

**Semi-Blind Source Separation via Sparse Representations  
and Online Dictionary Learning**

**A THESIS  
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF MINNESOTA  
BY**

**Sirisha Rambhatla**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
MASTER OF SCIENCE**

**Prof. Jarvis Haupt**

**December, 2012**

© Sirisha Rambhatla 2012  
ALL RIGHTS RESERVED

# Acknowledgements

It is only at the end of a journey that we realize how each event steered us forward towards our goal. I consider the time spent during my masters as transformative, both at personal and professional level.

Looking back, I consider myself fortunate to have the company and guidance of individuals who constantly encouraged me, most importantly supported me through the thick and the thins. To begin, I would like to extend my sincere thanks to my advisor Prof. Jarvis Haupt, who motivated me to take-up this fun, yet challenging problem. It was my pleasure to work and collaborate with him and I really cannot thank him enough for his encouragement and enthusiasm.

I am honored to have Prof. Zhi-Quan Luo and Prof. Arindam Banerjee on my committee and I would like to thank them for taking out time to guide me, despite their busy schedules.

Also, I take this opportunity to thank my mother Jayalakshmi, my father Dr. Rambhatla Gouri Shankar Sastry and of course my sister Sowmya, who have always surrounded me with immeasurable love and support. In addition, I would like to make special note of my friends Anki, Anuja, Aditi, Shivangi, Vikrant and Yash for being awesome, and my lab-mate and colleague Akshay for all the insightful discussions and much needed coffee breaks.

To Mom and Dad, who taught me:

*“Dhairye, Saahase lakshmi”.*

*“Courage and fortitude are the only wealth a person needs.”*

## Abstract

This work examines a semi-blind source separation problem having applications in audio, image, and video processing. The essential aim is to separate one source whose local structure is partially or approximately known from another *a priori* unspecified but structured source, given only a single linear combination of the two sources. We propose a novel separation technique based on local sparse approximations. A key feature of our procedure is the *online* learning of dictionaries (using only the data itself) to sparsely model the unknown structured background source. The performance of the proposed approach is demonstrated via simulation in a stylized application of audio source separation.

# Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>List of Tables</b>	<b>vi</b>
<b>List of Figures</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Background and Problem Formulation</b>	<b>3</b>
2.1 Prior Art . . . . .	4
2.1.1 Low Rank and Robust Low Rank Approximation . . . . .	4
2.1.2 Low Rank Plus Sparse in a Known Dictionary . . . . .	5
2.1.3 Morphological Component Analysis . . . . .	6
2.2 Our Contribution: “Semi-blind” Morphological Component Analysis . . .	6
<b>3 Semi-Blind Morphological Component Analysis (SBMCA)</b>	<b>7</b>
3.1 Dictionary Learning Stage . . . . .	8
3.2 Sparse Approximation Stage . . . . .	9
3.3 Initialization of SBMCA . . . . .	9
3.4 A Note on Debiasing . . . . .	9
3.5 SBMCA vs MCA : Is there a difference? . . . . .	10
<b>4 Evaluation: An Application in Audio Forensics</b>	<b>11</b>
4.1 Source Separation in the Frequency Domain . . . . .	13

4.2 Source Separation in the Time Domain . . . . .	25
<b>5 Conclusion and Discussion</b>	<b>37</b>
<b>References</b>	<b>39</b>
<b>Appendix A. Glossary and Acronyms</b>	<b>42</b>
A.1 Glossary . . . . .	42
A.2 Acronyms . . . . .	43

# List of Tables

4.1	Analysis of reconstruction SNR(in dB): Frequency Domain Separation . .	13
4.2	SBMCA Algorithm Parameters for the Frequency Domain Source Separation . . . . .	14
4.3	Analysis of reconstruction SNR(in dB): Time Domain Separation . . . .	25
4.4	SBMCA Algorithm Parameters for Time Domain Source Separation . . .	26
A.1	Acronyms . . . . .	43

# List of Figures

4.1	A segment of mixture components (noise free): (a) the nominally periodic signal $x_p$ (each segment is the discharge corresponding to one of the two resistive load states, randomly selected and with random timing offsets); (b) the background signal $x_u$ ; (c) the mixture $x$ . . . . .	12
4.2	Frequency Domain Source Separation using SBMCA. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal. . . . .	15
4.3	Frequency Domain Source Separation using MCA-DCT-Fourier. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal. . . . .	16
4.4	Frequency Domain Source Separation using MCA-Identity-Fourier. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal. . . . .	17
4.5	A segment of extracted nominally periodic signal $x_p$ (Frequency Domain Separation). (a) The original nominally periodic signal, (b) and (d) The extracted nominally periodic signal via SBMCA and MCA-Identity-Fourier respectively, (c) and (e) The absolute error between the extracted and original nominally periodic signal via SBMCA and MCA-Identity-Fourier respectively. . . . .	18
4.6	A segment of extracted speech signal $x_u$ (Frequency Domain Separation). (a) The original speech signal, (b) and (d) The extracted speech signal via SBMCA and MCA-Identity-Fourier respectively, (c) and (e) The absolute error between the extracted and original speech signal via SBMCA and MCA-Identity-Fourier respectively. . . . .	19

4.7	Histogram of normalized error-per-block of extracted nominally periodic signal $x_p$ and extracted speech signal $x_u$ with noise $\sigma = 0$ for frequency domain source separation using Semi-blind MCA, MCA-DCT-Fourier and MCA-Identity-Fourier for the audio forensic application. . . . .	21
4.8	Histogram of normalized error-per-block of extracted nominally periodic signal $x_p$ and extracted speech signal $x_u$ with noise $\sigma = 0.001$ for frequency domain source separation using Semi-blind MCA, MCA-DCT-Fourier and MCA-Identity-Fourier for the audio forensic application. . . . .	22
4.9	Histogram of normalized error-per-block of extracted nominally periodic signal $x_p$ and extracted speech signal $x_u$ with noise $\sigma = 0.01$ for frequency domain source separation using Semi-blind MCA, MCA-DCT-Fourier and MCA-Identity-Fourier for the audio forensic application. . . . .	23
4.10	Histogram of normalized error-per-block of extracted nominally periodic signal $x_p$ and extracted speech signal $x_u$ with noise $\sigma = 0.1$ for frequency domain source separation using Semi-blind MCA, MCA-DCT-Fourier and MCA-Identity-Fourier for the audio forensic application. . . . .	24
4.11	Time Domain Source Separation using SBMCA. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal. . . . .	27
4.12	Time Domain Source Separation using MCA-DCT. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal. . . . .	28
4.13	Time Domain Source Separation using MCA-Identity. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal. . . . .	29
4.14	A segment of extracted nominally periodic signal $x_p$ (Time Domain Separation). (a) The original nominally periodic signal, (b) and (d) The extracted nominally periodic signal via SBMCA and MCA-DCT respectively, (c) and(e) The absolute error between the extracted and original nominally periodic signal via SBMCA and MCA-DCT respectively. . . . .	30

4.15	A segment of extracted speech signal $x_u$ (Time Domain Separation). (a) The original speech signal, (b) and (d) The extracted speech signal via SBMCA and MCA-DCT respectively, (c) and(e) The absolute error between the extracted and original speech signal via SBMCA and MCA-DCT respectively. . . . .	31
4.16	Histogram of normalized error-per-block measured using the vector $l_2$ norm of extracted nominally periodic signal $x_p$ and extracted speech signal $x_u$ with zero mean gaussian noise $\sigma = 0$ for time domain separation using Semi-blind MCA and MCA with DCT and Identity basis for the audio forensic application. . . . .	32
4.17	Histogram of normalized error-per-block measured using the vector $l_2$ norm of extracted nominally periodic signal $x_p$ and extracted speech signal $x_u$ with zero mean gaussian noise $\sigma = 0.001$ for time domain separation using Semi-blind MCA and MCA with DCT and Identity basis for the audio forensic application. . . . .	33
4.18	Histogram of normalized error-per-block measured using the vector $l_2$ norm of extracted nominally periodic signal $x_p$ and extracted speech signal $x_u$ with zero mean gaussian noise $\sigma = 0.01$ for time domain separation using Semi-blind MCA and MCA with DCT and Identity basis for the audio forensic application. . . . .	34
4.19	Histogram of normalized error-per-block measured using the vector $l_2$ norm of extracted nominally periodic signal $x_p$ and extracted speech signal $x_u$ with zero mean gaussian noise $\sigma = 0.1$ for time domain separation using Semi-blind MCA and MCA with DCT and Identity basis for the audio forensic application. . . . .	35

# Chapter 1

## Introduction

The blind source separation (BSS) problem entails separating a collection of signals, each comprised of a superposition of some unknown sources, into their constituent components (A canonical example of the BSS task arises in the so-called *cocktail party problem*). A number of methods have been proposed to address this problem. Perhaps the most well-known among these is independent component analysis (ICA) [1], where the sources are assumed to be independent non-Gaussian random vectors. Other approaches entail more classical matrix factorization techniques like principal component analysis (PCA) [2–4], or, when appropriate for the underlying model, non-negative matrix factorization (NNMF) [5].

Here we focus on a slightly different, and often more challenging setting – the so-called *single channel* source separation problem – where only a single mixture of the source signals is observed. Single channel source separation problems require the use of some additional *a priori* knowledge about the sources and their structure in order to perform separation [6–9]. Here, we assume that the local structure of one of the source signals is *approximately* known, and our aim is to separate this partially known source from an unknown “background” source.

Our approach is based on local sparse approximations of the mixture data. A novel feature of our proposed method is in our representation of the unknown background source – we describe a technique for learning (from the data itself) a *model* that sparsely represents the unknown background source, using tools from the dictionary learning literature (see, eg., [10–12]).

Chapter 2 sets the stage for the problem we consider here more formally, and discusses the nature of our contributions in the context of existing works in sparse representation, dictionary learning, and low-rank modeling. In Chapter 3 we discuss our algorithm and describe the various facets of its implementation. We then examine the performance of our proposed approach in an audio source separation problem encountered in an audio forensics application in Chapter 4. Finally, we discuss conclusions and possible extensions in Chapter 5.

## Chapter 2

# Background and Problem Formulation

As discussed in the previous chapter, our effort here is motivated by a single-channel semi-blind source separation problem, in which the goal is to separate a nominally periodic and approximately known signal from unknown but structured background interference, given only a superposition of the two sources.

Let  $x \in \mathbb{R}^n$  represent our observed data, and suppose that  $x$  may be decomposed as a sum of two unknown sources – one of which ( $x_p \in \mathbb{R}^n$ ) exhibits local structure that is partially known, and the other ( $x_u \in \mathbb{R}^n$ ) is unknown. Our investigation is motivated by an audio source separation task, where  $x$  is comprised of samples of an underlying continuous time waveform, and we consider  $x_p$  to be samples of a source that is a nominally regular repetition of one of a small number of temporally local prototype signals. One example scenario where this model is applicable is the case where  $x_p$  is, except for some unknown offset jitter, periodic. Our aim is to separate the sources  $x_p$  and  $x_u$  from observations of  $x$ , which may be noisy or otherwise corrupted.

Our proposed approach is based on the principle of local sparse approximations. In order to state our overall problem in generality, we describe an equivalent model for our data  $x$  that facilitates the local analysis inherent to our approach. Let us suppose that  $m$  is an integer that divides  $n$  evenly, such that  $n/m = q$ , an integer. Then  $x \in \mathbb{R}^n$  may

be represented equivalently as a  $m \times q$  matrix  $X$ :

$$X = X_p + X_u, \tag{2.1}$$

where  $X_p$  is a matrix whose columns are non-overlapping length- $m$  segments of  $x_p$ , and similarly for  $X_u$ . The goal of our effort is, in essence, to separate  $X$  into its constituent matrices  $X_p$  and  $X_u$ .

As alluded above, our separation approach entails leveraging local structure in each of the components of  $X$ . Our main contribution comes in the form of a procedure that, given our “partial” information about the columns of  $X_p$ , enables us to learn, in an online fashion and from the data itself, a dictionary  $D$  such that columns of  $X_u$  are accurately expressed as linear combinations of (a small number of) columns of  $D$ . In a broader sense, our work is related to some classical approximation approaches as well as several recent works on matrix decomposition. We briefly describe these background and related efforts here, in an effort to put our main contribution in context.

## 2.1 Prior Art

### 2.1.1 Low Rank and Robust Low Rank Approximation

Consider the model (2.1) and suppose that the columns of  $X_p$  can each be represented as a linear combination of some  $r$  linearly independent vectors, implying that  $X_p$  is a matrix of rank  $r$ . Now, different separation techniques may be employed depending on our assumptions of  $X_u$ . Perhaps the simplest case is where  $X_u$  is random noise (e.g., having entries that are iid zero-mean Gaussian); in this case, the problem amounts to a denoising problem, which can be solved using ideas from low-rank matrix approximation. In particular, it is well-known that the approximation  $\widehat{X}_p$  obtained via the truncated (to rank  $r$ ) singular value decomposition (SVD) of  $X$  is a solution of the optimization

$$\widehat{X}_p = \arg \min_{L, \text{rk}(L) \leq r} \|X - L\|_F^2, \tag{2.2}$$

where  $\text{rk}(L)$  is the function that returns the rank of  $L$ .

It is well-known that certain (non-Gaussian) forms of interference  $X_u$  may cause the accuracy of estimators of the low-rank component obtained via truncated SVD to

degrade significantly. This is the case, for example, when  $X_u$  is comprised of sparse large (in amplitude) impulsive noise. In these cases, the low-rank approximation problem can be modified to its *robust* counterpart, which goes by the name of robust PCA in the literature [13, 14]. The robust PCA approach aims to simultaneously estimate both the low-rank  $X_p$  and the sparse  $X_u$ , by solving the convex optimization

$$\{\widehat{X}_p, \widehat{X}_u\} = \arg \min_{L, S} \|L\|_* + \lambda \|S\|_1 \quad \text{subject to } X = L + S, \quad (2.3)$$

where  $\lambda > 0$  is a regularization parameter. Here  $\|L\|_*$  denotes the *nuclear norm* of  $L$ , which is the sum of the singular values of  $L$ . The nuclear norm is a convex relaxation of the non-convex rank function  $\text{rk}(L)$ . Further,  $\|S\|_1$  is the sum of the absolute entries of  $S$  – essentially the  $\ell_1$  norm of a vectorized version of  $S$ , which is a convex relaxation of the non-convex  $\ell_0$  quasinorm that counts the number of nonzeros of  $S$ . Each of these procedures may be extended, in a natural way, to allow for noise or modeling error.

Here, of course, we explicitly assume that  $X_u$  is more highly structured, making the separation problem more well-suited to a new suite of techniques that explicitly exploit such structure.

### 2.1.2 Low Rank Plus Sparse in a Known Dictionary

A useful extension of the robust PCA approach arises in the case where  $X_u$  is not itself sparse, but possesses a sparse representation in some known dictionary or basis. One example is the case where the background source is locally smooth, implying it can be sparsely represented using a few low-frequency discrete cosine transform or Fourier basis elements. Formally, suppose that for some known matrix  $D$ , we have that  $X_u = DA_u$ , where the columns of  $A_u$  are sparse. The components of  $X$  can be estimated by solving the following optimization [15]

$$\{\widehat{X}_p, \widehat{A}_u\} = \arg \min_{L, A} \|L\|_* + \lambda \|A\|_1 \quad \text{subject to } X = L + DA \quad (2.4)$$

Note that an estimate  $\widehat{X}_u$  of  $X_u$  may be obtained directly as  $\widehat{X}_u = D\widehat{A}_u$ . This approach assumes (implicitly) *a priori* knowledge of a dictionary that sparsely represents the background signal, which may be a restrictive assumption in practice. Again, extensions to account for noise or modeling error are straightforward.

### 2.1.3 Morphological Component Analysis

A more general model arises when  $X_p$  is not low-rank, but instead, its columns are also sparsely represented in a known dictionary. Suppose that  $X_p$  and  $X_u$  are sparsely represented in some known dictionaries  $D_1$  and  $D_2$ , such that  $X_p = D_1 A_1$  and  $X_u = D_2 A_2$ , and that the columns of  $A_1$  and  $A_2$  are sparse. Such models were employed in recent work on Morphological Component Analysis (MCA) [16–18], which aimed to separate a signal into its component sources based on structural differences codified in the columns of the known dictionaries. The MCA decomposition can be accomplished by solving the following optimization

$$\{\widehat{A}_1, \widehat{A}_2\} = \arg \min_{A_1, A_2} \|A_1\|_1 + \lambda \|A_2\|_1 \text{ subject to } X = D_1 A_1 + D_2 A_2, \quad (2.5)$$

for some  $\lambda > 0$ ; the estimates of  $X_p$  and  $X_u$  may be formed as  $\widehat{X}_p = D_1 \widehat{A}_1$  and  $\widehat{X}_u = D_2 \widehat{A}_2$ , respectively. When  $X_p$  and  $X_u$  are each comprised of a single column, this optimization is equivalent to the so-called Basis Pursuit technique [19], which formed a foundation of much of the recent work in sparse approximation. Note that, as with the previously mentioned approach, this approach also assumes a priori knowledge of a dictionary that sparsely represents the background. Again, extensions to the noisy case are straightforward.

## 2.2 Our Contribution: “Semi-blind” Morphological Component Analysis

Our focus here is similar to the MCA approach above, but we assume only one of the dictionaries, say  $D_1$ , is known. In this case, the MCA approach transforms into a semi-blind separation problem where we try to also learn a dictionary  $D_2$  to represent the unknown signal. Our main contribution comes in the form of a “Semi-Blind” MCA procedure, designed to solve the following modified form of the MCA decomposition

$$\{\widehat{A}_1, \widehat{A}_2, \widehat{D}_2\} = \arg \min_{A_1, A_2, D_2} \|X - D_1 A_1 - D_2 A_2\|_F^2 + \widetilde{\lambda}_1 \|A_1\|_1 + \widetilde{\lambda}_2 \|A_2\|_1, \quad (2.6)$$

for  $\widetilde{\lambda}_1, \widetilde{\lambda}_2 > 0$ , and this problem forms the basis of the remainder of this thesis. Specifically, in Chapter 3 we propose a procedure, based on alternating minimization, for obtaining local solutions to optimizations of the form (2.6).

## Chapter 3

# Semi-Blind Morphological Component Analysis (SBMCA)

As described in Section 2.2 of Chapter 2, our model assumes that the data matrix  $X$  can be expressed as the superposition of two component matrices,  $X_p$  and  $X_u$ . Further, we assume that each of the component matrices possesses a sparse representation in some dictionary, such that  $X_p \approx D_1 A_1$  and  $X_u \approx D_2 A_2$ , where  $D_1$  is known a priori. Our essential aim, then, is to identify an estimate  $\widehat{A}_1$  of the coefficient matrix  $A_1$  and estimates  $\widehat{D}_2$  and  $\widehat{A}_2$  of the matrices  $D_2$  and  $A_2$ . Our estimates of the separated components are then given by  $\widehat{X}_p = D_1 \widehat{A}_1$ , and  $\widehat{X}_u = \widehat{D}_2 \widehat{A}_2$ .

We propose an approach to solve (2.6) that is based on alternating minimization, and is summarized here as Algorithm 1. We note that the lack of joint convexity makes the SBMCA algorithm sensitive to initialization. Therefore, any suitable initialization using sparse approximation techniques, depending upon the problem setting, can be employed. Let  $\lambda_1, \lambda_2, \lambda_3 > 0$  be user specified regularization parameters. Our initial estimate of coefficients  $A_1$ , corresponding to the coefficients of  $X_p$  in the known dictionary  $D_1$ , is obtained via

$$\widetilde{A}_1 = \arg \min_{A_1} \|X - D_1 A_1\|_F^2 + \lambda_1 \|A_1\|_1, \quad (3.1)$$

which is a simple LASSO-type problem. We then proceed in an iterative fashion, as outlined in the following sections, for a few iterations or until some appropriate convergence criteria is satisfied.

---

**Algorithm 1:** Semi-Blind MCA Algorithm
 

---

**Input:** Original Data  $X \in \mathbb{R}^{m \times q}$ , Known Dictionary  $D_1 \in \mathbb{R}^{m \times d}$ ,  
 Regularization parameters  $\lambda_1, \lambda_2, \lambda_3 > 0$ ,  
 Number of elements in unknown dictionary  $\ell$ .

**Initialize:**  $\tilde{A}_1 \leftarrow \arg \min_{A_1} \|X - D_1 A_1\|_F^2 + \lambda_1 \|A_1\|_1$   
 (or other suitable initialization depending on the problem.)

**Iterate (repeat until convergence):**

**repeat**

*Dictionary Learning:*

$$\{\tilde{D}_2, \tilde{A}_2\} \leftarrow \arg \min_{D_2, A_2} \|X - D_1 \tilde{A}_1 - D_2 A_2\|_F^2 + \lambda_2 \|A_2\|_1$$

*Coefficient Update:*

$$\begin{aligned} \tilde{D} &= [D_1 \ \tilde{D}_2] \\ [\tilde{A}_1^T \ \tilde{A}_2^T]^T &\triangleq \tilde{A} \leftarrow \arg \min_A \|X - \tilde{D} A\|_F^2 + \lambda_3 \|A\|_1 \end{aligned}$$

**until** convergence

**Output:** Learned dictionary  $\hat{D}_2 \leftarrow \tilde{D}_2$ ,  
 Coefficient estimates  $\hat{A}_1 = \tilde{A}_1$ ,  $\hat{A}_2 = \tilde{A}_2$ .

---

### 3.1 Dictionary Learning Stage

Given the estimate  $\tilde{A}_1$ , we can essentially “subtract” the current estimate of  $X_p$  from  $X$ , and apply a dictionary learning step to identify estimates of the unknown dictionary  $D_2$  and the corresponding coefficients  $A_2$ . In other words, we solve

$$\{\tilde{D}_2, \tilde{A}_2\} = \arg \min_{D_2, A_2} \|X - D_1 \tilde{A}_1 - D_2 A_2\|_F^2 + \lambda_2 \|A_2\|_1. \quad (3.2)$$

Now, given the estimate  $\tilde{D}_2$ , we update our current estimate of the *overall* dictionary  $\tilde{D} = [D_1 \ \tilde{D}_2]$ . We then update the *overall* coefficient matrix by solving another sparse approximation problem, as described next.

## 3.2 Sparse Approximation Stage

Given our current estimate of the overall dictionary, we update the corresponding coefficient matrices by solving the following LASSO-like problem:

$$[\tilde{A}_1^T \ \tilde{A}_2^T]^T \triangleq \tilde{A} = \arg \min_A \|X - \tilde{D}A\|_F^2 + \lambda_3 \|A\|_1. \quad (3.3)$$

Now, we extract the submatrix  $\tilde{A}_1$  from  $\tilde{A}$ , and repeat, beginning again with the dictionary learning step. These steps are iterated for a fixed number of iterations, or until some appropriate convergence criteria is satisfied, such as, when the difference between successive estimates of the quantities of interest are sufficiently small.

## 3.3 Initialization of SBMCA

As with any optimization problem of this nature (when the objective function is not jointly convex) initialization may play an instrumental role in ensuring the success of the procedure. Likewise, for the SBMCA approach, the lack of joint convexity makes the algorithm somewhat sensitive to the initialization. This can be observed from the dictionary learning step in Algorithm 1, where the elements of dictionary  $D_2$  depend on how efficiently the source one was extracted initially. Failure to initialize the algorithm properly may lead to interference of the sources and deteriorate the performance of the algorithm. To this effect, the initialization step in Algorithm 1 can be modified to any desirable initialization, such as OMP (Orthogonal Matching Pursuit) [20] formulation or Lasso, as the case may be.

## 3.4 A Note on Debiasing

Sparse approximation techniques such as Lasso, which employ  $l_1$  norm, the convex relaxation of  $l_0$  norm, are prone to “shrinkage” of coefficients. To reverse this effect, a debiasing step, which essentially is a least-square fit on the essential support (non-zero elements) of the coefficient matrix  $A$ , can be employed after every Lasso step or at the end of the SBMCA algorithm. However, there is a catch in terms of the sparsity basis learnt for the unknown background signal  $x_u$ . We note that debiasing

step involves choosing a single threshold to select the “essential” dictionary elements, because the relative sparsity of the representation of signals  $x_p$  and  $x_u$  in dictionaries  $D_1$  and  $D_2$  respectively, may vary (the magnitude of the coefficients in these dictionaries have different distribution), a chosen threshold for debiasing may be detrimental for one or both of the signals. Nevertheless, debiasing may be employed effectively, in cases where dictionaries  $D_1$  and  $D_2$  provide similar sparse representations for signals  $x_p$  and  $x_u$ , respectively, depending upon the problem setting

### 3.5 SBMCA vs MCA : Is there a difference?

At first glance the only difference between SBMCA and MCA appears to be the unknown dictionary  $D_2$ , in place of a known standard basis or dictionary. This sums up the algorithmic difference between the procedures, though a deeper difference lies in the interpretation of the problem.

Although a standard complete basis can always represent any signal, the sparsity of the corresponding representation is not ensured. For example, arbitrary waveforms such as speech can be represented by any orthonormal complete basis (the identity basis or the DCT basis, for example). Although speech is inherently bandlimited, its representation in DCT basis may not be overly sparse, due to occasional sharp transitions having wide-band frequency content. However, by using our online Dictionary Learning approaches, our aim is to learn frequently occurring patterns and to develop a data-dependent sparse representation. Further, the primary aim of using the dictionary learning procedure in SBMCA is to separate the sources, in the process we obtain the sparse basis for the unknown signal (an added bonus!).

Making an educated guess about the sparsity basis of a signal may not be possible in all the cases. In addition, as we will see in Chapter 4, the choice of basis effects the separation results. To this effect, an “online” methodology which does not require any knowledge about the unknown signal is attractive.

## Chapter 4

# Evaluation: An Application in Audio Forensics

We consider an audio source separation motivated by an application in audio forensics. Specifically, we examine a stylized version of a problem that arises during altercations where electroshock law enforcement devices are utilized.

For the sake of this example, we suppose that the electroshock devices discharge approximately 36 times per second, and that the waveforms generated by the device during the discharge may take one of two different forms depending on the level of resistive load encountered by the device. We suppose that we are able to obtain an audio recording of the altercation, which is comprised of audio corresponding to the nominally periodic discharge of the device, superimposed with background noise (eg., speech and/or other noises). Our aim is to separate the superposition into its components, with the ultimate goal of being able to determine whether the device encountered a highly resistive load or not (which correlates with whether the device is delivering current to the suspect, or not).

Figure 4.1 shows a segment of the signals used in our simulation. We simulate the form of the approximately periodic signals ( $x_p$ ), shown in Figure 4.1 (a), using two distinct exponentially decaying sinusoids (obtained using series RLC circuits with different parameters) to model the loaded and open circuit states. We generate two distinct waveforms, which correspond to the two states (high and low resistive load),

and form the overall signal  $x_p$  by concatenating randomly-selected versions of these prototype signals, each of which is subject to a few samples of timing offset in order to model the non-idealities of the actual electroshock device. A speech signal<sup>1</sup> shown in Figure 4.1 (b), was used to model background noise that may be present during the alteration. We simulate the overall raw audio data,  $x$  as a linear combination of  $x_p$ ,  $x_u$  and zero-mean random Gaussian noise vector whose entries are iid  $\mathcal{N}(0, \sigma^2)$ .

Figure 4.1 (c) depicts the signals of interest, along with their mixture, in the ideal case  $\sigma = 0$ . The data matrix  $X$  is formed from the signal  $x$  as discussed in Chapter 2 using non-overlapping segments with 400 samples each.

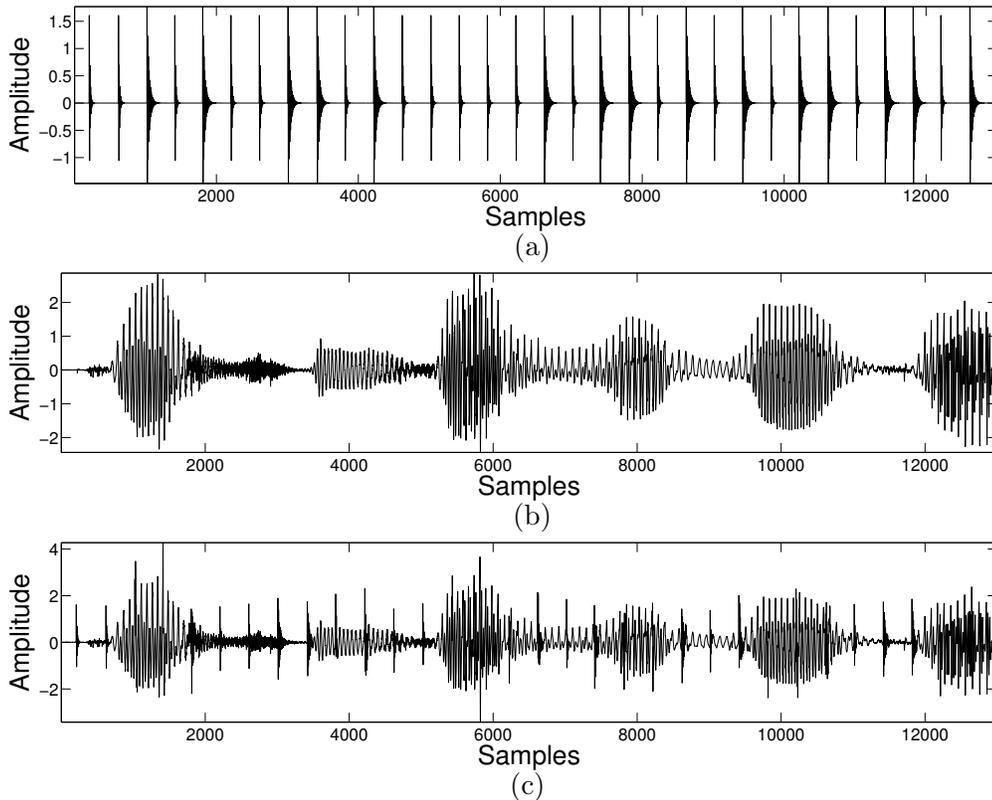


Figure 4.1: A segment of mixture components (noise free): (a) the nominally periodic signal  $x_p$  (each segment is the discharge corresponding to one of the two resistive load states, randomly selected and with random timing offsets); (b) the background signal  $x_u$ ; (c) the mixture  $x$ .

<sup>1</sup> Speech Samples obtained from VoxForge Speech Corpus: [www.voxforge.org/home](http://www.voxforge.org/home)

We now attempt to solve the source separation problem discussed above, by performing separation in both the frequency domain and the time domain, in Section 4.1 and Section 4.2 respectively. Further we analyze the performance of the SBMCA algorithm and compare it to the performance of MCA algorithm under different noise levels.

## 4.1 Source Separation in the Frequency Domain

Observing the signals shown in Figure 4.1, it is natural to first look at the possible separation in the frequency domain via SBMCA and MCA techniques. The frequency domain treatment may also help us to address the timing uncertainty in nominally periodic nature of  $x_p$  as the magnitude spectra of the constituent prototype signals remain invariant under time shifts. Therefore,  $D_1$  now consists of two elements corresponding to magnitude spectra of the two prototype signals.

We can now proceed to perform the morphological separation of the magnitude spectra of each column of the data matrix  $X$ . Owing to the conjugate symmetry of the magnitude spectra, we perform the analysis on the first  $(\frac{m}{2} + 1)$  elements of each column of  $X$ . The morphological separation by SBMCA, where the unknown dictionary  $D_2$  is learnt, MCA with  $D_2$  as the DCT basis elements (MCA-DCT-Fourier) of size  $(\frac{m}{2} + 1) \times (\frac{m}{2} + 1)$  and MCA with  $D_2$  as the Identity basis elements (MCA-Identity-Fourier) of size  $(\frac{m}{2} + 1) \times (\frac{m}{2} + 1)$  is then carried out to complete the analysis. The initial

Table 4.1: Analysis of reconstruction SNR(in dB): Frequency Domain Separation

Noise $\mathcal{N}(0, \sigma^2)$	$\sigma = 0$		$\sigma = 0.001$		$\sigma = 0.01$		$\sigma = 0.1$	
Method\Signal	$x_p$	$x_u$	$x_p$	$x_u$	$x_p$	$x_u$	$x_p$	$x_u$
<b>SBMCA</b>	<b>10.56</b>	<b>16.08</b>	<b>10.56</b>	<b>16.07</b>	<b>10.52</b>	<b>16.10</b>	<b>8.25</b>	12.51
<b>MCA-DCT-Fourier</b>	5.74	12.42	5.74	12.42	5.71	12.41	4.72	10.29
<b>MCA-Identity-Fourier</b>	10.05	15.32	10.05	15.32	9.99	15.30	6.83	<b>12.71</b>

estimate of the coefficient matrix  $A_1$  for the SBMCA algorithm is formed by a OMP step based on the extracted nominally periodic signal  $x_p$  using the MCA-Identity-Fourier

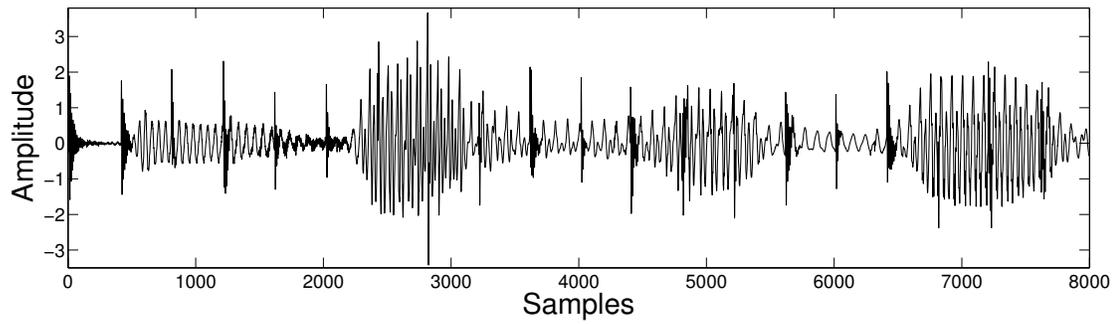
algorithm.

For the reconstruction of the signals  $x_p$  and  $x_u$ , we employ Wiener Filtering. We use the estimated magnitude spectra (after ensuring the positivity of extracted spectra and appropriate extension to form the full magnitude spectra) to estimate the corresponding power spectral densities of the two components to form our final estimates. The best reconstruction SNR(in dB) of extracted signals  $x_p$  and  $x_u$  at noise levels of  $\sigma = 0, 0.001, 0.01, 0.1$  is shown in Table 4.1. The *best achievable* SNRs (in dB) for MCA methods (MCA-DCT-Fourier and MCA-Identity-Fourier) are found by finding best parameters for  $x_p$  and  $x_u$  separately i.e. these SNR (in dB) levels may not be jointly achieved. In case of SBMCA we *clairvoyantly* tune the parameters (together) and the SNR (in dB) values are jointly achievable. It can be observed that SBMCA performs better than MCA-DCT-Fourier and MCA-Identity-Fourier, in terms of SNR (in dB), across various noise levels both on the nominally periodic signal,  $x_p$  and the unknown background noise,  $x_u$ , with an exception for the unknown background signal at  $\sigma = 0.1$ , where MCA-Identity-Fourier performs better than the other methods. Table 4.2 shows values of parameters (number of elements  $l$  in the dictionary  $D_2$ , the regularization parameter for the dictionary learning step  $\lambda_2$  and coefficient update step  $\lambda_3$ ) used for the SBMCA algorithm.

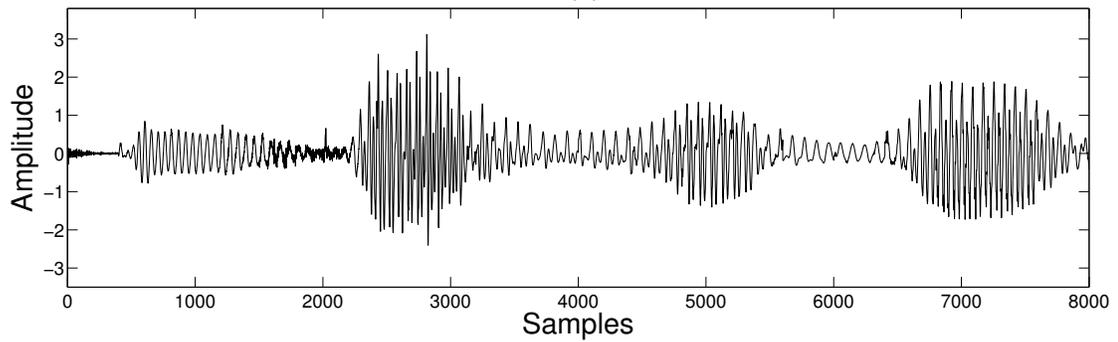
Table 4.2: SBMCA Algorithm Parameters for the Frequency Domain Source Separation

Noise $\mathcal{N}(0, \sigma^2)$	Parameters Values		
	Number of elements $l$ in $D_2$	$\lambda_2$	$\lambda_3$
$\sigma = 0$	110	0.03	0.01
$\sigma = 0.001$	110	0.03	0.01
$\sigma = 0.01$	110	0.03	0.01
$\sigma = 0.1$	60	0.05	0.1

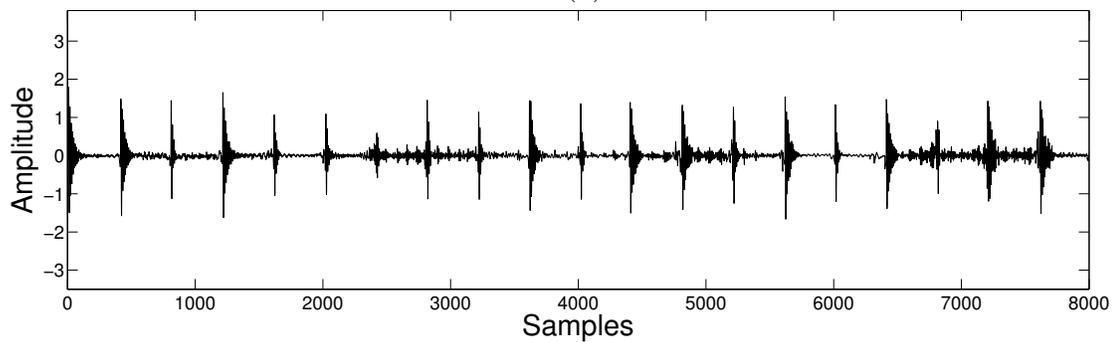
Figure 4.2, 4.3 and 4.4 show the results of source separation in the frequency domain obtained via SBMCA, MCA-DCT-Fourier and MCA-Identity-Fourier respectively at  $\sigma = 0$ . In each of these figures, (a) depicts the original mixture, (b) the extracted speech signal (unknown) and (c) the extracted nominally periodic signal. Specifically, Figure 4.5 and 4.6 provide insights about source separation performance over a short



(a)



(b)



(c)

Figure 4.2: Frequency Domain Source Separation using SBMCA. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal.

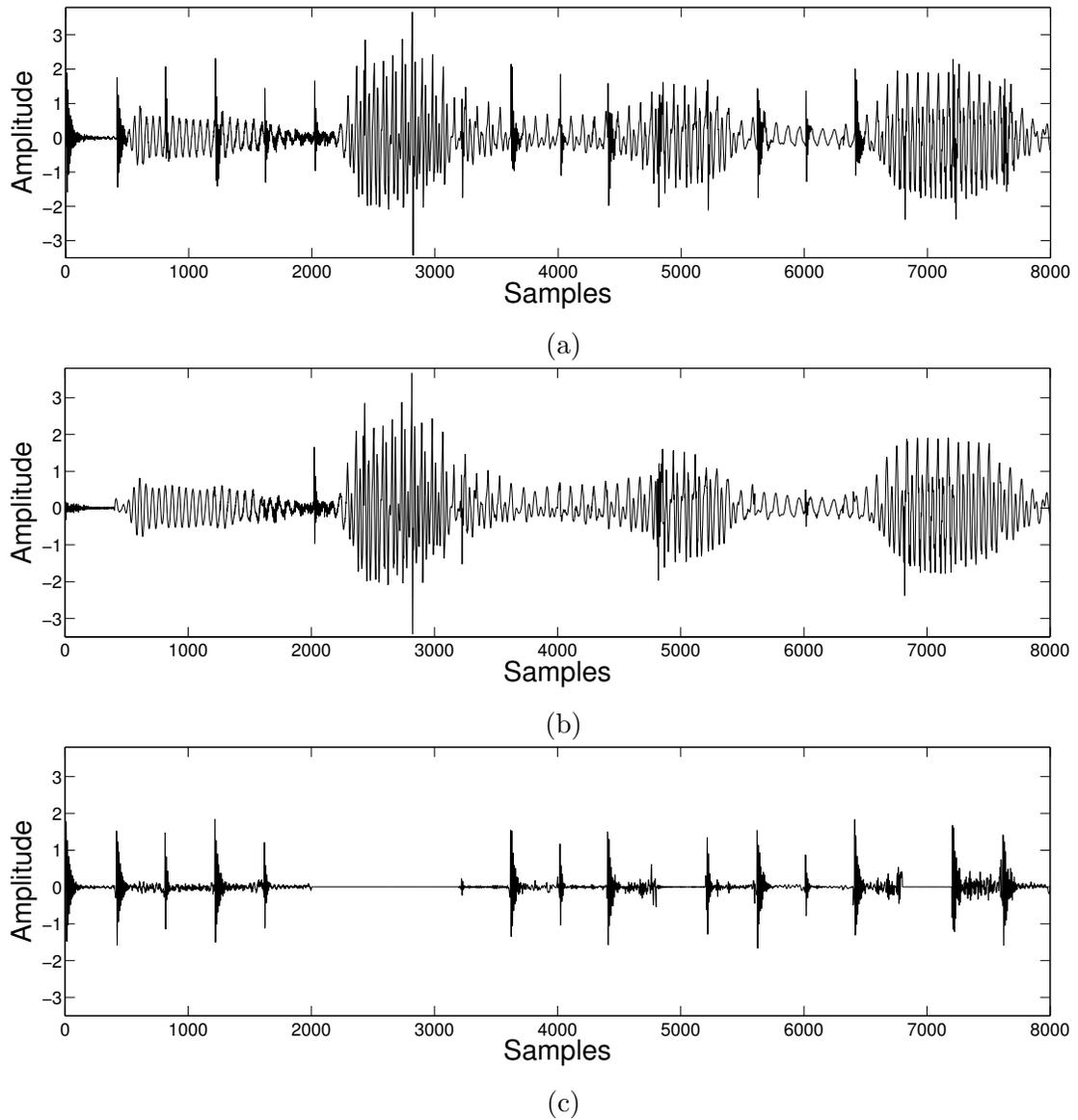


Figure 4.3: Frequency Domain Source Separation using MCA-DCT-Fourier. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal.

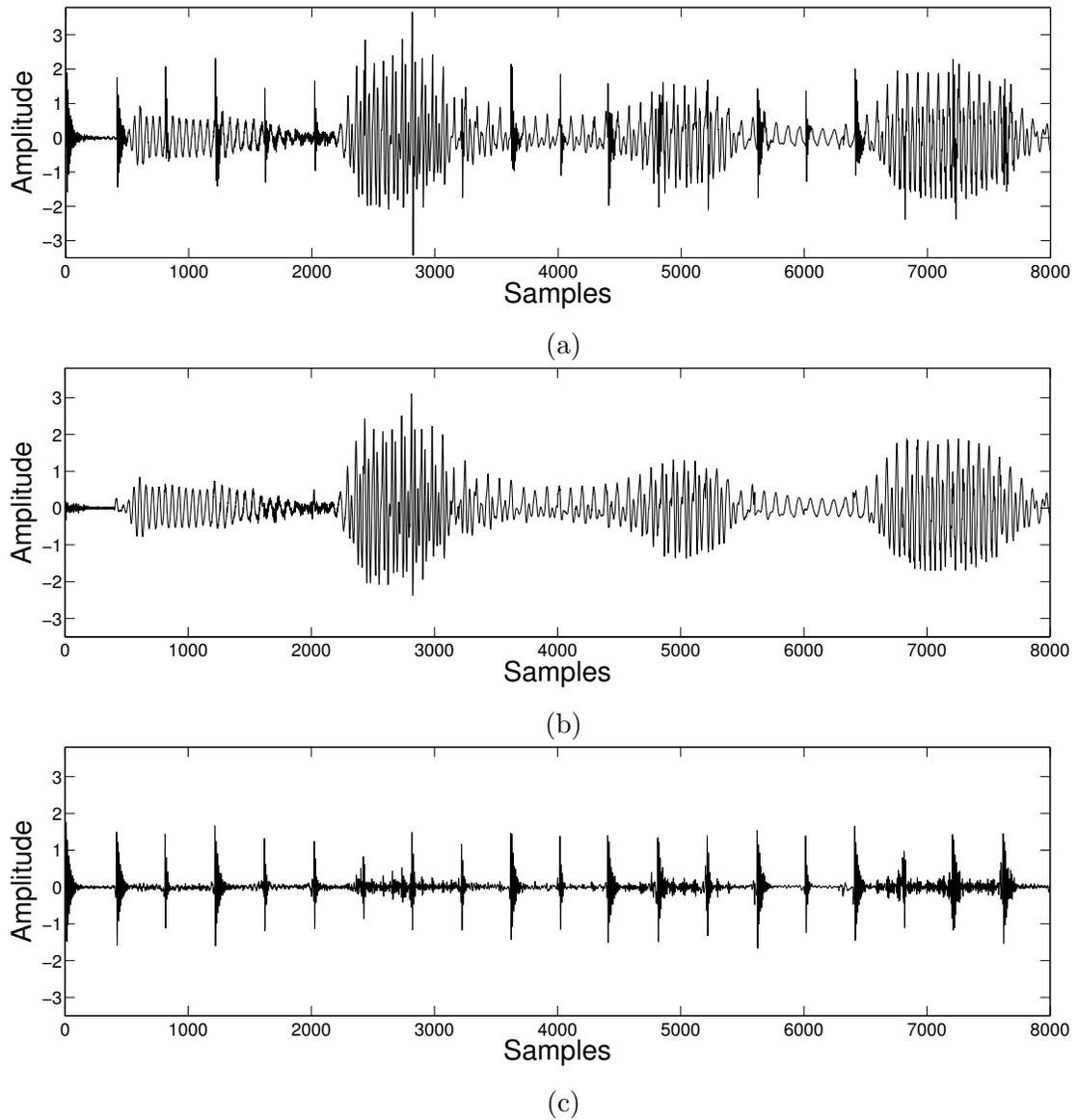
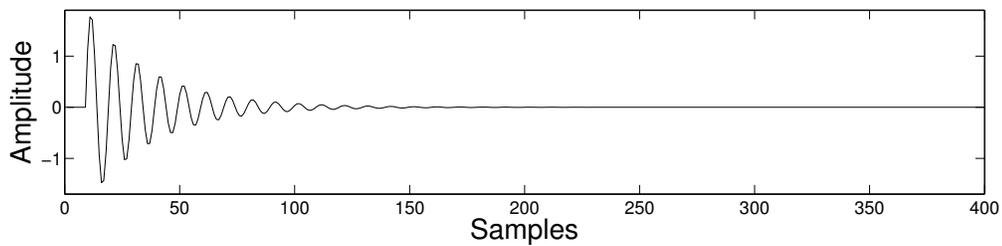
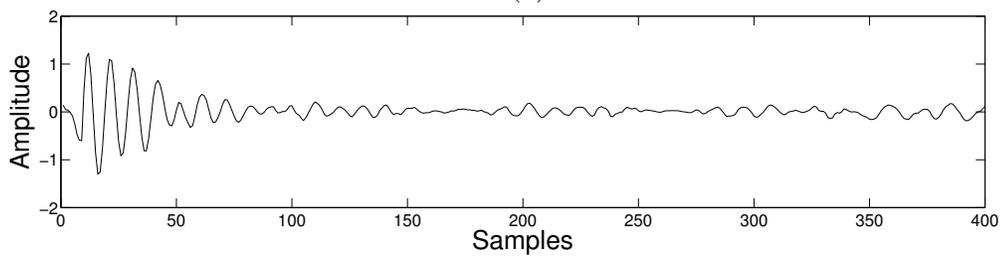


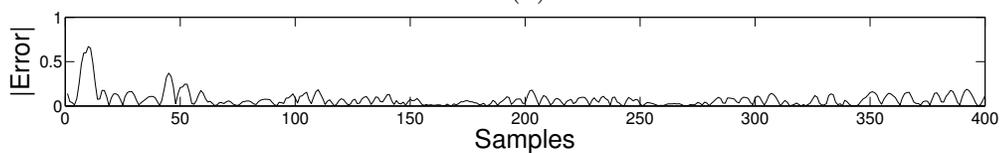
Figure 4.4: Frequency Domain Source Separation using MCA-Identity-Fourier. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal.



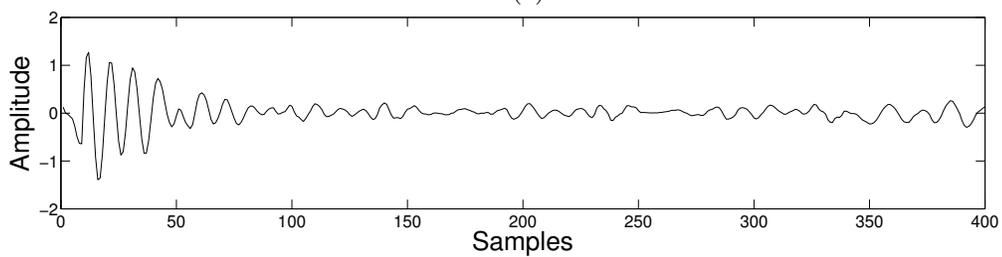
(a)



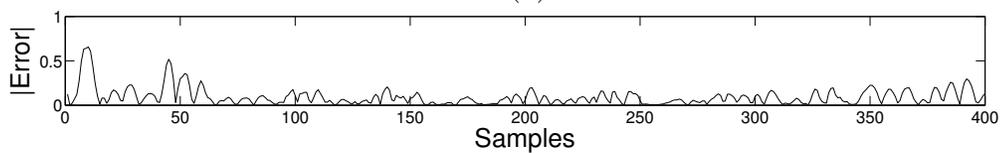
(b)



(c)



(d)



(e)

Figure 4.5: A segment of extracted nominally periodic signal  $x_p$  (Frequency Domain Separation). (a) The original nominally periodic signal, (b) and (d) The extracted nominally periodic signal via SBMCA and MCA-Identity-Fourier respectively, (c) and (e) The absolute error between the extracted and original nominally periodic signal via SBMCA and MCA-Identity-Fourier respectively.

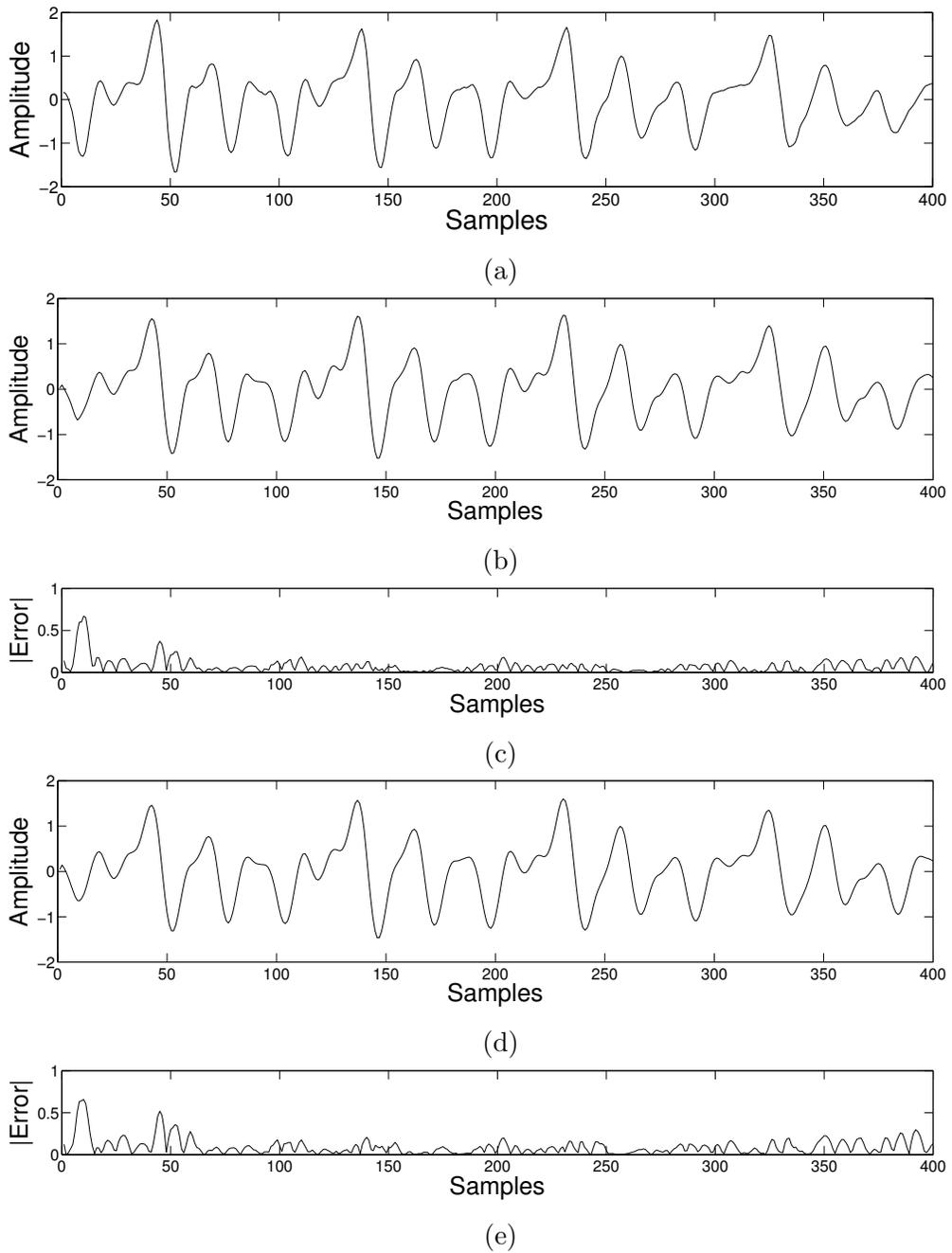


Figure 4.6: A segment of extracted speech signal  $x_u$  (Frequency Domain Separation). (a) The original speech signal, (b) and (d) The extracted speech signal via SBMCA and MCA-Identity-Fourier respectively, (c) and (e) The absolute error between the extracted and original speech signal via SBMCA and MCA-Identity-Fourier respectively.

period of time obtained via SBMCA and MCA-Identity-Fourier algorithms at  $\sigma = 0$ . Figure 4.5 shows extracted nominally periodic signal  $x_p$ , (a) depicts a segment of the original nominally periodic signal  $x_p$ , (b) and (d) show the separation achieved via SBMCA and MCA-Identity-Fourier (two best performing methods in terms of SNR(in dB)), respectively, (c) and (e) show the corresponding absolute errors via SBMCA and MCA-Identity-Fourier. Similarly, Figure 4.6 shows extracted speech signal  $x_u$ , (a) shows a segment of the original speech signal  $x_u$ , (b) and (d) show the separation achieved via SBMCA and MCA-Identity-Fourier, respectively, (c) and (e) show the corresponding absolute errors via SBMCA and MCA-Identity-Fourier. Comparing panels (c) and (e) in Figure 4.5 and Figure 4.6, we note that MCA-Identity-Fourier algorithm yields a little higher levels of errors as compared to the SBMCA, this is also evident from the corresponding SNR performance as shown in Table 4.1.

Another interesting metric to measure the performance of these algorithms is the the histogram of normalized errors-per-block shown in Figures 4.7-4.10, measured using the vector  $l_2$  norm, for each method at noise levels of  $\sigma = 0, 0.001, 0.01, 0.1$ , respectively. In each of these figures panels (a), (c) and (e) represent the histogram of normalized error-per-block for  $x_p$  and (b), (d) and (f) represent the histogram of normalized error-per-block for  $x_u$  via SBMCA, MCA-DCT-Fourier and MCA-Identity-Fourier respectively, with  $\sigma = 0$  in Figure 4.7,  $\sigma = 0.001$  in Figure 4.8,  $\sigma = 0.01$  in Figure 4.9 and  $\sigma = 0.1$  in Figure 4.10. We observe that both SBMCA and MCA-Identity-Fourier perform somewhat comparably. However, dip in the performance of MCA-DCT-Fourier can be attributed to the presence of a large number of blocks towards the high-error parts of the histogram.

By the analysis presented above we conclude that, although the frequency domain treatment allows us to tackle the timing uncertainty associated with the nominally periodic signal  $x_p$ , the phase ambiguity created due to lack of phase information in the separation stage, proves detrimental to the overall source separation performance.

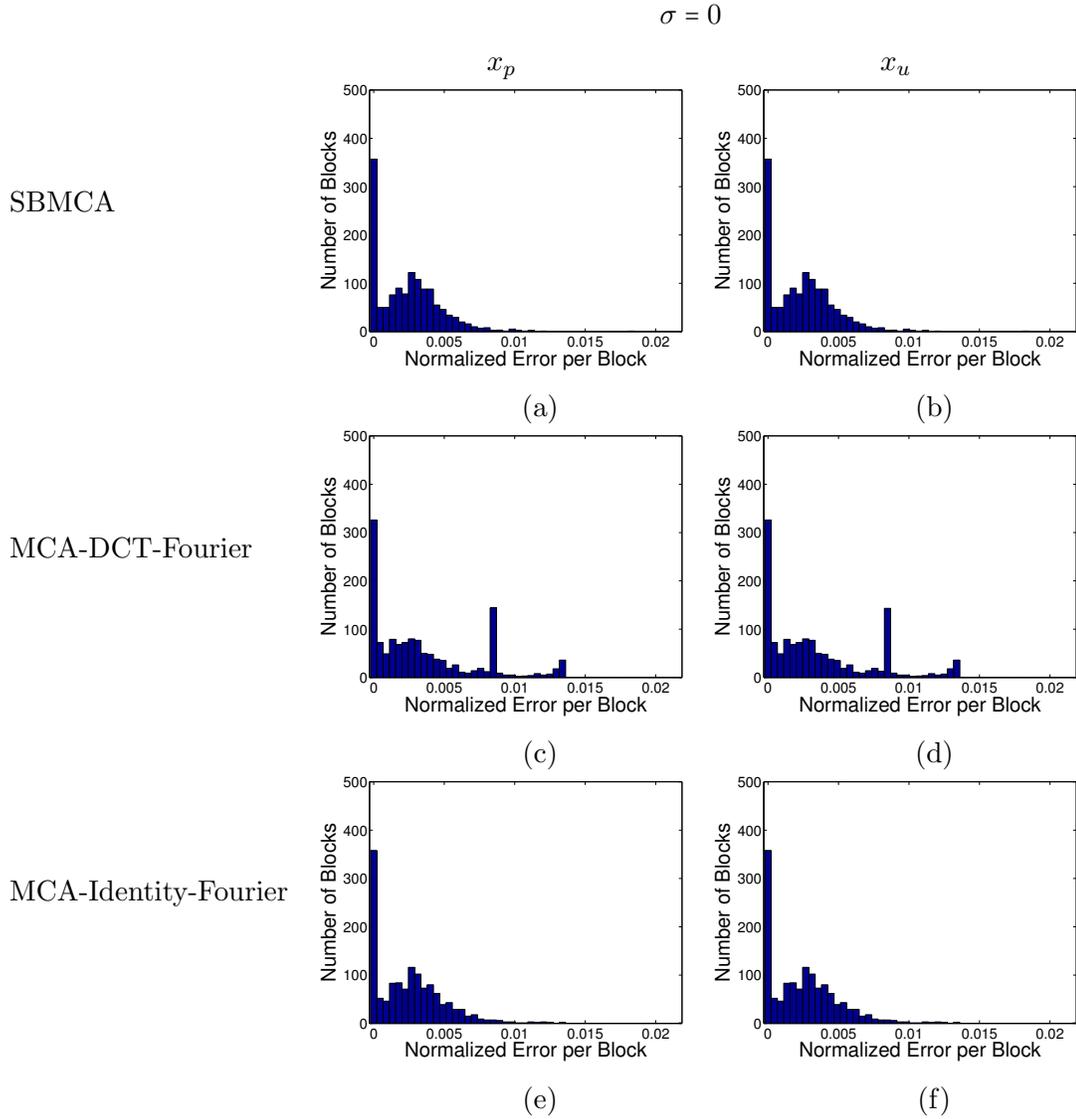


Figure 4.7: Histogram of normalized error-per-block of extracted nominally periodic signal  $x_p$  and extracted speech signal  $x_u$  with noise  $\sigma = 0$  for frequency domain source separation using Semi-blind MCA, MCA-DCT-Fourier and MCA-Identity-Fourier for the audio forensic application.

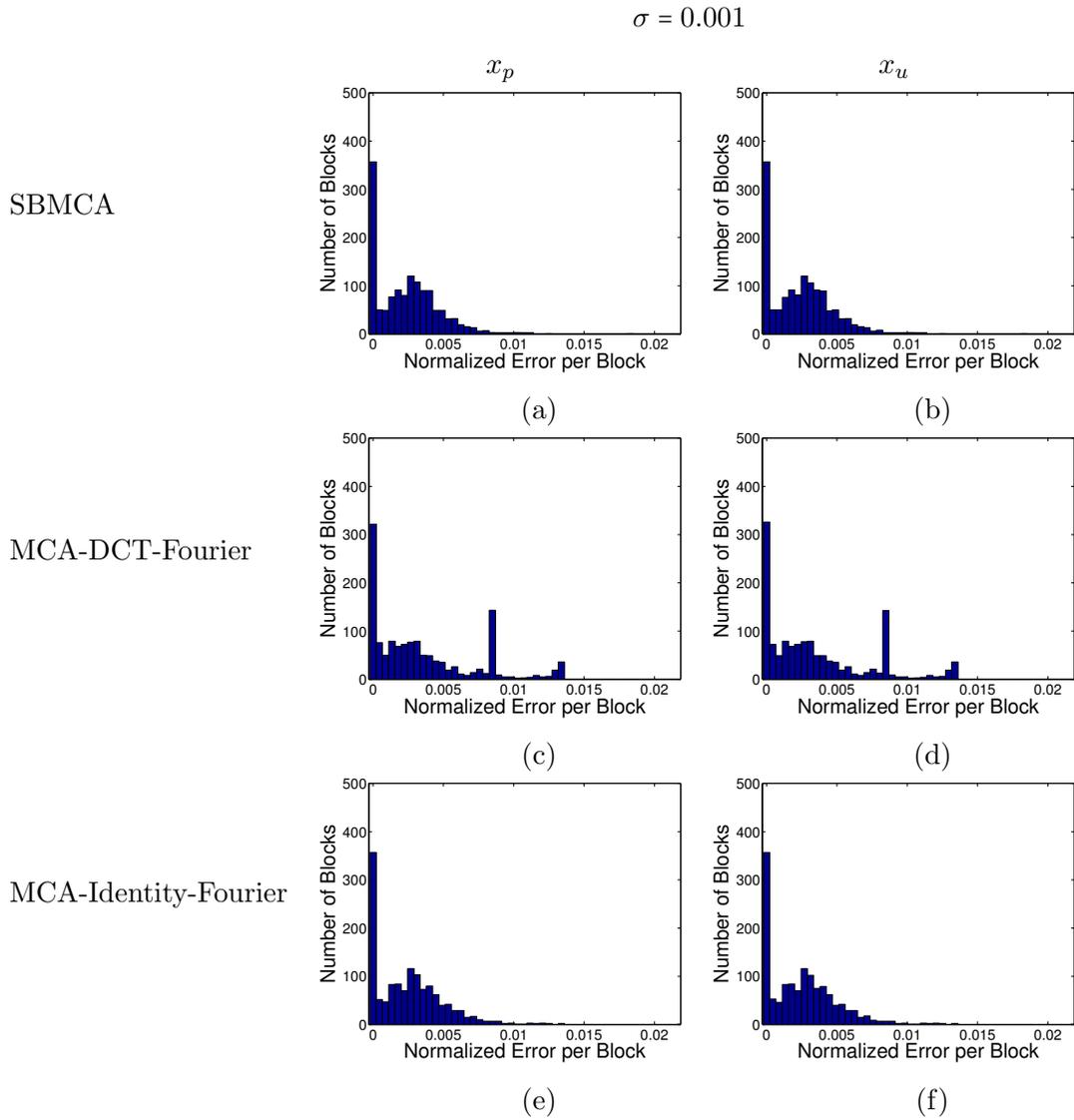


Figure 4.8: Histogram of normalized error-per-block of extracted nominally periodic signal  $x_p$  and extracted speech signal  $x_u$  with noise  $\sigma = 0.001$  for frequency domain source separation using Semi-blind MCA, MCA-DCT-Fourier and MCA-Identity-Fourier for the audio forensic application.

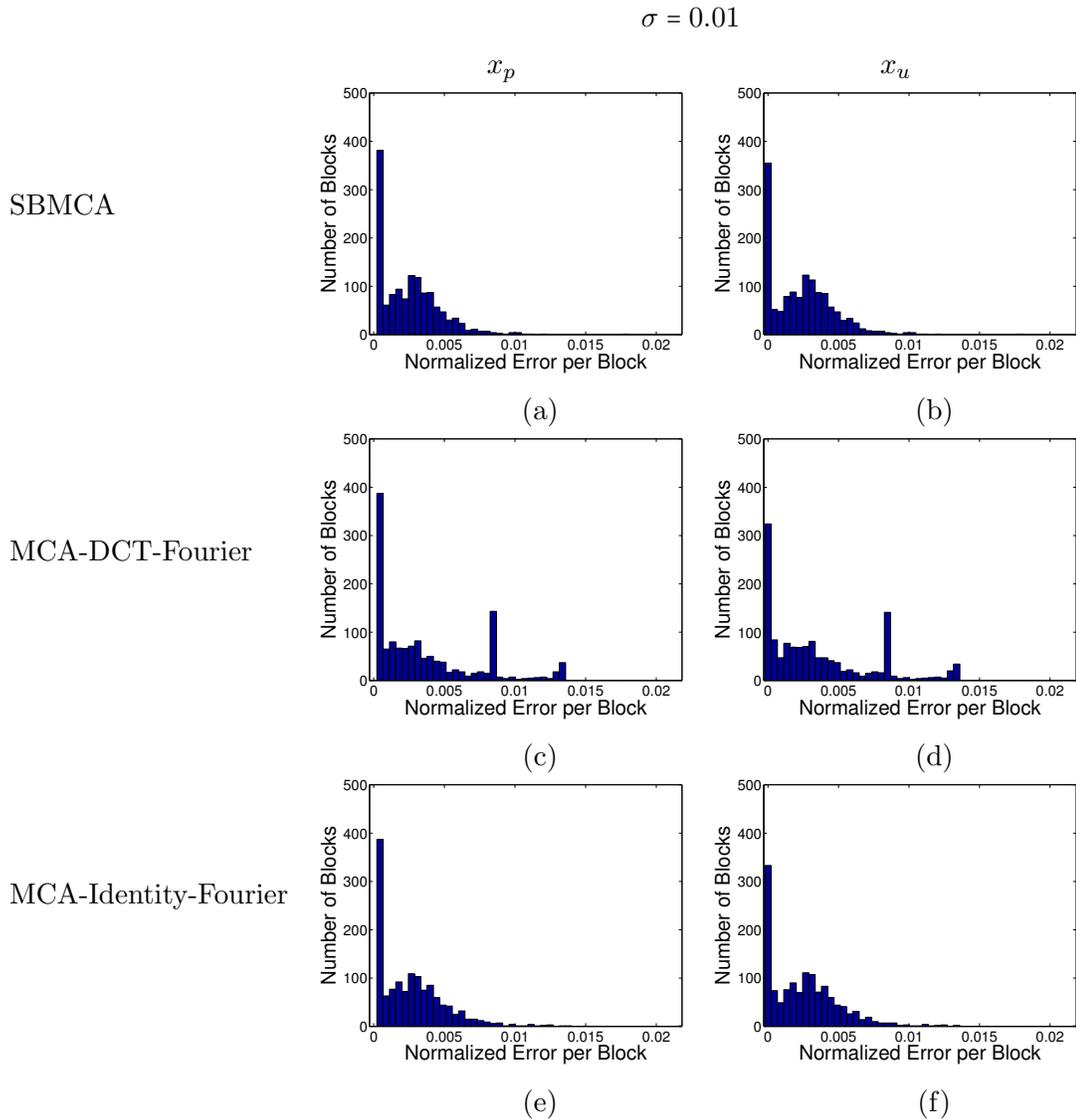


Figure 4.9: Histogram of normalized error-per-block of extracted nominally periodic signal  $x_p$  and extracted speech signal  $x_u$  with noise  $\sigma = 0.01$  for frequency domain source separation using Semi-blind MCA, MCA-DCT-Fourier and MCA-Identity-Fourier for the audio forensic application.

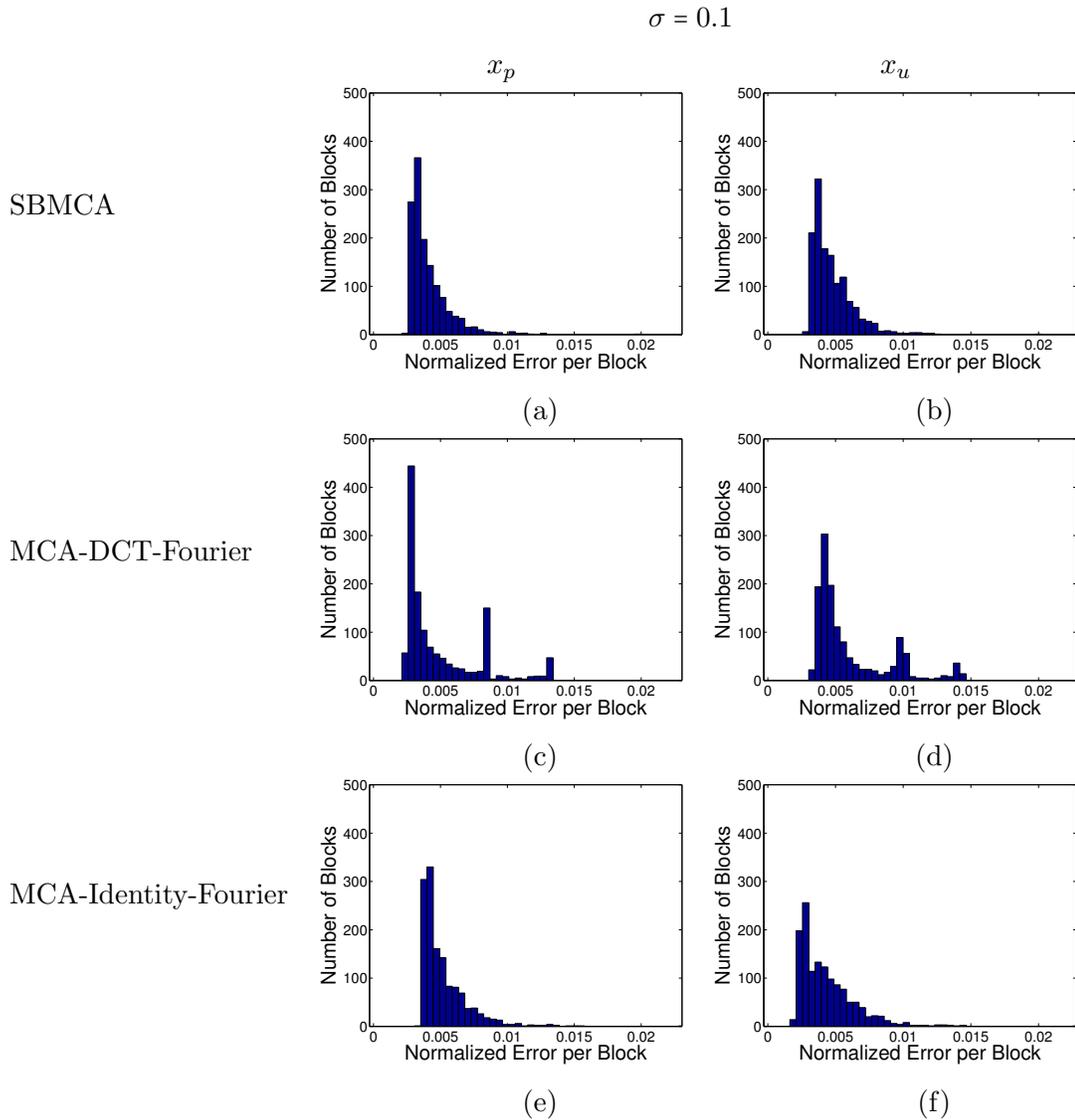


Figure 4.10: Histogram of normalized error-per-block of extracted nominally periodic signal  $x_p$  and extracted speech signal  $x_u$  with noise  $\sigma = 0.1$  for frequency domain source separation using Semi-blind MCA, MCA-DCT-Fourier and MCA-Identity-Fourier for the audio forensic application.

## 4.2 Source Separation in the Time Domain

Next we consider the source separation in time domain which was the topic of discussion in the publication associated with this thesis [21]. In case of time domain separation, we form the dictionary  $D_1$  by incorporating certain circular shifts of the nominal prototype pulses from which  $x_p$  was generated. We then employ the semi-blind MCA approach discussed in Chapter 3 to separate the background audio from the nominally known periodic portion.

We compare the performance of our approach with two versions of MCA, one using the DCT basis and the other using the identity basis to form the dictionary  $D_2$ . We use the estimated  $x_p$ , obtained via MCA-DCT procedure to initialize our approach by applying one step of orthogonal matching pursuit (OMP) [20] on the estimate of  $x_p$  obtained via MCA-DCT to form the initial (one component per column) estimate  $A_1$  for the SBMCA algorithm.

Table 4.3: Analysis of reconstruction SNR(in dB): Time Domain Separation

Noise $\mathcal{N}(0, \sigma^2)$	$\sigma = 0$		$\sigma = 0.001$		$\sigma = 0.01$		$\sigma = 0.1$	
Method\Signal	$x_p$	$x_u$	$x_p$	$x_u$	$x_p$	$x_u$	$x_p$	$x_u$
<b>SBMCA</b>	<b>23.72</b>	<b>29.32</b>	<b>23.73</b>	<b>29.32</b>	<b>23.08</b>	<b>27.40</b>	<b>19.72</b>	<b>16.84</b>
<b>MCA-DCT</b>	20.44	26.02	20.46	26.02	20.19	24.96	18.09	16.72
<b>MCA-Identity</b>	10.90	16.06	10.90	16.44	10.90	16.33	10.78	11.44

Table 4.3 lists the *best achievable* reconstruction SNRs (in dB) of each method. We note that our interest here is in comparing the best performances achieved by MCA and our proposed method, so we *clairvoyantly* tune the value(s) of the regularization parameter to give the lowest error for each task. Table 4.4 shows values of parameters (number of elements  $l$  in the dictionary  $D_2$ , the regularization parameter for the dictionary learning step  $\lambda_2$  and coefficient update step  $\lambda_3$ ) used for the SBMCA algorithm. (In general, a different regularization parameter may have been utilized to obtain the reconstruction SNRs of each signal component, even for the same method and same noise level – in other words, the SNRs listed may not be jointly achievable from a *single* implementation of any of the stated procedures).

Table 4.4: SBMCA Algorithm Parameters for Time Domain Source Separation

Noise $\mathcal{N}(0, \sigma^2)$	Parameters Values		
	Number of elements $l$ in $D_2$	$\lambda_2$	$\lambda_3$
$\sigma = 0$	1000	0.1	0.001
$\sigma = 0.001$	1000	0.1	0.001
$\sigma = 0.01$	900	0.2	0.01
$\sigma = 0.1$	110	0.1	0.1

Figure 4.11, 4.12 and 4.13 show the source separation results obtained via SBMCA, MCA-DCT and MCA-Identity respectively at  $\sigma = 0$ . In each of these figures, (a) depicts the original mixture, (b) the extracted speech signal (unknown) and (c) the extracted nominally periodic signal. Specifically, Figure 4.14 and 4.15 provide insights about the time domain source separation performance over a short period of time obtained via SBMCA and MCA-DCT (two best performing methods in terms of SNR (in dB)) at  $\sigma = 0$ . Figure 4.14 shows extracted nominally periodic signal  $x_p$ , (a) depicts a segment of the original nominally periodic signal  $x_p$ , (b) and (d) show the separation achieved via SBMCA and MCA-DCT, respectively, (c) and (e) show the corresponding absolute errors obtained via SBMCA and MCA-DCT. Similarly, Figure 4.15 shows extracted speech signal  $x_u$ , (a) shows a segment of the original speech signal  $x_u$ , (b) and (d) show the separation achieved via SBMCA and MCA-DCT, respectively, (c) and (e) show the corresponding absolute errors via SBMCA and MCA-DCT. Comparing panels (c) and (e) in Figure 4.14 and Figure 4.15, we note that the SBMCA algorithm yields lower errors as compared to the MCA-DCT algorithm; this feature is also highlighted in the histogram of error-per-block results discussed next.

A second, perhaps more interesting, performance comparison is shown Figures 4.16-4.19, which depicts the histogram of normalized errors-per-block, measured using the vector  $l_2$  norm, for each method. In each of these figures panels (a), (c) and (e) represent the histogram of normalized error-per-block for  $x_p$  and (b), (d) and (f) represent the histogram of normalized error-per-block for  $x_u$  via SBMCA, MCA-DCT and MCA-Identity respectively, with  $\sigma = 0$  in Figure 4.16,  $\sigma = 0.001$  in Figure 4.17,  $\sigma = 0.01$  in Figure 4.18 and  $\sigma = 0.1$  in Figure 4.19. We observe from the distribution of  $l_2$  errors

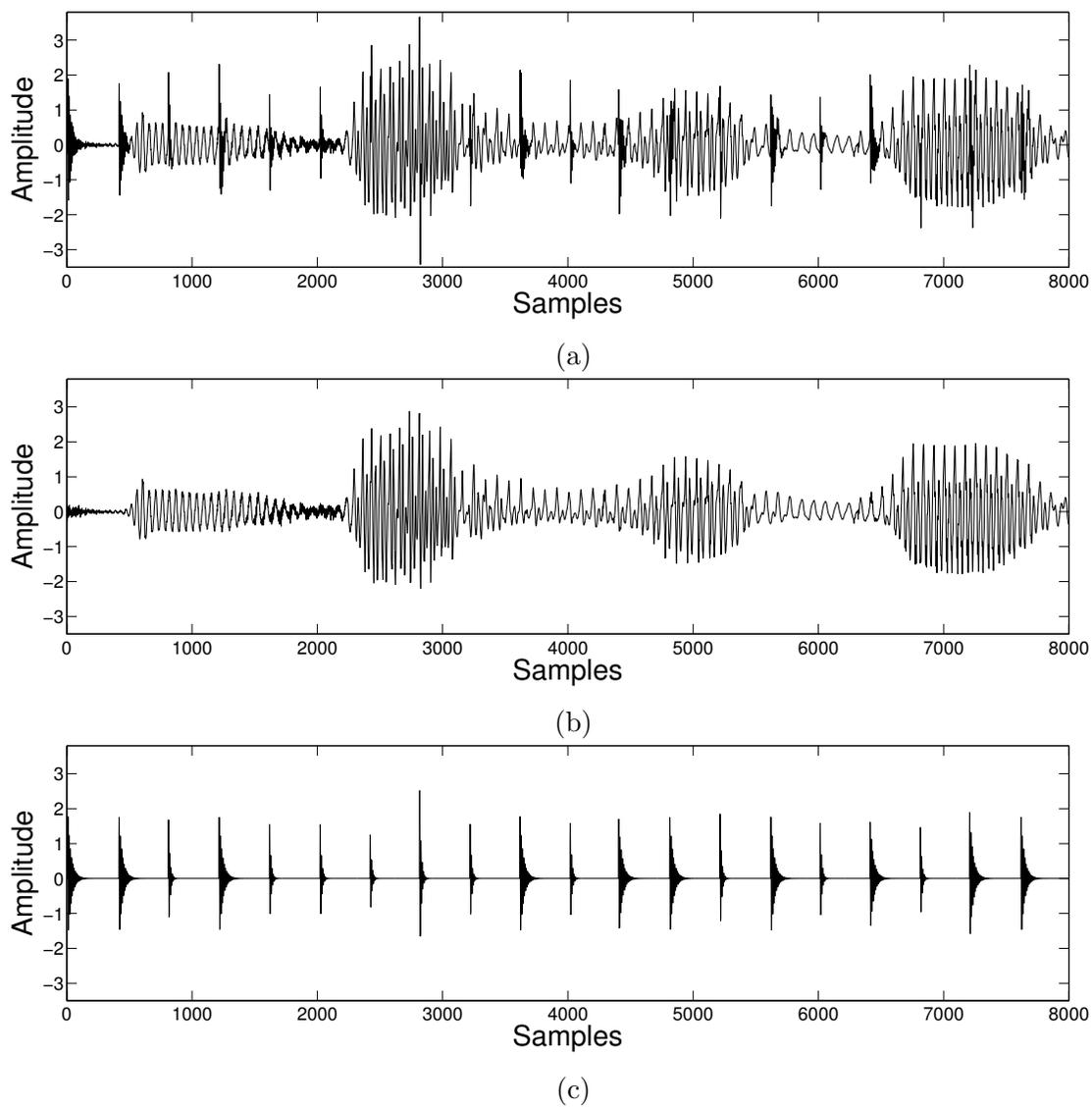


Figure 4.11: Time Domain Source Separation using SBMCA. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal.

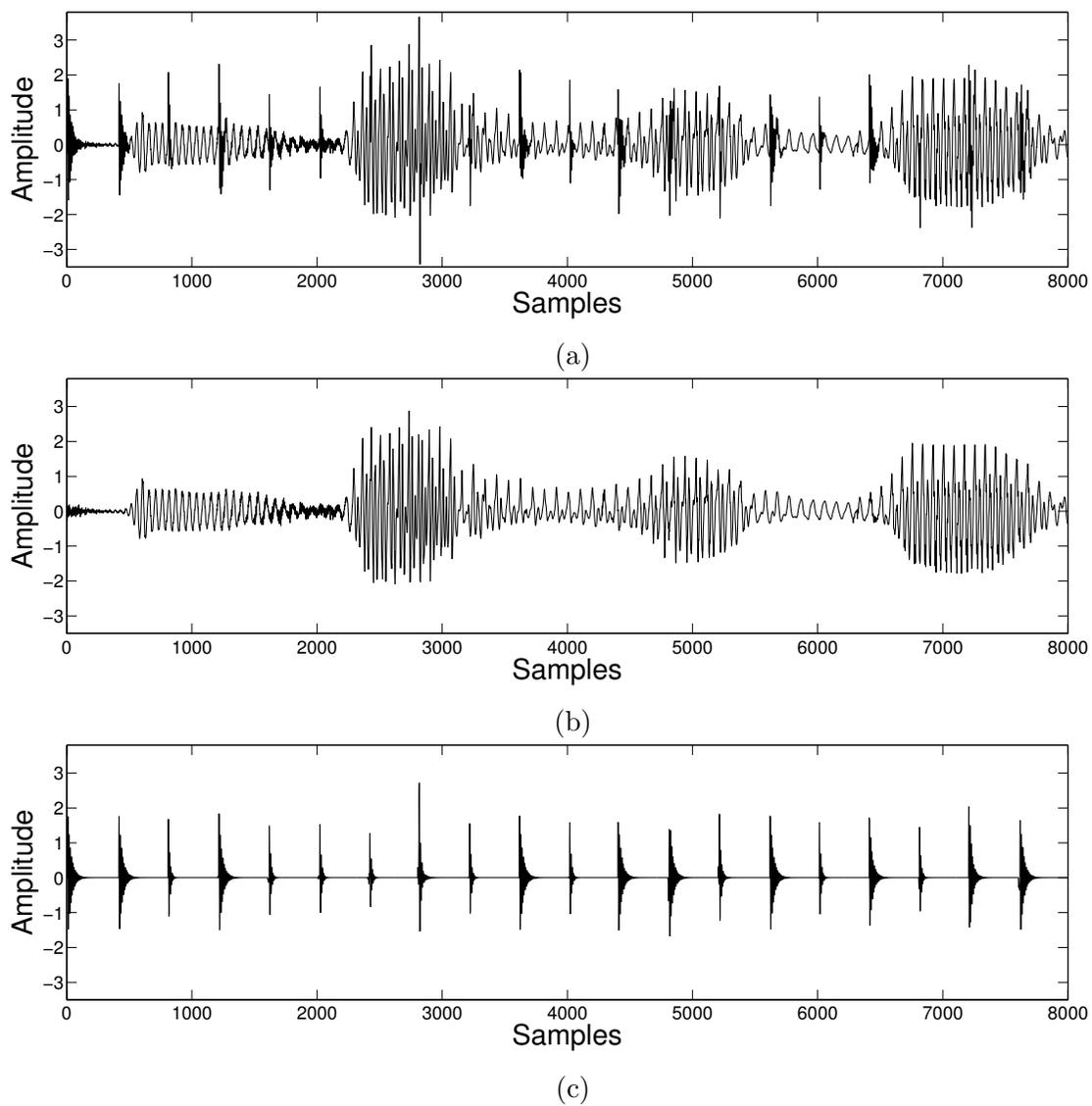


Figure 4.12: Time Domain Source Separation using MCA-DCT. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal.

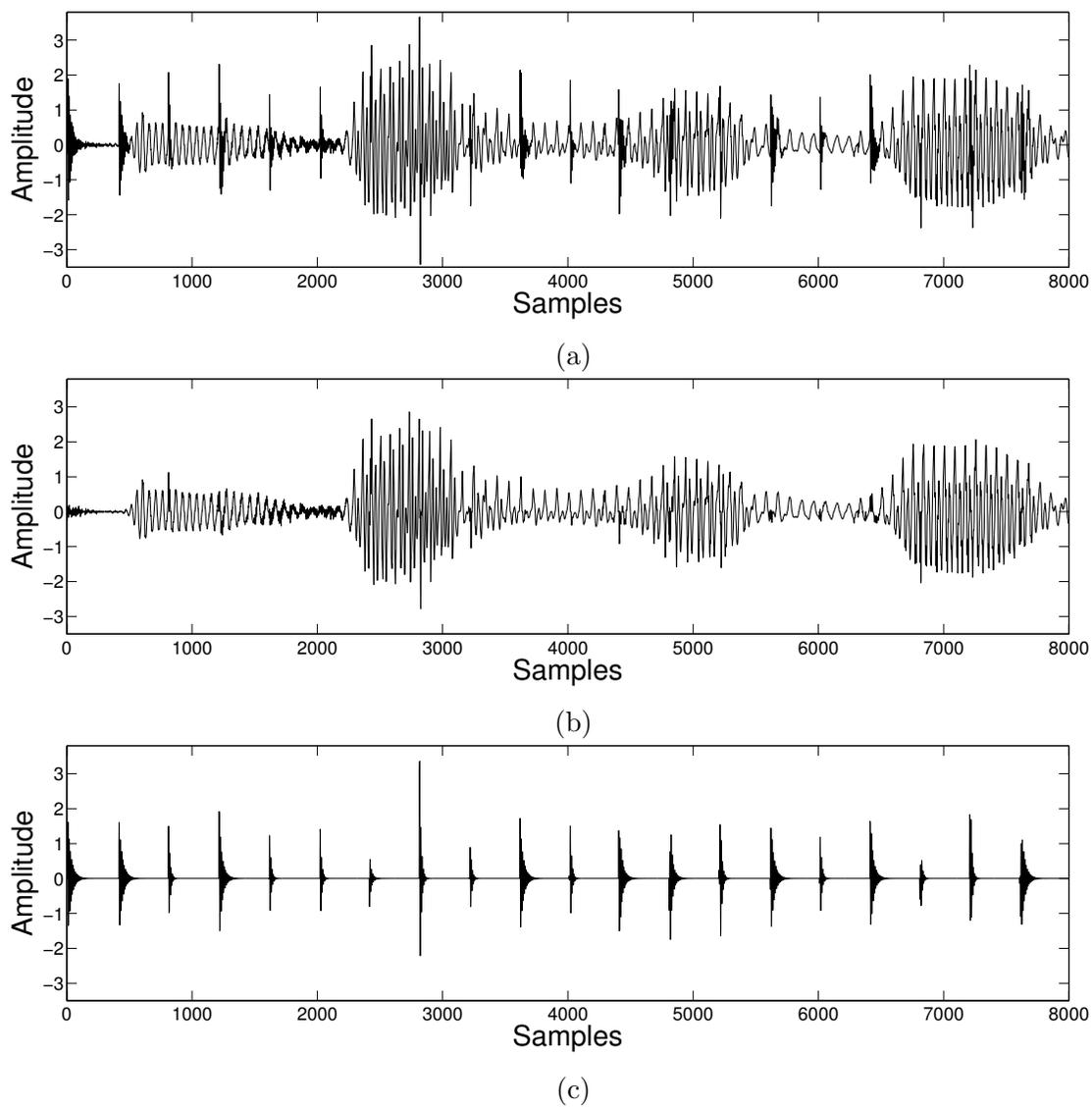


Figure 4.13: Time Domain Source Separation using MCA-Identity. (a) A segment of original mixture, (b) extracted speech signal and (c) extracted nominally periodic signal.

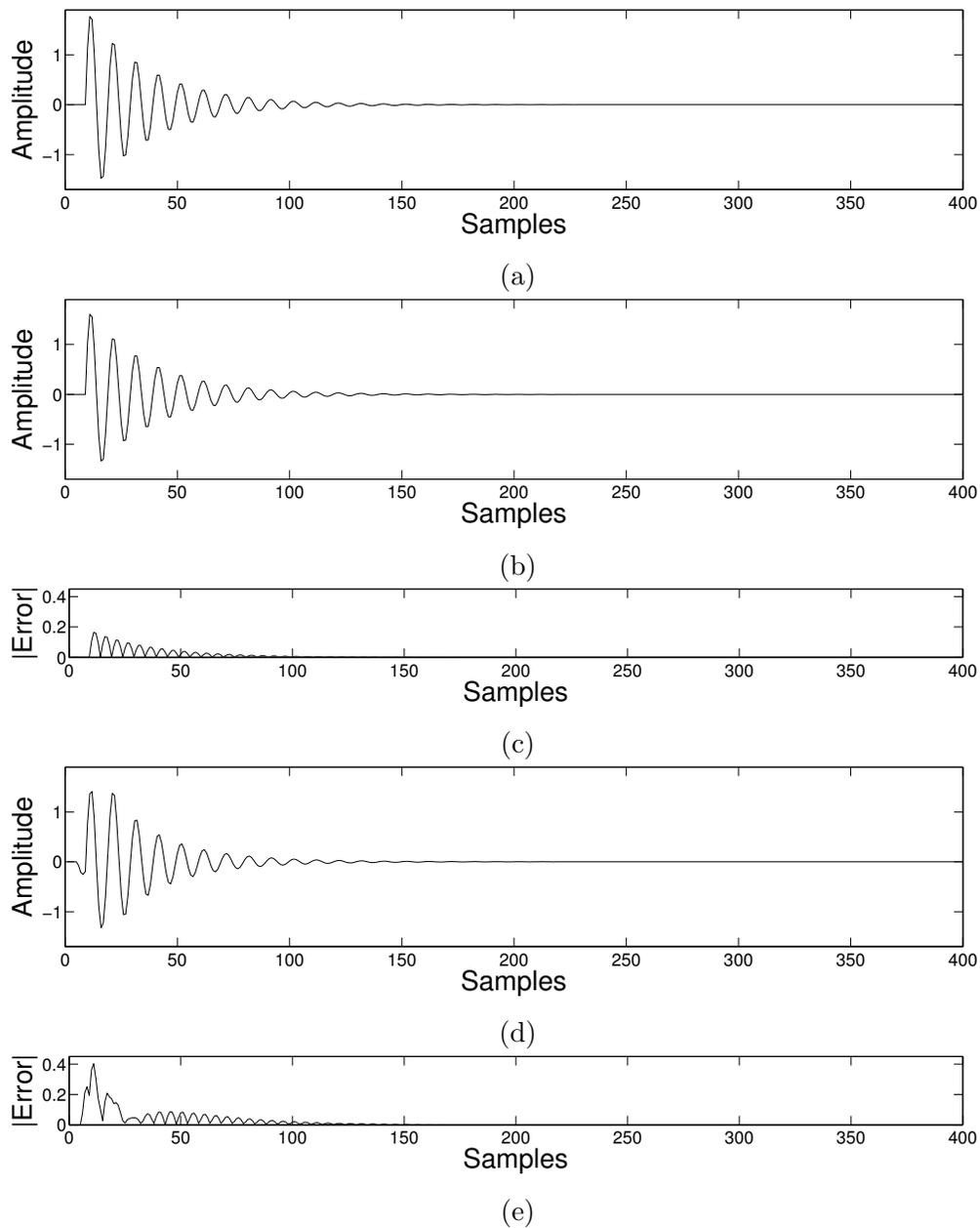


Figure 4.14: A segment of extracted nominally periodic signal  $x_p$  (Time Domain Separation). (a) The original nominally periodic signal, (b) and (d) The extracted nominally periodic signal via SBMCA and MCA-DCT respectively, (c) and (e) The absolute error between the extracted and original nominally periodic signal via SBMCA and MCA-DCT respectively.

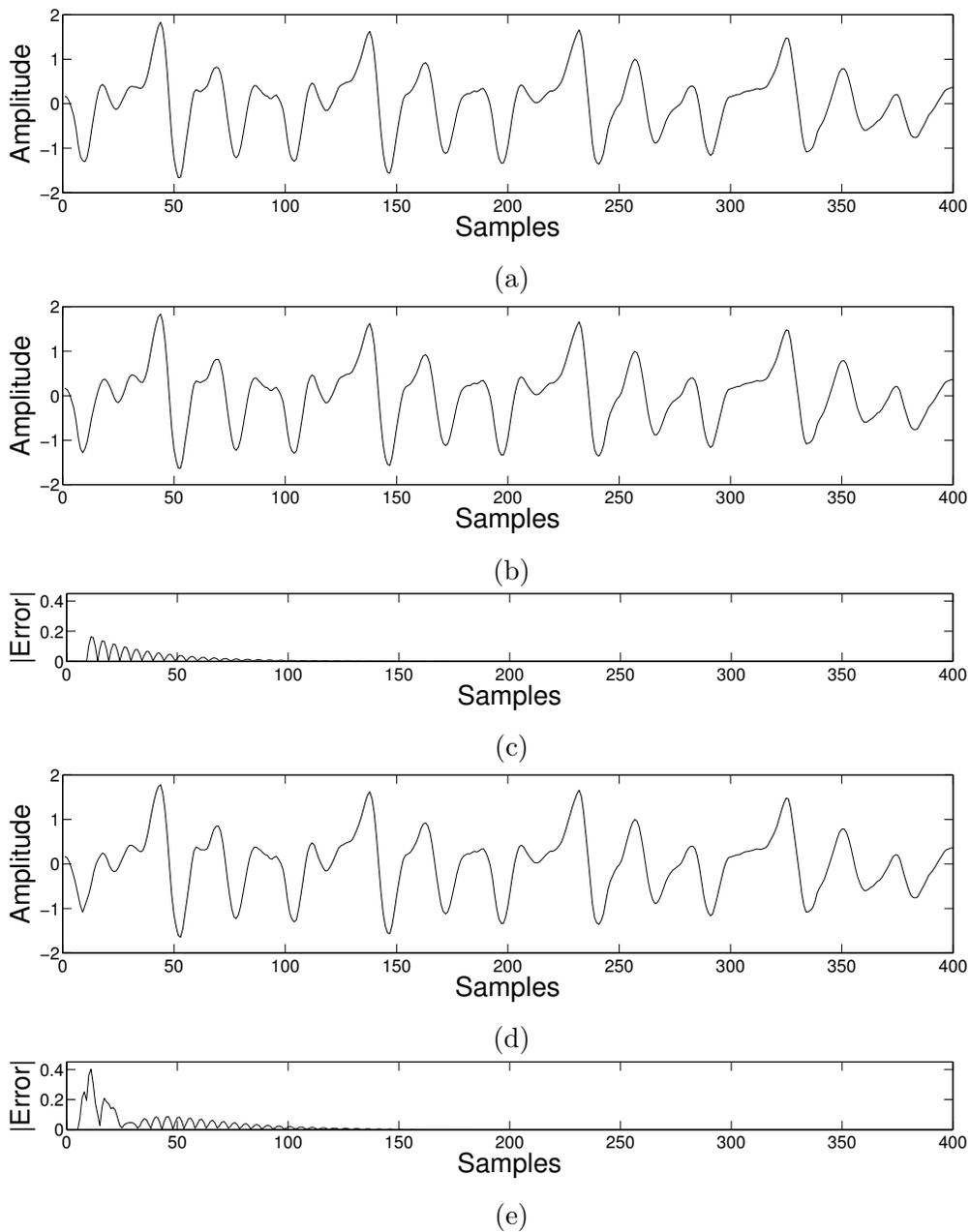


Figure 4.15: A segment of extracted speech signal  $x_u$  (Time Domain Separation). (a) The original speech signal, (b) and (d) The extracted speech signal via SBMCA and MCA-DCT respectively, (c) and (e) The absolute error between the extracted and original speech signal via SBMCA and MCA-DCT respectively.

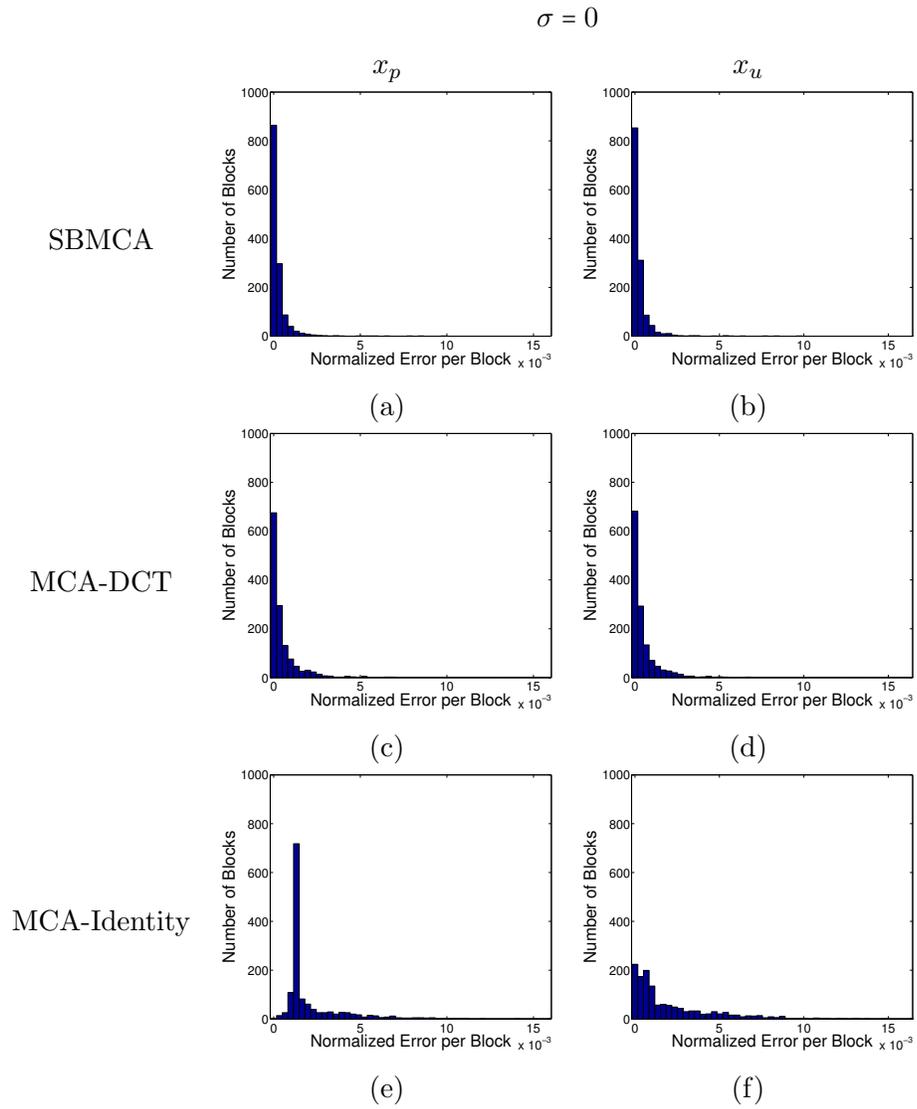


Figure 4.16: Histogram of normalized error-per-block measured using the vector  $l_2$  norm of extracted nominally periodic signal  $x_p$  and extracted speech signal  $x_u$  with zero mean gaussian noise  $\sigma = 0$  for time domain separation using Semi-blind MCA and MCA with DCT and Identity basis for the audio forensic application.

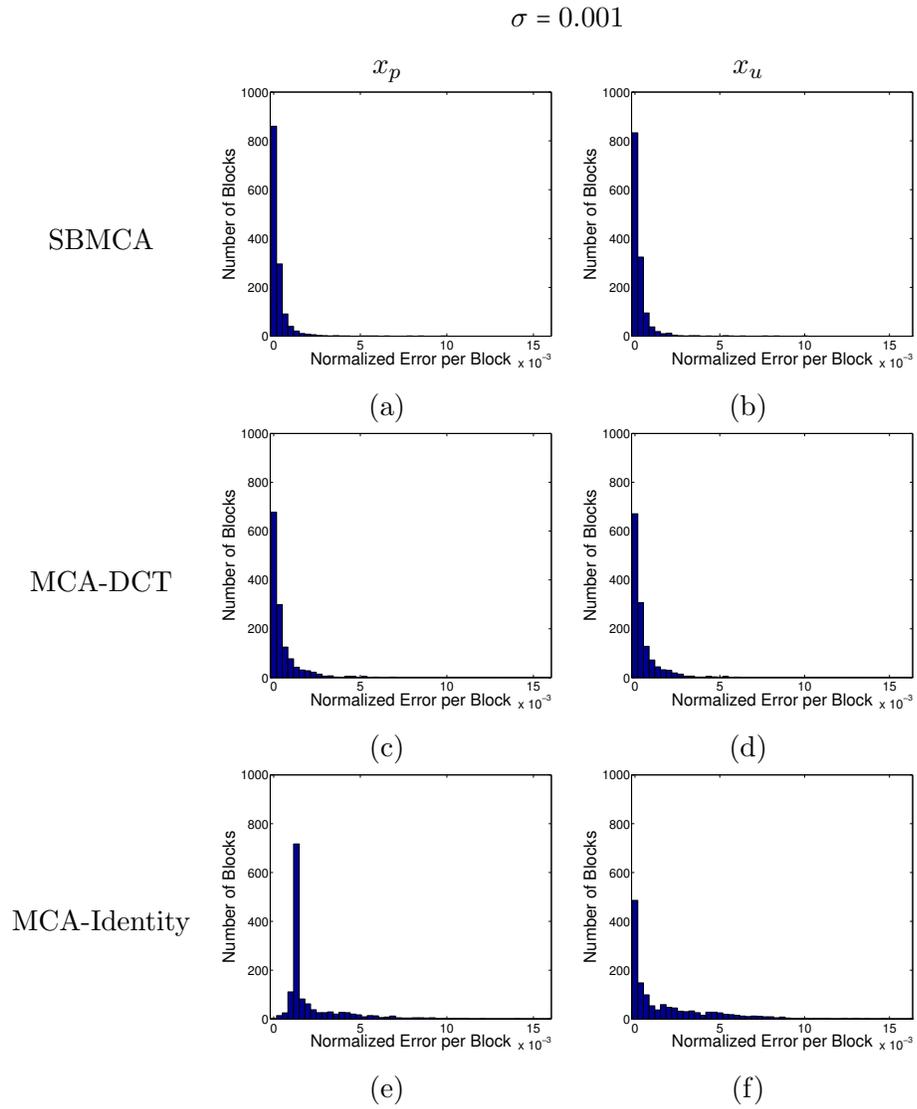


Figure 4.17: Histogram of normalized error-per-block measured using the vector  $l_2$  norm of extracted nominally periodic signal  $x_p$  and extracted speech signal  $x_u$  with zero mean gaussian noise  $\sigma = 0.001$  for time domain separation using Semi-blind MCA and MCA with DCT and Identity basis for the audio forensic application.

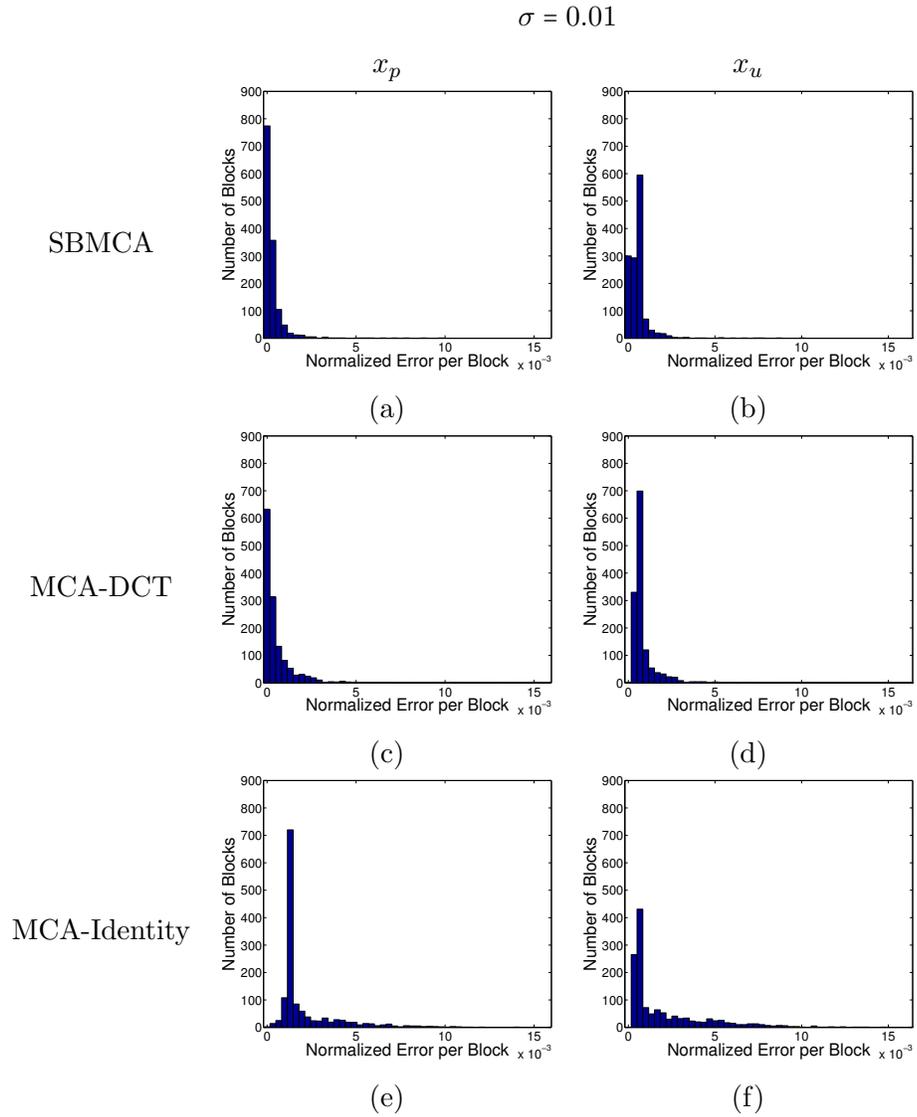


Figure 4.18: Histogram of normalized error-per-block measured using the vector  $l_2$  norm of extracted nominally periodic signal  $x_p$  and extracted speech signal  $x_u$  with zero mean gaussian noise  $\sigma = 0.01$  for time domain separation using Semi-blind MCA and MCA with DCT and Identity basis for the audio forensic application.

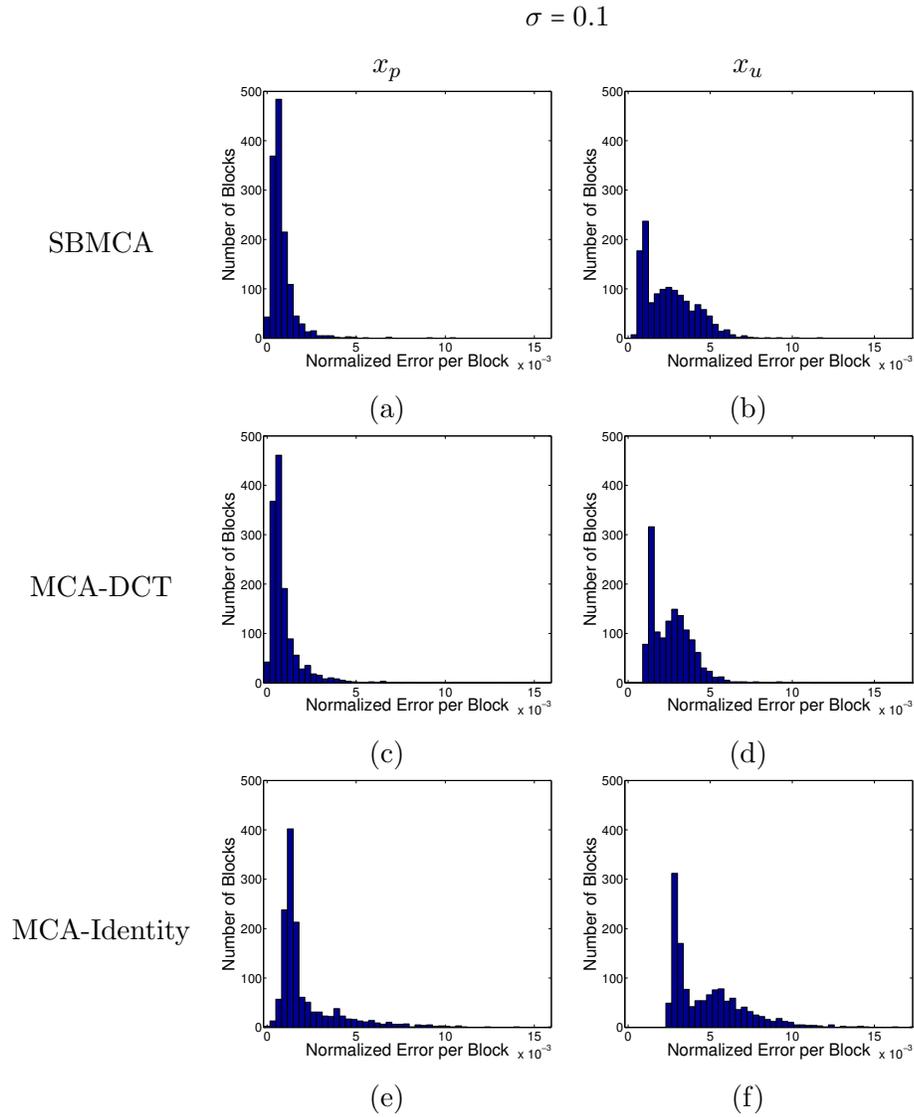


Figure 4.19: Histogram of normalized error-per-block measured using the vector  $l_2$  norm of extracted nominally periodic signal  $x_p$  and extracted speech signal  $x_u$  with zero mean gaussian noise  $\sigma = 0.1$  for time domain separation using Semi-blind MCA and MCA with DCT and Identity basis for the audio forensic application.

across blocks, that the SBMCA procedure (Figure 4.16- 4.19 (a) and (b)) results in larger number of blocks with lower errors as compared to the MCA-DCT (Figure 4.16- 4.19 (c) and (d)) and MCA-Identity (Figure 4.16- 4.19 (e) and (f)). This feature is of primary importance in the audio forensics application where classifying each period of the nominally periodic signal  $x_p$ , as one of the two prototype signals, is of interest.

From the analysis of MCA-DCT and MCA-Identity results, we observe that the separation results are depend on the choice of the basis employed. This sheds light onto yet another advantage of employing the SBMCA procedure i.e. *a priori* knowledge about the sparse basis of the unknown signal ( $x_u$ ) is not required for source separation.

It is important to note that the performance of the source separation via SBMCA rests heavily on the local structure of the unknown signal  $x_u$ . This scenario is analogous to any online Dictionary Learning based denoising problem. The effect of presence of structure is evident at higher levels of noise ( $\sigma = 0.1$ ), where sparsity of the signal plays an important role in rejecting noise.

## Chapter 5

# Conclusion and Discussion

We proposed a semi-blind source separation technique based on local sparse approximations. Our approach exploits partial prior knowledge of one of the sources, in the form of a dictionary which sparsely represents local segments of one of the sources. A key feature of our approach is the online learning of a dictionary (from the mixed source data itself) for representing the unknown background source. We posed the overall problem as an optimization problem and proposed a solution approach based on alternating minimization, and we verified its effectiveness via simulation in a stylized audio separation application.

We note that the timing uncertainty inherent in our application suggests that our approach may be combined with other existing *alignment* techniques. The task of aligning batches of images is well-studied in the computer vision literature. For example, techniques that aim to exploit the low-rankness of (properly aligned) images have been proposed in [22], [23], [24], and more recently [25] proposed a robust alignment procedure that can be viewed as an extension of robust PCA to the case where columns of the nominally low-rank component are each subjected to some (unknown) misalignment operation (translation, rotation) that must be learned from the data. Here, by contrast, we tolerate the unknown (one-dimensional) misalignment in a straight-forward manner, by either replicating columns of our known dictionary at a number of shifts (in the time domain analysis) or considering separation of frequency amplitude spectra (where the resulting spectra are invariant to the unknown timing offset – a phase shift). It would be interesting to see whether the robust alignment techniques of [25] may be extended to

our setting where the background signal (or corruption) is sparse in a perhaps unknown dictionary that is learned from the data itself. We defer these extensions, as well as the investigation of our approach to other applications (e.g., in image or video processing) to future efforts.

# References

- [1] C. Jutten and J. Herault. Blind Separation of Sources, Part I: An Adaptive Algorithm Based on Neuromimetic Architecture. *Signal Processing*, 24(1):1–10, July 1991.
- [2] H. Hotelling. Analysis of a Complex of Statistical Variables into Principal Components. *Journal of Educational Psychology*, 24:417–441, 1933.
- [3] C. Eckart and G. Young. The Approximation of One Matrix by Another of Lower Rank. *Psychometrika*, 1:211–218, 1936.
- [4] I.T. Jolliffe. *Principal Component Analysis*. Springer Verlag, 1986.
- [5] A. Cichocki, R. Zdunek, and S. Amari. New Algorithms for Non-Negative Matrix Factorization in Applications to Blind Source Separation. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5:V, May 2006.
- [6] G. J. Jang and T. W. Lee. A Maximum Likelihood Approach to Single-Channel Source Separation. *Journal of Machine Learning Research*, 4:1365–1392, December 2003.
- [7] T. P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. McKeown, V. Iragui, and T. J. Sejnowski. Removing Electroencephalographic Artifacts by Blind Source Separation. *Psychophysiology*, 37:163–178, 2000.
- [8] M. N. Schmidt and R. K. Olsson. Single-channel Speech Separation using Sparse Non-negative Matrix Factorization. *International Conference on Spoken Language Processing (INTERSPEECH)*, 2006.

- [9] M.E. Davies and C.J. James. Source Separation using Single Channel ICA. *Signal Processing*, 87(8):1819 – 1832, 2007.
- [10] B. A. Olshausen and D. J. Field. Sparse Coding with an Overcomplete Basis Set: A Strategy Employed by V1? *Vision Research*, 37(23):3311–3325, 1997.
- [11] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: Design of Dictionaries for Sparse Representation. *In Proceedings of SPARS05*, pages 9–12, 2005.
- [12] J. Mairal, F. Bach, J. Ponce, and G. Sapiro. Online Learning for Matrix Factorization and Sparse Coding. *Journal of Machine Learning Research*, 11:19–60, 2010.
- [13] E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust Principal Component Analysis? *Journal of the ACM*, 58(3):11, 2011.
- [14] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky. Rank-Sparsity Incoherence for Matrix Decomposition. *Society for Industrial and Applied Mathematics (SIAM) Journal on Optimization*, 21(2):572–596, 2011.
- [15] M. Mardani, G. Mateos, and G. B. Giannakis. Recovery of Low-Rank Plus Compressed Sparse Matrices with Application to Unveiling Traffic Anomalies. *IEEE Transactions on Information Theory*, 2012. Online: [arXiv:1204.6537v1](https://arxiv.org/abs/1204.6537v1) [cs.IT].
- [16] J. L. Starck, M. Elad, and D. L. Donoho. Image Decomposition via the Combination of Sparse Representations and a Variational Approach. *IEEE Transactions on Image Processing*, 14(10):1570–1582, 2005.
- [17] D. L. Donoho and G. Kutyniok. Microlocal Analysis of the Geometric Separation Problem. *CoRR*, abs/1004.3006, 2010.
- [18] J. Bobin, J. L. Starck, J. Fadili, Y. Moudden, and D. L. Donoho. Morphological Component Analysis: An Adaptive Thresholding Strategy. *IEEE Transactions on Image Processing*, 16(11):2675–2681, 2007.
- [19] S. Chen and D. Donoho. Basis Pursuit. *In Conference Record of the Twenty-Eighth Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 41–44. IEEE, 1994.

- [20] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad. Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition. In *Proceedings of the 27th Annual Asilomar Conference on Signals, Systems and Computers*, pages 40–44, 1993.
- [21] S. Rambhatla and J. Haupt. Semi-Blind Source Separation via Sparse Representations and Online Dictionary Learning. *submitted to IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2013. Online: [arXiv:1212.0451](https://arxiv.org/abs/1212.0451) [cs.SD].
- [22] B. J. Frey and N. Jovic. Transformed Component Analysis: Joint Estimation of Spatial Transformations and Image Components. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1190–1196, 1999.
- [23] B. J. Frey and N. Jovic. Transformation-Invariant Clustering using the EM Algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:1–17, 2003.
- [24] A. Vedaldi, G. Guidi, and S. Soatto. Joint Data Alignment Up to (Lossy) Transformations. In *CVPR*. IEEE Computer Society, 2008.
- [25] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma. RASL: Robust Alignment by Sparse and Low-Rank Decomposition for Linearly Correlated Images. In *CVPR*, pages 763–770. IEEE, 2010.

# Appendix A

## Glossary and Acronyms

Care has been taken in this thesis to minimize the use of jargon and acronyms, but this cannot always be achieved. This appendix defines jargon terms in a glossary, and contains a table of acronyms and their meaning.

### A.1 Glossary

- **Source Separation** – a classical signal processing problem entailing separation of constituent sources from one( referred to as Single Channel) or several( referred to as Multiple Channel) mixtures of individual sources.
- **Cocktail party problem** – the source separation problem, illustrated by an example of a cocktail party, where multiple speakers are simultaneously speaking and the aim is to separately comprehend each speaker.
- **Sparsity** – a property of a signal to achieve compact representation in some basis.

## A.2 Acronyms

Table A.1: Acronyms

Acronym	Meaning
BSS	Blind Source Separation
DCT	Discrete Cosine Transform
ICA	Independent Component Analysis
LASSO	Least Absolute Shrinkage and Selection Operator
MCA	Morphological Component Analysis
NNMF	Non-Negative Matrix Factorization
OMP	Orthogonal Matching Pursuit
PCA	Principal Component Analysis
RLC	Resistance-Inductance-Capacitance
RPCA	Robust Principal Component Analysis
SBMCA	Semi-Blind Morphological Component Analysis
SVD	Singular Value Decomposition