

On Least Squares Equalization and Channel Shortening for Speech Dereverberation

Ina Kodrasi*, *Student Member, IEEE*, Simon Doclo, *Senior Member, IEEE*

Abstract

In this letter, a generalized framework for least squares acoustic multichannel equalization techniques is established, which enables to analyze the properties (i.e., existence and uniqueness) of the resulting reshaping filters. It is shown that least squares equalization techniques yield reshaping filters that lie in the subspace spanned by the multiple solutions maximizing the so-called channel shortening (CS) cost function. Since least squares reshaping filters have been experimentally validated to yield a trade-off between reverberant tail suppression and perceptual speech quality preservation and they lie in the subspace of CS solutions, it can be said that the multiple CS solutions offer the potential to achieve a high performance both in terms of reverberant tail suppression and perceptual speech quality preservation. Hence, a novel subspace-based equalization (SuB) technique is proposed, which constrains the reshaping filter to be a linear combination of the multiple CS solutions. Simulation results illustrate the high dereverberation performance of the proposed technique.

Index Terms

least squares equalization, channel shortening, subspace-based equalization

I. INTRODUCTION

Acoustic multichannel equalization comprises an attractive approach to speech dereverberation due to its potential to theoretically achieve perfect dereverberation [1], [2]. However, in practice such an approach

The authors are with the Signal Processing Group, Department of Medical Physics and Acoustics, and Cluster of Excellence Hearing4All, University of Oldenburg, Oldenburg, Germany (e-mail: ina.kodrasi@uni-oldenburg.de; simon.doclo@uni-oldenburg.de).

This work was supported in part by a Grant from the GIF, the German-Israeli Foundation for Scientific Research and Development, the Cluster of Excellence 1077 “Hearing4All” funded by the German Research Foundation (DFG), and the Marie Curie Initial Training Network DREAMS (Grant no. 316969).

remains challenging given its sensitivity to estimation errors in the room impulse responses (RIRs) between the source and the microphone array [2], [3]. Over the past decades, several acoustic multichannel equalization techniques have been proposed [1], [2], [3], [4], [5], [6], [7], [8], [9]. On the one hand, least squares equalization techniques such as the multiple-input/output inverse theorem (MINT)-based technique [1], the partial multichannel equalization technique based on MINT (P-MINT) [2], the relaxed multichannel least squares technique (RMCLS) [3], and the relaxed multichannel least squares technique with constrained initial taps (RMCLS-CIT) [4], design reshaping filters such that the (weighted) system response equals a given (weighted) target response. On the other hand, channel shortening aims at an overall system response with maximum energy in the direct path and early reflections and minimum energy in the reverberant tail. In [3] it has been shown that multiple solutions exist which maximize the CS cost function. Furthermore, in [2] and [3] it has been proven that the reshaping filters designed using the MINT and P-MINT techniques can be expressed as linear combinations of these multiple CS solutions.

The aim of this letter is twofold. Firstly, a generalized framework for state-of-the-art least squares equalization techniques is established. A theoretical analysis based on the Rouché-Capelli theorem [10] is provided to determine the properties (i.e., existence and uniqueness) of the solution(s) for each technique. Using this analysis, it can be shown that state-of-the-art least squares equalization techniques yield reshaping filters which lie in the subspace spanned by the multiple CS solutions. Although this result may appear trivial, to the best of our knowledge it was not yet available in the literature. Secondly, the subspace of CS solutions is exploited to derive a novel subspace-based equalization (SuB) technique aiming at achieving a high reverberant tail suppression and perceptual speech quality preservation. Experimental results demonstrate the advantage of the proposed technique as compared to the least squares equalization techniques in the presence of RIR estimation errors.

II. ACOUSTIC MULTICHANNEL EQUALIZATION

A. Problem Formulation

Consider an acoustic system with a single speech source and M microphones as depicted in Fig. 1. The m -th microphone signal $x_m(n)$ at time index n is given by

$$x_m(n) = s(n) * h_m(n), \quad m = 1, \dots, M, \quad (1)$$

with $s(n)$ the clean speech signal, $h_m(n)$ the L_h taps long RIR between the source and the m -th microphone, i.e., $\mathbf{h}_m = [h_m(0) \ h_m(1) \ \dots \ h_m(L_h - 1)]^T$, and $*$ denoting convolution. Equalization

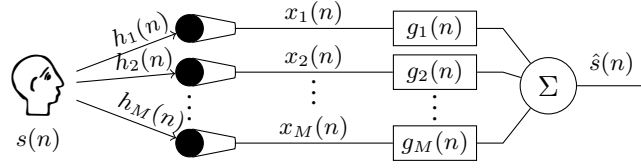


Fig. 1: Multichannel equalization system.

techniques apply filters \mathbf{g}_m of length L_g , i.e., $\mathbf{g}_m = [g_m(0) \ g_m(1) \ \dots \ g_m(L_g - 1)]^T$, such that the output of the system $\hat{s}(n)$ is given by

$$\hat{s}(n) = \sum_{m=1}^M x_m(n) * g_m(n) = s(n) * \underbrace{\sum_{m=1}^M h_m(n) * g_m(n)}_{c(n)}, \quad (2)$$

with $c(n)$ denoting the overall response between the source and the output, referred to as the equalized impulse response (EIR). The EIR of length $L_c = L_h + L_g - 1$ can be described in vector notation as $\mathbf{c} = [c(0) \ c(1) \ \dots \ c(L_c - 1)]^T$. Using the ML_g -dimensional reshaping filter

$$\mathbf{g} = [\mathbf{g}_1^T \ \mathbf{g}_2^T \ \dots \ \mathbf{g}_M^T]^T \quad (3)$$

and the $L_c \times ML_g$ -dimensional multichannel convolution matrix \mathbf{H} , i.e.,

$$\mathbf{H} = [\mathbf{H}_1 \ \mathbf{H}_2 \ \dots \ \mathbf{H}_M], \quad (4)$$

with \mathbf{H}_m the convolution matrix of $h_m(n)$, the EIR can be expressed as $\mathbf{c} = \mathbf{H}\mathbf{g}$. The reshaping filter can then be constructed based on different design objectives for the EIR \mathbf{c} . Since the true RIRs are typically not available in practice [11], [12], [13], equalization techniques design \mathbf{g} using the estimated multichannel convolution matrix $\hat{\mathbf{H}}$ (constructed from the estimated RIRs $\hat{h}_m(n)$), hence optimizing the estimated EIR

$$\hat{\mathbf{c}} = \hat{\mathbf{H}}\mathbf{g} \quad (5)$$

Under the assumptions typically made in acoustic multichannel equalization [1], i.e.,

- the estimated RIRs do not share any common zeros, and
- $L_g \geq \lceil \frac{L_h - 1}{M - 1} \rceil$,

the estimated multichannel convolution matrix $\hat{\mathbf{H}}$ is assumed to be a full row-rank matrix [14].

In the following, the optimization criteria of state-of-the-art equalization techniques are reviewed.

B. Least Squares Equalization Techniques

The aim of least squares equalization is to design a reshaping filter \mathbf{g} such that the (weighted) EIR equals a (weighted) target EIR \mathbf{c}_t . Formally, these techniques seek to compute a solution \mathbf{g} to the system of equations

$$\boxed{\mathbf{W}\hat{\mathbf{H}}\mathbf{g} = \mathbf{W}\mathbf{c}_t} \quad (6)$$

with \mathbf{W} being a diagonal weighting matrix. The definition of the weighting matrix \mathbf{W} and target EIR \mathbf{c}_t for MINT [1], P-MINT [2], RMCLS [3], and RMCLS-CIT [4] is presented in Tables I and II respectively, where \mathbf{I} is the $L_c \times L_c$ -dimensional identity matrix, τ denotes a delay in number of samples, L_d is the length in number of samples of the direct path and early reflections, L_r is the number of taps to constrain in RMCLS-CIT, and $p \in \{1, \dots, M\}$. From these definitions of \mathbf{W} and \mathbf{c}_t , it can be observed that on the one hand, techniques such as MINT and P-MINT aim at setting all taps of the resulting EIR to a desired target EIR, which results in a good perceptual speech quality preservation but low reverberant tail suppression [2], while on the other hand, techniques such as RMCLS and RMCLS-CIT do not constrain all taps of the EIR, which results in a high performance in terms of reverberant tail suppression, but lower perceptual speech quality preservation [2], [4].

TABLE I: Definition of the weighting matrix \mathbf{W} for state-of-the-art least squares equalization techniques.

Technique	Weighting matrix \mathbf{W}
MINT	\mathbf{I}
P-MINT	\mathbf{I}
RMCLS	$\mathbf{W}_c = \text{diag}[\underbrace{1 \dots 1}_\tau \underbrace{1 \ 0 \dots 0}_{L_d} 1 \dots 1]^T$
RMCLS-CIT	$\mathbf{W}_r = \text{diag}[\underbrace{1 \dots 1}_\tau \underbrace{1 \ 1 \dots 1}_{L_r} \underbrace{0 \dots 0}_{L_d - L_r} 1 \dots 1]^T$

TABLE II: Definition of the target EIR \mathbf{c}_t for state-of-the-art least squares equalization techniques.

Technique	Target EIR \mathbf{c}_t
MINT	$\mathbf{d} = [\underbrace{0 \dots 0}_\tau 1 \ 0 \dots 0]^T$
P-MINT	$\hat{\mathbf{h}}_p^d = [\underbrace{0 \dots 0}_\tau \underbrace{\hat{h}_p(0) \dots \hat{h}_p(L_d - 1)}_{L_d} 0 \dots 0]^T$
RMCLS	$\mathbf{d} = [\underbrace{0 \dots 0}_\tau 1 \ 0 \dots 0]^T$
RMCLS-CIT	$\hat{\mathbf{h}}_p^d = [\underbrace{0 \dots 0}_\tau \underbrace{\hat{h}_p(0) \dots \hat{h}_p(L_d - 1)}_{L_d} 0 \dots 0]^T$

C. Channel Shortening (CS)

Channel shortening [3], [5] aims at maximizing the energy in the first L_d taps of the EIR (i.e., direct path and early reflections), while minimizing the energy in the remaining taps (i.e., reverberant tail). This objective can be expressed as the maximization of a generalized Rayleigh quotient, i.e.,

$$J_{\text{CS}}(\mathbf{g}) = \frac{\|\mathbf{W}_d \hat{\mathbf{c}}\|_2^2}{\|\mathbf{W}_u \hat{\mathbf{c}}\|_2^2} = \frac{\|\mathbf{W}_d \hat{\mathbf{H}} \mathbf{g}\|_2^2}{\|\mathbf{W}_u \hat{\mathbf{H}} \mathbf{g}\|_2^2} = \frac{\mathbf{g}^T \hat{\mathbf{B}} \mathbf{g}}{\mathbf{g}^T \hat{\mathbf{A}} \mathbf{g}} \quad (7)$$

where $\mathbf{W}_d = \text{diag}\{\mathbf{w}_d\}$ and $\mathbf{W}_u = \text{diag}\{\mathbf{w}_u\}$, with

$$\mathbf{w}_d = [\underbrace{0 \dots 0}_\tau \underbrace{1 \dots 1}_{L_d} 0 \dots 0]^T, \quad \mathbf{w}_u = \mathbf{1} - \mathbf{w}_d, \quad (8)$$

and

$$\hat{\mathbf{B}} = \hat{\mathbf{H}}^T \mathbf{W}_d^T \mathbf{W}_d \hat{\mathbf{H}}, \quad \hat{\mathbf{A}} = \hat{\mathbf{H}}^T \mathbf{W}_u^T \mathbf{W}_u \hat{\mathbf{H}}. \quad (9)$$

Maximizing (7) is equivalent to solving the generalized eigenvalue problem $\hat{\mathbf{B}} \mathbf{g} = \lambda \hat{\mathbf{A}} \mathbf{g}$, where the optimal reshaping filter \mathbf{g}_{CS} is the generalized eigenvector corresponding to the largest generalized eigenvalue λ_{max} . In [3] it has been shown that L_d linearly independent generalized eigenvectors $\mathbf{g}_{\text{CS}}^i, i = 1, \dots, L_d$, exist, which maximize the CS cost function in (7) to $\lambda_{\text{max}} = \infty$. Hence, any reshaping filter \mathbf{g} in the subspace spanned by the L_d solutions \mathbf{g}_{CS}^i leads to an estimated EIR $\hat{\mathbf{c}}$ which satisfies the system of equations

$$\begin{cases} \|\mathbf{W}_d \hat{\mathbf{c}}\|_2^2 \neq 0 \\ \|\mathbf{W}_u \hat{\mathbf{c}}\|_2^2 = 0. \end{cases} \quad (10)$$

In the next section, it will be shown that all considered least squares equalization techniques yield an estimated EIR $\hat{\mathbf{c}}$ which satisfies (10).

III. INSIGHTS INTO LEAST SQUARES EQUALIZATION TECHNIQUES AND THEIR RELATION TO CS

In the following, the Rouché-Capelli theorem [10] is used to establish the existence and uniqueness of solutions to (6) for the different definitions of \mathbf{W} and \mathbf{c}_t .

Rouché-Capelli theorem: Consider the system of equations $\mathbf{A} \mathbf{x} = \mathbf{b}$ with \mathbf{A} a $p \times q$ matrix. Such a system has a solution if and only if the rank of the coefficient matrix \mathbf{A} is equal to the rank of the augmented matrix $[\mathbf{A}|\mathbf{b}]$. If a solution exists and $q = \text{rank}(\mathbf{A})$, this solution is unique, otherwise there are an infinite number of solutions.

Exploiting the fact that $\hat{\mathbf{H}}$ is a full row-rank matrix, Table III summarizes the rank of the coefficient and augmented matrix for each least squares technique. Since the rank of the coefficient matrix is equal

TABLE III: Rank of the coefficient and augmented matrix for state-of-the-art least squares equalization techniques.

Technique	$\mathbf{Ax}=\mathbf{b}$	$\text{rank}(\mathbf{A})$	$\text{rank}[\mathbf{A} \mathbf{b}]$
MINT	$\hat{\mathbf{H}}\mathbf{g}=\mathbf{d}$	L_c	L_c
P-MINT	$\hat{\mathbf{H}}\mathbf{g}=\hat{\mathbf{h}}_p^d$	L_c	L_c
RMCLS	$\mathbf{W}_c\hat{\mathbf{H}}\mathbf{g}=\mathbf{W}_c\mathbf{d}$	L_c-L_d+1	L_c-L_d+1
RMCLS-CIT	$\mathbf{W}_r\hat{\mathbf{H}}\mathbf{g}=\mathbf{W}_r\hat{\mathbf{h}}_p^d$	$L_c-L_d+L_r$	$L_c-L_d+L_r$

to the rank of the augmented matrix for all definitions of \mathbf{W} and \mathbf{c}_t , the system of equations in (6) is always solvable. For the MINT and P-MINT techniques we need to distinguish amongst the following two cases:

- if $L_c = L_h + L_g - 1 = ML_g$, i.e., $\hat{\mathbf{H}}$ is a square matrix and hence, $\text{rank}(\hat{\mathbf{H}}) = L_c = ML_g$, there is a unique solution to (6),
- otherwise there are an infinite number of solutions.

For the RMCLS technique there is always an infinite number of solutions since the number of columns in $\mathbf{W}\hat{\mathbf{H}}$ is always greater than its rank, i.e., $ML_g > L_c - L_d + 1$. Furthermore, also for the RMCLS-CIT technique there is always an infinite number of solutions since $ML_g > L_c - L_d + L_r$.

Given that (6) is solvable for all considered least squares techniques, applying all solutions \mathbf{g} to the estimated RIRs yields the respective target EIRs. Clearly, all target EIRs in Table II satisfy the system of equations in (10). Hence, all solutions \mathbf{g} of the least squares equalization techniques lie in the subspace spanned by the multiple CS solutions, i.e.,

$$\mathbf{g} = \mathbf{G}\boldsymbol{\alpha} \quad (11)$$

with $\mathbf{G} = [\mathbf{g}_{\text{CS}}^1 \ \dots \ \mathbf{g}_{\text{CS}}^{L_d}]$ being the $ML_g \times L_d$ -dimensional matrix whose columns are the L_d channel shortening solutions and $\boldsymbol{\alpha}$ being an L_d -dimensional linear combination vector.

In the following section, the subspace of CS solutions is exploited to derive a novel subspace-based equalization technique. Some preliminary results on exploiting the subspace of CS solutions were discussed in [15] and [16].

IV. SUBSPACE-BASED EQUALIZATION TECHNIQUE

In [2] it has been shown that RMCLS achieves a high reverberant tail suppression but a low perceptual speech quality preservation. On the other hand, P-MINT achieves a low reverberant tail suppression but

a high perceptual speech quality preservation. Since the theoretical analysis in Section III has shown that all least squares filters lie in the subspace spanned by the CS solutions, it can be said that this subspace offers the potential to achieve a high dereverberation performance, both in terms of reverberant tail suppression and perceptual speech quality preservation. Furthermore, in [2] it has been shown that incorporating regularization in equalization techniques in order to decrease the energy of the reshaping filter yields a significant increase in robustness in the presence of RIR estimation errors. Based on such theoretical and experimental observations, the following regularized subspace-based cost function (SuB) is proposed

$$J_{\text{SuB}}^{\text{R}}(\alpha) = \|\mathbf{W}(\hat{\mathbf{H}}\mathbf{G}\alpha - \mathbf{c}_t)\|_2^2 + \delta\|\mathbf{G}\alpha\|_2^2 \quad (12)$$

with δ being a regularization parameter. The minimization of (12) aims at computing the reshaping filter as a linear combination of the multiple CS solutions which yields a desired (weighted) target EIR \mathbf{c}_t , while still controlling the energy of the reshaping filter in order to increase robustness to RIR estimation errors. Minimizing (12) yields the linear combination vector

$$\alpha_{\text{SuB}}^{\text{R}} = [(\mathbf{W}\hat{\mathbf{H}}\mathbf{G})^T(\mathbf{W}\hat{\mathbf{H}}\mathbf{G}) + \delta\mathbf{I}]^{-1}(\mathbf{W}\hat{\mathbf{H}}\mathbf{G})^T(\mathbf{W}\mathbf{c}_t) \quad (13)$$

and the subspace-based reshaping filter

$$\mathbf{g}_{\text{SuB}}^{\text{R}} = \mathbf{G}\alpha_{\text{SuB}}^{\text{R}} \quad (14)$$

V. EXPERIMENTAL RESULTS

In this section, the performance of the proposed subspace-based technique will be investigated for different definitions of the weighting matrix \mathbf{W} and target EIR \mathbf{c}_t and compared to the performance of regularized least squares equalization. For an overview of the regularized least squares equalization techniques, the reader is referred to [2].

We have considered a measured 2-channel acoustic system with reverberation time $T_{60} \approx 450$ ms as the true system to be equalized. The sampling frequency is $f_s = 8$ kHz and the length of the RIRs is $L_h = 3600$. Furthermore, the simulation parameters are set to $\tau = 0$, $L_g = 3599$, $p = 1$, $L_d = 0.05f_s$ (i.e., 50 ms), and $L_r = \lfloor \frac{L_d}{3} \rfloor$. The estimated RIRs are simulated as in [2], i.e.,

$$\hat{h}_m(n) = h_m(n)[1 + e_m(n)], \quad (15)$$

with $e_m(n)$ an uncorrelated Gaussian noise sequence with zero mean and variance such that a normalized channel mismatch

$$E_m = 10 \log_{10} \frac{\|\mathbf{h}_m - \hat{\mathbf{h}}_m\|_2^2}{\|\mathbf{h}_m\|_2^2} \quad (16)$$

is generated. The considered normalized channel mismatch values are $E_m \in \{-33 \text{ dB}, -30 \text{ dB}, \dots, -15 \text{ dB}\}$.

The reverberant energy suppression is evaluated using the energy decay curve (EDC) of the resulting EIR [17], whereas the perceptual speech quality is evaluated using the objective speech quality measure PESQ [18]. The reference signal employed in PESQ is the clean speech convolved with the first part of the true first RIR, i.e., $s(n) * h_1^d(n)$.

Finally, the set of considered regularization parameters for all techniques is $\delta \in \{10^{-9}, 0.5 \times 10^{-8}, 10^{-8}, \dots, 10^{-1}\}$ and the used parameter is intrusively selected as the one leading to the highest PESQ score.

Fig. 2 depicts the EDC of the true RIR h_1 and the average EDCs (averaged over the different considered mismatch values) obtained using the regularized SuB technique with the weighting matrix \mathbf{W} and the target vector \mathbf{c}_t defined according to Table I and II. It appears that using the weighting matrix and the target vector of RMCLS and P-MINT yields the highest reverberant energy suppression in the CS subspace domain, with the reverberant tails being well below the reverberant tail of the true first RIR h_1 . However, it can be said that the performance of the regularized SuB technique for all definitions of \mathbf{W} and \mathbf{c}_t is quite similar in terms of reverberant energy suppression. On the other hand, significant performance differences can be noticed in terms of perceptual speech quality preservation for all considered mismatch values as depicted in Fig. 3. As illustrated in this figure, the regularized SuB technique with \mathbf{W} and \mathbf{c}_t defined as in P-MINT (i.e., $\mathbf{W} = \mathbf{I}$ and $\mathbf{c}_t = \hat{\mathbf{h}}_p^d$) yields a significantly higher PESQ score than using any other definition of the weighting matrix and target vector, offering a large improvement over the

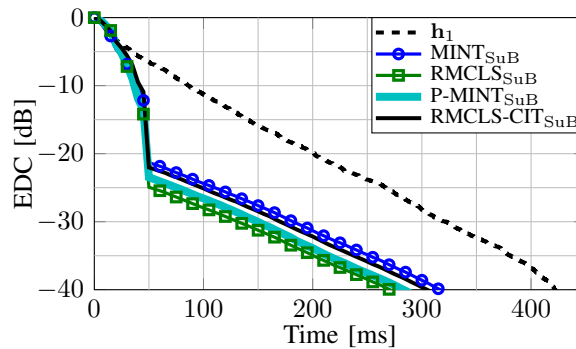


Fig. 2: EDC of the true RIR h_1 and average EDC (averaged over several channel mismatch values) of the EIR obtained using the regularized SuB technique with \mathbf{W} and \mathbf{c}_t defined as in MINT, RMCLS, P-MINT, and RMCLS-CIT.

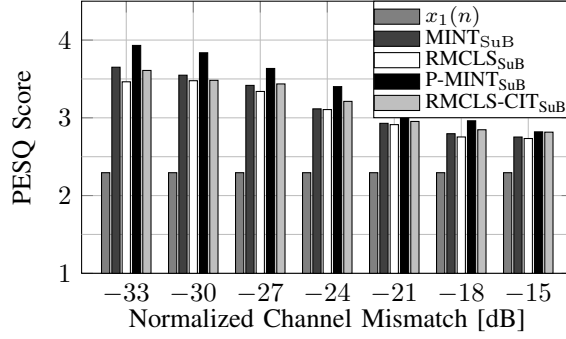


Fig. 3: PESQ score of the reverberant signal $x_1(n)$ and PESQ score of the output signal $\hat{s}(n)$ using the regularized SuB technique with \mathbf{W} and \mathbf{c}_t defined as in MINT, RMCLS, P-MINT, and RMCLS-CIT for several channel mismatch values.

perceptual speech quality of the unprocessed microphone signal $x_1(n)$. Based on these simulations, it can be concluded that the regularized SuB technique with the weighting matrix and target vector defined as in P-MINT leads to the highest dereverberation performance out of all considered subspace-based techniques. Since the regularized SuB technique based on P-MINT results in the highest performance, its performance is compared to the regularized least squares P-MINT technique in order to evaluate the advantages of doing channel reshaping in the CS subspace domain as compared to least squares

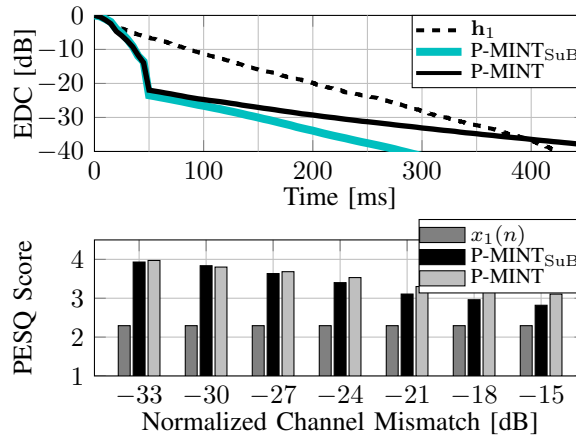


Fig. 4: The performance of the regularized SuB technique ($\mathbf{W} = \mathbf{I}$, $\mathbf{c}_t = \hat{\mathbf{h}}_p^d$) and of the regularized P-MINT technique. Top figure: Average EDC over several channel mismatch values. Bottom figure: PESQ score for several channel mismatch values.

equalization. Fig. 4 depicts the performance in terms of average EDC (top figure) and PESQ score (bottom figure) of the regularized SuB technique ($\mathbf{W} = \mathbf{I}$, $\mathbf{c}_t = \hat{\mathbf{h}}_p^d$) as well as of the regularized least squares P-MINT technique. As illustrated in this figure, the regularized SuB technique significantly outperforms the regularized P-MINT technique in terms of reverberant energy suppression. However, while the obtained PESQ scores for both techniques are very similar for moderate channel estimation errors, as the channel estimation errors increase, the regularized P-MINT technique appears to yield a higher perceptual speech quality.

VI. CONCLUSION

In this letter, we have theoretically analyzed the properties of the solutions of state-of-the-art least squares equalization techniques and showed that they belong to the subspace spanned by the multiple CS solutions. A novel regularized subspace-based equalization (SuB) technique has been proposed, where the reshaping filter is constrained to be a linear combination of the multiple CS solutions. Simulation results demonstrate that the regularized SuB technique with the weighting matrix and target vector defined as in P-MINT outperforms the regularized least squares P-MINT technique in terms of reverberant tail suppression in the presence of RIR estimation errors. Furthermore, the regularized SuB technique also yields a similar perceptual speech quality as the regularized P-MINT technique for moderate channel estimation errors.

REFERENCES

- [1] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 2, pp. 145–152, Feb. 1988.
- [2] I. Kodrasi, S. Goetze, and S. Doclo, "Regularization for partial multichannel equalization for speech dereverberation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 9, pp. 1879–1890, Sep. 2013.
- [3] W. Zhang, E. A. P. Habets, and P. A. Naylor, "On the use of channel shortening in multichannel acoustic system equalization," in *Proc. International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Tel Aviv, Israel, Sep. 2010.
- [4] F. Lim and P. A. Naylor, "Relaxed Multichannel Least Squares with Constrained Initial Taps for Multichannel Dereverberation," in *Proc. International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Aachen, Germany, Sep. 2012.
- [5] M. Kallinger and A. Mertins, "Multi-channel room impulse response shaping - a study," in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Toulouse, France, May 2006, pp. 101–104.
- [6] F. Lim and P. A. Naylor, "Robust low-complexity multichannel equalization for dereverberation," in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013, pp. 689–693.

- [7] F. Lim, M. R. P. Thomas, and P. A. Naylor, "MINTFormer: A spatially aware channel equalizer," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, New York, USA, Oct. 2013.
- [8] A. Mertins, T. Mei, and M. Kallinger, "Room impulse response shortening/reshaping with infinity- and p -norm optimization," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 2, pp. 249–259, Feb. 2010.
- [9] J. O. Jungmann, R. Mazur, and A. Mertins, "Joint time-domain reshaping and frequency-domain equalization of room impulse responses," in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Florence, Italy, May 2014, pp. 6692–6696.
- [10] I. R. Shafarevich and A. O. Remizov, *Linear algebra and geometry*. Springer Publishing Company, 2010.
- [11] B. D. Radlovic, R. C. Williamson, and R. A. Kennedy, "Equalization in an acoustic reverberant environment: robustness results," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 3, pp. 311–319, May 2000.
- [12] K. Hasan and P. A. Naylor, "Analyzing effect of noise on LMS-type approaches to blind estimation of SIMO channels: robustness issue," in *Proc. European Signal Processing Conference (EUSIPCO)*, Florence, Italy, Sep. 2006.
- [13] L. Xiang, A. W. H. Khong, and P. A. Naylor, "A forced spectral diversity algorithm for speech dereverberation in the presence of near-common zeros," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 3, pp. 888–899, Mar. 2012.
- [14] G. Harikumar and Y. Bresler, "FIR perfect signal reconstruction from multiple convolutions: minimum deconvolver orders," *IEEE Transactions on Signal Processing*, vol. 46, no. 1, pp. 215–218, Jan. 1998.
- [15] I. Kodrasi, S. Goetze, and S. Doclo, "A perceptually constrained channel shortening technique for speech dereverberation," in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013, pp. 151–155.
- [16] I. Kodrasi and S. Doclo, "Regularized Subspace-Based Acoustic Multichannel Equalization for Speech Dereverberation," in *Proc. European Signal Processing Conference (EUSIPCO)*, Marrakech, Morocco, Sep. 2013.
- [17] P. A. Naylor and N. D. Gaubitch, *Speech Dereverberation*, 1st ed. Springer Publishing Company, 2010.
- [18] ITU-T, *Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs P.862*, International Telecommunications Union (ITU-T) Recommendation, Feb. 2001.