EMBEDDED OPTIMIZATION ALGORITHMS FOR MULTI-MICROPHONE DEREVERBERATION

Toon van Waterschoot, Bruno Defraene, Moritz Diehl, and Marc Moonen

Dept. E.E./ESAT, SCD-SISTA, KU Leuven / iMinds-KU Leuven Future Health Department
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
{toon.vanwaterschoot,bruno.defraene,moritz.diehl,marc.moonen}@esat.kuleuven.be

ABSTRACT

In this paper we propose a new approach to multi-microphone dereverberation, based on the recent paradigm of embedded optimization. The rationale of embedded optimization in performing online signal processing tasks, is to replace traditional adaptive filtering algorithms based on closed-form estimators by fast numerical algorithms solving constrained and potentially non-convex optimization problems. In the context of dereverberation, we adopt the embedded optimization paradigm to arrive at a joint estimation of the source signal of interest and the unknown room acoustics. It is shown how the inherently non-convex joint estimation problem can be smoothed by including regularization terms based on a statistical late reverberation model and a sparsity prior for the source signal spectrum. A performance evaluation for an example multi-microphone dereverberation scenario shows promising results, thus motivating future research in this direction.

Index Terms— dereverberation, embedded optimization, nonlinear least squares, regularization, sparsity

1. INTRODUCTION

Dereverberation refers to the process of removing reverberation from microphone signals recorded in a reverberant room. Since reverberation often has a fundamental impact on the time-frequency signal characteristics, dereverberation has been found to be a crucial component in diverse speech and audio applications, such as hearing assistance, automatic speech recognition, voice communications, and acoustic surveillance. Despite its wide applicability, dereverberation is generally still considered one of the most challenging problems in the area of acoustic signal enhancement [1]. One of the major difficulties is that dereverberation is an *inverse problem*, i.e., one aims at inverting the room impulse response (RIR), which is typically non-minimum-phase and possibly time-varying. Furthermore, dereverberation is usually also a *blind problem*, in which both the sound source signal and the room acoustics are unknown.

The state of the art in speech dereverberation can be classified into three categories [1, Ch. 1]: (1) beamforming, (2) speech enhancement, and (3) blind system identification and inversion. Most of the existing methods rely on the use of *multiple microphones*. This is implictly the case for the beamforming approaches which are based on microphone array processing, see, e.g., [2]. Speech enhancement approaches to dereverberation have also been shown to benefit from the use of multiple microphones, e.g., for accurately estimating the late reverberant signal spectrum [3] or for enhancing a linear prediction residual by spatiotemporal averaging [4]. Finally, blind system identification is typically based on the cross-relation between different microphone signals [5], while the inversion of a non-minimum-phase system has been shown to be feasible only in the multi-channel case [6].

In this paper, a different approach to dereverberation is proposed, which does not exactly fit into one of the three categories mentioned earlier. The proposed approach is somehow related to the blind system identification and inversion approach, however, it differs in that it does not require an explicit system inversion. Indeed, the major weakness of the blind system identification and inversion approach is that the design of a (multi-channel) inverse filter often appears to be an ill-posed problem, which may be due to (near-)common zeros [7] or system identification errors [8] in the RIRs. Recent solutions to alleviate this weakness are based on modifications in the inverse filter design, such as subband inversion [9], regularization [10], and forced spectral diversity [11].

Instead, we propose to avoid an explicit system inversion by adopting a recent paradigm coined as *embedded optimization*. This paradigm is based on the observation that the field

This research work was carried out at the ESAT Laboratory of KU Leuven, and was supported by Research Council KUL: PFV/10/002 Optimization in Engineering Center OPTEC, GOA/10/09 MaNet, GOA/10/11 Global real- time optimal control of autonomous robots and mechatronic systems; Flemish Government: IOF / KP / SCORES4CHEM, iMinds 2013; FWO: PhD/postdoc grants, projects: G.0320.08 (convex MPC), G.0377.09 (Mechatronics MPC); IWT: PhD Grants, projects: SBO LeCoPro; Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO, Dynamical systems, control and optimization, 2012-2017), IUAP P7/23 (BESTCOM, Belgian network on stochastic modeling analysis design and optimization of communication systems, 2012-2017); EU: FP7-DREAMS (MC ITN-316969), FP7-EMBOCON (ICT-248940), FP7-SADCO (MC ITN-264735), ERC ST HIGHWIND (259 166), Eurostars SMART, ACCM. The scientific responsibility is assumed by its authors.

of numerical optimization has reached a degree of maturity and computational efficiency such that it can be applied to online signal processing problems that are traditionally solved using recursive implementations of "classical" estimators admitting a closed-form solution [12]. In particular, it allows to directly estimate a signal vector of interest, rather than taking a detour by designing a filter to recover a signal of interest from noisy or corrupted observations.

The outline of this paper is as follows. In Section 2 we propose a relevant signal model and formulate the multi-microphone dereverberation problem. In Section 3 we propose a number of embedded optimization algorithms for multi-microphone dereverberation. These algorithms are evaluated in Section 4 for a simple example scenario. Finally, Section 5 concludes the paper.

2. PROBLEM STATEMENT

Consider a point source emitting a sound signal $s_0(t),\ t = 1,\dots,N$, which propagates inside a room and is picked up by M microphones at different positions. The resulting microphone signals ($m = 1,\dots,M$) are defined as

$$y_m(t) = \mathbf{h}_{m,0}^T(t)\mathbf{s}_0(t) + e_{m,0}(t),\ t = 1,\dots,N \quad (1)$$

where the length-L RIR vector $\mathbf{h}_{m,0}(t)$ from the source to the mth microphone at time t is defined as

$$\mathbf{h}_{m,0}(t) = \begin{bmatrix} h_{m,0}^{(0)}(t) & \dots & h_{m,0}^{(L-1)}(t) \end{bmatrix}^T,\ t = 1,\dots,N \quad (2)$$

the length-L source signal vector $\mathbf{s}_0(t)$ at time t is defined as

$$\mathbf{s}_0(t) = \begin{bmatrix} s_0(t) & \dots & s_0(t-L+1) \end{bmatrix}^T,\ t = 1,\dots,N \quad (3)$$

and $e_{m,0}(t),\ t = 1,\dots,N$, denotes measurement noise.

In this paper, we make a number of assumptions that may not be valid in realistic sound acquisition scenarios, but which will allow us (1) to focus on the core issues encountered in the dereverberation problem, postponing some practical and implementation issues to future work (see Section 5), and (2) to investigate and interpret the proposed algorithms' behavior only w.r.t. these core issues, disregarding the potential impact of other issues on the algorithm performance. The assumptions are the following (with $m = 1,\dots,M$):

- microphone signals are available for the entire time window $t \in [1, N]$ under consideration;
- RIRs are time-invariant within the time window $t \in [1, N]$ under consideration, i.e., $\mathbf{h}_{m,0}(t) \equiv \mathbf{h}_{m,0}$;
- initial source signal conditions $s_0(t),\ t \leq 0$ are known (and assumed equal to zero for ease of notation);
- no measurement noise is present, i.e., $e_{m,0}(t) \equiv 0$;
- all RIRs have equal and known length $L \leq N$.

Based on these assumptions, the problem considered in this paper can be formulated as follows:

Problem 1 (Multi-microphone dereverberation) *Given a length-MN vector of microphone signals generated as*

$$\mathbf{y} = \mathbf{H}_0\mathbf{s}_0 \quad (4)$$

find the best possible estimate of the length-N source signal vector $\mathbf{s}_0$. Here, with $m = 1,\dots,M$,

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1^T & \dots & \mathbf{y}_M^T \end{bmatrix}^T,\ \mathbf{y}_m = \begin{bmatrix} y_m(1) & \dots & y_m(N) \end{bmatrix}^T \quad (5)$$

$$\mathbf{H}_0 = \begin{bmatrix} \mathbf{H}_{1,0}^T & \dots & \mathbf{H}_{M,0}^T \end{bmatrix}^T \quad (6)$$

$$\mathbf{H}_{m,0} = \begin{bmatrix} h_{m,0}^{(0)} & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ h_{m,0}^{(L-1)} & \dots & h_{m,0}^{(0)} & \ddots & 0 \\ \vdots & \ddots & \vdots & \ddots & 0 \\ 0 & \dots & h_{m,0}^{(L-1)} & \dots & h_{m,0}^{(0)} \end{bmatrix}_{N\times N} \quad (7)$$

$$\mathbf{s}_0 = \begin{bmatrix} s_0(1) & \dots & s_0(N) \end{bmatrix}^T. \quad (8)$$

Since the RIRs in $\mathbf{H}_0$ as well as the source signal vector $\mathbf{s}_0$ are unknown, we define the following parameter vectors,

$$\mathbf{h} = \begin{bmatrix} \mathbf{h}_1^T & \dots & \mathbf{h}_M^T \end{bmatrix}^T,\ \mathbf{h}_m = \begin{bmatrix} h_m^{(0)} & \dots & h_m^{(L-1)} \end{bmatrix}^T \quad (9)$$

$$\mathbf{s} = \begin{bmatrix} s(1) & \dots & s(N) \end{bmatrix}^T \quad (10)$$

$$\mathbf{e} = \begin{bmatrix} \mathbf{e}_1^T & \dots & \mathbf{e}_M^T \end{bmatrix}^T,\ \mathbf{e}_m = \begin{bmatrix} e_m(1) & \dots & e_m(N) \end{bmatrix}^T \quad (11)$$

and a data model admitting two equivalent formulations,

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{e} \quad (12)$$

$$= (\mathbf{I}_M \otimes \mathbf{S})\,\mathbf{h} + \mathbf{e} \quad (13)$$

where $\mathbf{H}$ is a $MN \times N$ matrix with the coefficients of the RIRs parameter vector $\mathbf{h}$ in a block Toeplitz structure as in (6)-(7), $\mathbf{I}_M$ is the $M \times M$ identity matrix, $\otimes$ denotes the Kronecker product, and $\mathbf{S}$ is a $N \times L$ Toeplitz matrix defined as

$$\mathbf{S} = \begin{bmatrix} s(1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ s(N-L+1) & \dots & s(1) \\ \vdots & \ddots & \vdots \\ s(N) & \dots & s(N-L+1) \end{bmatrix}. \quad (14)$$

The error signal parameter vector $\mathbf{e}$ is included to account for estimation errors in both $\mathbf{h}$ and $\mathbf{s}$.

3. EMBEDDED OPTIMIZATION ALGORITHMS

State-of-the-art multi-microphone dereverberation algorithms in the category of blind system identification and inversion approach Problem 1 using a two-step procedure. First, an estimate $\hat{\mathbf{H}}$ of the RIRs matrix is computed using a blind identification method that typically exploits the cross-relation between different microphone signals [5]. Second, an M-input, single-output inverse filter $\mathbf{g}$ is designed and an estimate of the source signal vector is obtained as

$$\hat{\mathbf{s}} = \mathbf{G}\mathbf{y} \tag{15}$$

where $\mathbf{G}$ is a block Toeplitz matrix of appropriate dimensions, containing the inverse filter coefficients in $\mathbf{g}$.

Instead, we propose to *jointly* estimate the RIRs parameter vector $\mathbf{h}$ and the source signal parameter vector $\mathbf{s}$. We derive three nonlinear least squares (NLS) optimization problems for estimating $\mathbf{h}$ and $\mathbf{s}$, and point out their strengths and weaknesses. More particularly, we consider NLS problems without regularization (**NLS**), with ℓ_2-norm regularization exploiting prior knowledge on $\mathbf{h}$ (**ℓ_2-RNLS**), and with ℓ_1-norm and ℓ_2-norm regularization exploiting prior knowledge on $\mathbf{s}$ and $\mathbf{h}$ (**ℓ_1/ℓ_2-RNLS**). A block coordinate descent (BCD) approach is adopted for solving these problems, resulting in three iterative algorithms in which $\mathbf{h}$ and $\mathbf{s}$ are estimated sequentially.

The sequential nature of the proposed algorithms shows a certain degree of similarity with the state-of-the-art two-step procedure for blind system identification and inversion. A crucial difference, however, is that the RIRs parameter vector $\mathbf{h}$ is not just estimated once, but its estimate is iteratively refined as improved estimates of the source signal parameter vector become available. Another similarity with the state of the art, is that the source signal parameter vector estimate resulting from the NLS and ℓ_2-RNLS problems is linearly related to the microphone signal vector $\mathbf{y}$, so that it can be interpreted as the result of an inverse filtering approach, even though an inverse filter is never explicitly designed or computed. When solving the ℓ_1/ℓ_2-RNLS problem, however, the source signal parameter vector estimate is not linearly related to the microphone signal vector and so an inverse filtering interpretation is not appropriate.

3.1. NLS problem

The starting point for the derivation of embedded optimization algorithms solving Problem 1, is the formulation of an NLS optimization problem for the data model (12)-(13),

$$\min_{\mathbf{h},\mathbf{s},\mathbf{e}} \quad \|\mathbf{e}\|_2^2 \tag{16}$$
$$\text{s. t.} \quad \mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{e} \tag{17}$$
$$= (\mathbf{I}_M \otimes \mathbf{S})\,\mathbf{h} + \mathbf{e}. \tag{18}$$

The proposed solution strategy consists in first minimizing (16) w.r.t. $\{\mathbf{s},\mathbf{e}\}$ for a fixed value of $\mathbf{h} = \hat{\mathbf{h}}$ using the equality constraints in (17), then minimizing (16) w.r.t. $\{\mathbf{h},\mathbf{e}\}$ for a fixed value of $\mathbf{s} = \hat{\mathbf{s}}$ using the equality constraints in (18), and repeating this procedure for a number of iterations (here fixed to $k_{\max}$). The resulting BCD algorithm is shown in Algorithm 1, where $(\cdot)^+$ denotes the Moore-Penrose pseudoinverse.

Algorithm 1 BCD algorithm for NLS problem

Input initial RIRs parameter vector estimate $\hat{\mathbf{h}}^{(0)}$
Output parameter vector estimates $\hat{\mathbf{s}} = \hat{\mathbf{s}}^{(k_{\max})}$, $\hat{\mathbf{h}} = \hat{\mathbf{h}}^{(k_{\max})}$
for $k = 1, \ldots, k_{\max}$ **do**
  $\hat{\mathbf{s}}^{(k)} = (\hat{\mathbf{H}}^{(k-1)})^+\mathbf{y}$
  $\hat{\mathbf{h}}^{(k)} = \left(\mathbf{I}_M \otimes (\hat{\mathbf{S}}^{(k)})^+\right)\mathbf{y}$
end for

3.2. ℓ_2-regularized NLS problem

It is well known that the NLS optimization problem in (16)-(18) generally has multiple local solutions, and the BCD algorithm will only converge to the global solution if the algorithm is properly initialized (i.e., if a good initial estimate for either $\mathbf{h}$ or $\mathbf{s}$ is available). An effective approach for smoothing the NLS objective function, and hence facilitating convergence to a meaningful local solution, is the addition of a regularization term incorporating prior knowledge on the unknown parameter vectors. A first approach to regularization consists in the addition of a weighted ℓ_2-norm of the RIRs parameter vector $\mathbf{h}$ to (16),

$$\min_{\mathbf{h},\mathbf{s},\mathbf{e}} \quad \|\mathbf{e}\|_2^2 + \lambda_1\|\mathbf{h}\|_{\mathbf{W}}^2 \tag{19}$$
$$\text{s. t.} \quad \mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{e} \tag{20}$$
$$= (\mathbf{I}_M \otimes \mathbf{S})\,\mathbf{h} + \mathbf{e} \tag{21}$$

A mean-square-error optimal choice for the weighting matrix $\mathbf{W}$ corresponds to the inverse covariance matrix of the true RIRs vector $\mathbf{h}_0 = \begin{bmatrix}\mathbf{h}_{1,0}^T & \ldots & \mathbf{h}_{M,0}^T\end{bmatrix}^T$, which is considered to be a random zero-mean variable having a Gaussian probability density function [13]. In the context of dereverberation, the use of a statistical model for the late reverberation component in the RIRs has been proven useful in a variety of algorithms. The most commonly used model is the so-called Polack's model [3], which approximates the temporal envelope of the RIRs using an exponential function with a fixed decay α. We will adopt this model in the proposed algorithm, and neglect any cross-correlations between the RIRs parameters, such that $\mathbf{W}$ can be defined as a diagonal matrix,

$$\mathbf{W} = \mathbf{I}_M \otimes \text{diag}\left\{1, e^{2\alpha}, \ldots, e^{2(L-1)\alpha}\right\}. \tag{22}$$

The resulting BCD algorithm is given in Algorithm 2.

Algorithm 2 BCD algorithm for ℓ_2-RNLS problem

Input initial RIRs parameter vector estimate $\hat{\mathbf{h}}^{(0)}$, Polack's model decay α, regularization parameter λ_1
Output parameter vector estimates $\hat{\mathbf{s}} = \hat{\mathbf{s}}^{(k_{\max})}$, $\hat{\mathbf{h}} = \hat{\mathbf{h}}^{(k_{\max})}$
$\bar{\mathbf{W}} = \text{diag}\left\{1, e^{2\alpha}, \ldots, e^{2(L-1)\alpha}\right\}$
for $k = 1, \ldots, k_{\max}$ **do**
  $\hat{\mathbf{s}}^{(k)} = (\hat{\mathbf{H}}^{(k-1)})^+\mathbf{y}$
  $\hat{\mathbf{h}}^{(k)} = \left(\mathbf{I}_M \otimes \left[(\hat{\mathbf{S}}^{(k)T}\hat{\mathbf{S}}^{(k)} + \lambda_1\bar{\mathbf{W}})^{-1}\hat{\mathbf{S}}^{(k)T}\right]\right)\mathbf{y}$
end for

3.3. ℓ_1/ℓ_2-regularized NLS problem

With the aim of obtaining an additional smoothing effect, prior knowledge on the source signal vector can also be incorporated in the ℓ_2-RNLS optimization problem. Building on the proven efficiency of sparse representations for speech and audio signals [14], an ℓ_1-norm regularization in a suitable spectral basis seems to be appropriate for this purpose. When combined with the ℓ_2-norm regularization of the RIRs parameter vector, this results in the following optimization problem,

$$\min_{\mathbf{h},\mathbf{s},\mathbf{e}} \quad \|\mathbf{e}\|_2^2 + \lambda_1\|\mathbf{h}\|_{\mathbf{W}}^2 + \lambda_2\|\mathbf{Ds}\|_1 \tag{23}$$

$$\text{s. t.} \quad \mathbf{y} = \mathbf{Hs} + \mathbf{e} \tag{24}$$

$$= (\mathbf{I}_M \otimes \mathbf{S})\,\mathbf{h} + \mathbf{e}. \tag{25}$$

Here, $\mathbf{D}$ is an $N \times N$ orthogonal matrix defining a spectral transform, such as the discrete Fourier or cosine transform (DFT/DCT). In contrast to the previous two problems, the optimization problem in (23)-(25) does not admit a closed-form solution when optimizing w.r.t. $\{\mathbf{s}, \mathbf{e}\}$. However, this particular subproblem is convex and can therefore be efficiently solved using existing software (e.g., we use CVX/SeDuMi). The resulting BCD algorithm is shown in Algorithm 3.

Algorithm 3 BCD algorithm for ℓ_1/ℓ_2-RNLS problem

Input initial RIRs parameter vector estimate $\hat{\mathbf{h}}^{(0)}$, Polack's model decay α, orthogonal spectral transform matrix $\mathbf{D}$, regularization parameters λ_1, λ_2

Output parameter vector estimates $\hat{\mathbf{s}} = \hat{\mathbf{s}}^{(k_{\max})}, \hat{\mathbf{h}} = \hat{\mathbf{h}}^{(k_{\max})}$

1: $\bar{\mathbf{W}} = \text{diag}\left\{1, e^{2\alpha}, \ldots, e^{2(L-1)\alpha}\right\}$

2: **for** $k = 1, \ldots, k_{\max}$ **do**

3: $\quad \hat{\mathbf{s}}^{(k)} = \arg\min_{\mathbf{s}} \|\mathbf{y} - \hat{\mathbf{H}}^{(k-1)}\mathbf{s}\|_2^2 + \lambda_2\|\mathbf{Ds}\|_1$

4: $\quad \hat{\mathbf{h}}^{(k)} = \left(\mathbf{I}_M \otimes \left[(\hat{\mathbf{S}}^{(k)T}\hat{\mathbf{S}}^{(k)} + \lambda_1\bar{\mathbf{W}})^{-1}\hat{\mathbf{S}}^{(k)T}\right]\right)\mathbf{y}$

5: **end for**

---

4. EVALUATION

The proposed embedded optimization algorithms for multi-microphone dereverberation are evaluated here by means of a simulation example. A microphone signal vector $\mathbf{y}$ is generated by filtering a source signal vector $\mathbf{s}_0$ of length $N = 1024$, corresponding to a quasi-stationary voiced segment of a male speech signal sampled at 8 kHz, using $M = 5$ synthetic RIRs of length $L = 100$. The RIRs are synthesized by shaping $M = 5$ different Gaussian white noise (GWN) sequences with an exponential envelope corresponding to Polack's model with $\alpha = 0.025$. The same envelope is used for designing the weighting matrix $\mathbf{W}$ in the ℓ_2-RNLS and ℓ_1/ℓ_2-RNLS problems. The regularization parameters have been chosen as $\lambda_1 = \lambda_2 = 0.1$, and $\mathbf{D}$ is the DCT matrix. All algorithms start from a random GWN initial RIRs parameter vector estimate $\hat{\mathbf{h}}^{(0)}$ and perform $k_{\max} = 10$ iterations. In the simulation results, the inherent scaling ambiguity has been removed by plotting $\hat{\mathbf{s}}/a$ and $a\hat{\mathbf{h}}$ rather than $\hat{\mathbf{s}}$ and $\hat{\mathbf{h}}$, with $a = \sqrt{\hat{\mathbf{s}}^T\hat{\mathbf{s}}/\mathbf{s}_0^T\mathbf{s}_0}$.

Fig. 1 shows the true and estimated RIR ($m = 2$), while Fig. 2 compares the magnitude spectrum of the true and estimated source signal. As expected, the BCD algorithm does not converge to the global NLS problem solution, and suffers from a severe overestimation of the coefficients in the RIR tail as well as large (≥ 10 dB) source spectrum estimation errors in some frequency regions. The ℓ_2-norm regularization is seen to have a beneficial effect on the *overall* estimation performance, yielding RIR and source spectrum estimates that follow the envelopes of the true RIR and source spectrum. In addition, the ℓ_1-norm regularization further increases the *local* estimation performance: the right plot in Fig. 1 shows an improved estimation of the early RIR coefficients, while the top right subplot in Fig. 2(c) illustrates the improved accuracy of the estimated quasi-harmonic speech components.

5. CONCLUSION

In this paper, we have introduced a new approach to multi-microphone dereverberation, based on a recent paradigm known as embedded optimization. Three sequential optimization algorithms have been proposed, which enable the joint estimation of the unknown source signal and room acoustics. By adopting an iterative numerical optimization strategy, the need for an explicit inverse filter design is avoided. However, the inclusion of appropriate regularization terms in the inherently non-convex optimization problem appears to be crucial for assuring convergence to a meaningful local solution. In particular, the addition of a weighted ℓ_2-norm of the RIRs parameter vector, based on a statistical model for late reverberation, leads to an improved overall estimation performance. In addition, the accuracy of the estimated early reflections and (quasi-)harmonic source signal components can be further increased by incorporating an ℓ_1-norm regularization term for the source signal parameter vector DFT/DCT.

The work presented in this paper is a first step towards the development of efficient and reliable embedded optimization algorithms for multi-microphone dereverberation. A number of challenges for future research remain, e.g.,

- to move from batch to online (frame-based) processing, properly managing initial/final conditions,
- to generalize the ℓ_2-norm regularization for dealing with realistic impulse responses,
- to take measurement noise into account,
- to arrive at autonomous optimization algorithms involving proper termination criteria and cross-validation procedures for adjusting the regularization parameters,
- to derive fast SQP/SCP algorithms exploiting the particular dereverberation problem structure,
- to use perceptual criteria in the problem formulations,
- to evaluate the resulting dereverberation performance.

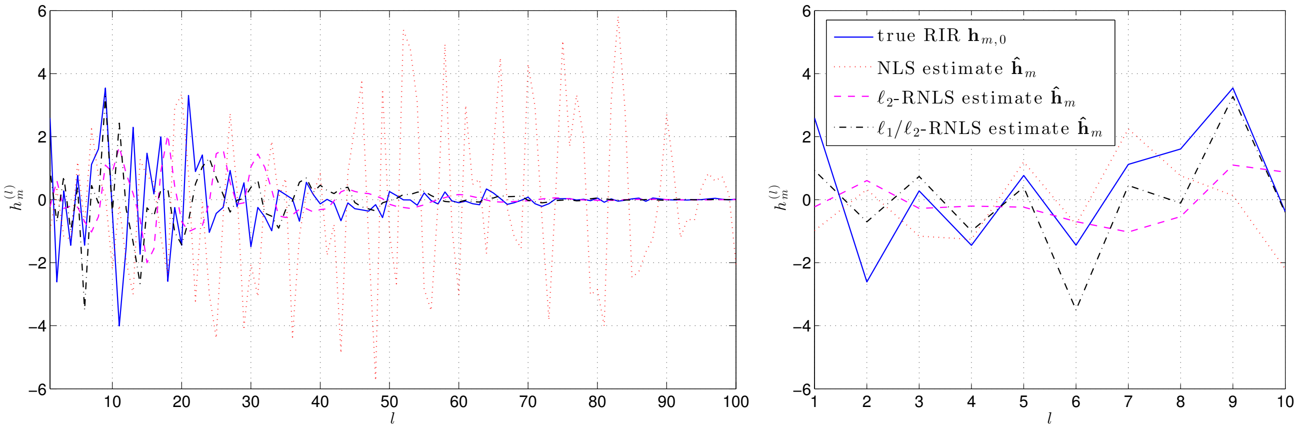
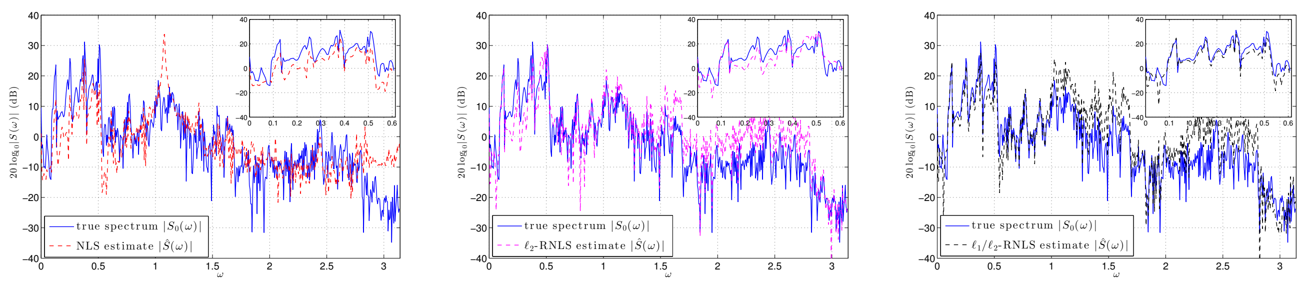


Fig. 1. Comparison of true RIR and RIR parameter vector estimates for $m = 2$.



(a) NLS

(b) ℓ_2-RNLS

(c) ℓ_1/ℓ_2-RNLS

Fig. 2. Comparison of magnitude spectra of true source signal and source signal parameter vector estimates.

6. REFERENCES

[1] P. A. Naylor and N. D. Gaubitch, Eds., *Speech dereverberation*, Springer, London, 2010.

[2] J. L. Flanagan, A. C. Surendran, and E. E. Jan, "Spatially selective sound capture for speech and audio processing," *Speech Commun.*, vol. 13, no. 1–2, pp. 207–222, Oct. 1993.

[3] E. A. P. Habets, "Multi-channel speech dereverberation based on a statistical model of late reverberation," in *Proc. 2005 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP '05)*, Philadelphia, PA, USA, Mar. 2005, vol. 4, pp. 173–176.

[4] N. D. Gaubitch and P. A. Naylor, "Spatiotemporal averaging method for enhancement of reverberant speech," in *Proc. 2007 IEEE Int. Conf. Digital Signal Process. (DSP '07)*, Cardiff, UK, June 2007, pp. 607–610.

[5] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Process.*, vol. 43, no. 12, pp. 2982–2993, Dec. 1995.

[6] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-36, no. 2, pp. 145–152, Feb. 1988.

[7] W. Zhang and P. A. Naylor, "An experimental study of the robustness of multichannel inverse filtering systems to near-common zeros," in *Proc. 17th European Signal Process. Conf. (EUSIPCO '09)*, Glasgow, Scotland, UK, Aug. 2009, pp. 194–198.

[8] W. Zhang and P. A. Naylor, "A system-identification-error-robust method for equalization of multichannel acoustic systems," in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP '10)*, Dallas, TX, USA, Mar. 2010, pp. 109–112.

[9] N. D. Gaubitch and P. A. Naylor, "Equalization of multichannel acoustic systems in oversampled subbands," *IEEE Trans. Audio Speech Lang. Process.*, vol. 17, no. 6, pp. 1061–1070, Aug. 2009.

[10] I. Kodrasi, S. Goetze, and S. Doclo, "Increasing the robustness of acoustic multichannel equalization by means of regularization," in *Proc. 2012 Int. Workshop Acoustic Signal Enhancement (IWAENC '12)*, Aachen, Germany, Sept. 2012, pp. 161–164.

[11] X. S. Lin, A. W. H. Khong, and P. A. Naylor, "A forced spectral diversity algorithm for speech dereverberation in the presence of near-common zeros," *IEEE Trans. Audio Speech Lang. Process.*, vol. 20, no. 3, pp. 888–899, Mar. 2012.

[12] J. Mattingley and S. Boyd, "Real-time convex optimization in signal processing," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 50–61, May 2010.

[13] T. van Waterschoot, G. Rombouts, and M. Moonen, "Optimally regularized adaptive filtering algorithms for room acoustic signal enhancement," *Signal Processing*, vol. 88, no. 3, pp. 594–611, Mar. 2008.

[14] M. D. Plumbey, T. Blumensath, L. Daudet, R. Gribonval, and M. E. Davies, "Sparse representations in audio and music: from coding to source separation," *Proc. IEEE*, vol. 98, no. 6, pp. 995–1005, 2010.