receiver is illustrated in Fig. 1.

To make the ICA-solution more reliable, a novel strategy during the BSS phase was considered. Recall that having a processing window of length \( 2C \), usually three symbols (or at least two) fall into the window; the early, middle and the late symbol, see Eq (7). Considering conventional receivers, the detection of the early and late symbols might not be that useful since one should disregard them by using only a partial code, resulting in much lower SNR at the despreader output. Considering ICA, these partial symbols are seen just another independent components with lower power, and consequently, can be estimated likewise. Therefore, a maximal ratio combining of them (with appropriate delaying) is proposed to get better SNR for the hard decision.
Fig. 1. Successive ICA-assisted receiver structure. $r_i(t)$ is received data after the $(i-1)$th IC round ($r_1(t) = r(t)$). $R = [r_1 \ldots r_K]$ is matrix containing sampled data. Vectors $z_n$ are source components by FastICA separation. Number of these vectors decreases in each IC round. Soft and hard decisions of users' symbols are denoted by $\hat{z}_b$ and $\hat{\delta}_b$, respectively. $I_i(t)$ denotes interference due to users detected in $i$th round.

Fig. 2. Bit-error-rates as a function of number of users. All users in system had equal powers. SNR was fixed to 20 dB and external jammer was absent.

Fig. 3. Bit-error-rates as a function of number of users. The users were split in two power groups. SNR was fixed to 20 dB wrt. the weakest users and external jammer was absent.

5. NUMERICAL EXPERIMENTS

A performance of the successive ICA receiver was studied numerically. Each user was supposed to send data blocks of $M = 5000$ QPSK symbols. Symbols were spread using Gold codes of length $C = 31$. All results are based on average bit error rates (BER) over 1000 independent repetition. BER values are computed with respect to all $K$ users. During FastICA, the iteration was terminated if any of the elements of a matrix $W_kW_k^H - I$ was less than a predetermined threshold. Here $I$ is an identity matrix and $W_k$ is the ICA de-mixing matrix after $k$th iteration.

According to the results, the successive ICA receiver provides clear improvements to the system capacity with respect to basic ICA as well as the SIC and PIC receivers. Figures 2 and 3 show bit-error-rates as a function of number of users ($K$) in the cases of users with equal powers and two power groups (power ratio of 10 dB between the groups). An external jammer was absent in these experiments. For instance, a capacity gain is at least six users on the BER-level of $10^{-5}$ compared to the basic ICA receiver and PIC receiver with five cancellation stages in the case of equal powers.

Recall that standard ICA assumes the number of independent components to be less than the number of observations. Here theoretical number of components is $3K$ (or $3(K+1)$ if jammer is present), since each symbol of $K$ users is included in three successive sampled data vectors. On the other hand, length of sampling window, $2C$, defines the number of observations. Thus, theoretical maximum to the number of users is $(3C \approx 20)$ (or $(3C - 1 \approx 19)$. Considering this, the performance of successive ICA receiver is outstanding also with the greatest values of $K$. In fact, the performance improves slightly in range from 18 to 20 users in the case of equal powers. This is due to increasing number of interference cancellations.

The effect of noise was examined in the system with $K = 16$ users. The successive BSS/ICA receiver outperformed reference methods evidently on a moderate SNR range (SNR $\geq 10\, dB$) even...
in systems without an external jammer. E.g. the case of users with equal powers is depicted in Fig. 4.

The systems with a jammer was examined, too. The number of users was again $K = 16$ and SJR was varied from -20 to 30 dB. The big values of SJR (i.e. cases where the spread signal is stronger than the jammer) was included to correspond the situation where some other jammer suppression method is first used but some part of jammer’s power is still remaining. Performances of the successive ICA receiver and basic ICA receiver were nearly at the same level when the jammer was stronger than users in system. However, the SIC and PIC receivers had clearly worse performance in this SJR range. The successive ICA receiver is, nevertheless, superior to all reference methods when SJR is positive. Fig. 5 shows the results in the case of users with equal power.

6. CONCLUSIONS

In this paper we adapted an ideology of successive interference cancellation to BSS/ICA-assisted uplink reception. The effectiveness of the new method was evaluated against existing basic BSS/ICA-assisted reception and conventional nonlinear interference cancellation schemes (SIC and PIC). The systems with and without additional external interference were considered. Numerical experiments indicated a clear performance gain in comparison to reference methods. Above all, the successive BSS/ICA method outperformed the conventional schemes clearly also when basic BSS/ICA method barely outdid them. All this was achieved even though ICA processing was terminated with a very mild condition, resulting in low computational load.

7. REFERENCES


DS-CDMA CAPACITY ENHANCEMENT USING BLIND SOURCE SEPARATION BASED GROUP-WISE SUCCESSIVE INTERFERENCE CANCELLATION

by

Toni Huovinen and Tapani Ristaniemi

DS-CDMA Capacity Enhancement Using Blind Source Separation Based Group-Wise Successive Interference Cancellation

Toni Huovinen
Department of Mathematical Information Technology, University of Jyväskylä
P.O. Box 35 (Agora), FIN-40041, Univ. Of Jyväskylä, Finland
Email: tonih@cc.jyu.fi

Tapani Ristaniemi
Institute of Communications Engineering, Tampere University of Technology
P.O.Box 553, FIN-33101, Tampere, Finland
Email: Tapani.Ristaniemi@tut.fi

Abstract — Recent studies have found successive interference cancellation (SIC) schemes employing blind source separation (BSS) methods to be promising competitors for conventional IC schemes. The main reason for this is the inherent ability of BSS to mitigate many kinds of interference sources e.g. multi-access, multi-path and out-of-cell interferences.

In this paper we propose a modification of a BSS-SIC-type receiver, and high-light its nice property of being able to drastically enhance the number of users a system can support, given a certain bit-error-rate requirement, compared to that of conventional parallel and successive interference cancellation (PIC and SIC, respectively) and BSS-alone schemes. The paper also shows how the proposed scheme relaxes the inherent weakness of standard BSS. Namely, the inadequateness in highly loaded systems due to the “more sources than observations”-problem, which can be somewhat circumvented in BSS-SIC-type receivers.

I. INTRODUCTION

The capacity of a Direct-Sequence Code Division Multiple Access (DS-CDMA) system is limited by the multiple access interference (MAI), which arises due to the non-ideal crosscorrelations between the spreading sequences. Conventional detection considers this interference as background noise and thus becomes inadequate in highly loaded systems. The other extreme from computational and performance points of view is optimal detection [1], followed by numerous suboptimal solutions with lower complexity.

Interference rejection/cancellation has been considered one of the most attractive class of suboptimal solutions, and has been studied extensively in the past [2]–[5]. Parallel and successive interference cancellation (PIC and SIC, respectively) are the main categories within this class, describing the procedure by which the interference is subtracted from the original data, after regenerating the interference from the tentatively estimated data. This procedure can, naturally, be repeated many times resulting in multi-stage IC schemes.

One relatively new idea is to employ blind source separation (BSS) techniques [6] on interference cancellation. Applications have been found e.g. in MIMO systems [7], I/Q processing receivers [8], DS-CDMA blind multi-user detection [9] and DS-CDMA out-of-cell interference cancellation [10]. What makes BSS techniques attractive is their ability to separate signals from a mixture of original source signals in a blind manner i.e. without explicit knowledge of signals’ waveforms. Consequently, many types of interferences can be inherently mitigated.

A major drawback for standard BSS is the case where the number of source signals (to be blindly extracted from the received data) is greater than the number of observations made, in which case standard BSS model doesn’t hold anymore. This problem is commonly known as a “more sources than sensors”-problem, or that the data have an over-complete basis. For example, from the single antenna DS-CDMA reception point of view, in order to have a standard BSS model available in the receiver, the number of users $K$ can be at most $K \leq \frac{N}{C}$ in a single-path channel, where $C$ is the spreading factor and $N$ is the observation interval (in symbols). Considering $L$ independent multipaths, the requirement even gets harder, $K \leq \frac{N-1}{C}$. Hence, the number of users in the system must be rather low to make standard BSS applicable.

In this paper we somewhat circumvent the “more sources than sensors”-problem by combining BSS and successive interference cancellation in a meaningful way. This is an extension of the previous studies [11]–[13] on the problem at hand into a more user-intensive systems (that is, $K \rightarrow C$) where standard BSS doesn’t work anymore. The BSS-SIC scheme works roughly speaking as follows: First, BSS is used to estimate as many users as it can. Here, some of the users are hence implicitly considered as noise. Among the set of estimated user-signals, the most properly estimated user-signals are subtracted from the received original signal. This procedure is repeated as many times as there exists users to be detected. Hence, this procedure is an iterative BSS scheme supported by a group-wise successive interference cancellation (or vice versa). The potential of the new receiver structure is evaluated with respect to existing BSS receivers and conventional SIC and PIC in a highly loaded ($K \rightarrow C$) DS-CDMA uplink.

II. SIGNAL MODEL

A considered channel model is a DS-CDMA uplink with additive white gaussian noise (AWGN) [14]. The well-known form for one-path signal, $r(t)$, is assumed:

$$r(t) = \sum_{m=1}^{M} \sum_{k=1}^{K} b_{km} a_k s_k(t - mT - d_k T_c/T).$$

(1)

Here $b_{km}$ is the $m$th symbol sent by $k$th user. The complex coefficient of the $k$th user’s channel is denoted by $a_k$, which is assumed to remain the same during the data block of, say, $M$ symbols. $s_k(\cdot)$ is $k$th user’s binary chip sequence, supported by $[0,T_c]$, where $T$ is the symbol duration. $T_c$ is the chip duration. For notational simplicity, the user delays are assumed to be discretized, and hence $d_k \in \{0,\ldots,C-1\}$, where $C$ is number of chips in the spreading code. The delays are assumed to remain
constant during the block of \( M \) data symbols. \( n(t) \) denotes noise.

The received continuous-time signal is assumed to be sampled by chip-matched filtering, and using processing window size of two symbols. The sampled data has the form [9], [15]

\[
r_{m} = \sum_{k=1}^{K} a_{k} (b_{m-k-1} g_{k} + b_{m} g_{k} + b_{m+k-1} g_{k}) + n_{m}
\]

Here \( n_{m} \) denotes noise vector and the code vectors of length \( 2C \) are defined as

\[
g_{k} = \begin{bmatrix} g_{k} (d_{k}) \\ g_{k} (d_{k}) \\ g_{k} (d_{k}) \end{bmatrix} = \begin{bmatrix} s_{k} (C - d_{k}) \\ s_{k} (C - d_{k}) \\ s_{k} (C - d_{k}) \end{bmatrix}^{T}
\]

With a simple manipulation, we can get a compact representation for the data,

\[
r_{m} = G b_{m} + n_{m}.
\]

The \( 2C \times 3K \) dimensional code matrix \( G \) contains the code vectors and path strengths, while the \( 3K \)-vector \( b_{m} \) contains the symbols:

\[
G = \begin{bmatrix} \ldots a_{k} (d_{k}) a_{k} (d_{k}) a_{k} (d_{k}) \ldots \end{bmatrix}
\]

\[
b_{m} = \begin{bmatrix} \ldots b_{m-k-1} b_{m} b_{m+k-1} \ldots \end{bmatrix}^{T}
\]

With a simple manipulation, we can get a compact representation for the data,

\[
r_{m} = G b_{m} + n_{m}.
\]

The \( 2C \times 3K \) dimensional code matrix \( G \) contains the code vectors and path strengths, while the \( 3K \)-vector \( b_{m} \) contains the symbols:

\[
G = \begin{bmatrix} \ldots a_{k} (d_{k}) a_{k} (d_{k}) a_{k} (d_{k}) \ldots \end{bmatrix}
\]

\[
b_{m} = \begin{bmatrix} \ldots b_{m-k-1} b_{m} b_{m+k-1} \ldots \end{bmatrix}^{T}
\]

Notice that the code vectors are included in three successive vectors \( b_{m-1}, b_{m}, b_{m+1} \). Hence, we say that these vectors are early, middle and late parts of \( M \) symbols, respectively.

### III. INDEPENDENT COMPONENT ANALYSIS AND DS-CDMA

Independent component analysis (ICA) [6] is a statistical technique where the goal is to represent a set of random variables as a linear transformation of statistically independent component variables. The main application of ICA is blind source separation (BSS) problem, which has become an attractive field of research in the statistical signal processing and neural network communities. The growing interest in ICA is mainly due to emerging new practical application areas, where the assumption of independence is both realistic and powerful.

In standard linear ICA, the \( m \) observed signals \( x_{1}(t), \ldots, x_{n}(t) \) at the time instant \( t \) are assumed to be linear combinations of \( n \) unknown but statistically independent source signals \( s_{1}(t), \ldots, s_{n}(t) \) at the time \( t \). By introducing the data vector \( x(t) = [x_{1}(t), \ldots, x_{m}(t)]^{T} \) for the observed signals, and the source vector \( s(t) = [s_{1}(t), \ldots, s_{n}(t)]^{T} \) for the source signals, the instantaneous noisy linear ICA mixture model is given by

\[
x(t) = A s(t) + n(t)
\]

Here the \( n \times n \) unknown but constant mixing matrix \( A \) contains the mixing coefficients, and \( n(t) \) denotes the additive noise vector at time \( t \). The ICA model (5) is readily a DS-CDMA signal model as shown in [9] for the downlink case.

In ICA, the source signals \( s(t) \) are estimated using only the observations \( x(t) \) by finding an \( n \times m \) unmixing matrix \( W \). In other words, a set of \( n \) filters \( w_{k} \) of length \( m \) should be found and stacked into a matrix \( W \). This matrix should be such that the \( n \)-vector \( W x(t) \) results in a set of mutually independent signals, or in practise, a set of signals which are mutually as independent as possible. Hence, it also recovers the set of original sources as well as possible. Because of the blindness of the problem, only the waveforms of the sources can be estimated. For estimating the unmixing (separating) matrix \( W \), many different methods have been proposed [6]. Most of these ICA methods exploit the statistical independence of the sources, but there exist other approaches, too, which utilize e.g. either temporal correlations or non-stationarity of the sources. The mutual performance of these methods depends largely on the validity of the assumptions made on them in the problem at hand.

Among many suitable applications for ICA, recent studies have considered BSS/ICA based interference cancellation, by which we loosely speaking mean an iterative multi-user receiver, where the estimated interference is subtracted from the received signal prior to the estimation of a particular user. Previous studies on BSS/ICA based interference cancellation have considered standard ICA method, which set a strict upper bound to number of independent components. Standard ICA assumes the number of them to be less or equal than the number of observations. Here theoretical number of components is \( 3K \), since each sample vector contains portions of 3 consecutive symbols of each user. On the other hand, as the length of processing window is \( 2C \), it also defines the number of observations. Thus, theoretical maximum to the number of users is \( \frac{3}{2} C \). If number of users is greater than \( \frac{3}{2} C \), we say that system is oversaturated (with respect of ICA). Hence, standard ICA could not be used.

A basic BSS/ICA-assisted uplink receiver was considered recently for non-oversaturated DS-CDMA systems in [11] and [12]. The receiver separates all the users simultaneously using a FastICA-algorithm [16]. After separation, users are identified by their spreading codes. Hence, the receiver is semi-blind. Strictly speaking, each user’s spreading code is translated by the user’s path delay and multiplied by the user’s path coefficient. Then the correlations between this re-built spreading code, say \( c_{k} \), and a columns of the mixing matrix got from the ICA, say \( w_{k} \), are computed for each user. That is to say, \( \rho(k,k') = \frac{c_{k} w_{k}^{H}}{|c_{k}||w_{k}|} \) is computed for each \( k, k' \). Finally, the source which corresponds to the best correlation is chosen to be that user’s data. The weakness of this method, however, is that it is becomes inadequate in highly loaded systems.

### IV. BSS/ICA BASED INTERFERENCE CANCELLATION IN OVEVSATURATED SYSTEMS

Separating source signals in oversaturated system is more a difficult task, since mixing model is not invertible anymore [6]. Consequendly, typical ICA methods for oversaturated systems are computationally more expensive compared to standard ICA. From the computational complexity point of view, a modification of a FastICA algorithm [17] is one of the most attractive ICA algorithm designed to operate in oversaturated systems. The modification is based on concept of quasiorthogonality, which roughly speaking means that we are now considering a set of nearly orthogonal basis for the representation of the data, thus enabling the increase in dimensionality compared to strictly orthogonal basis. As will be shown in the numerical experiments, the use of quasiorthogonal basis in ICA helps to achieve tolerable performance. However, the performance is still only moderate.

The standard BSS/ICA-receiver can be extended to work in oversaturated systems much more effectively. Since identification procedure is based purely on cross-correlation properties of the spreading codes, changes are needed only in ICA-separation. A successive BSS/ICA-assisted receiver structure
was introduced in [13]. That receiver combines the concepts of standard BSS/ICA and successive interference cancellation. The combination was found to improve performance in non-oversaturated systems. In this paper the method is modified to work also in oversaturated systems.

The principle of successive BSS/ICA receiver is quite straightforward. The received signal is first separated by FastICA algorithm with the quasiorthogonal basis in ICA. The correlations \( \rho(k, k') \) are calculated like in the basic receiver, and thus each user can be identified from the set of FastICA outputs. Second, a threshold, \( \rho_1 \), is set to indicate proper detection. This is needed because subtraction of erroneously detected signal will actually enhance interference. Thus, only the signal corresponding to the user \( k \) with \( \rho(k, k') > \rho_1 \) for some \( k' \) is subtracted from the received signal. After subtraction of all such users’ signals, the next ICA separation is performed for the interference-subtracted signal. The order of ICA model decreases in subtraction, which help ICA to recover the remaining users. The procedure is repeated successively until all users are detected. The successive ICA receiver is illustrated in Fig. 1.

Note, that the core of FastICA is the following:

1. Whiten the original data (or the data after subtraction) as follows: \( y_m = T r_m \), such that \( E(y_m y_m^H) \) equals the identity matrix. Such \( T \) can be simply found in many ways [6].
2. Take a random \( 2C \times 3K \) matrix \( W = [w_1 \ldots w_{3K}] \).
3. Quasiothogonalize matrix \( W \) as in step 5 and normalize its columns.
4. FastICA: For all \( i = 1, \ldots, 3K \), let
   \[
   w_i \leftarrow E\{y_m(w_i^H y_m)^*y_m^H y_m\} - 2w_i
   \]
5. Quasiothogonalization: \( W \leftarrow \frac{1}{2} W - \frac{1}{2} WW^H W \)
6. Normalize columns of matrix \( W \).
7. Go to FastICA-step until converged.

Above \( * \) denotes a complex conjugate of complex number \( z \) and \( M^H \) stands for a Hermitian (i.e. an element wise complex conjugate and transpose) of matrix \( M \).

To make the ICA-solution more reliable, a novel strategy during the BSS phase was considered. Recall that having a processing window of length \( 2C \), usually three symbols (or at least two) fall into the window; the early, middle and the late symbol, see Eq (5). Considering conventional receivers, the detection of the early and late symbols might not be that useful since one should despread them by using only a partial code, resulting in much lower SNR at the despreader output. Considering ICA, these partial symbols are seen just another independent components with lower power, and consequently, can be estimated likewise. Therefore, a maximal ratio combining of them (with appropriate delaying) is used to get better SNR for the hard decision.

V. NUMERICAL EXPERIMENTS

The performance of BSS-SIC -type receiver in oversaturated systems was studied with numerical experiments. Each of \( K \) users was assumed to transmit data blocks of \( M = 5000 \) QPSK symbols. Symbols were spread using Gold codes of length \( C = 31 \). Recall that a system is oversaturated if \( \frac{2}{3}C \leq K \leq C \). Hence, \( K \) was varied in the interval from 21 to 31. The users were split in two power groups. Inside each group all users had

\[ r(t) \]

\[ r_1(t) \]

\[ I_i(t) \]

Figure 1: Successive ICA-assisted receiver structure. \( r_i(t) \) is received data after the \((i-1)\)th IC round \((r_1(t) = r(t))\). \( R = [r_1 \ldots r_M] \) is matrix containing sampled data. Vectors \( \tilde{z}_n \) are source components by FastICA separation. Number of FastICA iterations was 110 in ICA based receivers. SNR was fixed to 20 dB wrt. the weakest users.

Figure 2: Bit-error-rates as a function of number of users. Number of FastICA iterations was 110 in ICA based receivers. SNR was fixed to 20 dB wrt. the weakest users.
equal power and power ratio between groups was 10 dB. All results are based on average bit error rates (BER) over 1000 independent repetition. BER values have been computed with respect of all $K$ users.

The results demonstrate that ICA-assisted methods maintain their competence also with great numbers of users. Especially the successive ICA-receiver outperforms reference methods, SIC and PIC, clearly. This can be seen e.g. in Fig. 2. The figure depicts BERs of the ICA-assisted methods as well as the reference methods as a function of $K$ (number of users). Number of FastICA-iterations was fixed to be 110 in both ICA-assisted receivers.

The number of FastICA-iterations plays a significant role in ICA-assisted methods. On one hand, too small number of iterations impairs a performance and, on the other hand, increasing the number increases a computational load of the methods. The experiments were repeated with different numbers of FastICA-iterations to find out an optimal number of iterations. Figs. 3 and 4 illustrate performances of basic and successive receiver, respectively. PIC-receiver is used again as a reference. The figures reveals another favorable feature of successive ICA-receiver compared to basic ICA-receiver. In addition to clear performance improvement, the successive receiver converges (wrt. FastICA-iterations) even faster than the basic receiver. Fig. 5 demonstrates how many iterations successive ICA-receiver needs in order to achieve BER-levels of $10^{-2}$ and $10^{-3}$.

Bit-error-rates of the successive ICA-receiver seem to deteriorate slightly after certain number of iterations. This is best seen in cases of 24–28 users (Fig. 4). Deterioration is due to increasing number of interference subtractions after first ICA-separation. Like in standard SIC and PIC-methods, detection errors in tentative symbol estimates inflict on extra interference into following separation rounds.
VI. Conclusions

In this paper we proposed a modification of a BSS-SIC -type receiver, which, unlike standard BSS, is able to operate in highly loaded systems. We high-lighted its ability to drastically enhance the number of users a system can support, given a certain bit-error-rate requirement, compared to that of conventional parallel and successive interference cancellation (PIC and SIC) schemes and standard BSS-alone receiver. The paper also showed how the proposed scheme relaxes the inherent weakness of standard BSS. Namely, the inadequateness in highly loaded systems due to the “more sources than observations”-problem, which was somewhat circumvented in the proposed successor ICA receiver.

References

Joint Delay Tracking and Interference Cancellation in DS-CDMA Using Successive ICA for Oversaturated Data

by

Tapani Ristaniemi and Toni Huovinen

Joint Delay Tracking and Interference Cancellation in DS-CDMA Systems Using Successive ICA for Oversaturated Data

Tapani Ristaniemi and Toni Huovinen
Institute of Communications Engineering, Tampere University of Technology
P.O.Box 553, FIN-33101, Tampere, Finland
{Tapani.Ristaniemi,Toni.Huovinen}@tut.fi

Abstract. Recent studies have found successive interference cancellation (SIC) schemes employing blind source separation (BSS) methods to be promising competitors for conventional interference cancellation schemes in DS-CDMA systems. One reason for this is the inherent ability of BSS to mitigate many kinds of interference sources e.g. multi-access, multi-path and out-of-cell interferences. In addition, “more sources than observations”-problem can be somewhat circumvented by combining the SIC-ideology to BSS. Hence, BSS-SIC-type receivers can be used also in highly loaded systems where both conventional ICA and conventional interference cancellation usually fails. Recently proposed BSS-SIC receivers have needed accurate estimates for the user delays. This is mandatory for any interference canceller, since subtraction of a inaccurately estimated source from the original data actually enhances interference. In this paper we propose a scheme in which the ICA solution is used to tracking of users’ delays, too, which helps to avoid the performance losses due to the inaccuracies in conventional delay tracking circuitry.

1 Introduction

The capacity of a Direct-Sequence Code Division Multiple Access (DS-CDMA) system is in practice limited by the multiple access interference (MAI), which arises due to the non-ideal crosscorrelations between the spreading sequences. Conventional detection considers this interference as background noise and thus becomes inadequate in highly loaded systems. The other extreme from computational and performance points of view is individually optimum detection [1], which has been followed by numerous sub-optimum solutions with much lower complexity.

Interference rejection/cancellation has been considered one of the most attractive class of suboptimal solutions, and has been studied extensively in the past [2]–[5]. Parallel and successive interference cancellation (PIC and SIC, respectively) are the main categories within this class, describing the procedure by which the interference is subtracted from the original data, after regenerating
the interference from the tentatively estimated data. This procedure can, naturally, be repeated many times resulting in multi-stage interference cancellation schemes. Needless to say, any interference subtractive receiver performs the better the more accurately tentative decisions are being made. This is because the interference level is then reduced the most. Otherwise, from a DS-CDMA signal point of view, two different situations will occur:

(a) The particular user signal is not completely taken out from the original data or it is even strengthened; or in addition to that,

(b) a fictitious multipath component is generated and added to the original data.

The former happens e.g. if all the other parameters except the (complex-valued) amplitude is correctly estimated. The latter happens if the phase of the spreading code is inaccurately estimated. To mild the consequences of the situation (a) one can e.g. estimate the reliability of tentative decision from the soft decisions. Doing this way one actually chooses in favor of accepting only partial mitigation (which happens almost surely) rather than accepting occasional interference enhancement. In this paper we consider the avoidance of the situation (b), for which the most natural solution is more accurate delay estimation.

One relatively new idea is to employ blind source separation (BSS) techniques [6] in interference subtractive receivers. In [7] it was shown how the parametric form of the mixing matrix can efficiently be used to refine the ICA solution, and hence avoid interference enhancement while subtracting that source from the original data. The BSS-SIC -type receiver was further developed in [8] for highly loaded systems. Recall that a major drawback for standard BSS is the case where the number of source signals (to be blindly extracted from the received data) is greater than the number of observations made. This is a commonplace situation in communications applications in which cases standard BSS model doesn’t hold anymore. The key finding in [8] was the ability of a BSS-SIC -type receiver structure to somewhat circumvent the “more sources than observations”-problem in the sense that adequate performance (in terms of bit-error probability) is still achievable even in extremely highly loaded system, whereas conventional parallel and successive interference cancellation only remain at a moderate level.

In this paper we further develop the BSS-SIC -type receiver (at the expense of negligible increase in computation) to cope with erroneous timing estimates. Recall that the timing information, that is, the phase of the band-spreading code, should be known accurately to avoid interference enhancement during the subtraction phase. Roughly speaking, the ICA solution implicitly includes the timing information, thus giving a possibility to estimate it due to the fact that the parametric form of the mixing matrix is known. Numerical examples are given to show how the joint delay tracking and interference cancellation based on BSS-SIC -type receiver tolerates quite inaccurate initial parameter estimates, unlike conventional parallel and successive interference cancellation schemes as well as advanced LMMSE-PIC [9] receiver. The latter receiver estimates each user with an optimal linear MMSE detector, after which subtraction follows.
2 Signal Model

Consider a DS-CDMA uplink channel with additive white gaussian noise (AWGN) [10]. The well-known form for the single-path data, \( r(t) \), is assumed:

\[
r(t) = \sum_{m=1}^{M} \sum_{k=1}^{K} b_{km} a_k s_k (t - mT - \tau_k T_c / T).
\]

(1)

Here \( b_{km} \) is the \( m \)th symbol sent by \( k \)th user. The complex coefficient of the \( k \)th user’s channel is denoted by \( a_k \), which is assumed to remain the same during the data block of, say, \( M \) symbols. \( s_k(\cdot) \) is \( k \)th user’s binary chip sequence, supported by \([0, T)\), where \( T \) is the symbol duration. \( T_c \) is the chip duration. The user delay is denoted \( \tau_k = d_k + \delta_k \), where \( d_k \in \{0, \ldots, C - 1\} \), \( C \) is number of chips in the spreading code and \( \delta \in (-\frac{1}{2}, \frac{1}{2}) \). The delays are assumed to remain constant during the block of \( M \) data symbols. \( n(t) \) denotes noise.

The received continuous-time signal is assumed to be sampled by chip-matched filtering, and using processing window size of two symbols. The sampled data has the form [11], [12]

\[
r_m \overset{\text{def}}{=} \sum_{k=1}^{K} a_k (b_{km-1} g_k + b_{km} g_k + b_{km+1} \bar{g}_k) + n_m
\]

(2)

Here \( n_m \) denotes noise vector and the code vectors of length \( 2C \) are defined as

\[
\begin{align*}
g_k &= g_k(\delta_k, d_k) = (1 - |\delta_k|) c_k(d_k) + |\delta_k| c_k(d_k + \text{sign}(\delta_k)) \\
g\_k &= g_k(\delta_k, d_k) = (1 - |\delta_k|) c_k(d_k) + |\delta_k| c_k(d_k + \text{sign}(\delta_k)) \\
g\_k &= g_k(\delta_k, d_k) = (1 - |\delta_k|) \bar{c}_k(d_k) + |\delta_k| \bar{c}_k(d_k + \text{sign}(\delta_k))
\end{align*}
\]

(3)

where

\[
\begin{align*}
c_k(d_k) &\overset{\text{def}}{=} [s_k[C - d_k + 1] \ldots s_k[C] \ 0^T_{2C - d_k}]^T \\
c\_k(d_k) &\overset{\text{def}}{=} [0^T_{d_k} \ s_k[1] \ldots s_k[C] \ 0^T_{C - d_k}]^T \\
\bar{c}_k(d_k) &\overset{\text{def}}{=} [0^T_{C+d_k} \ s_k[1] \ldots s_k[C - d_k]]^T.
\end{align*}
\]

(4)

With a simple manipulation, we can get a compact representation for the data,

\[
r_m = G b_m + n_m.
\]

(5)

The \( 2C \times 3K \) dimensional code matrix \( G \) contains the code vectors and path strengths, while the \( 3K \)-vector \( b_m \) contains the symbols:

\[
G \overset{\text{def}}{=} \begin{bmatrix}
\cdots a_k g_k & a_k g_k & a_k \bar{g}_k & \cdots
\end{bmatrix}^T
\]

(6)

\[
b_m \overset{\text{def}}{=} \begin{bmatrix}
\cdots b_{km-1} b_{km} b_{km+1} \cdots
\end{bmatrix}^T.
\]
The DS-CDMA signal model (5) is readily a linear noisy ICA model with \( m = 2C \) observations of \( n = 3K \) source components.

Without losing any generality we assume that the delay of a user \( k \) is pre-estimated to be equal to \( d_k \), that is, \( \hat{\tau}_k = d_k \), even though \( d_k + \delta_k \) would be the correct one. Consequently, \( \delta_k \) is considered as a residual delay estimation error after delay tracking circuitry. It is thus typical to consider this error to obey zero mean Gaussian distribution.

Recall that conventional single user detection makes a decision according to

\[
\hat{b}_k = \text{sgn}(a_k^* r_{m}^H c_k(\hat{\tau}_k))
\]

from which we see that there is a timing misalignment equal to \( \delta_k \) between the local replica of the spreading code and that of received data.

### 3 BSS/ICA Based Successive Interference Cancellation

By an interference subtractive receiver we loosely speaking mean an iterative multi-user receiver, where the estimated interference is subtracted from the received signal prior to the estimation of a particular user. The principle of this kind of receiver [7, 8] utilizing BSS/ICA is quite straightforward and is here only shortly revisited. Namely, the received signal is first separated by ICA. After separation, users are identified by their spreading codes. Hence, the receiver is semi-blind. Strictly speaking, for the user identification a correlation

\[
\rho(k, k') = \frac{c_k(\hat{\tau}_k) w_{k'}^H}{\|c_k(\hat{\tau}_k)\| \|w_{k'}\|}
\]

is computed for each \( k, k' \), where \( w_{k'} \) corresponds to the ICA basis vector of \( k' \):th source. Next, a threshold, \( \rho_1 \), is set to indicate proper detection. This is needed because subtraction of erroneously detected signal would actually enhance interference. Thus, only the signal corresponding to the user \( k \) with \( |\rho(k, k')| > \rho_1 \) for some \( k' \) is subtracted from the receiver signal. After subtraction of all such users’ signals, the next ICA separation is performed for the interference-subtracted signal. The order of ICA model decreases in subtraction, which help ICA to recover the remaining users. The procedure is repeated successively until all users are detected. Furthermore, one can easily show, that phase of complex number \( \rho(k, k') \) in (8) equals to phase shift produced by ICA. This makes it possible to correct ICA’s phase ambiguity. The successive ICA receiver is illustrated in Fig. 1.

It is of primary importance to see that what is actually subtracted from the original data is a tentative decision of the form \( a_k c_k(\hat{\tau}_k) \hat{b}_k \) rather than \( w_k \hat{b}_k \). Hence the ICA solution is first refined (according to the knowledge of the parametric form of the mixing matrix) before subtraction is performed.

### 4 Joint Delay Tracking and Interference Cancellation Using BSS/ICA

The receiver structure described above is now developed to cope with inaccuracies in delay estimation. Recall that ICA performs purely in blind manner.
Fig. 1. Group-wise successive ICA receiver structure. $r_i(t)$ is received data after the $(i-1)^{th}$ interference subtraction round ($r_1(t) = r(t)$). $R = [r_1 \ldots r_M]$ is matrix containing sampled data. Vectors $\tilde{z}_n$ are source components by FastICA separation. Number of these vectors decreases in each round. Soft and hard decisions of users’ symbols are denoted by $\hat{z}_k$ and $\hat{b}_k$, respectively. $I_i(t)$ denotes re-generated interference due to the users detected in $i^{th}$ round.

The first occasion where some a priori knowledge of the users is needed is the user identification phase (8). Naturally, the timing information can also be used to generate a good initial value for ICA iterations, and hence speed up the separation [11]. Anyway, what lousy timing estimate ultimately does is that it worsen the user identification. More importantly, given that a user is nevertheless identified having an erroneous timing estimate, the subtraction of that user enhance interference the more the bigger was the timing inaccuracy. To avoid that situation the delay of each identified user could first be refined according to

$$\hat{\delta}_k \leftarrow \arg \max_{\delta} \frac{g_k(\delta, \hat{\tau}_k)w_{k'}^H}{\|g_k(\delta, \hat{\tau}_k)\| \|w_{k'}\|}$$

(9)
where the maximization of the correlation is performed in a close neighborhood of $\hat{\tau}_k$, e.g., $\delta \in (-1/2, 1/2)$. The signal corresponding to the user $k$ is re-built after delay refinement, only after which it is beneficial to go to the subtraction phase:

$$r_m \leftarrow r_m - a_k (b_{k,m-1} g_k(\hat{\delta}_k, \hat{\tau}_k) + b_{km} g_k(\hat{\delta}_k, \hat{\tau}_k) + b_{k,m+1} g_k(\hat{\delta}_k, \hat{\tau}_k)) \quad (10)$$

5 Numerical Experiments

The performance of the proposed receiver is studied with numerical experiments. Each of $K$ users is assumed to transmit data blocks of $M = 5000$ QPSK symbols. Symbols are spread using Gold codes of length $C = 31$. Two service classes are assumed, which is modelled as a power difference of $10 \text{ dB}$ between the two user groups. Inside both groups all the users are assigned the same power. All the results are based on average bit error rates (BER) over 1000 independent repetition. BER values are computed with respect of all $K$ users.

The length of the all the receivers is $2C$ and hence both the LMMSE-PIC and LMMSE detectors are truncated to that length. Recall that the optimal length for asynchronous data would be $MC$ (the whole block of symbols) which is not a sensible choice for the receiver length in practise [1].

As an “Basic ICA” (one of the reference methods) we used the version of FastICA algorithm, which is able to operate also in oversaturated systems [13]. This was used also in successive ICA receivers.

The results demonstrate that the proposed receiver is clearly more immune to delay estimation errors than the reference ones. This can been seen in Fig. 2. If the delays are known perfectly, LMMSE-PIC-receiver outperforms successive ICA-receiver (Fig. 2 (a)). However, the proposed receiver is superior to LMMSE-PIC, if there is even a slight error in delay estimation (Fig. 2 (b)). Fig. 3 depicts the performance of successive ICA-receiver as a function of the number of FastICA iterations performed. The figure shows that both successive ICA-receivers overtake e.g. LMMSE-PIC-receiver very fast when system have delay estimation errors.

Finally, we also notice from the experiments that the FastICA with successive interference cancellation performs clearly better than FastICA alone, given that the delays are tracked, too.

6 Conclusions

In this paper we considered successive interference cancellation (SIC) schemes employing blind source separation (BSS) methods in DS-CDMA systems. Especially, a BSS-SIC -type receiver structure where the users’ delays are simultaneously tracked was proposed and evaluated via numerical examples. The main finding was that SIC-ideology combined with ICA for oversaturated systems is quite beneficial given that certain key parameters like users’ delays are tracked simultaneously. The example cases assumed $m = 2C = 62$ observations from the mixture of $n = 3K = 66 – 90$ sources.
Fig. 2. (a) Bit-error-rates as a function of number of users in case of perfect delay estimation. (b) Bit-error-rates as a function of variance of delay estimation error in a system of $K = 22$ users. In both figures, users are split in two power groups and SNR is fixed to 20 dB wrt. the weakest users. In addition, 110 FastICA iterations are used in successive ICA receiver.

Fig. 3. Bit-error-rates as a function of number of FastICA iterations. The $K = 22$ users are split in two power groups. SNR is fixed to 20 dB wrt. the weakest users and users’ estimated delays are assumed to have an error $\delta_k$ with variances $\sigma^2 = 0.01$ (a) and $\sigma^2 = 0.05$ (b) (unit of error corresponds to one chip time).
References

INDEPENDENT COMPONENT ANALYSIS USING SUCCESSIVE INTERFERENCE CANCELLATION FOR OVERSATURATED DATA

by

Toni Huovinen and Tapani Ristaniemi

Signal Processing

Independent component analysis using successive interference cancellation for oversaturated data

Toni Huovinen* and Tapani Ristaniemi

Institute of Communications Engineering, Tampere University of Technology, P.O. Box 553, FIN-33101, Tampere, Finland

SUMMARY

In this paper, advanced interference cancellation strategies are considered and applied in DS-CDMA uplink receivers. Namely, we introduce a new successive interference cancellation (SIC) scheme in which blind source separation (BSS) techniques based on independent component analysis (ICA) are considered. Our main emphasis is on the oversaturated data in which situation easily arises for example in DS-CDMA systems when the number of active connections/users is high compared to number of chips per bit being used, that is, a highly loaded system. Proposed receiver structures combine the main benefits of pure BSS/ICA and SIC methods: (i) inherent mitigation of various types of interference sources by BSS/ICA, (ii) robustness against parameter estimation errors due to BSS/ICA, (iii) greatly improved interference suppression capability due to novel combination of SIC ideology and ICA for oversaturated data. The last item in the list can also be seen as a way to somewhat circumvent the ‘more sources than sensors’—problem, which is a challenging problem in BSS area. Numerical experiments with DS-CDMA uplink data are given to illustrate the achieved gains in capacity and bit-error-rate performance compared to conventional successive and parallel interference cancellation (PIC) schemes as well as more an advanced LMMSE-PIC receiver and variants of BSS/ICA techniques alone. Copyright © 2005 AEIT.

1. INTRODUCTION

In practice, the multiple access interference (MAI) due to the non-ideal cross-correlations between the spreading sequences limit the performance of a direct sequence code division multiple access (DS-CDMA) system. Conventional detection considers this interference as a background noise and thus becomes inadequate in highly loaded systems. Conventional detection is also quite sensitive to external interference sources like adjacent channel interference or jamming which has lead to development of numerous interference rejection techniques [1].

Successive and parallel interference cancellation schemes (abbreviated SIC and PIC respectively) [2–4] are traditional methods in the category of sub-optimum multiuser detection (MUD) schemes. These methods are often considered as reasonable alternatives to optimum MUD [5] since they have essentially lower computational complexity. In SIC and PIC, the interference is regenerated according to the tentative estimates of data and channel parameters. Then the regenerated interference is subtracted from the received signal either in serial or parallel manner, which results in a signal with lower level of interference and thus more reliable demodulation. This procedure can, naturally, be repeated many times resulting in multistage interference cancellation schemes.

Any interference subtractive receiver performs understandably the better with the more accurate tentative decisions being made before interference subtraction. This is because the interference level is then reduced the most. Otherwise, from a DS-CDMA signal point of view, two different situations will occur:
(a) The particular user signal is not completely cancelled from the original data or it is even strengthened; or in addition to that,
(b) a fictitious multipath component is generated and added to the original data.

The former happens for example if all the other parameters except the (complex-valued) amplitude are correctly estimated. The latter happens if the phase of the spreading code is inaccurately estimated. A presence of external interference naturally worsen the tentative decision making, thus, resulting in both situations, (a) and (b), to occur more likely. To mild the consequences of the situation (a), one can estimate the reliability of tentative decision from the soft decisions [5]. For example, one could choose the amplitude of a particular user to minimise the energy of cancellation error in the subtraction phase. Doing this way one actually chooses in favour of accepting only partial mitigation (which happens almost surely) rather than accepting occasional interference enhancement. In this paper, we mainly focus on the avoidance of the situation (b), for which the most natural solution is more accurate delay estimation. We also consider the suppression of external interference.

One relatively new idea in interference rejection is to employ blind source separation (BSS) techniques [6] in DS-CDMA reception. What makes BSS techniques attractive is their ability to separate signals from a mixture of original source signals in a completely blind manner, that is, without explicit knowledge of signals’ waveforms. Consequently, many types of interference sources, for example internal interferences due to multiple access and out-of-cell interferences (intentional and unintentional) due to cellular network can be mitigated in DS-CDMA systems. In addition, typical BSS methods like independent component analysis (ICA) relies solely on higher-order statistical properties of data, which makes the methods very robust against problems related to incomplete cross-correlation properties of users and a near-far problem as well. Communication-related applications of BSS have been found in MIMO systems [7], I/Q processing receivers [8], DS-CDMA blind multi-user detection [9] and DS-CDMA out-of-cell interference cancellation [10].

A recent study has also found BSS techniques to be beneficial in interference subtractive receivers [11]. In that paper an iterative least squares technique for digital signal separation was enhanced by an interference cancellation stages, a procedure similar to conventional successive or parallel interference cancellation schemes. Namely, the method takes advantage of an alternating estimation of the mixing matrix and user symbol streams in until converged, while employing an interference cancelling stages during the symbol streams’ estimation.

Another type of BSS-assisted interference cancellation schemes has been considered recently in References [12, 13]. These approaches obey more the fundamental procedures of conventional successive interference cancellation which are the estimation of interference, regeneration of interference and finally subtraction of the regenerated interference from the original data. What makes the schemes advanced compared to conventional SIC is the way of estimating the interference using a higher-order statistics by means of ICA and a sensitive selection of the interference sources to be subtracted.

In this paper, we present the ideas of preliminary conference papers [12, 13] in an unified manner and propose further enhancements. In these ICA-based interference cancellation schemes, the parametric form of the mixing matrix is efficiently used to refine the interference estimate offered by ICA, and hence avoids interference enhancement while subtracting that source from the original data. Second, group-wise successive interference subtraction is employed, which is beneficial in highly loaded systems when ICA is employed. Recall that a major drawback for standard BSS/ICA is the case where the number of source signals (to be blindly extracted from the received data) is greater than the number of observations being made. This is a commonplace situation in communications applications in which cases standard BSS/ICA model does not hold anymore. The key finding in Reference [13] was the ability of a ICA-SIC-type receiver structure to somewhat circumvent the ‘more sources than observations’—problem in the sense that adequate performance (in terms of bit-error probability) is still achievable even in extremely highly loaded system, whereas conventional parallel and successive interference cancellation only remain at a moderate level.

An important property of ICA-assisted receiver is its inherent capability to cope with erroneous parameter estimates. Roughly speaking, the estimate of interference by ICA implicitly includes the timing information of the user signals, thus giving a possibility to refine preliminary delay estimates before regeneration of interfering signals. The refinement of delay estimates is not inherently possible in conventional interference subtractive receivers, but additional delay tracking circuitry is needed. In this paper, we propose a modification of ICA-SIC which is, in practice, somewhat immune to delay estimation errors. The new modification constitute a two-step delay tracker within ICA processing where a coarse pre-tracking helps

Copyright © 2005 AEIT
the identification procedure during ICA and a finer post-tracking improves the interference regeneration, and hence, more effective interference cancellation. Numerical examples are given to validate the effectiveness of the method especially in highly loaded system where traditional PIC and SIC, as well as ICA alone provides inadequate performance. In addition, robustness against delay estimation errors is demonstrated and the performance is compared to commonly known and more an advanced LMMSE-PIC-type receivers [14]. As a by-product it is shown how the proposed receiver structures somewhat circumvent the known ‘more sources than sensors’—problem.

The rest of the paper is organised as follows. Section 2 introduces the systems characteristics, continued by Section 3 which deals with independent component analysis, a technique applied later on. In Section 4, we introduce the ICA-based interference cancellation schemes and numerical examples are given in Section 5.

2. SYSTEM CHARACTERISTIC

2.1. Notations

The following notations are used throughout this paper:

- \((\cdot)^*\) complex conjugate;
- \((\cdot)^T\), \((\cdot)^H\) transpose and Hermitian transpose;
- \(\Re\{\cdot\}, \Im\{\cdot\}\) real and imaginary part;
- \(i\) imaginary unit;
- \(E\{\cdot\}\) expected value;
- \(\sign(\cdot)\) sign function;
- \(\|\cdot\|\) Euclidean norm;
- \(\mathbf{0}_n\) \(n\)-dimensional vector of zeros

In addition, lower and upper case boldface symbols stand for vectors and matrices respectively.

2.2. Signal model

A DS-CDMA uplink channel with additive white Gaussian noise (AWGN) [15] is studied. In addition, channel is assumed to contain an external interference signal \(j(t)\), which is modelled as a continuous wave (possibly pulsed at the symbol level) with a frequency, phase offsets and power equal to \(\Delta f\), \(\theta\) and \(J\) respectively:

\[
j_p(t) = \delta_p \sqrt{J} e^{(2\pi n \Delta f t + \theta)}
\]  

Here \(\delta_p(t) = 1\) with a probability \(p\) during a symbol. Hence, the interference corresponds to a continuous wave when \(p = 1\) and pulsed wave at the symbol level otherwise. The above two cases are used as examples of narrowband and wideband interference.

Hence, the received signal has the form

\[
r(t) = \sum_{m=1}^{M} \sum_{k=1}^{K} b_{km} a_k \delta_k \left( t - mT - \frac{T_c}{T} \right) + n(t) + j_p(t)
\]

in which \(b_{km}\) is the \(m\)th symbol sent by \(k\)th user. The complex coefficient of the \(k\)th user’s channel is denoted by \(a_k\), which is assumed to remain the same during the data block of, say, \(M\) symbols. \(s_k(\cdot)\) is \(k\)th user’s binary chip sequence, supported by \([0, T]\), where \(T\) is the symbol duration. \(T_c\) is the chip duration. The user delay is denoted \(\tau_k = d_k + \delta_k\), where \(d_k \in \{0, \ldots, C-1\}\), \(C\) is number of chips in the spreading code and \(\delta_k \in (-1/2, 1/2)\). The delays are assumed to remain constant during the block of \(M\) data symbols. \(n(t)\) denotes noise.

The received continuous-time signal is assumed to be sampled by chip-matched filtering. Using processing window size of two symbols, the sampled data can be collected into vectors of length \(2C\) as follows [9, 16]:

\[
r_m \equiv \sum_{k=0}^{K} a_k (b_{km-1} \mathbf{g}_k + b_{km} \mathbf{g}_k + b_{km+1} \mathbf{g}_k) + n_m 
\]

+ \(j_p((m-1)T)1 + j_p(mT)1 + j_p((m+1)T)1\)

Without loss of generality, it is here assumed that the interference remains constant during a symbol. Here \(n_m\) denotes noise vector and the code vectors of length \(2C\) are defined as

\[
\mathbf{g}_k \equiv \mathbf{g}_k(\delta_k, d_k) = (1 - |\delta_k|) \mathbf{e}_k(d_k) + |\delta_k| \mathbf{e}_k(d_k + \sign(\delta_k))
\]

\[
\mathbf{b}_k \equiv \mathbf{b}_k(\delta_k, d_k) = (1 - |\delta_k|) \mathbf{c}_k(d_k) + |\delta_k| \mathbf{c}_k(d_k + \sign(\delta_k))
\]

\[
\mathbf{b}_k \equiv \mathbf{b}_k(\delta_k, d_k) = (1 - |\delta_k|) \mathbf{e}_k(d_k) + |\delta_k| \mathbf{e}_k(d_k + \sign(\delta_k))
\]

\[
\mathbf{1} \equiv [1, \ldots, 1] \mathbf{0}_{2C-d_k}^T
\]

\[
\mathbf{1} \equiv [0_{d_k}^T, 1, \ldots, 1] \mathbf{0}_{C-d_k}^T
\]

\[
\mathbf{T} \equiv [0_{C-d_k}^T, 1, \ldots, 1]^T
\]

where

\[
\mathbf{g}_k(d_k) \equiv \begin{bmatrix} s_k[C - d_k + 1], \ldots, s_k[C] \mathbf{0}_T^T \end{bmatrix}^T
\]

\[
\mathbf{c}_k(d_k) \equiv \begin{bmatrix} 0_{d_k}, s_k[1], \ldots, s_k[C] \mathbf{0}_T^T \end{bmatrix}^T
\]

\[
\mathbf{c}_k(d_k) \equiv \begin{bmatrix} 0_{C-d_k}, s_k[1], \ldots, s_k[C - d_k] \mathbf{0}_T^T \end{bmatrix}^T
\]
With a simple manipulation, we can get a compact representation for the data

\[ r_m = Gb_m + n_m \]  

(6)

The \( 2C \times (K+1) \) dimensional code matrix \( G \) contains the code vectors and path strengths, while the \((3K+1)\)-vector \( b_m \) contains the symbols and external interference:

\[ G \overset{\text{def}}{=} [ \cdots a_k \mathbf{g}_k(d_k) a_k \mathbf{g}_k(d_k) a_k \mathbf{g}_k(d_k) \cdots ]^{\top} \]  

(7)

\[ b_m \overset{\text{def}}{=} [ \cdots b_{k(m-1)} b_{km} b_{k(m+1)} \cdots ]^{\top} \]  

(8)

Notice that \( m \)th symbols are included in three successive vectors \( b_{m-1}, b_m \) and \( b_{m+1} \). Hence, we say that these vectors are ‘early’, ‘middle’ and ‘late’ parts of \( m \)th symbols respectively. Recall that without jamming, the dimensions of the code matrix \( G \) and symbols vector \( b_m \) would be equal to \( 2C \times 3K \) and \( 3K \) respectively.

Without losing any generality, we assume that the delay of a user \( k \) is pre-estimated to be equal to \( d_k \), that is, \( \tau_k = d_k \), even though \( d_k + \delta_k \) would be the correct one. Consequently, \( \delta_k \) is considered as an residual delay estimation error after delay tracking circuitry. It is thus typical to consider this error to obey zero mean Gaussian distribution.

Recall that conventional single user detection makes a decision according to

\[ \hat{b}_k = \text{sign}(a_k^{\top}r_{k}^{\top}c_k(\tau_k)) \]  

(9)

from which we see that there is a timing misalignment equal to \( \delta_k \) between the local replica of the spreading code and that of received data.

3. INDEPENDENT COMPONENT ANALYSIS

Independent component analysis (ICA) [6] is a statistical method for searching independent random variables from a set of observed linear combinations of them. One of the main applications of ICA is BSS, which has become an attractive field of research in statistical signal processing and neural network communities. ICA performs purely in a blind manner, that is, without any explicit knowledge of original source variables or a mixing process. ICA relies on assumption of statistical independency of source variables or signals. Although independency is a very strong assumption from a theoretical point of view, it is often quite a realistic assumption in practice. Consequently, ICA has drawn a lot of attention in various application fields lately.

3.1. Standard linear ICA for non-oversaturated data

In complex valued linear ICA, the \( m \) observed signals \( x_1(t), \ldots, x_m(t) \) at the time instant \( t \) are assumed to be linear combinations of \( n \) unknown but statistically independent, complex valued source signals \( s_1(t), \ldots, s_n(t) \) at the time \( t \). By introducing the data vector \( x(t) = [x_1(t), \ldots, x_m(t)]^T \) for the observed signals, and the source vector \( s(t) = [s_1(t), \ldots, s_n(t)]^T \) for the source signals, the instantaneous noisy linear ICA model is given by

\[ x(t) = As(t) + n(t) \]  

(10)

Here the unknown but constant \( m \times n \) mixing matrix \( A \) contains the complex mixing coefficients, and \( n(t) \) denotes the additive noise vector at time \( t \). In addition, \( A \) is assumed to have a full rank, which also implies that the number of sources \( n \) should be less than or equal to the number of observations \( m (n \leq m) \).

Notice that the DS-CDMA signal model Equation (6) is readily an ICA model with \( m = 2C \) observations of \( n = 3(K+1) \) source components. This is due to fact that it is reasonable to assume mutually and temporally independent symbols, as well as, that the code matrix \( G \) has a full rank.

In ICA, the source signals \( s(t) \) are estimated using only the observations \( x(t) \) by finding an \( n \times m \) unmixing matrix \( W \). This matrix should be such that the \( n \)-vector \( W x(t) \) recovers the set of original sources as well as possible. Because of the blindness of the problem, the original sources \( s_k \) cannot be recovered fully uniquely. To see this, recall firstly that the source components can only be determined up to some multiplicative constant, since for all complex constants \( \lambda_k \neq 0 \)

\[ x_k(t) = \sum_{k=1}^{K} a_k s_k(t) + n_k(t) \]  

(11)

In other words, only the waveforms of the sources can be estimated. Another indeterminacy of ICA is the order of the estimated independent components. Since a sum is
independent from the order of summation, the source components $s_k$ can be directly estimated only up to some permutation. Thus, with ICA it is possible to find a random vector

$$s' = [\lambda_1 s_{p(1)}, \lambda_2 s_{p(2)}, \ldots, \lambda_K s_{p(K)}]^T \quad (12)$$

in which $\lambda_k, k = 1, 2, \ldots, K$, are some arbitrary complex numbers, $p: \{1, 2, \ldots, K\} \rightarrow \{1, 2, \ldots, K\}$ is some unknown permutation and $s = [s_1, s_2, \ldots, s_K]^T$ is an original source vector. This is proven with mathematical exactness in Reference [18].

For estimating the unmixing (separating) matrix $W$, many different methods have been proposed [6]. Most of these are ICA methods exploiting the statistical independence of the sources, but there also exist other BSS approaches that utilise temporal correlations or non-stationarity of the sources. The mutual performance of these methods depends largely on the validity of the assumptions made on them in the problem at hand. One very promising solution for the linear ICA is a so-called FastICA algorithm [6, 9] due to its simplicity and fast convergence. FastICA is a fixed-point gradient method, and its basic form relies on the sample fourth-order statistics, kurtosis. However, other forms of the algorithm employing more robust non-linear transformations of higher-order statistics have been developed also [6].

### 3.2. ICA for oversaturated data

A major drawback for standard ICA methods is the case where the number of source signals is greater than the number of observations made ($n > m$), in which case standard ICA model does not hold anymore. This problem is commonly known as a ‘more sources than sensors’—problem, or that the data is oversaturated. For example, from the single antenna DS-CDMA reception point of view, in order to have a standard ICA model available in the receiver, the number of users ($K$) can be at most $(2/3)C$, assuming processing window size of two symbols.\(^1\) This holds for a single-path channel without external jammer signals. Considering $L$ independent multipaths, the requirement gets even harder, $K \leq (2/3L)C$. Hence, the number of users in the system must be rather low to make standard ICA applicable.

Separating source signals from oversaturated data is more a difficult task, since mixing model is not invertible anymore [6]. Consequently, typical ICA methods for oversaturated systems are computationally more expensive compared to standard ICA. From the computational complexity point of view, a modification of a FastICA algorithm [19] is one of the most attractive ICA algorithm designed to operate in oversaturated systems. The modification is based on concept of quasiorthogonality, which roughly speaking means that we are now considering a set of nearly orthogonal basis for the representation of the data, thus enabling the increase in dimensionality compared to strictly orthogonal basis. Since the only change to the standard FastICA algorithm is in orthogonalisation, the modified algorithm have the same computational complexity than the standard one.

The iterative algorithm of symmetric FastICA for complex valued data is the following [20]:

1. Center the original data (i.e. make its mean zero).
2. Whiten the centred data as follows: $y_m = \mathbf{Tr}_m$, such that $E\{y_my_m^H\}$ equals the identity matrix.
3. Take a random $2C \times 3(K+1)$ matrix $W = [w_1 \ldots w_3(K+1)]$.
4. Quasiorthogonalise matrix $W$ as in step 6 and normalise its columns.
5. FastICA: For all $i = 1, \ldots, 3(K+1)$, let

$$w_i \leftarrow E\{y_m (w_i^H y_m)^{-1} w_i^H y_m^2\} - 2w_i$$

6. Quasiorthogonalisation: $W \leftarrow \frac{1}{\sqrt{2}} W - \frac{1}{\sqrt{2}} WW^H W$

7. Normalize columns of matrix $W$.
8. Go to FastICA-step until converged.

A whitening (item 2) is a common pre-processing task in BSS algorithms, which makes the remaining separation procedure simpler. It is a linear transform $T$ which de-correlates the observed mixtures, and normalise the components to have unit variances. Such $T$ can be simply found in many ways [6]. One way is use a principal component analysis (PCA), which gives the whitening transform as:

$$T = E^{-1/2} D^H$$

in which matrix $D$ have the principal eigenvectors of the data correlation matrix estimate $R = 1/M \sum_{m=1}^M \mathbf{r}_m \mathbf{r}_m^H$ as columns, and the diagonal matrix $E$ contains the corresponding eigenvalues on its diagonal. Since relative magnitudes of the eigenvalues are, in practice, directly proportional to the strengths of independent components,\(^5\) the whitening by PCA tends to prefer the strong

---

\(^1\)In general, using processing window size of $N$ symbols the upper bounds for $K$ become $(N/(N+1))C$ and $N/(N+1)L)C$.

\(^5\)Recall that ICA model Equation (10) is not unique with respect to scaling. Consequently, a definition of independent component’s strength is not self-explanatory. Here, however, the strength is essentially defined by received power of the user. See Reference [21] for a rigorous definition of component strength.
components, and some weak components might be lost [21]. In general, this is an undesirable phenomenon, since BSS algorithms should not be sensitive to the differences in component strengths. Nevertheless, a subtractive BSS structure, which is considered in this paper, benefits of this phenomenon as will be seen in the numerical experiments.

In addition, the numerical experiments show that the use of quasiorthogonal basis in ICA, alone, helps to achieve tolerable performance. However, the performance improvement is still only moderate, and additional means are needed. In the following, we propose one such mean after which the performance is greatly improved.

4. SUCCESSIVE INTERFERENCE CANCELLATION USING BSS/ICA

By an interference subtractive receiver, we loosely speaking mean an iterative multiuser receiver, where the estimated interference is subtracted from the received signal prior to the estimation of a particular user. In conventional successive interference cancellation, users are detected one by one and subtraction of interference due to the preceding user is performed before the next user is detected. In the following, more advanced group-wise SIC scheme utilising ICA is considered.

4.1. Receiver structure

The principle of ICA-assisted SIC receiver is quite straightforward. The received signal Equation (6) is first separated using ICA. Recall that a single-path channel was assumed because of simplicity. However, it was shown in Reference [17] that increasing the number of propagation paths does not have an essential influence on ICA separation. In fact, it was found out that a greater number of paths result in a slightly more reliable separation compared to single-path scenario.

After separation, a user identification is needed, since users can be separated only up to unknown permutation by ICA. The users are identified according to their spreading codes. Hence, the receiver is semi-blind. Strictly speaking, for the user identification a correlation metric

$$\rho(k,k') = \frac{c_k(\hat{\tau}_k)^H w_{k'}}{|c_k(\hat{\tau}_k)||w_{k'}|}$$

(14)

is computed for each $k, k'$, where $w_{k'}$ corresponds to the ICA basis vector of $k'$th source. Next, a threshold, $\rho_t$, is set to indicate proper detection. This is needed because subtraction of erroneously detected signal would actually enhance interference. Thus, only the signal corresponding to the user $k$ with $|\rho(k,k')| > \rho_t$ for some $k'$ is subtracted from the received signal.

If some independent external interference signals are present, ICA considers them as additional source components and separates them as well. Understandably, these components cannot be identified, since known code sequences are not associated with them. Nevertheless, identification of undesired interference components is not necessary. Matter of substance is that the desired components, that is, components due to users’ data, are not affected by external interference thanks to ICA separation.

Notice that correlation $\rho(k,k')$ includes also a phase ambiguity (coefficient $\lambda_k$ in Equation (12)) produced by ICA. Namely, assuming that index $k'$ corresponds to $k$th user’s source component, we have

$$\rho(k,k') = \frac{c_k(\hat{\tau}_k)^H w_{k'}}{|c_k(\hat{\tau}_k)||w_{k'}|} = \frac{c_k(\lambda_k^{-1} a_k c_k)}{|c_k(\lambda_k^{-1} a_k c_k)|} = \frac{|\lambda_k| a_k}{|\lambda_k| a_k} \rho(k,k')$$

(15)

Hence, ICA’s phase ambiguity can be corrected. The soft decision of $k$th user’s $m$th symbol is obtained as

$$z_{km} = a_k^* \rho(k,k') z_{km}^{ICA} = |\lambda_k| a_k b_{km}$$

(16)

in which $z_{km}^{ICA}$ denotes separated source component which has been identified to correspond to $k$th user.

It is of primary importance to see that what is actually subtracted from the original data is a tentative decision of the form $a_k c_k(\hat{\tau}_k)b_{km}$ rather than $w_{k'} b_{km}$, in which $b_{km} = \text{sign}(\text{Re}\{z_{km}\}) + i \text{sign}(\text{Im}\{z_{km}\})$. Hence the ICA solution is first refined (according to the knowledge of the parametric form of the mixing matrix) before subtraction is performed.

After subtraction of all qualified users’ signals, the next ICA separation is performed for the interference-subtracted signal. The order of ICA model decreases in subtraction helping ICA to recover the remaining users. The procedure is repeated successively until all users are detected. The successive ICA receiver is illustrated in Figure 1.

4.2. Tracking delays in ICA-SIC

Recall that ICA performs purely in blind manner. The first occasion where some prior knowledge of the users is needed is the user identification phase Equation (14). Naturally, the timing information can also be used to generate a good initial value for ICA iterations, and hence speed up the separation [9]. Anyway, what a lousy timing estimate ultimately does is that it worsen the user
identification process. More importantly, given that a user is, nevertheless, identified having an erroneous timing estimate at the receiver, the subtraction of that user will enhance the interference the more the bigger is the timing inaccuracy.

To avoid the performance degradation during the subtraction phase, the delay of each identified user could first be refined (assuming single-path channel) according to

$$\hat{\delta}_k \leftarrow \operatorname{arg\,max}_{\delta} \frac{g_k(\delta, \tilde{\tau}_k)H w_{\delta}}{\|g_k(\delta, \tilde{\tau}_k)\| \|w_{\delta}\|}$$

(17)

where the maximisation of the correlation is performed in a close neighbourhood of $\tilde{\tau}_k$ for example $\delta \in (-1/2, 1/2)$. This is a sensible way to track delays since the timing information is implicitly included in the ICA solution $w_{\delta}$. In other words, also the delay tracking procedure takes advantage of interference suppression due to ICA separation. After tracking, the signal corresponding to the user $k$ is rebuilt and subtracted:

$$r_m \leftarrow r_m - a_k \left( b_{k,m-1}g_k(\hat{\delta}_k, \tilde{\tau}_k) + b_{k,m+1}g_k(\hat{\delta}_k, \tilde{\tau}_k) \right)$$

(18)

Notice that delay tracking Equation (17) is performed after user identification. It is thus of interest whether the user identification may fail even if ICA would give a adequate estimate $w_{\delta}$. This kind of failure in the identification process may easily happen if the timing error at the receiver is too large, giving a low correlation value for the user identification metric Equation (14). To illustrate that, Figure 2 presents values of

$$\xi_k(\delta) = \frac{|c_k(\delta_k)H g_k(\delta, \delta_k)|}{|c_k(\delta_k)||g_k(\delta, \delta_k)||}$$

(19)

in case of 31-Gold codes [22]. The figure shows that, with $\rho_k = 0.9$, a timing error of an order of one-third of the chip duration results in identification failure in ideal case. This means, for example, that in case of normally distributed error with variance 0.05, an average of 10% of identification attempts fails due to the timing error at the receiver. Consequently, corresponding users’ components are discarded in the subtraction.

A straightforward modification to the user identification procedure is proposed in order to prevent the user identification failures due to inaccuracies in timing at the receiver.
As already mentioned, the separation is not affected by inaccurate timing estimates since ICA is totally blind. This is also seen from the fact that the sources $b_n$ does not include timing information. Hence, the exact knowledge of the timing is included implicitly in the ICA solution $\mathbf{w}_k$. Therefore, the user identification procedure is modified simply by replacing Equation (14) with a so-called code pre-tracking:

$$\rho(k, k') = \max_{\delta^m \in \Delta} \frac{\mathbf{r}_k^H (\delta^m, \tau_k) \mathbf{w}_{k'}}{||\mathbf{g}_k (\delta^m, \tau_k)|| ||\mathbf{w}_k||}$$

in which $\Delta \subset (-1/2, 1/2)$ is some finite set. (Recall that actual delay estimation error $\delta_k$ is assumed to be in $(-1/2, 1/2)$ for all $k$.) Selection of the set $\Delta$ depends on the used spreading codes and the value of identification threshold $\rho_1$. For example, in the case of the 31-Gold codes and $\rho_1 = 0.9$, a reasonable choice is $\Delta = \{-1/4, 0, 1/4\}$. This prevents identification failures due to timing errors resulting yet only insignificant increase in computational load. In addition, the computational increment can be easily compensated in the code post-tracking Equation (17), since now it is enough to perform the maximisation of correlation in $(\delta_k^{pre} - 1/4, \delta_k^{pre} + 1/4)$ only, where $\delta_k^{pre}$ is an argument of the maximum in Equation (20).

5. NUMERICAL EXPERIMENTS

The successive ICA receiver (ICA-SIC) is evaluated numerically against conventional SIC and PIC receivers as well as more an advanced LMMSE and LMMSE-PIC receivers [14]. Also, so called 'Basic ICA' receiver [23] is used as a reference. In that receiver users are separated using ICA only once and identified by their spreading codes, but no interference subtractions are performed. The quasi-orthogonalisation based version of the FastICA algorithm is used in both ICA-SIC and basic ICA, since it is able to operate in over-saturated systems.

Each user is supposed to send data blocks of $M = 5000$ QPSK symbols. Symbols are spread using Gold codes of length $C = 31$. SNR is fixed to 20 dB with respect to the weakest user. In ICA-SIC, value $\rho = 0.9$ is used for the identification threshold. All results are based on average bit error rates (BER) over 1000 independent repetitions. BER values are computed with respect to all $K$ users.

5.1. Effect of user load

The first result demonstrates that ICA-assisted methods maintain their competence also in highly loaded systems. Especially, the proposed ICA-SIC outperforms conventional SIC and PIC clearly even if users’ timing information is perfectly known. This can be seen in Figure 3. ICA-SIC also performs better than non-subtractive LMMSE receiver as seen in Figure 4. On the other hand, with accurate timing, LMMSE-PIC receivers outperforms ICA-SIC. However, the performance gain decreases as the number of users increases.

Especially, the proposed ICA-SIC outperforms conventional SIC and PIC clearly even if users’ timing information is perfectly known. This can be seen in Figure 3. ICA-SIC also performs better than non-subtractive LMMSE receiver as seen in Figure 4. On the other hand, with accurate timing, LMMSE-PIC receivers outperforms ICA-SIC. However, the performance gain decreases as the number of users increases.

The number of FastICA-iterations plays a significant role in ICA-assisted methods. On one hand, too small number of iterations impairs a performance, and on the other hand, increasing the number of iterations increases the computational load of the methods. The experiments were repeated with different numbers of FastICA-iterations to find out an optimal number of iterations. Figures 5 and 6 illustrate performances of basic and successive ICA receivers respectively. Conventional PIC is used again as a reference. The figures reveal an other favourable feature of successive ICA-receiver compared to basic ICA-receiver. In addition to clear performance improvement, the successive receiver converges (with respect to FastICA-iterations) even faster than the basic receiver. Figure 7 demonstrates how many iterations...
Figure 4. BERs as a function of number of users in case of perfect delay estimation. The \( K \) users are assumed to be in two equal sized service classes, which is modelled as a power difference of 10 dB between the two user groups. Inside both groups all the users are assigned the same power. In addition, 110 FastICA iterations are used in successive ICA receiver.

Figure 5. BERs of Basic ICA-receiver (solid line) and conventional PIC receiver with five cancellation stages (dashed line) as a function of number of FastICA iterations. The \( K \) users are assumed to be in two equal sized service classes, which is modelled as a power difference of 10 dB between the two user groups. Inside both groups all the users are assigned the same power. No delay estimation error.

Figure 6. BERs of Successive ICA-receiver (solid line) and conventional PIC receiver with five cancellation stages (dashed line) as a function of number of FastICA iterations. The \( K \) users are assumed to be in two equal sized service classes, which is modelled as a power difference of 10 dB between the two user groups. Inside both groups all the users are assigned the same power. No delay estimation error.

Figure 7. Number of FastICA iterations needed for certain BER-level in successive ICA-receiver. The \( K \) users are assumed to be in two equal sized service classes, which is modelled as a power difference of 10 dB between the two user groups. Inside both groups all the users are assigned the same power. The figure is obtained by linear interpolation from Figure 6. No delay estimation error.
successive ICA-receiver needs in order to achieve BER levels of $10^{-2}$ and $10^{-3}$.

BERs of the successive ICA-receiver seem to deteriorate slightly after certain number of iterations. This is best seen in cases of 24–28 users (Figure 6). Deterioration is due to increasing number of interference subtractions after first ICA-separation. Like in standard SIC and PIC-methods, detection errors in tentative symbol estimates inflict extra interference into following separation rounds.

5.2. Effect of timing inaccuracies

The following results demonstrate that the proposed ICA-SIC receiver is eminently immune to delay estimation errors. This can be seen in Figures 8 and 9. If the delays are known perfectly, LMMSE-PIC-receiver outperforms successive ICA-receiver. However, the performance of LMMSE-PIC weakens very fast as variance of delay error increases. On the contrary, the performance of ICA-SIC receiver does not suffer from the increment. A comparison of these two figures reveals an influence of pre-delay tracking. The pre-tracking is not performed in Figure 8 which causes a deterioration of performance due to user identification failures if timing inaccuracy is great enough. In Figure 9, the pre-tracking is performed. Consequently,

Figure 8. BERs as a function of variance of delay estimation error. The $K=22$ users are assumed to be in two equal sized service classes, which is modelled as a power difference of 10 dB between the two user groups. Inside both groups all the users are assigned the same power. In addition, 110 FastICA iterations are used in successive ICA receiver.

Figure 9. BERs as a function of variance of delay estimation error in a system of $K=30$ users. Two equal sized service classes are assumed, which is modelled as a power difference of 10 dB between the two user groups. Inside both groups all the users are assigned the same power. One hundred forty FastICA iterations are used in successive ICA receiver.

Figure 10. BERs as a function of number of FastICA iterations. The $K=22$ users are assumed to be in two equal sized service classes, which is modelled as a power difference of 10 dB between the two user groups. Inside both groups all the users are assigned the same power. Users’ estimated delays are assumed to have an error $\delta_t$ with variance $\sigma_t^2 = 0.01$ (unit of error corresponds to one chip time).

Copyright © 2005 AEIT

Euro. Trans. Telecomms. 2006; 17:577–589
the performance does not fall off even with the greatest values of variance. This shows that also the user identification is now accurate despite of timing errors.

Figure 10 depicts the performance of successive ICA-receivers as a function of the number of FastICA iterations performed. The figure shows that both successive ICA-receivers overtake for example LMMSE-PIC-receiver very fast when timing inaccuracies exist. Finally, we notice from the experiments that the FastICA with successive interference cancellation performs clearly better than FastICA alone also after the delays are tracked. This reveals the importance of additional parameter tracking loops in successive ICA separations.

5.3. Effect of received power distribution

The influence of users’ power distribution is shown in Figures 11 and 12. First, the case of two different service classes, modelled as a power difference of 10 dB between the two user groups is depicted in Figure 11. ICA-SIC receiver provides best performance when the power groups have approximately the same size (in number of users). When the number of either stronger or weaker users increases, the performance decreases rapidly. Second, in case of uniformly distributed received powers, ICA-SIC performs better with greater ratio between the strongest and the weakest user (Figure 12).

Pre-processing the received data by PCA explains the behaviour of performance in these two example. As stated above, if users have different powers, PCA processing prefers relatively strong users. This means, in practice, that these users are detected more likely after ICA separation. On the other hand, subtractions are performed for original, un-preprocessed data, which is not affected by PCA. Hence, originally relatively weak users are intact after the subtractions. How these users can be detected after following ICA separations depend on their relative strengths and number. This all is to say, that in a sense, PCA pre-processing hides the weak users in early stages whereas subtractions reveals them again in the later stages. Therefore, ICA-SIC receiver provides the best performance gain when number of users having (nearly) equal received powers is small enough.

5.4. Effect of external interference

Finally, the system with external jammer signal is considered. The jammer \( f_p(t) \) is modelled as in Equation (1). Figure 13 shows that ICA-SIC outperforms LMMSE-PIC receivers clearly. Notice that a performance gain is significant also with relatively weak jammer, say \( \text{SIR} \in \{10, 20\} \).
This could correspond the situation where some other narrowband interference suppression method is first used but some residual part of interferer’s power is still present. Beyond the SJR of 20 dB, the performance of ICA-SIC settles on its interference-free level.

A bad performance of ICA-SIC in a presence of very strong jammer signal ($SJR \approx 2 \frac{C_0}{C_{138}}$) is again explained by PCA pre-processing. In the first stage, ICA is likely able to recover only the source components due to jammer, since user components are relatively weak and lost in PCA processing. Because jammer components cannot be identified as explained in Section 4, nor can they be reconstructed and subtracted. Consequently, performing more ICA separations does not help in this case.

6. CONCLUSIONS

In this paper, we introduced a new SIC scheme in which BSS techniques based on independent component analysis (ICA) were considered. Our main emphasis was on the situation of oversaturated data which easily arises for example in DS-CDMA systems when the number of active connections/users is high compared to number of chips per bit being used, that is, a highly loaded system. It was shown how inherent properties of BSS/ICA help in a mitigation of both internal multiaccess interference and external interference sources, like adjacent-cell-interference or (intentional) jamming. It was also shown how BSS/ICA solution can be used to refine tentative timing estimates by adding a novel delay tracking circuitry between source separations and subtractions procedures. This results in a receiver structure which is, in practice, an immune to inaccurate timing estimation.

Most important result was greatly improved interference suppression capability due to novel combination of SIC ideology and ICA in case of oversaturated data. This can also be seen as a way to somewhat circumvent the ‘more sources than sensors’ problem, which is a challenging problem in BSS area. The best performance was obtained when users’ received powers were unevenly distributed. This is due to pre-processing by principal component analysis, which in a sense hides the weak users in early separation stages, and successive subtractions, which reveals them again in the later stages.

Numerical experiments with DS-CDMA uplink data were given to illustrate the achieved gains in capacity and BER performance compared to conventional successive and parallel interference cancellation schemes as well as more an advanced LMMSE-PIC receiver and variants of BSS/ICA techniques alone.

REFERENCES


Copyright © 2005 AEIT

Euro. Trans. Telecomms. 2006; 17:577–589


**AUTHORS’ BIOGRAPHIES**

**Toni Huovinen** was born in Riihimäki, Finland, in 1977. He received his M.Sc. in Telecommunications in 2003 from the University of Jyväskylä (JYU). During 2003–2004, he worked as a researcher in Department of Mathematical Information Technology at JYU. Currently he is working as a researcher in Institute of Communications Engineering at Tampere University of Technology, where he is aiming to doctoral degree. His research interest includes statistical signal processing for communications and especially for interference cancellation.

**Tapani Ristaniemi** was born in Kauhava, Finland, in 1971. He received his M.Sc. in Mathematics in 1995, Ph. Lic. in Applied Mathematics in 1997, and Ph.D. in Telecommunications in 2000 from the University of Jyväskylä, Jyväskylä, Finland. During 2001–2003 he was a professor of telecommunications at the Department of Mathematical Information Technology, University of Jyväskylä. In 2003, he joined the Institute of Communications Engineering in the Tampere University of Technology, Finland, where he is a professor of wireless data communications. His research interests include signal processing for communications and radio resource management for wireless networks.
BLIND INTERFERENCE CANCELLATION SCHEMES FOR DS-CDMA SYSTEMS

by

Kartikesh Raju, Toni Huovinen and Tapani Ristaniemi

Blind Interference Cancellation Schemes for DS-CDMA Systems

Karthikesh Raju¹, Toni Huovinen², and Tapani Ristaniemi²

¹ Neural Networks Research Centre, HUT, Finland
² Institute of Communications Engineering, TUT, Finland

Abstract

In this paper, we consider the use of Independent Component Analysis (ICA) [1] to circumvent the main bottlenecks of conventional methods for the DS-CDMA interference cancellation problem. In the uplink, we present a single antenna interference cancellation scheme [2] wherein internal MAI as well as external interference is mitigated using a combination of ICA and the principle of conventional successive interference cancellation (SIC). For the downlink case, we look at an antenna array scheme [3], where ICA acts as a pre-filter to conventional detection, separating the information component from the interference component and thus providing an interference-free signal to conventional receivers.

Introduction

Interference cancellation is a vital task in Direct Sequence Code Division Multiple Access (DS-CDMA) systems. Successive and parallel interference cancellation schemes (SIC and PIC, resp.) are well-known conventional methods to mitigate the effects of multiple access interference (MAI) in the uplink receivers, where the spreading codes of the interfering users are known. These suboptimal multiuser detection (MUD) solutions are often considered as reasonable alternatives to optimal MUD [4] since they are of lower complexity. However, in the case of high systems load the amount of MAI makes their performance poor. Delay estimation errors form another problem. Recall that SIC and PIC are subtractive interference rejectors and hence they require accurate timing estimation to regenerate the interfering users to prevent interference enhancement. In the downlink, on the other hand, conventional detection inherently mitigates interference by the amount of a processing gain. An increase in the processing gain, however, leads to bandwidth expansion. The need for increased processing gains can be relaxed by the use of multiple antenna sensors by which beamforming techniques can be used for cancelling the interfering signal. Unfortunately, most beamforming methods require that the look direction is well separated and that the knowledge of the array response in the desired signal direction is known. They also fail to perform optimally when the interfering signals are coherent with the desired signal [5].

ICA Based Uplink Interference Cancellation

Among many applications for ICA [1], recent studies have considered ICA based successive interference cancellation (ICA-SIC) [2], by which we loosely speaking mean an iterative multi-user receiver, where the estimated interference is subtracted from the received signal prior to the estimation of a particular user.
Here we briefly summarize the algorithm. For more details, see [2]. First, the received signal is separated by ICA, say, FastICA algorithm, with the quasiorthogonal basis in ICA. The concept of quasiorthogonality is needed, since highly loaded system in terms of DS-CDMA is actually over-complete in terms of ICA (more sources than observations). After separation, users are identified by the help of their spreading codes. Hence, the receiver is semi-blind.

Next, a pre-determined threshold is set to indicate proper tentative detection. This is needed because subtraction of erroneously detected signal will actually enhance interference. Thus, the signal corresponding to a particular user is subtracted from the received signal only if the identification is proper. After subtraction of all such users’ signals, the next ICA separation is performed for the interference-subtracted signal. The order of ICA model decreases in subtraction, which help ICA to recover the remaining users. The procedure is repeated successively until all users are detected.

Recall that ICA performs purely in blind manner. The first occasion where some prior knowledge of the users is needed is the user identification phase. In addition, given that a user is, nevertheless, identified having an erroneous timing estimate at the receiver, the subtraction of that user will enhance the interference the more the bigger is the timing inaccuracy. However, a correct timing information is implicitly included in the ICA solution and can be, therefore, tracked in a sensible way. We term the methods for such tracking as a coarse code pre-tracking and a finer code post-tracking, see [6] for the details.

**Post-Switch ICA-RAKE for downlink interference cancellation**

Another attractive application of ICA is its use in the downlink array receivers [3]. In such receivers, ICA is first performed for the array data to separate the desired information from the other interferences. After that, conventional detection is performed for selected source. This leads to the basic ICA-RAKE receiver structure.

The basic ICA-RAKE receiver, however, is still somewhat impractical as such. This is because it employ an ICA block even though it may not always be desirable. In fact, if the interference or jammer signal is weak or even absent, the additional ICA jammer suppression block might even cause additional interference to the information bearing signal. Therefore, we improve the basic ICA-RAKE receiver by introducing two a novel switching strategy, called a post-switching scheme, which select either ICA-assisted or conventional RAKE detection in the array receiver chain. Shortly, the final output of the hybrid receiver is chosen to be that branch which results in higher correlation with the training symbols.

**Numerical Experiments**

In this section we give a couple of numerical examples of the performance of the receiver structures shown above. In the uplink case, the system had $K = 30$ users, all of whose data were spread using Gold Codes of length $C = 31$. In the downlink
Figure 1: Performance characteristics of ICA-SIC uplink receiver.

case, the system had up to $K = 31$ users, though at any point in time only a few (8-16) were simulated. The data block was of size $M = 5000$ QPSK symbols. All the results are based on average bit error rates (BER) over 1000 independent runs. BER values are computed with respect to all $K$ users. Additionally, in the downlink the channel was either a $L = 1$ or $L = 2$ path channel.

The ICA-SIC method was compared to LMMSE-PIC [7] receivers. It estimates each user with an optimal linear MMSE detector, after which subtraction follows. First of all, it was observed (Fig. 1) that the ICA-SIC receiver is eminently immune to delay estimation errors. If the delays are known perfectly, LMMSE-PIC-receiver outperforms successive ICA-receiver. However, the performance of LMMSE-PIC weakens very fast as variance of delay error increases. On the contrary, the performance of the proposed BSS-SIC receiver does not suffer from the increment at all. This shows that the user identification is accurate despite of timing errors.

The system with external jammer signal was also considered. Fig. 1 shows that ICA-SIC outperforms LMMSE-PIC receivers clearly. Notice, that a performance gain is significant also with relatively weak jammer, say SJR $\in [10,20]$. This could correspond the situation where some other narrowband interference suppression method is first used but some residual part of interferer’s power is still present. Beyond the SJR of 20 decibels, the performance of ICA-SIC settles on its interference-free level.

Next the results for interference suppression in the downlink case are shown. Figure 2 shows that even when the jammer power is quite low, ICA offers about 2.5dB gain compared to RAKE (e.g. 10-15dB SJR). In the case of multipath, the jammer induced interference component is separated by ICA. Since final post processing is done by a conventional RAKE receiver, there is a certain loss in performance. The effect of an additional multipath can be seen only after $-15$dB, before which the presence of a strong jammer helps in offsetting the inter path interferences.

Fig. 2 also shows the distribution of correct bits in an $L = 1$ and an $L = 5$ coherent path channel. The jammer power was $-30$dB with SNR=$20$dB. The block size was $M = 5000$ QPSK symbols. With a single path pre-processing ICA almost gets rid
of the jammer. Most of the blocks have at least 95% correct bits. Without pre-
processing in the RAKE receiver most blocks have at most 10% correct bits, with
the maximum number of correct bits at around 65%. When the number of coherent
paths is increased to $L = 5$, most blocks have at least 70% correct bits in the case of
ICA-RAKE while conventional detection is not capable of handling a strong jammer
along with interpath interference.

![Figure 2: Performance characteristics of ICA-RAKE downlink array receiver.](image)

**References**


beamformer in the presence of correlated sources and its behavior under spatial

[6] T. Ristaniemi and T. Huovinen, “Joint delay tracking and interference cancel-
lation in DS-CDMA using successive ica for oversaturated data,” in *Proc. 5th

[7] J. Lilleberg M. Latva-aho, “Parallel interference cancellation in multiuser detec-
JOINT MULTIPATH DELAY TRACKING AND INTERFERENCE CANCELLATION IN DS-CDMA SYSTEMS USING SUCCESSIVE ICA FOR OVERSATURATED DATA

by

Tapani Ristaniemi and Toni Huovinen

Joint Multipath Delay Tracking and Interference Cancellation in DS-CDMA Systems Using Successive ICA for Oversaturated Data

Tapani Ristaniemi and Toni Huovinen
Institute of Communications Engineering, Tampere University of Technology
P.O.Box 553, FIN-33101, Tampere, Finland
Email: {Tapani.Ristaniemi,Toni.Huovinen}@tut.fi

Abstract—A recent idea is to use blind source separation (BSS) methods within successive interference cancellation (SIC) in DS-CDMA reception. This has been found to result in an interference cancellation solution, that is able to i) mitigate different kinds of interference sources (e.g. multi-access, multipath and out-of-cell interferences) simultaneously, ii) tackle "more sources than observations"-problem and, consequently, operate in highly loaded systems and iii) deal with imperfect channel estimation better than conventional BSS receiver and conventional interference cancellation. Thus far however, a single path channel has been assumed in studies behind these findings. In this paper we extend our previous studies to multipath channels where inter-path interference hardens the joint delay tracking and interference cancellation problem even more.

I. INTRODUCTION

Interference rejection/cancellation has been considered one of the most attractive class of suboptimal solutions, and has been studied extensively in the past [1]–[5]. Parallel and successive interference cancellation (PIC and SIC, respectively) are the main categories within this class, describing the procedure by which the interference is subtracted from the original data, after regenerating the interference from the tentatively estimated data. Any interference subtractive receiver performs the better the more accurately tentative decisions are being made. This is because the interference level is then reduced the most. Otherwise, from a DS-CDMA signal point of view, two different situations will occur:

(a) The particular user signal is not completely taken out from the original data or it is even strengthened; or in addition to that,
(b) a fictitious multipath component is generated and added to the original data.

The former happens e.g. if all the other parameters except the (complex-valued) amplitude is correctly estimated. The latter happens if the phase of the spreading code is inaccurately estimated. To mild the consequences of the situation (a) one can e.g. estimate the reliability of tentative decision from the soft decisions. Doing this way one actually chooses in favor of accepting only partial mitigation (which happens almost surely) rather than accepting occasional interference enhancement. In this paper we consider the avoidance of the situation (b), for which the most natural solution is more accurate delay estimation. We also consider the suppression of external interference sources (like adjacent cell interference or intentional jammer) which conventionally is not possible with subtractive interference cancellators.

One relatively new idea is to employ blind source separation (BSS) or, especially, Independent Component Analysis (ICA) [6] in the interference subtractive receivers. In [7] it was shown how the parametric form of the mixing matrix can efficiently be used to refine the ICA solution, and hence avoid interference enhancement while subtracting that source from the original data. The benefit of BSS/ICA processing is especially emphasized in highly loaded systems. Recall that a major drawback for standard ICA is the case where the number of source signals (to be blindly extracted from the received data) is greater than the number of observations made. This is a commonplace situation in communications applications in which cases standard ICA model doesn’t hold anymore. Nevertheless, ICA-SIC -type receiver structure is able to somewhat circumvent the “more sources than observations”- problem in the sense that adequate performance (in terms of bit-error probability) is still achievable even in extremely highly loaded system, whereas conventional parallel and successive interference cancellation only remain at a moderate level [8]. Further, ICA algorithms see possible external interference sources as additional source components, thus also their contribution tends to be separated during estimation of interfering users. Consequently, ICA assisted receivers tolerate inherently external interference.

Further, BSS/ICA-assisted receivers can also cope with erroneous timing estimates (at the expense of negligible increase in computation) [8]. Recall that the timing information, that is, the phase of the band-spread code, should be known accurately to avoid interference enhancement during the subtraction phase. Roughly speaking, the ICA solution implicitly includes the timing information, thus giving a possibility to estimate it due to the fact that the parametric form of the mixing matrix is known.

In this paper we extend the receiver to operate in multipath channels which makes the problem even more difficult. However, from ICA point of view, introducing (fixed) multipaths just change a mixing process in the data model which is not expected to be a too serious problem for overcomplete ICA.
In fact, it was found out in [9] that for standard ICA (i.e. non-overcomplete) model a greater number of paths results in a slightly more reliable ICA separation compared to single-path scenario. Hence, using overcomplete ICA within SIC is expected to perform well also in highly loaded multipath system. However, conventional parallel and successive interference cancellation schemes as well as advanced LMMSE-PIC [10] receiver are expected to clearly suffer from multipath channel.

II. SIGNAL MODEL

Studied channel is DS-CDMA uplink with fixed multipaths and additive white gaussian noise (AWGN) [11]. In addition, channel is assumed to contain an external (non-Gaussian) interference signal \( j_p(t) \), which is pulsed at the symbol level with probability \( p \). (Probability \( p = 0 \) coincides absence of the external interference.) Hence, the received signal has the form

\[
r(t) = \sum_{k=1}^{K} \sum_{m=1}^{M} b_{km} \sum_{l=1}^{L} a_{kl} s_k(t - mT - \tau_{kl} T_c) + \eta(t) + j_p(t),
\]

in which \( b_{km} \) is the \( m \)th symbol sent by user \( k \in \{1, \ldots, K\} \). Each symbol travels through \( L \) paths. The complex coefficient of the \( k \)th user’s \( l \)th path is denoted by \( a_{kl} \), which is assumed to remain the same during the data block of, say, \( M \) symbols. \( s_k(t) \) is \( k \)th user’s binary chip sequence, supported by \([0, T)\), in which \( T \) is the symbol duration. \( T_c \) is the chip duration. The user delay is denoted by \( \tau_{kl} = d_{kl} + \delta_{kl} \), in which \( d_{kl} \in [0, \ldots, C - 1] \), \( C \) is number of chips in the spreading code and \( \delta_{kl} \in (-\frac{1}{2}, \frac{1}{2}] \). Also the delays are assumed to remain constant during the block of \( M \) data symbols. \( \eta(t) \) denotes noise.

The discrete-time counterpart of the data is well known to be of the form

\[
r_m = G b_m + \eta_m,
\]

in which \( 2C \times 3(K + 1) \) dimensional code matrix \( G \) contains the code sequences and path strengths, while the \( 3(K + 1) \)-vector \( b_m \) contains the symbols and a contribution of the external jammer signal. That is, using processing window size of two symbols we have that

\[
G \overset{\text{def}}{=} \begin{bmatrix} \cdots & \sum_{l=1}^{L} a_{kl} g_{kl} & \cdots \\ \sum_{l=1}^{L} a_{kl} g_{kl} & \cdots & \sum_{l=1}^{L} a_{kl} g_{kl} \\ \sum_{l=1}^{L} a_{kl} g_{kl} & \cdots & \sum_{l=1}^{L} a_{kl} g_{kl} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots 
\end{bmatrix} \quad \text{and} \quad b_m \overset{\text{def}}{=} \begin{bmatrix} \cdots & b_{km-1} b_{km} b_{km+1} \cdots \\ \cdots & j_p((m-1)T) j_p(mT) j_p((m+1)T) \cdots 
\end{bmatrix}.
\]

In (3), column vectors of length \( 2C \) are defined as

\[
\begin{align*}
\mathbf{g}_{kl} & \overset{\text{def}}{=} (1 - |d_{kl}|) \mathbf{g}_s(\hat{\tau}_{kl}) + |d_{kl}| \mathbf{g}_s(\tau_{kl} + \text{sign}(d_{kl})) \\
\mathbf{g}_s & \overset{\text{def}}{=} (1 - |d_{kl}|) \mathbf{g}_s(\hat{\tau}_{kl}) + |d_{kl}| \mathbf{g}_s(\tau_{kl} + \text{sign}(d_{kl})) \\
\mathbf{g}_{kl} & \overset{\text{def}}{=} (1 - |d_{kl}|) \mathbf{g}_s(\tau_{kl}) + |d_{kl}| \mathbf{g}_s(\hat{\tau}_{kl} + \text{sign}(d_{kl})) \\
1 & \overset{\text{def}}{=} \begin{bmatrix} 1 & \cdots & 1 \\
1 & \cdots & 1 \\
1 & \cdots & 1 \\
\end{bmatrix}
\end{align*}
\]

where, \( \hat{\tau}_{kl} \in \{0, 1, \ldots, C\} \) is a delay estimate of \( k \)th user’s \( l \)th path and \( \delta_{kl} = \hat{\tau}_{kl} - \tau_{kl} \) is an estimation error in delay. In addition, code vectors of length \( 2C \) are given as

\[
\begin{align*}
\mathbf{c}_{kl}(d) & \overset{\text{def}}{=} [s_k(C - d + 1) \ldots s_k(C)]^T (d) \\
\mathbf{c}_s & \overset{\text{def}}{=} [0_3^T \ s_k(C)]^T \\
\mathbf{c}_{kl} & \overset{\text{def}}{=} [0_3^T s_k(C)]^T
\end{align*}
\]

Hence in model (2), the effect of multipaths is absorbed in matrix \( G \). Notice also that \( m \)th symbols are included in three successive vectors \( b_{m-1}, b_m \) and \( b_{m+1} \) in (4). Hence, we say that these vectors are “early”, “middle” and “late” parts of \( m \)th symbols, respectively. Recall that without jamming, the dimensions of the code matrix \( G \) and symbols vector \( b_m \) would be equal to \( 2C \times 3K \) and \( 3K \), respectively.

Without loosing any generality we assume that the delay of \( k \)th user’s \( l \)th path is pre-estimated to be equal to \( d_{kl} \), that is, \( \hat{\tau}_{kl} = d_{kl} \), even though \( d_{kl} + \delta_{kl} \) would be the correct one. Consequently, \( \delta_{kl} \) is considered as an residual delay estimation error after delay tracking circuitry. It is thus typical to consider this error to obey zero mean Gaussian distribution.

The DS-CDMA signal model (2) is readily a linear noisy ICA model with \( m = 2C \) observations of \( n = 3(K + 1) \) source components [6].

III. BSS/ICA BASED SUCCESSIVE INTERFERENCE CANCELLATION

By an interference subtractive receiver we loosely speaking mean an iterative multi-user receiver, where the estimated interference is subtracted from the received signal prior to the estimation of a particular user. The principle of this kind of receiver [8] utilizing BSS/ICA is straightforward. First, the received signal is separated by ICA. After separation, users are identified by correlating their spreading codes with ICA basis vectors \( u_k \). Hence, the receiver is semi-blind. Next, a threshold, \( \rho_1 \), is set to indicate proper detection. This is needed because subtraction of erroneously detected signal would actually enhance interference. Thus, only the signal corresponding to the user \( k \) with identification metric \( \rho(k, k') > \rho_1 \) for some \( k' \) is subtracted from the receiver signal. After subtraction of all such users’ signals, the next ICA separation is performed for the interference-subtracted signal. The order of ICA model decreases in subtraction, which help ICA to recover the remaining users. The procedure is repeated successively until all users are detected. It is of primary importance to see that the ICA solution is first refined (according to the knowledge of the parametric form of the mixing matrix) before subtraction is performed.

IV. DELAY TRACKING IN SUCCESSIVE ICA-RECEIVER

The receiver structure described above can be developed to cope with inaccuracies in delay estimation. Recall that ICA performs purely in blind manner. The first occasion where some a prior knowledge of the users is needed is the user identification phase. Naturally, the timing information can also be used to generate a good initial value for ICA iterations, and hence speed up the separation [12]. Anyway,
what lousy timing estimate ultimately does is that it worsen the user identification. More importantly, given that a user is nevertheless identified having an erroneous timing estimate, the subtraction of that user enhance interference the more the bigger was the timing inaccuracy. To avoid that situation in one path channel, the delay of each identified user could first be refined according to

\[
\hat{\delta}_k \leftarrow \arg \max_{\delta} \frac{|g_k(\hat{\tau}_k, \delta)w_{k'}^H|}{|g_k(\hat{\tau}_k, \delta)||w_{k'}|}
\]

where the maximization of the correlation is performed in a close neighborhood of \(\hat{\tau}_k\), e.g. \(\delta \in (-1/2, 1/2)\). The signal corresponding to the user \(k\) is re-built after delay refinement, only after which it is beneficial to go to the subtraction phase. This is a sensible way to track delays since the timing information is implicitly included in the ICA solution \(w_{k'}\). In other words, also the delay tracking procedure takes advantage of interference suppression due to ICA separation.

V. ICA/SIC for Multi-path Channels

A straightforward extension to multipath channels is got by computing the correlation metric either for each multipath component separately (in case of time-variant channel) or for the composite of multipath components (in case of time-invariant channel).

In case of a time-variant multipath channel, we first try to identify each user’s each path separately using the correlation metric. Each user-path which is properly identified will be delay-tracked before subtraction:

\[
\hat{\delta}_{kl} \leftarrow \arg \max_{\delta} \frac{|g_{kl}(\hat{\tau}_{kl}, \delta)w_{k'}^H|}{|g_{kl}(\hat{\tau}_{kl}, \delta)||w_{k'}|}
\]

In another word, each multipath component of each user are separately tested whether it is beneficial to subtract it or not.

In case of a time-invariant channel, we first try to identify a user using the replica of received composite chip sequences \(\sum_{l=1}^{L} a_{kl}g_{kl}\) and the correlation metric. In case of proper identification, all the user-paths are tracked one-by-one (in the order of decreasing path gain) according to the following principle:

1. For each user \(k\) identified by the correlation metric, set \(l = 1\) and go to step 2.
2. Define

\[
z_1(\delta) = a_{kl}g_{kl}(\hat{\tau}_{kl}, 0) + \sum_{l' \neq 1} a_{kl}g_{kl}(\hat{\tau}_{kl}, 0).
\]

Then,

\[
\hat{\delta}_{kl} \leftarrow \arg \max_{\delta} \frac{|z_1(\delta)w_{k'}^H|}{|z_1(\delta)||w_{k'}|}
\]

3. Set \(l \leftarrow l + 1\).
4. Define

\[
z_l(\delta) = a_{kl}g_{kl}(\hat{\tau}_{kl}, \delta) + \sum_{l' \neq l} a_{kl}g_{kl}(\hat{\tau}_{kl}, \chi),
\]

where

\[
\chi = \begin{cases} 
0, & \text{if } l' > l \\
\hat{\delta}_{kl}, & \text{otherwise}
\end{cases}
\]

Then,

\[
\hat{\delta}_{kl} \leftarrow \arg \max_{\delta} \frac{|z_l(\delta)w_{k'}^H|}{|z_l(\delta)||w_{k'}|}
\]

VI. Numerical Experiments

A performance evaluation of the receiver was studied numerically. Each of the \(K=30\) users was supposed to send data blocks of \(M = 5000\) QPSK symbols. Symbols were spread using Gold codes of length \(C = 31\). Hence, strong MAI is preset and the data is over-saturated (from ICA point of view).

All the results are based on average bit error rates (BER) over 1000 independent repetition. BER values are computed with respect to all \(K = 30\) users. The following fixed multipath profiles were tested: two-path (10 dB power difference), three-path (equal powers) and four-path (3 dB differences). Also, typical ITU Pedestrian A and Vehicular A channels were tested.

Figures 1-5 represent the BER curves as a function of delay estimation error variance. It is seen that regardless of the multipath profile, ICA/SIC results in quite a constant \((10^{-2}\) or better) BER for a wide range of delay estimation errors.

Figure 6 represents the case where user identification is made more difficult due to the presence of external jammer signal. It is seen that the jammer affect the conventional schemes much more than ICA/SIC, as expected.

VII. Conclusions

In this paper we considered blind source separation (BSS) methods within successive interference cancellation (SIC) in a DS-CDMA reception with multi-paths. Numerical experiments revealed the robustness of the proposed scheme against the delay estimation errors which are known to be one of the expected imperfections at the DS-CDMA receivers.

REFERENCES

Fig. 1. Bit-error-rates as a function of error variance in delays. \( K = 30 \) asynchronous users were spread with 31-Gold codes. Two equal sized service classes were assumed, which was modeled as a power difference of 10 dB between the two user groups. Inside both groups all the users were assigned the same power. Channel of each user had two propagation paths with power difference of 10 dB. SNR was fixed to 20 dB wrt. the weakest users.

Fig. 2. Bit-error-rates as a function of error variance in delays. \( K = 30 \) asynchronous users were spread with 31-Gold codes. Two equal sized service classes were assumed, which was modeled as a power difference of 10 dB between the two user groups. Inside both groups all the users were assigned the same power. Channel of each user had four propagation paths with powers of 0 dB, -3 dB, -6 dB and -9 dB. SNR was fixed to 20 dB wrt. the weakest users.

Fig. 3. Bit-error-rates as a function of error variance in delays. \( K = 30 \) asynchronous users were spread with 31-Gold codes. Two equal sized service classes were assumed, which was modeled as a power difference of 10 dB between the two user groups. Inside both groups all the users were assigned the same power. Channel of each user had three propagation paths with equal powers. SNR was fixed to 20 dB wrt. the weakest users.

Fig. 4. Bit-error-rates as a function of error variance in delays. \( K = 30 \) asynchronous users were spread with 31-Gold codes. Two equal sized service classes were assumed, which was modeled as a power difference of 10 dB between the two user groups. Inside both groups all the users were assigned the same power. ‘ITU pedestrian A’ channel profile was assumed for each user. SNR was fixed to 20 dB wrt. the weakest users.
Fig. 5. Bit-error-rates as a function of error variance in delays. $K = 30$ asynchronous users were spread with 31-Gold codes. Two equal sized service classes were assumed, which was modeled as a power difference of 10 dB between the two user groups. Inside both groups all the users were assigned the same power. 'ITU vehicular A' channel profile was assumed for each user. SNR was fixed to 20 dB wrt. the weakest users.

Fig. 6. Bit-error-rates as a function of signal to jammer ratio. $K = 30$ asynchronous users were spread with 31-Gold codes. Two equal sized service classes were assumed, which was modeled as a power difference of 10 dB between the two user groups. Inside both groups all the users were assigned the same power. Channel of each user had two propagation paths with power difference of 10 dB. System contained an external jammer signal which was pulsed on symbol level with probability $p = 0.5$. SNR was fixed to 20 dB wrt. the weakest users. No delay estimation error.
BLIND DIVERSITY RECEPTION AND INTERFERENCE CANCELLATION USING ICA

by

Toni Huovinen, Aali Shahed hagh ghadam and Mikko Valkama

ABSTRACT

In this paper, we consider blind diversity reception and interference rejection in multi-antenna communications context, in terms of maximizing the output signal-to-interference-and-noise ratio (SINR). More specifically, we demonstrate that independent component analysis (ICA), although originally designed for noise-free linear models, is able to provide essentially the best possible output SINR among all linear transformations of received data in noisy linear models. In particular, our experiments indicate that one of the most widely applied ICA algorithms, equivariant adaptive source identification (EASI) algorithm, is, in practice, identical with SINR maximizing generalized eigenfilter in terms of SINR, even though it does not use explicit knowledge of the channel states and noise statistics. We also show that, in a special case of interference-free (that is, noise only) system, the EASI algorithm attains the greatest diversity gain blindly, i.e., performs as a blind maximal ratio combiner (MRC).

Index Terms— Blind diversity reception, blind interference cancellation, independent component analysis, multi-antenna communications

1. INTRODUCTION

Multiple transmit and/or receiver antennas are generally seen as one of the key elements in future wireless communications system developments [1]. In terms of the overall link quality, multiple antennas can be used for diversity purposes to mitigate the fading characteristics of the individual links (diminish the effects of noise), and also for removing other interfering signal components falling on top of the desired signal. This is also the central theme in this paper. In general, we consider the previous challenging system scenario in which both additive noise and interference are present in the received signals. Assuming multiple receiver antennas, the purpose is then to push down the noise and interference as much as possible using linear signal processing techniques. Furthermore, the focus here in general is on blind signal processing and reception in the sense that the noise statistics and channel state information are assumed unknown.

One relatively new idea in interference rejection is to employ blind source separation (BSS) techniques [2]. What makes BSS techniques attractive is their ability to separate signals from a mixture of original source signals in a completely blind manner, i.e., without an explicit knowledge of waveform structure (modulation) or mixing coefficients. In addition, typical BSS methods rely solely on higher-order statistical properties of data in the temporal domain, which makes the methods also very robust against possible spectral distortions. Communications related applications of BSS have been found, e.g., in MIMO systems [3], I/Q processing receivers [4], DS-CDMA blind multi-user detection [5] and DS-CDMA out-of-cell interference cancellation [6]. Independent component analysis (ICA) [2] is a fairly new statistical technique by which BSS can be performed. In ICA a set of observed signals or random variables are basically expressed as linear combinations of statistically independent components, which are often called sources or source signals. The ICA problem is blind, because not only the source signals but also the mixing coefficients are unknown. Many different methods have been proposed to solve the ICA problem [2]. Most of these are proper ICA methods exploiting the statistical independence of the sources, but there exist also other approaches which utilize temporal correlations or nonstationarity of the sources. In general, the mutual performance of these methods depends largely on the validity of the above assumptions.

Typically, a noise-free linear mixing model is assumed in derivation of ICA algorithms in the literature. Needless to say, the noise-free model is unrealistic in most of the applications, especially, in telecommunications. Consequently, applications of ICA often assume a noisy linear model, but exploit one of the ICA algorithms developed for noise-free models. Thus, a presence of reasonable level of additive noise is thought to cause “only” some feasible distortion due to the model mismatch. In this paper, we demonstrate that, although noise can never be suppressed completely by any linear technique, the performance gain (in terms of input-output SINR) obtained using ICA is practically identical to that of the optimum linear receiver utilizing known channel and noise statistics. In other words, ICA performs blind SINR maximization. We also bear out that ICA provides the greatest diversity gain, i.e., maximizes output SNR, among linear receivers assuming an interference-free model. That is to say, ICA acts as a blind maximal ratio combiner (MRC).

In numerical experiments, we have selected a popular equivariant adaptive source identification (EASI) algorithm to represent ICA.

2. SYSTEM MODEL AND LINEAR RECEIVERS

The basic system model used in the following assumes one transmit antenna and $M \geq 2$ receiver antennas used for diversity reception and interference rejection. Thus the received signal at the $m$-th antenna is of the form

$$x_m(t) = h_{m,u}u(t) + h_{m,v}v(t) + n_m(t)$$  

(1)

in which $u(t)$ and $v(t)$ denote the desired and interfering signals, respectively, $n_m(t)$ models additive channel noise, and $h_{m,u}$ and

---

*This work is financially supported by the Finnish graduate school TISE, the Academy of Finland and Nokia Foundation.*
are the corresponding channel coefficients. Assuming that (a) the source component processes, \( u(t) \) and \( v(t) \) are mutually uncorrelated, and (b) that the noise components, \( n_m(t) \), \( m = 1 \ldots M \), are temporally white, Gaussian and mutually uncorrelated, this yields an \( M \times 2 \) linear model,

\[
x(t) = As(t) + n(t),
\]

in which \( s(t) = [u(t) \ v(t)]^T \) is a source vector with zero mean and \( \mathbb{E}[s(t)s(t)^H] = I \), \( n(t) = [n_1 \ n_2 \ldots n_M]^T \) is a zero mean Gaussian noise vector with \( \mathbb{E}[n(t)n(t)^H] = \sigma^2 I \) \( (\sigma^2 > 0) \) and \( A = [h_u \ h_v] \in \mathbb{C}^{M \times 2} \) (with \( h_u = [h_{1,u} \ h_{2,u} \ldots h_{M,u}]^T \) and \( h_v = [h_{1,v} \ h_{2,v} \ldots h_{M,v}]^T \)) is a full-rank mixing (or channel) matrix, which is assumed to stay constant during one processing block of data. Provided that it is assumed, in addition to (a) and (b), for all \( t \) that source component processes, \( u(t) \) and \( v(t) \), are i.i.d non-Gaussian (or at least that one of them is) and \( v \) mutually statistically independent and, also, that \( \epsilon(n(t)) \) is independent from \( s(t) \), then (2) equals also a noisy ICA model [25].

The purpose in the following is to blindly estimate \( u(t) \), i.e., suppress noise and interference as much as possible under the assumptions (a) - (c). Notice also that the assumption that \( u(t) \) is the desired signal and \( v(t) \) is the interfering one is completely formal and any distinct assumptions between the source components are not made.

Let \( w \in \mathbb{C}^M \) be an arbitrary linear filter and \( y_w = w^H x \) the corresponding output of a linear receiver. Signal-to-interference-and-noise ratio (SINR) of the output \( y_w \) is then defined as

\[
\eta(w) = \frac{\mathbb{E}[w^H h_u | u|^2 h_u^H w]}{\mathbb{E}[w^H (h_v + n)(h_v^H v^* + n^H)w]} = \frac{w^H R_u w}{w^H (R_v + \sigma^2 I) w}
\]

in which \( R_u = h_u h_u^H \) and \( R_v = h_v h_v^H \). Now, as seen in (3), maximizing SINR among all linear transformations of received data, i.e., maximizing \( \eta(w) \), equals to solving the generalized eigenvalue problem [7] associated with matrix pair \( (R_u, R_v + \sigma^2 I) \). Hence, in

\[
\max_{w \in C^M} \eta(w) = \lambda \text{ and } \arg \max_{w \in C^M} \eta(w) = e,
\]

in which \( \lambda \) stands for the greatest eigenvalue of hermitian matrix \( (R_u + \sigma^2 I)^{-1} R_v \) and \( e \) for any (even vector) in the corresponding eigenspace. The vector \( e \) is called SINR maximizing generalized eigenfilter (M-GEF) wrt. the source component \( u \) (i.e., the desired signal here). This solution assumes the knowledge of the channel coefficients and noise variance, and forms a natural reference technique for the forthcoming blind ICA developments.

It is also interesting to note that a linear minimum mean square estimator (LMMSE) of a source can be shown to maximize SINR among linear transformations. This is basically stated in [8] and in references therein. Some earlier works have used the LMMSE technique as the reference method in their numerical evaluations (see, e.g., [9]). However, the generalized eigenfilter based approach shown in this paper gives certain benefits, especially, in analytical studies and comparisons of methods.

While the above eigenfilter based solution gives the optimum reference technique, the other two natural reference receivers are obtained by simply considering (i) only noise or (ii) only interference. A conventional diversification reception, or maximal ratio combining (MRC) as often referred to, is actually closely related to M-GEF reception. Nevertheless, in MRC, the main objective is to maximize the signal-to-(AWG)-noise ratio (SNR) and, consequently, MRC ignores possible interfering signal components which results in a sub-optimum SINR performance, in general. However, M-GEF reception is consistent with MRC in the sense that, without interference, \( v \), these two methods coincide. This is easy to see by setting \( v = 0 \) and constraining \( w \) to have, say, a unit norm in which case (3) reduces to ordinary eigenvalue problem, which, for one, is well known to yield the MRC solution [10]. In a noise-free model, as the other extreme, infinite SINR is naturally obtained by inverting \( A \). However in general, the inversion is not equal to M-GEF solution in the noisy model, due to arbitrary noise amplification. These two methods (conventional MRC and pure system inversion) are used also in the following as reference. Notice that channel knowledge is also needed in these reference techniques.

### 3. ICA AND BLIND SINR MAXIMIZATION

In basic ICA, the goal is essentially to invert the model (2) blindly, that is, to find an unmixing matrix \( W \) such that \( W A \) is as close to identity as possible by using only the observations \( x(t) \). Because of the blindness, a solution of the ICA problem, \( W \), can be unique only up to left multiplication by arbitrary permutation and diagonal matrices. Identifiability of such a \( W \) is guaranteed in theory only for noise-free linear models and, consequently, basic ICA algorithms can not produce exactly an inverse of matrix \( A \) (not even up to the indeterminacies) if additive noise is present. Nevertheless, inverting \( A \) does not lead to the best SINR gain in a noisy system anyway, as stated above. For this reason, it is well-advised to compare the performance of noise-free ICA algorithms to the above-mentioned M-GEF bound if noisy model (2) is used.

Some of the recent studies have also proposed so called noisy ICA algorithms [25] that assume a presence of additive noise and tries to take it into account. However, finding an inverse of \( A \) is usually a main objective also in these algorithms instead of, e.g., maximizing SINR. Other ICA related blind algorithms, that assume the noisy model, try to de-noise received data either before or after the ICA separation. These approaches, nevertheless, lead necessarily to non-linear (affine, at least) processing of data.

In the following numerical experiments, we assume \( M = 2 \) receiver antennas and use the widely applied EASI algorithm [11], which is originally intended to perform noise-free ICA. EASI is an online algorithm which operates on individual samples of received data. One step of the EASI algorithm is given as

\[
B(t + 1) = B(t) - \mu U(t) B(t),
\]

in which \( \mu \) is a scalar step size and the update matrix, \( U(t) \), is defined as

\[
U(t) = y(t)y(t)^H - I + g(y(t))y(t)^H - y(t)g(y(t))^H.
\]

Here \( y(t) = B(t) x(t) \), \( I \) stands for identity matrix and \( g : \mathbb{C}^2 \to \mathbb{C}^2 \) is an arbitrary nonlinear function. Since only the current sample is used in each step of the algorithm, the update matrix (7) does not vanish asymptotically. Instead, a stability point of the algorithm is defined stochastically to be a matrix \( B' \) for which the mean of the update term (7) is zero (i.e., \( \mathbb{E}(U(t)) = 0 \)).

Fig. 1 depicts an example behavior of elements of matrix \( B \) under a significant noise level (SNR=5 dB). In this 2 × 2 example, coefficients converge after twenty thousand EASI updates.

Results in the next section show that the output SINR of the EASI algorithm wrt. to the desired source is almost identical to M-GEF bound. In theory, a small difference exists [12], but SINR figures are in practice almost indistinguishable according to the results.
Especially, SINR gain of EASI is significant compared to SINR of plain inversion of \( \mathbf{A} \) and, on the other hand, to MRC bound when both noise and interference are present.

Here, as also in the continuation, signal-to-noise ratio (SNR) is defined as the average ratio of the received desired signal power and additive noise power. The signal-to-interference ratio (SIR), in turn, is the average ratio of the received desired signal and interference powers. Given the power normalization of the formal sources stated below (2), the SIR values other than 0 dB are implemented by corresponding scaling of the interference channel coefficients.

An important particular case is a system with finite SNR, but in which interference is absent, thus, the case in which M-GEF and MRC bounds coincide. Interestingly enough, also the EASI algorithm provides exactly the same output SINR in this case. More precisely, assuming an interference-free \( 2 \times 2 \)-model, \( \mathbf{x} = \mathbf{h}_u \mathbf{u} + \mathbf{n} \) and a vector \( \mathbf{h}_u \in \mathbb{C}^2 \) such that \( \mathbf{h}_u^H \mathbf{h}_u = 0 \), a matrix \( \mathbf{B}' = [\alpha \mathbf{h}_u \beta \mathbf{h}_u]^{H} \in \mathbb{C}^{2 \times 2} \) is a stability point of the EASI algorithm, i.e.,

\[
E \left\{ \mathbf{y} \mathbf{y}^H - I + g(\mathbf{y}) \mathbf{y} - y g(y)^H \right\} = 0,
\]

for \( \mathbf{y} = \mathbf{B}' \mathbf{x} \) and for appropriate complex scalars \( \alpha \) and \( \beta \). Recall, that \( \mathbf{h}_u \) is now the MRC filter \[10\]. We prove the tenability of (8) rigorously in \[12\]. Here, simulation results support the claim above. Fig. 2 plots the difference between simulated SINR of EASI and M-GEF bound as a function of SIR. In the figure, the difference decreases to negligible level (of roughly 0.01 dB-unit) when SIR increases (i.e., when interference goes down). A difference in the order of 0.01 dB-unit is basically explained by the finite sample statistics.

4. FURTHER NUMERICAL EXPERIMENTS

Numerical results in this section set against the performance of EASI algorithm and SINR maximizing M-GEF approach under noisy environment. Also SINR performances of plain maximal ratio combining (MRC) and inversion of the channel matrix, the matrix \( \mathbf{A} \) in model (2), are simulated in the experiments. Both of the latter methods are, thus, suboptimum since both an interfering source component and additive noise are present.

In the experiments, two receiver antennas are used. Desired signals are QPSK signals where as interfering signals are 16-QAM signals. The selection of source constellations is more or less arbitrary, and it should not affect the general validity of the results. Channel coefficients, i.e., elements of the matrix \( \mathbf{A} \), are drawn randomly from zero mean Gaussian distributions (one distribution for each source component) for each processing block of \( N = 50000 \) symbols of data. Variances of these distributions are selected such that received SNR and SIR values correspond to given values on average. M-GEF bound is evaluated directly from the data model for each block. Hence, the bounds are not affected by finite sample statistics and, more importantly, they are the absolute upper bounds among all linear transformations of received data in case of each realization. Also output SINR’s of the MRC and inversion of \( \mathbf{A} \) are evaluated from the model.

A simple third-order nonlinearity \[11\]

\[
\mathbf{g} = [g_1, g_2]^T : \mathbb{C}^2 \rightarrow \mathbb{C}^2, \quad g_i(z) = |z_i|^2 z_i, \quad i = 1, 2,
\] (9)

is used in the EASI algorithm in the experiments. A permutation ambiguity of EASI outputs is circumvented by, first, evaluating the output SINR wrt. the desired source component for both outputs and, then, selecting the maximum one. Practical ways to select the desired output component are not concerned in this paper. All the gains plotted are wrt. the received SINR.

Figs. 3 and 4 show SINR gains as a function of SIR with fixed received SNR. The figures illustrate that the gains of EASI algorithm are, in practice, undistinguishable from M-GEF bounds. A difference is less than 0.1 dB-unit in whole SIR range plotted in the both figures. Figs. 5 and 6 give two examples of SINR gain vs. received SINR with fixed received SIR. Again, performances of EASI and M-GEF bound are essentially identical.

5. CONCLUSIONS

In this paper, we illustrated that basic independent component analysis (ICA) designed for noise-free linear models is able to provide essentially the best possible output SINR among all linear transformations of received data, in the challenging case of having both additive noise and interference disturbing the desired signal observation in a multi-antenna receiver context. Thus in effect, the ICA is able to do joint diversity reception and interference cancellation in a blind manner, such that the output SINR is maximized. In particular, our experiments indicated that one of the most widely applied ICA algorithms, EASI algorithm, is, in practice, identical with SINR maximizing generalized eigenfilter (M-GEF) in terms of SINR. In theory, EASI can not attain exactly the M-GEF bound when both noise
and interference are present, but difference was negligible (< 0.1 dB-unit) in all of our experiments. We also showed that, in an important special case of interference-free (i.e., noise only) system, the EASI algorithm provides precisely the greatest linear diversity gain blindly, i.e., performs as a blind maximal ratio combiner (MRC).

The observed output SINR behavior almost identical to the theoretical upper bound rises a question whether the EASI algorithm, or ICA in general, could be further fine-tuned at the algorithm level to attain exactly the best linear SINR blindly also in theory. Such an algorithm would readily be a generalization of both conventional blind diversity reception and blind interference cancellation. In this paper, we left this question open to be dealt with in future studies.

6. REFERENCES


DYNAMIC OFFSET MITIGATION IN DIVERSITY RECEIVERS USING ICA

by

Ali Shahed hagh ghadam, Toni Huovinen and Mikko Valkama

DYNAMIC OFFSET MITIGATION IN DIVERSITY RECEIVERS USING ICA

Ali Shahed hagh ghadam, Toni Huovinen and Mikko Valkama
Institute of Communications Engineering
Tampere University of Technology
Tampere, Finland

ABSTRACT

In this article, we devise a blind digital signal processing (DSP) method for mitigating the dynamic offset interference due to self-mixing of radio frequency (RF) interferers in multiantenna diversity receiver context, utilizing the direct-conversion radio architecture. The proposed method is based on independent component analysis (ICA) and does not require any channel state information or knowledge about the physical self-mixing or interference characteristics. Moreover, we demonstrate through computer simulations that the overall achievable signal-to-interference-plus-noise ratio (SINR) obtained using the proposed technique is extremely close to the maximum SINR achievable by any linear method. Thus the ICA processing yields optimum compromise between the processing of noise and offset interference in this sense. We also demonstrate through concrete design examples how the proposed method helps relaxing certain constraints on the receivers’ RF components.

I. INTRODUCTION

One important practical problem in direct-conversion type receivers is the self-mixing products of strong out-of-band RF signals falling on top of the desired channel at the RF mixer outputs [1-2]. This type of interference, called dynamic offset in this paper, is physically caused by the finite isolation between the receiver local oscillator (LO) and RF input [2]. The self-mixing of the input RF blocker signals decreases the obtainable SINR of the desired channel, especially in case of direct-conversion radio architectures in which the relatively weak desired signal is located around DC, i.e., at the same band as the resulting offset interference components. The effect of this interference on the overall desired channel SINR is thus especially troublesome in case of weak desired signal and strong out-of-band signals [1].

There have been several methods proposed in the literature to eliminate the dynamic offset problem, see e.g. [1-6]. The early solutions of this problem are implemented, at least partly, in analog part of the receiver [1-3]. Thus these solutions can be seen to increase the cost and volume of the receiver analog front-end. However, some of the more recent methods are devised entirely in the digital domain [4-6]. These digital methods proposed earlier can relax constrains, costs and volume of the RF parts of the receivers to some extent. In this paper, we present a new blind compensation technique for dynamic offset mitigation, being specifically targeted for multiantenna diversity receivers of future wireless systems. In effect, the proposed signal processing based on ICA is simultaneously providing diversity gain over individual fading links while also effectively removing the offset interference due to the receivers’ analog RF imperfections. To the best knowledge of the authors, such an approach has not been proposed earlier in the literature.

As discussed above, direct-conversion radio architecture suffers inherently from the dynamic offset interference. On the other hand, multiantenna diversity receivers provide us with additional possibilities to attenuate any interference and noise by properly combining the observed signals. More precisely, consider a general multiantenna receiver with M antennas. Then the overall observation vector \( \mathbf{x} = [x_1, x_2, \ldots, x_M]^T \) is of the form

\[
\mathbf{x} = \mathbf{A}s + \mathbf{n},
\]

where \( s = [s_1, s_2, \ldots, s_M] \) with \( s_d \) denoting the desired signal while \( s_2, \ldots, s_M \) model the possible interfering signals on top of the desired one. Here \( \mathbf{A} \) is a complex matrix representing the channel state and \( \mathbf{n} \) is additive white Gaussian noise (AWGN) vector. Then knowing \( \mathbf{A} \) and the noise covariance, \( \mathbb{E}[|\mathbf{n}|^2] = \sigma^2 I \), it is possible to find the optimum coefficients \( \mathbf{w}_{\text{opt}} \) in

\[
y = \mathbf{w}_{\text{opt}}^H \mathbf{x} = \mathbf{w}_{\text{opt}}^H \mathbf{A} s + \mathbf{w}_{\text{opt}}^H \mathbf{n},
\]

such that the overall SINR of the combined signal \( y \) is maximized with respect to \( s_d \) [7-9]. Thus it is possible to mitigate any interference (including the dynamic offset considered in this paper) in diversity receivers by properly combining the observations. At general level, this is of course well-known.

In the ordinary SINR maximization studies, the channel coefficients as well as the noise variance are assumed known [7]. In practice the receiver can estimate the channel state by transmitting training sequence known to the transmitter and receiver. Of course transmitting these extra symbols reduces the capacity of the communication link. Moreover the receiver somehow should be able to estimate the power of the noise \( \sigma^2 \). To avoid these complications, devising a method to estimate \( \mathbf{w}_{\text{opt}} \) blindly (i.e. without any knowledge of channel and power of noise) is highly desirable. This problem is closely related to the theory of independent component analysis (ICA) in noisy environment [10].

ICA theory and algorithms provide us with powerful tools to recover number of independent source signals up to scale and permutation by just observing linear mixtures of them. In the other words if we have a model,

\[
\mathbf{x} = \mathbf{A}s,
\]

by applying one of the ICA algorithms on a set of observed signals \( \mathbf{x} \) it is possible to separate the interfering signals \( s_2, \ldots, s_M \) from desired signal \( s_d \) entirely if they are statistically independent. But adding the noise to (3) (i.e., model presented in (1)) degrades the performance of ordinary ICA. In fact ac-
cording to the theory of ICA complete separation of interfering and desired signal is not possible in the noisy environment as shown in [9]. However recent studies in [8,9] show that although the complete separation of the sources is not possible in noisy environments, the SINR provided by ICA, as such, is extremely close to the maximum achievable SINR by any linear method [9]. Thus in the context of diversity receivers, ICA can be considered as a practical method of joint interference rejection and diversity combining.

In this article, we use this general ICA framework for devising appropriate receiver signal processing to blindly mitigate dynamic offset interference in addition to obtaining diversity against individually fading transmitter-receiver links. In Section II, we present essential signal and system models for dynamic offset interference due to RF self-mixing. In Section III, we propose applying ICA for joint offset interference mitigation and diversity combining. In Section IV, computer simulations results are shown to illustrate the good performance of the proposed approach, and also a practical receiver RF dimensioning example is given. Finally, conclusions are drawn in Section V.

II. DYNAMIC OFFSET INTERFERENCE IN DIRECT-CONVERSION RECEIVERS

The spurious signal components at or around DC generated by the self-mixing of the LO signal or the input RF signal is called offset (static and/or dynamic) [1-3]. These components can easily degrade the quality of the weak desired signal if located at the same frequency range. Most typically this is the case in the basic direct-conversion scenarios where the desired signal appears at the baseband after front-end downconversion. In the following we give more insights into the offset generation mechanisms and the significance of this type of interference in the context of direct-conversion receivers, as well as derive appropriate signal models describing the interference to be then used in the interference mitigation.

A. Static Offset Due to LO Leakage

Consider a direct-conversion receiver which suffers from static offset due to self-mixing of LO signals (Fig.1). Received RF signal, \( x_{RF}(t) \), can in general be written as

\[
x_{RF}(t) = \text{Re} \left\{ (k_u(t) + k_v(t)e^{-j\omega_{RF}t})e^{j\omega_{LO}t} \right\} + [k_u(t)e^{-j\omega_{LO}t} + k_v(t)e^{j\omega_{LO}+\omega_{RF}t}]
\]

where \( u(t) \) is the baseband equivalent desired signal and \( v(t) \) is the corresponding baseband equivalent of the RF blocker located \( \omega_{LO} \) away from the desired signal. \( k_u \) and \( k_v \) represent the relative complex channel gains of the two signals at \( \omega_{RF} + \omega_{LO} \) respectively.

To model the finite isolation between the mixers LO and RF ports, we use the leakage coefficients \( L_I \) and \( L_Q \). In the ideal case, these leakage coefficients are zero representing infinite attenuation while in practice, the isolation is in the order of 40 to 60dB [1]. Considering then first the leakage of the LO signal(s) into the RF mixer input port(s), the down-converted I/Q signal \( x_{down}(t) \) can be written as

\[
x_{down}(t) = ((x_{RF}(t) + L_I x_{RF}(t))\cos(\omega_{RF}t)) - j((x_{RF}(t) + L_Q \sin(\omega_{RF}t)\sin(\omega_{RF}t))
\]

Subsequently, lowpass-filtered signal \( x_{base}(t) \) can be written as

\[
x_{base}(t) = k_u(t) + (L_I - jL_Q).
\]

It is clear from (6) that the LO leakage generates static DC-offset on top of the desired signal \( u(t) \) at the baseband. There are several methods available in the literature for mitigating this type of interference [1-3]. In this paper, the main focus is on handling the more challenging case of dynamic offset interference explored next.

B. Dynamic Offset Due to RF Leakage

The process which yields the dynamic offset is presented in Fig.2. In this case, the finite isolation between the RF and LO causes the self-mixing of RF signals. The down-converted signal \( x_{down}(t) \) in this case is

\[
x_{down}(t) = x_{RF}(t)[\cos(\omega_{RF}t) + L_I x_{RF}(t)] \\
+ jx_{RF}(t)[\sin(\omega_{RF}t) + L_Q x_{RF}(t)]
\]

where now \( L_I \) and \( L_Q \) represent the leakage attenuation of the RF into the mixer LO port. Applying lowpass filtering on the above signal in (7) yields then

\[
x_{base}(t) = k_u(t) + (L_I + jL_Q)\left\{ |\sin(\omega_{RF}t)| + |\cos(\omega_{RF}t)| \right\}.
\]
Thus the desired signal $u(t)$ is clearly interfered by its own squared-envelope as well as by the squared-envelope of the RF blocker. Assuming next that the desired signal $u(t)$ is significantly weaker than $v(t)$ and including also the additive white Gaussian noise (AWGN) element, (8) can be written as

$$x_{\text{base}}(t) = k_u u(t) + |k_v|^2 (L + jL_Q) |v(t)|^2 + n(t).$$

(9)

Thus contrary to the static offset, dynamic offset shown in (8) and (9) in terms of the blocker squared-envelope, can easily degrade the quality of the desired signal $u(t)$, specially in cases of strong RF blocker. This is because the interference in (8) and (9) is proportional to $|v(t)|^2$ and thus the interference power in then proportional to $|v(t)|^2$. Based on this quadratic relation between the RF blocking signal and the generated dynamic offset interference component, it is clear that the interference effect all and all is very strongly dependent of the RF power of the original blocker, as well as of course on the leakage coefficients $L_1$ and $L_Q$. Fortunately, in most practical systems like GSM, the specified spectral masks limit the power of out-of-band blocking signals as a function of the frequency separation from the target carrier [11]. Some concrete examples of receiver RF dimensioning from the RF filtering and offset interference point of views are given in Section IV.

In the next section, we address the dynamic offset mitigation problem in multiantenna diversity receiver context, and formulate an efficient ICA-based blind mitigation technique. To keep the presentation practically oriented, a two-antenna receiver case is assumed in the following.

III. Dynamic Offset in Multiantenna Diversity Receivers and Its ICA-Based Compensation

Dynamic offset interference is addressed here for a two-antenna diversity receiver system. This case is chosen, without loss of generality, to keep the discussion practically oriented and applicable in mobile devices, where the two-antenna receiver case is indeed feasible. Based on the previous section, the overall signal model for the two-receiver case under RF self-mixing can essentially be written as

$$\begin{align*}
x_{\text{base},1}(t) &= \begin{bmatrix} k_{u,1} & k_{v,1} \end{bmatrix} \begin{bmatrix} L_1 + jL_Q \end{bmatrix} u(t) + \begin{bmatrix} n_{1}(t) \end{bmatrix}, \\
x_{\text{base},2}(t) &= \begin{bmatrix} k_{u,2} & k_{v,2} \end{bmatrix} \begin{bmatrix} L_2 + jL_Q \end{bmatrix} v(t) + \begin{bmatrix} n_{2}(t) \end{bmatrix},
\end{align*}$$

(10)

where the subscripts 1 and 2 refer to the two antennas. To state the above in a more compact form, (10) is written as

$$x_{\text{base}} = \begin{bmatrix} h_u & h_v \end{bmatrix} \begin{bmatrix} u(t) \\ |v(t)|^2 \end{bmatrix} + n,$$

(11)

where the effective system basis vectors read $h_u = [k_{u,1}, k_{u,2}]^T$ and $h_v = [k_{v,1}^2(L_1 + jL_Q), k_{v,2}^2(L_2 + jL_Q)]^T$. As in the single-antenna case, it is also clear from (10)-(11) that the SINR of the baseband observations are strongly affected by the offset interference relative to $|v(t)|^2$. Considering then the ICA terminology [10,13], the model (11) fits directly to the general instantaneous mixing model, in which the two formal source signals are $u(t)$ and $|v(t)|^2$. Thus in diversity receivers, the self-mixing products of the RF blockers falling on top of the desired signal can be modeled using a linear signal mixture model as in (11). Furthermore, assuming different channels and/or different leakage coefficients for the different receivers, the model in (10)-(11) is always identifiable.

In searching next for appropriate linear transformations of the observation $x_{\text{base}}$ to recover the desired signal $u(t)$, the natural reference in all performance evaluations is given by the best linear filter maximizing the overall SINR at the filter output. This is formally given by

$$y_{\text{opt}} = w_{\text{opt}}^H x_{\text{base}},$$

(12)

and the solution $w_{\text{opt}}$ maximizing the SINR with respect to $u(t)$ in $y_{\text{opt}}$ can be written as [9]

$$w_{\text{opt}} = \arg \max_w \frac{E[\hat{R}_u w]}{E[w^H (R_v + \sigma^2 I) w]},$$

(13)

in which

$$\begin{align*}
\hat{R}_u &= h_u h_u^T, \\
R_v &= h_v h_v^T.
\end{align*}$$

(14)

In practice, $w_{\text{opt}}$ can be calculated using the generalized eigenvalue filtering approach [9,12], i.e., $w_{\text{opt}}$ is equal to the eigenvector of the matrix $(R_v + \sigma^2 I) \hat{R}_u$, corresponding to the largest eigenvalue. The largest eigenvalue itself is the maximum SINR achievable by any linear filter [9,12]. We call this method SINR maximizing generalized eigenvector (M-GEF) and it forms a natural reference against which to compare the performance of any other linear technique.

It is clear from above discussion that obtaining the largest eigenvalue of the matrix $(R_v + \sigma^2 I) \hat{R}_u$ (i.e. maximum attainable SINR) and the corresponding eigenvector (i.e. $w_{\text{opt}}$) requires the knowledge of the channel state and noise variance which in most practical cases is not available to the receiver. Therefore devising a blind method to provide maximum SINR without knowledge of channel information and noise power levels is desirable.

It is demonstrated recently in [8,9] through simulations that certain popular ICA algorithms can meet the above criteria, i.e., they are blind and their performance is very close to maximum achievable SINR. Therefore, by applying one of the ICA algorithms on a set of vectors $x_{\text{base}}$ in (10), it is possible to also blindly mitigate the dynamic offset interference in multiantenna diversity receivers. This is very interesting in the sense that ICA provides automatically an optimum compromise in rejecting interference on one side and carrying out diversity combining against noise on the other. The only formal requirement is that the two source signals in (10), the desired signal $u(t)$ and the squared-envelope of the blocker $|v(t)|^2$, are statistically independent [10,13]. This, in turn, is plausible especially in cases where the desired signal and the RF blocker originate from physically separate sources. One practical complication in applying ICA is that the correct order and scaling of the estimated source signals cannot be blindly identified [10]. Hence to distinguish between the desired signal and dynamic offset component, some additional information of the desired signal waveform structure is required. Other than that, the processing is totally blind.
In the following section, we design a simulation experiment to illustrate the effectiveness of the proposed compensation method. The ICA results are compared against the M-GEF solutions and also the implications on RF design in terms of tolerable RF blocker levels are addressed.

IV. SIMULATION RESULTS

A. Obtained SINR Performance

Here we assess the obtainable offset interference rejection capability of the ICA-based approach using computer simulations. A similar two-receiver case as in the previous section is assumed, and two different complex modulations are deployed, QPSK and 16QAM. In the simulation setup, channel coefficients \( k_{u1}, k_{u2}, k_{v1} \), and \( k_{v2} \) are drawn randomly from the complex Gaussian distribution with zero mean and variance of one. The additive noise sequences \( n_1(t) \) and \( n_2(t) \) are white Gaussian noise with given power levels defined in the Table I, depending on the desired signal modulation type. Each experiment consists of receiving the two receivers’ signals for 50,000 symbol intervals, over which the ICA is then applied for mitigating the offset interference. The practical ICA algorithm used here is the so-called equivariant adaptive separation via independence (EASI) algorithm [13] with simple third order nonlinearity and adaptation step-size corresponding to convergence in the stated block length of 50,000 symbol intervals. The average output SINR is calculated for comparing the performance of the ICA-based processing against M-GEF reference (forming a theoretical bound for any linear interference cancellation method). In addition, assuming known channel state, also zero-forcing (ZF) estimator and maximum ratio combining (MRC) are also implemented for comparison, in which either the interference (ZF) or additive noise (MRC) is conceptually taken into account. The average SINR performances are obtained by averaging over 1000 different realization of channel coefficients. The RF-LO leakage coefficients are in the order of -50 to -56 dB, which represent state-of-the-art. The simulation parameters are summarized in Table I.

Table I. Simulation parameters used for experiments.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Case 1</th>
<th>Case 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desired Signal</td>
<td>QPSK, 0 dBm</td>
<td>16QAM, 0 dBm</td>
</tr>
<tr>
<td>Blocker</td>
<td>QPSK, [10...48] dBm</td>
<td>16QAM, [10...48] dBm</td>
</tr>
<tr>
<td>RF-LO Leakage</td>
<td>[-54 to -56] dB</td>
<td>[-54 to -56] dB</td>
</tr>
<tr>
<td>In-band SNR</td>
<td>5 dB</td>
<td>10 dB</td>
</tr>
</tbody>
</table>

Fig. 3. Case 1: The desired signal and the blocker are both QPSK and the in-band SNR for desired signal is 5 dB. Averaging is performed over 1000 realizations of the channel coefficients.

B. RF Dimensioning Example

Next we address the question how much the ICA-based mitigation can relax certain RF constraints in the receiver design. Again we consider a diversity receiver with two antennas, and for target SINR definitions, QPSK desired signal waveform is assumed. The target is to yield a minimum of 5 dB in-band SINR at compensator output (detector input) corresponding to a raw detection error rate in the order of \( 10^{-1} \) to \( 10^{-2} \). In the front-end design, the RF low noise amplifier...
The offset interference for ICA-based receiver can be calculated as SIR values, the maximum allowable power of the dynamic splitting the signal into the I and Q paths. Now Fig. 3 shows (ICA-based method and MRC are (roughly) -30dB and 0dB. From (15) and these input SIR values, the maximum allowable power of the dynamic offset interference for ICA-based receiver can be calculated as

\[ P_{vt}(ICA) = -82 + 30 = -52 \text{ [dBm]} \]  

(16)

and for MRC,

\[ P_{vt}(MRC) = -82 + 0 = -82 \text{ [dBm]} \]  

(17)

Then the relation between the power of the blocking signal \( vt(t) \) at RF LNA input and the power of the dynamic offset component at baseband is of the form

\[ P_{vt} = P_{vt} - \frac{L}{2} - G + 3, \]  

(18)

where \( L \) is the leakage coefficient. Therefore from (16), (17) and (18) we can calculate the maximum tolerable RF power of blocking signal at the RF LNA input in case of ICA-based receiver and MRC method as

\[ P_{vt}(ICA) = \frac{-52 + 60}{2} - 20 + 3 = -13 \text{ [dBm]}, \]  

\[ P_{vt}(MRC) = \frac{-82 + 60}{2} - 20 + 3 = -28 \text{ [dBm]} \]  

(19)

Thus based on (19), we can conclude that by implementing the ICA-based method for mitigating the dynamic offset it is possible to relax the attenuation constrains of the RF bandpass filter up to 15dB when compared to ordinary MRC processing. Furthermore, it should also be kept in mind that using the ICA-based method, the overall receiver DSP functionalities are simplified, compared to MRC-based receiver, in the sense that no channel estimation is needed in the ICA-based receiver. This is also an important practical benefit.

V. CONCLUSIONS

In this paper, a novel DSP method for mitigating the dynamic offset interference due to RF blocker self-mixing in direct-conversion diversity receivers was proposed. The proposed technique is stemming from modeling the dynamic offset interference as a linear signal mixture model, and then applying ICA on the set of observed signal vectors. Computer simulations were used to assess the achievable compensation performance, and based on the obtained results, the offset interference can be efficiently mitigated using the ICA-based approach, while also simultaneously obtaining diversity gain against ordinary channel noise. Based on the simulated interference tolerance, we also discussed through an RF dimensioning example, how the ICA-based compensator can relax the RF filtering requirements for out-of-band blocker signals in practical receiver design, when compared to ordinary MRC processing. The other main advantage of ICA-based method over the other existing techniques is that the proposed method in this paper does not need the channel state information, and thus channel estimation can be basically avoided.

VI. REFERENCES

HIGHER-ORDER BLIND ESTIMATION OF GENERALIZED EIGENFILTERS USING ICA

by

Toni Huovinen, Aali Shahed hagh ghadam and Mikko Valkama

Tampereen teknillinen yliopisto. Tietoliikennetekniikan laitos.
Tutkimusraportti 2008:1
Tampere University of Technology. Department of Communications Engineering.
Research Report 2008:1

Toni Huovinen, Ali Shahed hagh ghadam & Mikko Valkama

Higher-Order Blind Estimation of Generalized Eigenfilters Using ICA
Higher-Order Blind Estimation of Generalized Eigenfilters Using Independent Component Analysis

Toni Huovinen, Ali Shahed hagh ghadam and Mikko Valkama

Abstract

Assuming a noisy linear mixing model of source random variables or signals, maximizing the output signal-to-interference-and-noise-ratio (SINR) among linear transformations of observed data leads to solving the generalized eigenvalue problem. The explicit solution of the problem assumes the knowledge of the mixing coefficients and noise variance and, for this reason, is not a blind method as such. However, we show in this paper that the solution can be estimated blindly and directly using basic independent component analysis (ICA) designed for noise-free linear models. In addition, the theoretical and numerical results of the paper show that one of the most widely applied ICA algorithms, the equivariant adaptive source identification (EASI) algorithm, is, in practice, identical with SINR-maximizing generalized eigenfiltering, even though it does not use explicit knowledge of the mixing coefficients nor source and noise statistics. We also prove that, in the special case of interference-free (that is, noise only) system, the EASI algorithm is able to attain exactly the greatest diversity gain blindly, i.e., to perform as a blind maximal ratio combiner (MRC).

I. INTRODUCTION

Independent component analysis (ICA) [1] is a statistical signal processing technique which has attracted a lot of attention recently. Especially, it has been applied successfully to solving blind source separation (BSS) problem. Typically, a noise-free linear mixing model is assumed in derivation of ICA algorithms in the literature [1]. Needless to say, the noise-free model is unrealistic in most of the practical applications. Consequently, applications of ICA often assume a noisy linear model, but exploit one of the ICA algorithms developed for noise-free models. Thus, the presence of reasonable level of additive noise is thought to cause “only” some feasible distortion due to the model mismatch. However, numerical experiments reported in [2] indicate that, although noise can never be suppressed completely by any linear technique, the performance gain in terms of input-output signal-to-interference-and-noise-ratio (SINR) obtained using ICA is practically identical to that of the optimum (i.e., SINR-maximizing) linear transformation utilizing known channel and noise statistics. That paper is, nevertheless, restricted to a
certain telecommunications related interference suppression application in which it is enough to assume (and is assumed) rather simple $2 \times N \ (N \geq 2)$ noisy mixing model. Also some other earlier ICA works (see, e.g., [3]) have found out (by numerical experiments) similar results.

In this paper, our main objective is to generalize the experimental findings of [2] to $M \times N \ (M \geq N)$ models in the mathematically rigorous way. In particular, the forthcoming analytical study shows that conceptually ICA can indeed identify a linear input-output SINR maximizing transformation directly. Recall, that identifying the mixing matrix and inverting it do not lead to the linear SINR maximizing transformation as such due to noise enhancement. We also give conditions under which one of the most widely applied ICA algorithms, the equivariant adaptive source identification (EASI) algorithm [4], is, in theory, identical with SINR-maximizing generalized eigenfiltering [5]. More specifically, we give the necessary and sufficient conditions under which the matrix of the generalized eigenfilters is a stationary point of the EASI algorithm. We also show some numerical results verifying convergence of the EASI algorithm to the SINR-maximizing linear solution and showing that the performance of the EASI algorithm is remarkably close to the optimal (i.e., the maximal output SINR among all linear transforms of observed data) also in cases in which the above mentioned theoretical optimality condition is not met. In addition, we prove that, in the special case of interference-free (that is, noise only) system, the EASI algorithm can attain exactly the greatest diversity gain blindly, i.e., perform as a blind maximal ratio combiner (MRC).

It is also noteworthy to mention that a linear minimum mean square estimator (LMMSE) of a source component can be shown to maximize SINR among linear transformations. This is basically stated in [6] and in references therein. Some earlier ICA works have used the LMMSE reception as the reference method in their numerical evaluations (see, e.g., [3]). However, the generalized eigenfilter based approach used in this paper gives certain benefits (as will be seen) in the analytical studies and comparisons of methods.

II. NOTATIONS

In this paper, vectors and matrices are denoted by lower and upper case boldface symbols, respectively. In addition, the following notations are used:

- $(\cdot)^*$: complex conjugate;
- $(\cdot)^T, (\cdot)^H$: transpose and Hermitian transpose;
- $\Re\{\cdot\}, \Im\{\cdot\}$: real and imaginary part;
- $i$: imaginary unit;
- $E\{\cdot\}$: expected value;
- $\text{Tr}(\cdot)$: trace of matrix;
- $\text{diag}(\mathbf{v})$: diagonal matrix with elements of $\mathbf{v}$ on the diagonal;
- $\| \cdot \|$: Euclidean norm;
- $\mathbf{0}$: origin (vector of zeros).
### III. NOISY ICA MODEL AND OPTIMAL LINEAR FILTERING

A complex valued, linear, $M \times N$ ($M \geq N$) ICA model [1] with additive white Gaussian noise (AWGN) is assumed throughout this paper. Thus, the $M$-dimensional random observation vector, $\mathbf{x}$, is given as

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{\eta}, \quad (1)$$

in which $\mathbf{s} = [s_1 \ s_2 \ldots s_N]^T$ is a random source vector with complex valued, mutually independent components and $\mathbf{\eta} = [\eta_1 \ \eta_2 \ldots \eta_M]^T$ is a zero mean complex valued Gaussian noise vector (independent from $\mathbf{s}$) with a strictly positive variance $\sigma^2$ and the covariance $\mathbb{E}\{\mathbf{\eta}\mathbf{\eta}^H\} = \sigma^2\mathbf{I}$. Further, $\mathbf{A} = [\mathbf{h}_1 \ \mathbf{h}_2 \ldots \mathbf{h}_N] \in \mathbb{C}^{M \times N}$ is a full rank mixing (or channel) matrix. In an analytical part of the paper, the source and noise signal are considered plain random variables, which requires ergodicity of the actual time processes. Without loss of generality (see, [1]), it is assumed that the source components, $s_1, s_2, \ldots$ and $s_N$, have zero mean and unit variance. Together with the independence assumption, this implies that $\mathbb{E}\{\mathbf{ss}^T\} = \mathbf{I}$. In ICA literature, the complex valued sources are usually assumed also to be circularly symmetric at least up to second-order. A complex valued random variable, say $\mathbf{z}$, is said to be second-order circularly symmetric, if $\mathbb{E}\{\mathbf{z}^2\} = 0$ [7]. Circular symmetry of the source components is assumed also here. This assumption yields (again together with independency and zero mean assumptions) that $\mathbb{E}\{\mathbf{ss}^T\} = 0$. In the analysis, the second-order circular symmetry property is assumed also for the noise vector $\mathbf{\eta}$.

**Definition 1** Let $\mathbf{w} \in \mathbb{C}^{M\setminus\{0\}}$ be an arbitrary linear filter and $y_{\mathbf{w}} = \mathbf{w}^H\mathbf{x}$ the corresponding linear output. Signal-to-interference-and-noise ratio (SINR) wrt. the $n$-th source component, $s_n$, at the output $y_{\mathbf{w}}$ is then defined as

$$\rho_n(\mathbf{w}) = \frac{\mathbf{w}^H \mathbf{R}_n \mathbf{w}}{\mathbf{w}^H \mathbf{R}'_n \mathbf{w}}, \quad (2)$$

in which

$$\mathbf{R}_n := \mathbb{E}\{\mathbf{h}_n s_n s_n^* \mathbf{h}_n^H\} = \mathbf{h}_n \mathbf{h}_n^H \quad (3)$$

and

$$\mathbf{R}'_n := \mathbb{E}\left\{ \left( \sum_{k \neq n} \mathbf{h}_k s_k + \mathbf{\eta} \right) \left( \sum_{k \neq n} \mathbf{h}_k s_k + \mathbf{\eta} \right)^H \right\} = \sum_{k \neq n} \mathbf{h}_k \mathbf{h}_k^H + \sigma^2 \mathbf{I}. \quad (4)$$

**Remark 1** Since all eigenvalues of the Hermitian matrix $\mathbf{R}'_n$ are greater than or equal to $\sigma^2$, the Hermitian form in the denominator of (2) is positive definite, i.e., strictly positive for all $\mathbf{w} \in \mathbb{C}^{M\setminus\{0\}}$, provided that $\sigma^2 > 0$. Consequently, (2) is well-defined for all $\mathbf{w} \in \mathbb{C}^{M\setminus\{0\}}$.

Now, as seen in (2), maximizing SINR among all linear transformations of observed data, i.e., maximizing $\rho_n(\mathbf{w})$, equals to solving the generalized eigenvalue problem [5] associated with matrix pair $(\mathbf{R}_n, \mathbf{R}'_n)$. Hence,

$$\max_{\mathbf{w} \in \mathbb{C}^{M\setminus\{0\}}} \rho_n(\mathbf{w}) = \lambda_n \quad (5)$$
\[ \arg \max_{\mathbf{w} \in \mathbb{C}^{M}} \rho_n(\mathbf{w}) = \mathbf{e}_n, \tag{6} \]

in which \( \lambda_n \) stands for the greatest eigenvalue of the Hermitian matrix \((R'_n)^{-1}R_n\) (the matrix inverse exists, see Remark 1) and \( \mathbf{e}_n \) for the corresponding eigenvector. To be specific, since SINR \( \rho_n(\mathbf{w}) \) is scale invariant, \( \mathbf{e}_n \) can be any vector in one-dimensional eigensubspace corresponding to the eigenvalue \( \lambda_n \).

Also the linear minimum mean square error (LMMSE) estimator of a source can be shown to yield the maximum SINR among linear transformations. This is basically stated in [6] and in references therein. The LMMSE transformation for \( n \)-th source in model (1) is essentially given as
\[ \mathbf{w}'_n = C_x^{-1}h_n, \tag{7} \]

in which \( C_x = \mathbb{E}[\mathbf{x}\mathbf{x}^H] \) is the observation covariance. This linear transformation, thus, gives an explicit solution to the generalized eigenvalue problem above, i.e., \( \mathbf{e}_n = \mathbf{w}'_n \). Nevertheless, the solution assumes the knowledge of the mixing coefficients and noise variance and, for this reason, is not a blind method as such. However, it will be shown that the solution can be estimated blindly using ICA.

Next some further notations used throughout this paper are defined and the most important properties of filters maximizing the linear SINR are given.

**Definition 2** (i) The vector \( \mathbf{e}_n \) is called SINR-maximizing generalized eigenfilter (M-GEF) wrt. the source component \( s_n \). (ii) The matrix \( \mathbf{E} := [\mathbf{e}_1 \mathbf{e}_2 \ldots \mathbf{e}_N] \in \mathbb{C}^{N \times M} \) is called M-GEF transformation. Recall, that it also has the LMMSE characterization as
\[ \mathbf{E} = C_x^{-1}A. \tag{8} \]

(iii) The corresponding filtered output \( \mathbf{y} := \mathbf{E}^H\mathbf{x} \) is called M-GEF output.

**Lemma 1** If the columns of mixing matrix \( \mathbf{A} \) are mutually orthogonal, i.e., \( \mathbf{A}^H\mathbf{A} = \mathbf{D}_A \) for some diagonal matrix \( \mathbf{D}_A \in \mathbb{C}^{N \times N} \), then the M-GEF transformation \( \mathbf{E} \) has mutually orthogonal columns and, moreover, \( \mathbf{E} = \mathbf{A}\mathbf{D}_E \) for some diagonal matrix \( \mathbf{D}_E \in \mathbb{C}^{N \times N} \).

**Proof:** Let the columns of \( \mathbf{A} = [\mathbf{h}_1 \mathbf{h}_2 \ldots \mathbf{h}_N] \) be orthogonal and \( \mathbf{R}_n \) and \( \mathbf{R}'_n \), \( n = 1, 2, \ldots, N \), be defined as in (3) and (4), respectively. Now, it is enough to prove that the mixing coefficient vector \( \mathbf{h}_n \) is the M-GEF vector wrt. the source component \( s_n \) for all \( n \in \{1, 2, \ldots, N\} \). First, due to orthogonality of the columns of \( \mathbf{A} \), \( \mathbf{R}_n h_k = 0 \) for all \( k \neq n \), which implies that \( \mathbf{R}'_n h_n = \sigma^2 \mathbf{h}_n \). Consequently,
\[ \mathbf{R}_n \mathbf{h}_n = ||\mathbf{h}_n||^2 \mathbf{h}_n = \frac{||\mathbf{h}_n||^2}{\sigma^2} \mathbf{R}'_n \mathbf{h}_n, \tag{9} \]
i.e., \( \mathbf{h}_n \) is an eigenvector of \((\mathbf{R}'_n)^{-1}\mathbf{R}_n =: \mathbf{M}_n \) and the corresponding eigenvalue is \( ||\mathbf{h}_n||^2/\sigma^2 =: \lambda_n \).

Second, it is still needed to show that \( \lambda_n \) is the greatest eigenvalue of \( \mathbf{M}_n \), or equally, that \( \lambda_n \) really is the maximum value of the linear output SINR \( \rho_n(\mathbf{w}) \) in \( \mathbb{C}^M \setminus \{0\} \). But, using Schwarz’s inequality, it is seen that, for all \( \mathbf{w} \in \mathbb{C}^M \setminus \{0\} \),
\[ \rho_n(\mathbf{w}) \leq \frac{||\mathbf{h}_n||^2}{\sigma^2}. \tag{10} \]
Hence, this concludes the proof. ■

**Lemma 2** The following assertions are equivalent:

1) The mixing matrix $A = [h_1 \; h_2 \ldots h_N]$ has mutually orthogonal columns.
2) $E^HA = D := \text{diag}(d_1, d_2, \ldots, d_N)$ for some complex numbers $d_1, d_2, \ldots, d_N$.
3) The M-GEF output $y = [y_1 \; y_2 \ldots y_N]$ has uncorrelated components, i.e., $E\{y_i y_j^*\} = 0$ for all $i \neq j$.

**Proof:** First of all, notice that the complex number $d_n$ in the second assertion can be written as $d_n = e_n^H h_n$ for any $n \in \{1, 2, \ldots, N\}$. For this reason, $d_n$ is necessarily non-zero for any $n$, since otherwise $\lambda_n$ would be zero. Now, (i) The first assertion implies the second one according to Lemma 1. (ii) Also the inverse is correct, since assuming that the second assertion is true and using the definition of M-GEF filter $e_n$ we have for all $n$, that

$$d_n^* h_n = R_n e_n = \lambda_n R'_n e_n = \lambda_n \sigma^2 e_n.$$  (11)

Hence, $A = E\tilde{D}$ with appropriate diagonal matrix $\tilde{D}$ and, further, $A^H A = \tilde{D}^H E^H A = \tilde{D}^H D$. Finally, (iii) the second and the third assertions are equivalent, since the covariance of the observation is

$$E\{xx^H\} = AA^H + \sigma^2 I = \sum_{k=1}^{M} R_k + \sigma^2 I = R_n + R'_n \quad \text{for any } n.$$  (12)

Hence, the correlation between the $i$-th and $j$-th ($i \neq j$) output components, $E\{y_i y_j^*\}$, reduces now as follows:

$$E\{y_i y_j^*\} = e_i^H (R_j + R'_j) e_j = (1 + \frac{1}{\lambda_j}) e_i^H R_j e_j.$$  (13)

Here, $\lambda_j$ is strictly positive by definition, thus, the correlation is zero if and only if

$$0 = e_i^H R_j e_j = e_i^H h_j h_j^H e_j.$$  (14)

Therefore and because the mixing matrix $A = [h_1 \; h_2 \ldots h_N]$ is assumed to be full rank, all the output components are mutually uncorrelated if and only if $E^HA = D$. ■

A conventional diversity combining, or maximal ratio combining (MRC) as often referred to, is closely related to M-GEF. Nevertheless, in MRC, the main objective is to maximize the signal-to-(AWG)-noise ratio (SNR) and, consequently, MRC ignores possible interfering source components which results in a suboptimum SINR performance, in general. However, M-GEF wrt., say, the $n$-th source component is consistent with MRC (wrt. the same component) in the sense that, without interfering source components, $s_k$, $k \neq n$, these two methods coincide. This is easy to see by setting interfering sources to zero and constraining $w$ to have, say, a unit norm in which case (2) reduces to

$$\bar{\rho}_n(w) = \frac{w^H R_n w}{\sigma^2}.$$  (15)
Thus now, maximizing (15) is an ordinary eigenvalue problem, which, for one, is well known to yield the MRC solution [8], i.e.,
\[
\arg \max_{w \in \mathbb{C}^M} \hat{\rho}_n(w) = ch_n \text{ for any } c \in \mathbb{C}.
\]

An important difference between MRC and M-GEF is, nevertheless, the order of observation statistics needed to solve the problems blindly. Estimation of the covariance of the observation vector \( x \), \( C_x = \mathbb{E}\{xx^H\} = R_n + \sigma^T I \) (in noise-only system), is sufficient to solve the MRC problem, since
\[
\arg \max_{w \in \mathbb{C}^M} \hat{\rho}_n(w) = \arg \max_{w \in \mathbb{C}^M} w^H C_x w.
\]

Hence, the MRC solution is obtainable blindly using only second-order statistics of \( x \). Also SINR is a second-order measure of output’s “goodness” under general model (1), or more precisely, M-GEF depends only on second-order statistics of the contribution of the desired source, \( R_n \), and second-order statistics of the contribution of interference and noise, \( R'_n \), in the observation \( x \). However, \( R_n \) and \( R'_n \) can not be separated blindly from second-order statistics of the observation (\( C_x = R_n + R'_n \) for all \( n \)), but both are, anyway, needed separately when solving M-GEF problem. This suggests that blind linear output SINR maximization is rather a generalization of higher-order (noise-free) blind source separation than of MRC.

IV. INDEPENDENT COMPONENT ANALYSIS

In basic ICA, the goal is essentially to invert the model (1) blindly, that is, to find a demixing matrix \( B \in \mathbb{C}^{N \times M} \) such that \( BA \) is as close to identity as possible by using only the observations \( x \). Because of the blindness, a solution of the ICA problem, \( B \), can be unique only up to left multiplication by an arbitrary permutation and diagonal matrices. Blind identifiability of such a \( B \) is proved for the noise-free ICA model in [9] and for the noisy model in [10]. Nevertheless, in general, transforming the observation by inverse of \( A \) does not lead, as such, to the best linear SINR gain (i.e., to M-GEF solution) in a noisy system due to arbitrary noise amplification. (This is also seen in Lemma 2.) Naturally, a linear SINR maximizing transformation can be constructed as the LMMSE matrix (8) after the identification of \( A \) given that the observation covariance, \( C_x \), is also estimated. However, experimental results in [2] suggest that some ICA algorithms developed for noise-free models are able to provide directly input-output SINR gains very close to the best linear gain possible, in particular, the gains clearly better than with inverse transform of \( A \).

Good performance under noise can be basically explained by whitening which is accomplished in typical ICA algorithms, that is, the observed signal, \( x \), is first transformed linearly to vector \( z = V x \) such that \( \mathbb{E}\{zz^H\} = I \) [1]. In some algorithms (in EASI algorithm, for instance), linear whitening is performed implicitly during the actual ICA separation procedure [4]. The whitening transformation, \( V \), is not unique, however, one popular way is to use \( V = C_x^{-\frac{1}{2}} \). The whitening changes the ICA model (1) into
\[
z(t) = V x = V A s(t) + V \eta(t).
\]
Hence, the new mixing matrix is $\hat{A} := VA$. Now, since covariance of $z$ equals to the identity matrix ($C_z = I$), $\hat{A}$ is exactly the LMMSE matrix, $C_z^{-1} \hat{A}$, for $z$. Consequently, after whitening, ICA algorithms using an optimization criterion that is invariant to additive Gaussian noise [1] actually estimate the M-GEF transformation (or inverse of it) directly.

Majority of the well known ICA algorithms assume, that the mixing matrix after whitening is orthogonal (or unitary in the complex valued case), which is valid, in general, only if the linear model is noise-free. In the noisy model (1), this assumption is not true, or in other words, the matrix $\hat{A}$ is not orthogonal (unitary) in general. For this reason, the algorithms using the orthogonality constraint can not attain the M-GEF solution exactly in theory, but one can think that they tend to produce an orthogonalized estimate of the M-GEF (or LMMSE) transformation. We demonstrate this in the following, by considering the widely applied EASI algorithm as an example of the algorithms that are originally intended to perform noise-free ICA and have the orthogonality constraint. Notice, that there exists also ICA algorithms that do not use the orthogonality constraint (see, e.g., [11]).

V. EASI ALGORITHM AND NOISY MODEL

EASI algorithm is a recursive online algorithm which operates on individual samples of observed data. One recursion step of the EASI algorithm, i.e., of searching the demixing matrix $B \in \mathbb{C}^{N \times M}$, is given as

$$B_{t+1} = B_t - \mu U_t(y_t)B_t,$$

in which $\mu$ is a scalar step size and the update matrix, $U_t(y_t) \in \mathbb{C}^{N \times N}$, is defined as

$$U_t(y_t) = y_t y_t^H - I + g(y_t)y_t^H - y_t g(y_t)^H.$$  \hspace{1cm} (19)

Here $y_t = B_t x_t$, $I$ stands for identity matrix and $g : \mathbb{C}^N \rightarrow \mathbb{C}^N$ is an arbitrary nonlinear function. On right-hand side of (19), two first terms tends to whiten the output $y_t$, thus, the algorithm uses implicitly the orthogonality constraint discussed in the previous section. Since only the current sample is used in each step of the algorithm, the update matrix (19) does not vanish asymptotically. Instead, a stationary point of the algorithm is defined stochastically as follows: [4]

**Definition 3** An $N \times M$ matrix $B'$ is a stationary point of the EASI algorithm if the expected value of the update term, $U(y)$, is zero, i.e.,

$$E\{yy^H - I + g(y)y^H - yg(y)^H\} = 0.$$ \hspace{1cm} (20)

for $y = B'x$.

The convergence of the EASI algorithm is given as stability of the stationary point in noise-free model [4]. In other words, at least local convergence is guaranteed in theory. In the following theoretical analysis, we do not pay attention to the issue of convergence but solely show that the EASI stationary point is very closely related to the M-GEF solution. The convergence is verified with numerical experiments in the section VI.
A. Blind Maximal Ratio Combining

It was seen in the section III that, in the particular case of an interference-free linear model (i.e., the model with one source and noise only), blind maximization of the linear output SNR (or equally SINR) can be carried out using second-order observation statistics. In this section, we show that also the EASI algorithm can maximize SNR in interference-free system. Of course, using a higher-order statistical method to solve the second-order statistical problem is not advisable in practise, if one knows beforehand that the system is interference-free. Sometimes the possible absence of interfering source components is not known in which case using the higher-order ICA method can be reasonable. The following result is also a natural starting point to the more general analysis in the next section.

Strictly speaking, we prove here that, in the particular case of the interference-free model,

\[ \tilde{x} = h s + \eta, \]  

(21)

the MRC filter, which thus maximizes linear output SNR, is related to stationary point of EASI algorithm provided that a nonlinearity function \( g = [g_1 \ g_2 \ldots g_N]^T : \mathbb{C}^N \rightarrow \mathbb{C}^N; \)

\[ g_n(z) = |z_n|^2 z_n, n = 1, \ldots, N, \]

(22)
is used. This is formulated rigorously in the following proposition.

**Proposition 1** Let \( \tilde{N} \in \{1, 2, \ldots, M\} \) be arbitrary and assume the model (21). Let \( \{\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_{\tilde{N}}\} \) be a set of orthonormal vectors in \( \mathbb{C}^M \backslash \{0\} \) such that \( \tilde{e}_1 \) equals to normalized MRC filter, i.e., \( \tilde{e}_1 = \frac{h}{\|h\|}. \)

Further, let \( g : \mathbb{C}^{\tilde{N}} \rightarrow \mathbb{C}^{\tilde{N}} \) be defined as in (22) and \( \tilde{E} := [\tilde{e}_1 \ \tilde{e}_2 \ldots \tilde{e}_{\tilde{N}}]. \) Now, \( \exists \alpha = [\alpha_1 \ \alpha_2 \ldots \alpha_{\tilde{N}}]^T \in (\mathbb{C} \backslash \{0\})^{\tilde{N}} \) such that

\[ \tilde{B}_\alpha := \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_{\tilde{N}})\tilde{E}^H \]

(23)
is a stationary point of the EASI algorithm, i.e.,

\[ \mathbb{E} \{ \tilde{y}_\alpha^H \tilde{y}_\alpha - I + g(\tilde{y}_\alpha)\tilde{y}_\alpha^H - \tilde{y}_\alpha g(\tilde{y}_\alpha)^H \} = 0, \]

(24)
in which \( \tilde{y}_\alpha := \tilde{B}_\alpha \tilde{x}. \)

The following lemma is needed in proof of the proposition.

**Lemma 3** Let \( A = [a_1 \ a_2 \ldots a_M]^T \in \mathbb{C}^{M \times M} \) be arbitrary and \( \eta = [\eta_1 \ \eta_2 \ldots \eta_M]^T \) be a zero mean, complex valued Gaussian random vector with \( \mathbb{E} \{ \eta^H \eta \} = \sigma^2 I \) and \( \mathbb{E} \{ \eta^T \eta \} = 0. \) Then

\[ \mathbb{E} \{ \eta^H A \eta^H \} = \sigma^4 (A + \text{Tr}(A) I). \]

(25)

**Proof of the lemma:** Let \( i, j \) and \( k \in \{1, 2, \ldots, M\} \) and \( \epsilon_m, \ m = 1, 2, \ldots, M, \) denote the \( m \)-th Euclidean basis vector of \( \mathbb{R}^M. \) Then, first,

\[ \mathbb{E} \{ \eta_i \eta_j^* \eta_k^* \eta \} = \begin{cases} 0 & \text{if } i \neq j \text{ and } i \neq k \\ \sigma^4 \epsilon_j & \text{if } i \neq j \text{ and } i = k \\ \sigma^4 \epsilon_k & \text{if } i = j \neq k \\ 2\sigma^4 \epsilon_j & \text{if } i = j = k. \end{cases} \]

(26)
Consequently, the expected value of \((i, j)\)-th element of the matrix \(\eta^H A \eta^H\) is
\[
\mathbb{E}\left\{ (\eta^H A \eta^H)_{i,j} \right\} = \mathbb{E}\left\{ \eta_i \eta_j^* \eta^H A \eta \right\} = \sum_{k=1}^{M} a_k^T \mathbb{E}\left\{ \eta_k \eta_k^* \eta \right\}
\]
\[
= \begin{cases} 
\sigma^4 a_{ij} & \text{if } i \neq j \\
\sigma^4 (a_{ij} + \sum_{k=1}^{M} a_k^T \epsilon_k) & \text{if } i = j 
\end{cases} \tag{27}
\]

**Proof of Prop. 1:** To begin with, let us prove that the higher-order term of (24),
\[
\tilde{H}_\alpha := \mathbb{E}\left\{ g(\tilde{y}_\alpha) \tilde{y}_\alpha^H - \tilde{y}_\alpha g(\tilde{y}_\alpha)^H \right\} \in \mathbb{C}^{N \times \bar{N}},
\tag{28}
\]
vanesishes for all \(\alpha \in (\mathbb{C}\setminus\{0\})^M\). First, diagonal elements of \(\tilde{H}_\alpha\) are zeros due to skew-symmetry. Second, the \((i, j)\)-th \((i \neq j)\) off-diagonal element of \(\tilde{H}_\alpha\) is given as
\[
(\tilde{H}_\alpha)_{i,j} = \begin{cases} 
\mathbb{E}\left\{ \tilde{y}_i \tilde{y}_j (|\tilde{y}_i|^2 - |\tilde{y}_j|^2) \right\} & \text{if } i < j \\
\mathbb{E}\left\{ \tilde{y}_i \tilde{y}_j (|\tilde{y}_i|^2 - |\tilde{y}_j|^2) \right\} & \text{if } i > j 
\end{cases} \tag{29}
\]
in which \(\tilde{y}_n\) is the \(n\)-th component of \(\tilde{y}_\alpha\). Hence, it is enough to show that the off-diagonal elements are zeros for all \(i < j\). Now, since \(h^H e_j = 0\),
\[
\mathbb{E}\left\{ |\tilde{y}_i \tilde{y}_j|^2 \right\} = \mathbb{E}\left\{ \tilde{e}_i^H \tilde{y}^H \tilde{e}_j \tilde{y}^H \tilde{e}_j \right\} \tag{30}
\]
\[
= \mathbb{E}\left\{ s |s|^2 \right\} \tilde{e}_i^H h \mathbb{E}\left\{ \eta^H \right\} \tilde{e}_j \tilde{e}_i^H h^H \tilde{e}_j \tag{I}
\]
\[
+ \mathbb{E}\left\{ |s|^2 \right\} \mathbb{E}\left\{ \tilde{e}_i^H h \eta^H \tilde{e}_j \tilde{e}_i^H h^H \tilde{e}_j \right\} \tag{II}
\]
\[
+ \mathbb{E}\left\{ |s| \right\} \mathbb{E}\left\{ \tilde{e}_i^H h \eta \tilde{e}_j \tilde{e}_i^H h^H \tilde{e}_j \right\} \tag{III}
\]
\[
+ \mathbb{E}\left\{ s^2 \right\} \mathbb{E}\left\{ \tilde{e}_i^H \eta \tilde{e}_j \tilde{e}_i^H \eta^H \tilde{e}_j \right\} \tag{IV}
\]
\[
+ \mathbb{E}\left\{ s \right\} \mathbb{E}\left\{ \tilde{e}_i^H \eta \tilde{e}_j \tilde{e}_i^H \eta^H \tilde{e}_j \right\} \tag{V}
\]
\[
+ \mathbb{E}\left\{ s^2 \right\} \mathbb{E}\left\{ \tilde{e}_i^H \eta \tilde{e}_j \tilde{e}_i^H \eta^H \tilde{e}_j \right\} \tag{VI}
\]
\[
+ \mathbb{E}\left\{ s \right\} \mathbb{E}\left\{ \tilde{e}_i^H \eta \tilde{e}_j \tilde{e}_i^H h \eta \tilde{e}_j \right\} \tag{VII}
\]
\[
+ \mathbb{E}\left\{ s^2 \right\} \mathbb{E}\left\{ \tilde{e}_i^H \eta \tilde{e}_j \tilde{e}_i^H h \eta \tilde{e}_j \right\} \tag{VIII}
\]
Above, for all \(i < j\),
\begin{itemize}
  \item \(a\) terms \(I, II, IV, VI\) and \(VII\) are zero, since \(s\) is zero mean and second-order circular symmetric \((\mathbb{E}\left\{ s \right\} = \mathbb{E}\left\{ s^2 \right\} = 0\) and \(\mathbb{E}\{\eta\} = 0\);
  \item \(b\) terms \(III\) and \(V\) are zero, since \(\mathbb{E}\left\{\eta^H\right\} = \sigma^2 I\) and \(\tilde{e}_i^H \tilde{e}_j = 0\) and
  \item \(c\) term \(VIII\) is zero, since using Lemma 3 we have that
\end{itemize}
\[
\mathbb{E}\left\{ \tilde{e}_i^H \eta \tilde{e}_j \tilde{e}_i^H \eta \tilde{e}_j \right\} = \tilde{e}_i^H (\tilde{e}_j \tilde{e}_i^H + \text{Tr}(\tilde{e}_j \tilde{e}_i^H) I) \tilde{e}_j = 0. \tag{31}
\]
Hence, $\mathbb{E}\left\{\tilde{y}_i\tilde{y}_j^*|\tilde{y}_j|^2\right\} = 0$ and, similarly, also $\mathbb{E}\left\{\tilde{y}_i\tilde{y}_j^*\right\} = 0$ for all $\alpha \in (\mathbb{C}\backslash \{0\})^N$. Therefore, the off-diagonal terms and, consequently, the entire higher-order term $\tilde{H}_\alpha$ vanish for all $\alpha \in (\mathbb{C}\backslash \{0\})^N$.

To finish the proof, we need that also the second-order term of (24), $\mathbb{E}\{\tilde{y}_\alpha\tilde{y}_\alpha^H - I\}$, vanishes for appropriate selection of the vector $\alpha \in (\mathbb{C}\backslash \{0\})^\tilde{N}$. Now, denoting $D_\alpha = \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_\tilde{N})$, the covariance term is

$$
\mathbb{E}\{\tilde{y}_\alpha\tilde{y}_\alpha^H\} = \tilde{B}_\alpha h h^H \tilde{B}_\alpha^H + \sigma^2 \tilde{B}_\alpha \tilde{B}_\alpha^H = D_\alpha \tilde{E}^H h h^H \tilde{E} \tilde{D}_\alpha^H + \sigma^2 D_\alpha \tilde{E}^H \tilde{E} \tilde{D}_\alpha^H
$$

(32)

From which we see that the components of $\tilde{y}_\alpha$ are uncorrelated for all $\alpha \in (\mathbb{C}\backslash \{0\})^\tilde{N}$ and, surely, it is possible to choose $\alpha_1, \alpha_2, \ldots$ and $\alpha_\tilde{N} \in \mathbb{C}$ such that the covariance matrix equals unity.

Remark 2 (i) In Prop. 1, the selection that the vector $\tilde{e}_1$ is the MRC filter, is fully technical. Selecting any vector $\tilde{e}_n, n \in \{1, 2, \ldots, \tilde{N}\}$, to be the MRC filter leads to the same conclusion. This is, of course, consistent with permutation unambiguity of the EASI algorithm.

(ii) The output dimension, $\tilde{N}$, of transformation $\tilde{E}$ can be chosen freely (as far as $\tilde{N} \leq M$). Consequently, the use of the higher-order EASI algorithm to solve the second-order MRC problem is sensible also in practise, if it is not known beforehand wether the system is interference-free.

B. Blind SINR Maximization

The numerical experiments in the next section and also in [2] show that the EASI algorithm provides almost identical SINR performance with M-GEF in practice also under the general model (1). Unfortunately, in theory, the performances are not exactly equal in general, which is seen as follows. Since the higher-order term, $g(y)y^H - yg(y)^H$, in the EASI algorithm (19) is skew-Hermitian, the output, $\hat{y}$, corresponding to any stationary point of the algorithm is necessarily white, i.e., $\mathbb{E}\{\hat{y}\hat{y}^H\} = I$, due to the structure of the second-order term, $yy^H - I$. On the other hand, according to Lemma 2, the M-GEF output, $y$, has correlated components unless the mixing matrix $A$ has orthogonal columns. Therefore, the M-GEF transformation, $E$, can not be, in general, exactly a stationary point of the EASI algorithm.

It is, however, interesting to notice that the M-GEF transformation is, indeed, a stationary point of the EASI algorithm, if the mixing matrix has orthogonal columns. This is stated in the following proposition.

Proposition 2 Assume the model (1) with a mixing matrix $A$ for which $A^HA = D_A$ for some diagonal matrix $D_A \in \mathbb{C}^{N \times N}$ and let the nonlinearity $g$ be defined again as in (22). Now, $\exists \alpha = [\alpha_1 \alpha_2 \ldots \alpha_\tilde{N}]^T \in (\mathbb{C}\backslash \{0\})^N$ such that

$$
B_\alpha := \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_\tilde{N})E^H
$$

(33)

is a stationary point of the EASI algorithm, i.e.,

$$
\mathbb{E}\{y_\alpha y_\alpha^H - I + g(y_\alpha)y_\alpha^H - y_\alpha g(y_\alpha)^H\} = 0,
$$

(34)

in which $y_\alpha := B_\alpha x$. 
Fig. 1. Output SINR gains as a function of SIR wrt. the strongest component. $N = 4$ source components as well as $M = 4$ observations are used. Observed SNR is fixed to 5 dB.

Fig. 2. Output SINR gains as a function of observed SNR. All $N = 4$ sources have equal observed power. Number of observations is $M = 4$.

Proof: In essence, the proof is similar to the proof of Prop. 1. Namely, Lemma 1 implies that, for all $n \in \{1, 2, \ldots, N\}$, $e_n^H A s = c_n s_n$ for some complex constant $c_n$, from which it follows easily that the higher-order term in (34) vanishes for all selections of coefficient vector $\alpha \in (C \setminus \{0\})^N$. In addition, it is also straightforward to see that it is possible to choose such a vector $\alpha \in (C \setminus \{0\})^N$ that the covariance of $y_\alpha$ equals to unity. (The covariance is diagonal for all the selections according to Lemma 2.) \hfill \blacksquare

Remark 3 The Hermitian transpose of the M-GEF transformation with appropriate scale of columns is exactly a stationary point of the EASI algorithm if and only if the mixing matrix $A$ has orthogonal columns (i.e., orthogonality is the necessary and sufficient condition). This is a direct consequence of Prop. 2 and Lemma 2.

VI. NUMERICAL EXPERIMENTS

Numerical results in this section set against the performance of EASI algorithm and SINR-maximizing M-GEF approach under noisy environment. Also SINR performances of ordinary maximal ratio combining (MRC) and inversion of the mixing matrix, the matrix $A$ in model (1), are simulated in the experiments. Both of the latter methods are, thus, suboptimum since both interfering source components and additive noise are present. Here, signal-to-noise ratio (SNR) wrt. $n$-th source signal is defined as the average ratio of the power of $n$-th source signal’s contribution and additive noise power in the observed signals. The signal-to-interference ratio (SIR) wrt. $n$-th source signal, in turn, is the average ratio of the power of $n$-th source signal’s contribution and the powers of other components’ contribution in observed signals. Given the power normalization of the formal sources stated below (1), the SIR values other than 0 dB are implemented by appropriate scaling of the mixing coefficients.

In the experiments, four QPSK sources ($N = 4$) and four mixtures ($M = 4$) of them are used. The selection of source constellations is more or less arbitrary, and it should not affect the general
validity of the results. Mixing coefficients, i.e., elements of the matrix $A$, are drawn randomly from zero mean Gaussian distributions (one distribution for each source component) for each processing block of $N = 50000$ symbols of data. Variances of these distributions are selected such that observed SNR and SIR values correspond to given values on average. M-GEF bound is evaluated directly from the data model for each block. Hence, the bounds are not affected by finite sample statistics and, more importantly, they are the absolute upper bounds among all linear transformations of received data in case of each realization. Also output SIR’s of the MRC and inversion of $A$ are evaluated from the model.

The third-order nonlinearity (22) is used in the EASI algorithm in all the following experiments. A permutation ambiguity of EASI outputs is circumvented by, first, evaluating the output SIR wrt. all the source components for all EASI outputs and, then, selecting the maximum ones. Practical ways to identify the output components are not considered in this paper. All the gains plotted are average gains wrt. the received SIR over one thousand mixture realization.

In Fig. 1, it is assumed that one of the sources is the desired one and the other three are interfering ones. The figure shows SIR gains for the desired source as a function of SIR with fixed observed SNR.
The figure illustrates that the gain of EASI algorithm is nearly identical with M-GEF bound. A difference is roughly 0.25 dB-unit for low SIR values and tends to zero as SIR increases. This is well-consistent with the theoretic results, and especially with Prop. 1, since asymptotically (i.e., as SIR → ∞) the system tends to noise-only system. Notice also that asymptotic SINR gain for the EASI algorithm is roughly 6 dB (∼ 10\log_{10}(4) dB) which equals to theoretic MRC gain with four observations.

Figs. 2 and 3 give two examples of SINR gain vs. observed SNR with fixed received SIR. In the former figure, all the source components have equal observed power on average, which results in approximately SIR of -4.8 dB. Performance is plotted only wrt. one source component, since SINR gains wrt. the other components are, naturally, similar in this case. The latter figure gives an example of system with unequal observed average powers between the sources. Observed average powers of the three weakest sources are -15, -10, and -5 dB wrt. the strongest source. Both figures indicate that the performance behavior of EASI is almost identical with the M-GEF bound in practice, although, in theory, EASI is not able to attain exactly the bound in general.

VII. CONCLUSIONS

In this paper, we illustrated that basic independent component analysis (ICA) designed for noise-free linear models is able to solve, blindly and directly, the generalized eigenvalue problem, i.e., to provide essentially the best possible output SINR among all linear transformations of observed data, in the challenging case of having both additive noise and interference disturbing the desired signal observation. However, ICA algorithms constraining the estimated de-mixing matrix to be orthogonal (or unitary) cannot exactly attain the optimal solution in general, but in a sense they produce an orthogonalized version of the solution. In addition, the theoretical and numerical results of the paper showed that one of the most widely applied ICA algorithms, the equivariant adaptive source identification (EASI) algorithm, is, in practice, identical with SINR-maximizing generalized eigenfiltering. Strictly speaking, we gave the necessary and sufficient condition under which the stationary point of the EASI algorithm maximizes the linear output SINR. We also proved that, in the special case of interference-free (that is, noise only) system, the EASI algorithm can attain exactly the greatest diversity gain blindly, i.e., performs as a blind maximal ratio combiner (MRC). In addition, the numerical results were given to show that the performance of the EASI algorithm is remarkably close to the optimal (i.e., the maximal output SINR among all linear transforms of observed data) also in cases in which the above mentioned theoretical optimality condition is not met.

REFERENCES


