SINGLE CHANNEL REVERBERATION SUPPRESSION BASED ON SPARSE LINEAR PREDICTION

Nicolás López[1,2], *Yves Grenier*[2], *Gaël Richard*[2], *Ivan Bourmeyster*[1]

[1] Arkamys - 31 rue Pouchet, 75017 Paris, France

[2] Institut Mines-Télécom - Télécom ParisTech - CNRS-LTCI - 37/39 rue Dareau, 75014 Paris, France

ABSTRACT

Reverberation degrades speech intelligibility in telecommunications as well as it increases the word error rate in automatic speech recognition tasks. Several dereverberation methods have been proposed recently in order to counter these effects. In the single microphone case, the dereverberation problem is underdetermined and reverberation suppression approaches are preferred. In this paper we propose a novel method for single channel reverberation suppression. Late reverberation is estimated in the time-frequency domain as a sparse linear combination of previous frames. The predictors associated to the model are determined in a Lasso framework and a spectral subtraction filter is designed to produce the enhanced signal. This model does not require any additional information about the room acoustics and it is well suited for real-time applications. The method has state-of-the-art performance in terms of both reverberation suppression and spectral distortion.

Index Terms— Single Channel Speech Enhancement, Late Reverberation Estimation, Lasso, Sparse Linear Prediction

1. INTRODUCTION

The speech enhancement community has focused for a long time on noise reduction tasks, giving rise to several very efficient methods. Recently, the rapid development of mobile technologies and the use of hands-free devices in various (possibly big) enclosures has raised the problem of room reverberation. Reverberation affects telecommunications as it degrades speech intelligibility. It also affects vocal based Human-Machine Interfaces (HMI) by increasing the word error rate in Automatic Speech Recognition (ASR) tasks. Reverberation is commonly decomposed into early reflections and late reverberation. It has been shown that early reflections are sufficiently close to the direct sound to be integrated by the ear and improve intelligibility [1]. On the counterpart, late reverberation degrades intelligibility by smearing the time-frequency support of speech [2].

Several single channel dereverberation algorithms have been proposed in the last years. Cepstral approaches transform a deconvolution problem in the time domain into a simple subtraction in the cepstral domain [3]. These methods are effective for short reverberation filters and are widely used in speech recognition context as they allow to reduce the effect of the transmission channel. However, they cannot tackle the tail of reverberation when the filter is longer than the cepstral analysis window, rendering them impractical for usual reverberation.

Inverse filtering techniques exploit the effect of reverberation on the Linear Prediction (LP) residual of the signal. The inverse early reflections filter is found by adaptively maximizing the kurtosis [4, 5] or the skewness [6] of the LP residual. Late reverberation is further suppressed by spectral subtraction techniques [5, 6]. These techniques suffer from slow convergence rates and introduce pre-echo artifacts that need to be compensated in a postprocessing stage, adding some computational burden to the system.

Late reverberation is commonly addressed with spectral subtraction techniques as suggested in [2] and the Maximum Sparsity Power Prediction method in [7]. The late reverberation power spectral density (*psd*) is usually estimated as a delayed and damped version of the observed signal. The damping factor is defined as a function of the reverberation time (T_{60}) of the enclosure. If T_{60} is known, we obtain a reliable estimator of the reverberation *psd* that is used to design a time-frequency dereverberation filter. However, the accurate estimation of T_{60} is a research problem itself [8, 9] and needs important computational ressources. Late reverberation can also be predicted by exploiting the long term redundancies of reverberant signals as presented in [10].

In this paper late reverberation is modeled in the frequency domain as a linear combination of previously observed signal frames. We impose a sparsity constraint on the linear combination and propose a reverberation suppression algorithm based on the Lasso [11]. We design a time-frequency dereverberation filter based on Ephraim and Malah's spectral subtraction rule [12] to produce high quality dereverberated signals. The presented algorithm compares to state-of-the art dereverberation methods for a large range of T_{60} without needing any additional adaptation of its parameters. This leads to a fast and robust method that is suitable for real-time applications.

This paper is organized as follows: in Section 2 we introduce a sparse prediction model for late reverberation. In Section 3 we propose some strategies to reduce the complexity of the method. Experimental results are presented in Section 4 and some conclusions are drawn in Section 5.

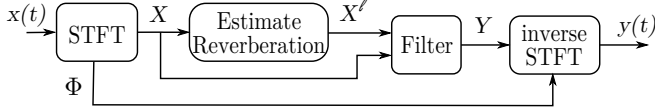## 2. FRAMEWORK FOR LATE REVERBERATION SUPPRESSION

The proposed method is based on a speech enhancement framework as illustrated in Figure 1. First, we will introduce our model for the estimation of late reverberation before we briefly discuss the choice of a spectral filter.

2.1. Sparse linear prediction model for late reverberation

Let $x(t)$ be the time domain reverberated signal. The signal is passed through a Short Time Fourier Transform (STFT) filterbank and we denote X the magnitude of the STFT. The phase matrix Φ is stored

Fig. 1. Reverberation suppression framework

for the reconstruction of the filtered signal. $X_{k,n}$ represents the element belonging to the kth frequency channel and the nth time frame of the matrix X.

In the frequency domain the reverberated signal can be written as:

$$X_{k,n} = X_{k,n}^e + X_{k,n}^\ell \, , \tag{1}$$

where $X_{k,n}^e$ and $X_{k,n}^\ell$ represent respectively the early and late reverberation terms [2]. In this paper we only address the estimation of $X_{k,n}^\ell$. Reverberation is produced by delayed and damped replicas of the direct sound. We propose to predict $X_{k,n}^\ell$ in each frequency channel as a linear combination of L signal frames that precede the current frame:

$$\hat{X}_{k,n}^\ell = \sum_{i=0}^{L-1} \alpha_i X_{k,n-i-\delta}. \tag{2}$$

A delay of δ frames is introduced in order to separate the effects of early and late reflections for the prediction. This results in the following model for the observed signal:

$$X_{k,n} = X_{k,n}^e + \sum_{i=0}^{L-1} \alpha_i X_{k,n-i-\delta}. \tag{3}$$

Late reverberation is modeled as a redundancy term that can be linearly predicted from past observed frames whereas the early component $X_{k,n}^e$ is the residual of the prediction. This model has been suggested in [13] where every past frame contributes to the estimation based on the long term correlations of the reverberated signal. In this paper, we assume that only a few past frames significantly contribute to the late reverberation estimate. In other words, we assume a sparse predictor $\alpha = [\alpha_0 \ \ldots \ \alpha_{L-1}]^T$. In a convex optimization framework, sparsity can be promoted by constraining the ℓ_1 norm of the predictor. Under this assumption we formulate our dereverberation problem as an instance of the Lasso [11]:

$$\underset{\alpha}{\text{minimize}} \ ||X_{k,n} - \boldsymbol{D_{k,n}} \, \alpha||^2 \quad \text{s.t.} \quad ||\alpha||_1 \leq \lambda, \tag{4}$$

For each time frame n and each frequency channel k, we solve (4) for the sparse predictor α that best explains the current observation $X_{k,n}$ as a linear combination of a certain signal-based dictionary $\boldsymbol{D_{k,n}}$ given a regularization parameter λ. The Lasso is solved using the Least Angle Regression (LARS) algorithm [14] which is known to be very efficient as long as the dimension of the problem is kept small.

Given the predictor α, late reverberation is estimated as:

$$\hat{X}_{k,n}^\ell = \boldsymbol{D_{k,n}} \, \alpha. \tag{5}$$

Using (2) and (5) it is clear that the signal-based dictionary $\boldsymbol{D_{k,n}}$ corresponding to this model is given by:

$$\boldsymbol{D_{k,n}} \triangleq \left[\ X_{k,n-\delta}, \quad \ldots \quad , X_{k,n-\delta-L+1} \ \right] \tag{6}$$

Note that if we set $L = 1$ the estimator in (2) becomes $X_{k,n}^\ell = \alpha X_{k,n-\delta}$ as proposed in [2] and [7]. Our model extends these approaches and selects the elements that are most relevant for the linear prediction. The proposed reverberation model does not rely on a physical model. Instead, we use a learning approach to obtain the parameter λ yielding the best reverberation suppression in a given acoustic condition. Our approach is different from the method in [15]. This technique estimates the clean speech spectrogram by maximizing the sparsity of the reverberated one while our method only assumes the sparsity of the linear predictor. In addition we proposed a framework suitable for *online* processing while [15] is oriented for *batch* processing.

2.2. Spectral filtering

Once we have estimated the *psd* of late reverberation we design a spectral filter G based on Ephraim and Malah's MMSE-log spectral amplitude estimator [12] aimed to filter X^ℓ out of X. We use the so called *decision directed approach* [16] to get the *a priori* and *a posteriori* Signal to Interference Ratios. Both are needed to compute G as described in [12]. In order to avoid annoying musical noise artifacts, we introduce a lower bound G_{min} to the values taken by G. Finally, we obtain the dereverberated spectrogram Y by element-wise multiplication:

$$Y = G \, X \tag{7}$$

We finally apply the phase of the reverberated signal Φ to the magnitude matrix Y and compute an inverse STFT to obtain the time domain dereverberated signal $y(t)$.

3. REDUCING THE COMPLEXITY OF THE ESTIMATOR

Late reverberation is estimated on the STFT magnitude matrix $X \in \mathbb{R}^{K \times M}$ composed of K frequency channels and M time frames. According to the model introduced in the previous section, one must solve problem (4) for each of the $K \times N$ time-frequency bins. This leads to a high computational burden. We propose in this section to reduce the complexity of the method through to block-wise and subband processing.

3.1. Block-wise processing

First we reduce the number of times problem (4) is solved by working in a block by block basis. Let us introduce the observation vector $V_{k,n} \in \mathbb{R}^N$ given by:

$$V_{k,n} \triangleq \left[\ X_{k,n} \quad \ldots \quad X_{k,n-N+1}. \ \right]^T. \tag{8}$$

For each frequency channel k, the N element vector $V_{k,n}$ is used to estimate simultaneously N frames of late reverberation. To this aim, successive observation vectors $V_{k,n}$ are concatenated to form a dictionary $\boldsymbol{D_{k,n}} \in \mathbb{R}^{N \times L}$ associated to the current observation and defined by:

$$\boldsymbol{D_{k,n}} \triangleq \left[\ V_{k,n-\delta} \quad V_{k,n-\delta-1} \quad \ldots \quad V_{k,n-\delta-L+1} \ \right] \tag{9}$$

We use (8) and (9) to compute the late reverberation predictor α by solving the Lasso problem:

$$\underset{\alpha}{\text{minimize}} \ ||V_{k,n} - \boldsymbol{D_{k,n}} \, \alpha||^2 \quad \text{s.t.} \quad ||\alpha||_1 \leq \lambda \tag{10}$$

Given the current predictor α, we can estimate a vector of late reverberation, denoted by $V_{k,n}^\ell \in \mathbb{R}^N$ and given by:

$$V_{k,n}^\ell = \boldsymbol{D_{k,n}} \alpha. \tag{11}$$

As we work with non overlapping blocks, the Lasso must only be solved $K \times \frac{M}{N}$ times. However increasing N reduces the temporal resolution of the estimator. According to our experiments, a good trade-off between complexity and resolution is obtained by choosing N such that $N\frac{R}{f_s} < 64$ms, where R denotes the hop size of the STFT and f_s the sampling rate.

3.2. Subband processing

The *psd* of reverberation is frequency dependent but varies slowly between neighbor frequencies. Hence we can reasonably reduce the frequency resolution of the late reverberation estimator by passing the magnitude matrix X through an arbitrary filterbank. This procedure is depicted on the left of Figure 2. First, we define a J-segments partition $\mathcal{P}$ of the interval $[1, K]$. For every segment of $\mathcal{P}$, we compute the average of its elements to produce the jth channel of the subsampled matrix $\tilde{X} \in \mathbb{R}^{J \times M}$. Then we build the corresponding observation vector $\tilde{V}_{k,n} \triangleq \begin{bmatrix} \tilde{X}_{k,n} & \dots & \tilde{X}_{k,n-N+1} \end{bmatrix}^T$ and the subsampled dictionary $\tilde{\boldsymbol{D}}_{\boldsymbol{k,n}}$, obtained by concatenation of adjacent observation vectors $\tilde{V}_{k,n}$. We solve the Lasso and get J predictors α associated to each subband. Late reverberation is then estimated with the dictionary introduced in Eq. (9). To achieve this we must assign the J predictors to the K frequency channels as shown on the right of Figure 2. Finally, we solve Equation (11) to recover the estimate.
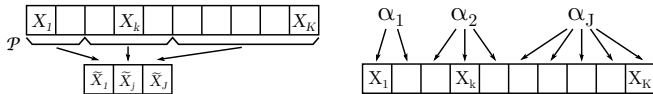
Fig. 2. Subband processing. *Left:* Building $\tilde{X}$ for the estimation of J predictors. *Right:* Assigning a predictor to each channel of X.

Our experiments show that the nature of the partition $\mathcal{P}$ is not critical. Even if we work with very few subbands ($J = 10$ instead of $K = 257$), we do not observe any significant degradation when compared to the method presented in Section 2.

The subsampling along the time and frequency axis reduces greatly the computation time because problem (4) must be solved only $J \times \frac{M}{N}$ times.

4. EXPERIMENTS AND EVALUATION

4.1. Settings

For the evaluation, we use anechoic speech samples taken from the TIMIT database. We use a subset of the database with 10 female and 10 male speakers, each one pronouncing one sentence. These signals are then convolved with two different sets of Room Impulse Responses (RIR). The first set is intended to evaluate the algorithm in realistic situations and contains measured RIRs taken from the AIR database [17]. The selected impulse responses correspond to a hands free use of a mock up mobile telephone in 10 different rooms. For the second set, we use the Fast Image-Source Method [18] to simulate the RIRs of a room with dimensions [3x4x5]m and T_{60} ranging from 100ms to 1.2s. This set will be used to evaluate the performance of the method as a function of the T_{60}.

The reverberated signals $x(t)$, sampled at 16 kHz, are processed with the proposed algorithm to produce the dereverberated signal $y(t)$. We evaluate the method using the Signal to Reverberation Modulation Ratio (SRMR [19]) and the Log Spectral Distorsion (LSD [2]) measures. For each speech sample we compute the SRMR on $x(t)$ and $y(t)$ and study the SRMR improvement defined as:

$$\Delta\text{SRMR} \triangleq \text{SRMR}\,[y(t)] - \text{SRMR}\,[x(t)] \qquad (12)$$

To evaluate the spectral distortions introduced by the processing we compute the LSD of $y(t)$ related to $d(t)$, the early echoes signal. We obtain $d(t)$ by filtering the anechoic signal with the RIR truncated 80 ms after the arrival of the direct sound.

We analyze each signal using a STFT filterbank with a 32 ms Hamming window and a hop size of 8 ms. For the subband processing, we use an octave filterbank to build a subsampled spectrogram with $J = 10$ frequency channels instead of the $K = 257$ available from the STFT. The octave filterbank is obtained by recursively performing a diadic partition of the available frequency bins. We performed a grid search on each parameter introduced in Section 3 and selected the value yielding the maximum SRMR on $y(t)$. From this analysis, the dictionary length is set to $L = 10$ and the delay is set to $\delta = 5$ frames. This delay corresponds to 40 ms of speech which is sufficient to remove the direct signal from the dictionary. For the block processing we use an observation length of $N = 8$ frames, corresponding to 64 ms long segments of speech signal. We solve problem (10) using the MATLAB's *mexLasso* function from the SPAMS optimization toolbox[1]. The estimated late reverberation is smoothed with a single pole low-pass filter with time constant $\tau = 32$ ms to compensate the discontinuities introduced by the block-wise processing. The smoothing constant for the *decision directed approach* and the spectral floor for the filter are set to $\beta = 0.98$ and $G_{min} = -12$ dB respectively.

4.2. Dereverberation experiments

#### 4.2.1. *Choice of the subsampling scheme*

In a first experiment, we evaluate the influence of the two subsampling strategies presented in Section 3 when used individually and together. In addition, we run 100 iterations of each approach on the whole database and we evaluate the average CPU time needed for the execution. We use a computer with an Intel Core i7-640M processor at 2.8GHz and 4 GB RAM. We analyze the average Real Time Factor (RTF) defined as the ratio between the processing time and the total length of the speech samples.

Subsampling	ΔSRMR	LSD[dB]	RTF[%]
No	**3.98** $\pm$ 0.66	3.16 $\pm$ 0.63	124.6
Time	2.96 $\pm$ 0.64	3.17 $\pm$ 0.61	25.3
Frequency	3.45 $\pm$ 0.60	**2.90** $\pm$ 0.58	8.7
Both	3.11 $\pm$ 0.60	3.29 $\pm$ 0.62	**3.7**

Table 1. Average scores and standard deviations in different subsampling configurations.

The results of the evaluation are summarized in Table 1. When we do not apply any subsampling, the proposed method yields the best results in terms of reverberation suppression but it is also very

---

[1]http://spams-devel.gforge.inria.fr/

(a) SRMR improvement (b) LSD
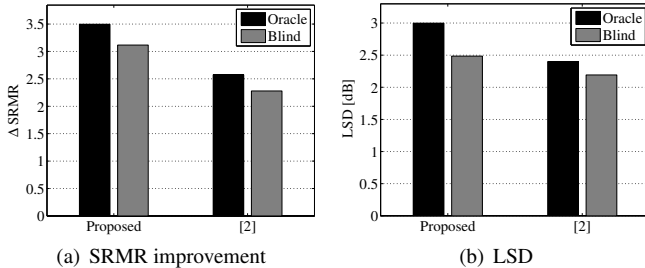
Fig. 3. Objective evaluation with recorded RIRs in *Oracle* and *Blind* conditions.

(a) SRMR improvement (b) LSD

Fig. 4. Objective evaluation with simulated RIRs as a function of T_{60}.

slow and impractical for real-time applications. We also observe that subsampling along the frequencies states for the major reduction of the complexity of the method. In addition, the estimated late reverberation introduces less spectral distortion that any other approach without significantly degrading the $\Delta SRMR$. The temporal subsampling degrades the reverberation suppression because of the reduced time resolution. Moreover the improvement of the RTF is limited because N must be kept small. Finally, with both time and frequency subsampling we have the fastest configuration but also the one introducing the more spectral distortion. It is interesting to notice that the scores are not significantly different and thus we can choose the subsampling scheme according to the available ressources. In the following we will only use the frequency subsampling as it keeps the average spectral distorsion low.
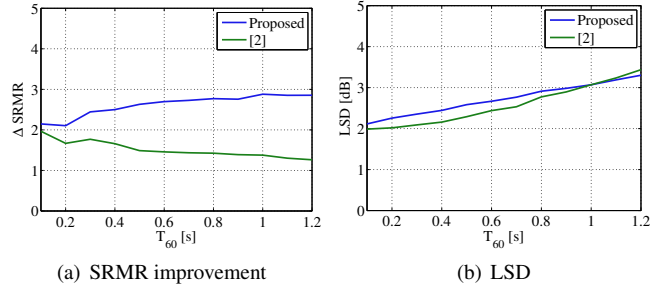
4.2.2. Comparison to the state-of-the-art

We compare our method to the efficient approach proposed by Habets in [2]. The same spectral filter with the same settings is used for both methods. Each method is steered by a single hyperparameter: T_{60} for [2] and λ for the proposed method. We consider two situations for the evaluation. In a first configuration, the optimal hyperparameters are found for each room by grid search and we evaluate the *oracle* performance of the algorithms. Then, we consider the *blind* case, where the hyperparameters are kept constant for every room. For this simulation we set $T_{60} = 300$ ms and $\lambda = 0.65$, which correspond to the optimal parameters in a room with $T_{60} = 300$ ms. This experiment is intended to evaluate the sensitivity of the algorithms to errors on the estimation of their hyperparameters. Figure 3 shows the average SRMR improvement and the LSD for both methods in the oracle and blind case.

We observe in Figure 3(a) a positive improvement of the SRMR for both methods which confirms a reduction of late reverberation. As expected, the oracle case leads to better dereverberation compared to the blind case. In both situations, the proposed method performs better than [2]. However, this increase in the dereverberation performance is obtained at the cost of additional spectral distortion as depicted in Figure 3(b). The proposed method introduces in average 0.6 dB of additional distortion compared to [2]. According to our informal listening tests, this does not affect the perceptual quality[2].

Now we compare the scores between the blind and oracle cases. In blind conditions, the reverberation suppression is less effective for both methods. As a consequence of this, less distortion is in-

troduced. However, the slight loss in $\Delta SRMR$ observed with the proposed method yields a more significant reduction of the LSD leading to only 0.3dB of additional distortion with respect to Habets method. From this analysis we argue that the proposed method can work in blind conditions without any significant loss in terms of reverberation reduction compared to the ideal case. By avoiding the estimation of the hyperparameter, we save important computational resources. The proposed method has an average RTF of 8.7% while our implementation of the method from Habets has an RTF of 1.8%. The competing method is clearly faster but it needs additional resources for the estimation of T_{60}. Our method is fast enough to work in real-time conditions even if it is slower than [2].

Finally, in Figure 4 we evaluate both methods in blind conditions with the simulated RIRs. The SRMR improvement is confirmed for all the considered T_{60} and the proposed method achieves better reverberation suppression. Regarding the LSD, our method introduces slightly more distortion than the competing one but the gap between them is reduced when T_{60} increases. The proposed method shows satisfying reverberation suppression capabilities for every T_{60} without setting a room dependent hyperparameter λ. Moreover, the spectral distortion is bounded to levels that compare with the state of the art even for short T_{60}.

5. CONCLUSION

In this paper we proposed a new algorithm for the suppression of reverberation in the frequency domain. We modeled late reverberation as a linear combination of previous observations as suggested in [13]. By constraining this linear model to be sparse our problem fits into a Lasso framework that can be efficiently solved with sparse optimization techniques. The estimated reverberation was filtered in a spectral subtraction framework adapted to this particular problem. We also proposed two strategies to reduce the complexity of the estimator. The proposed method performs slightly better than the state of the art algorithm of [2] in terms of SRMR without introducing much additional distortion . We tested our method in oracle and blind conditions and found that the dereverberation performance of our method is not significantly affected when we do not estimate the optimal hyperparameters for the model. This allows the proposed method to perform blind dereverberation at least in a certain range of reverberation times. In addition, the proposed algorithm is sufficiently fast for real time applications.

---

[2]Audio examples are available online:
http://perso.telecom-paristech.fr/~nlopez

6. REFERENCES

[1] A. K. Nábêlek, T. R. Letwoski, and F. M. Tucker, "Reverberant Overlap-and Self-Masking in Consonant Identification." *Journal of the Acoustical Society of America*, 1989.

[2] E. A. P. Habets, S. Gannot, and I. Cohen, "Late Reverberant Spectral Variance Estimation Based on a Statistical Model," *Signal Processing Letters, IEEE*, vol. 16, no. 9, pp. 770–773, 2009.

[3] D. Bees, M. Blostein, and P. Kabal, "Reverberant Speech Enhancement Using Cepstral Processing," in *Proc. International Conference on Acoustics, Speech and Signal Processing, 1991*, Toronto, Canada, 1991, pp. 977–980.

[4] B. W. Gillespie, H. S. Malvar, and D. A. F. Florêncio, "Speech Dereverberation via Maximum-Kurtosis Subband Adaptive Filtering," in *Proc. International Conference on Acoustics, Speech and Signal Processing, 2001*, Salt Lake City, USA, 2001, pp. 3701–3704.

[5] M. Wu and D. L. Wang, "A Two-Stage Algorithm for One-Microphone Reverberant Speech Enhancement," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 14, no. 3, pp. 774–784, 2006.

[6] S. Mosayyebpour, M. Esmaeili, and T. Gulliver, "Single-Microphone Early and Late Reverberation Suppression in Noisy Speech," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 21, no. 2, pp. 322–335, 2013.

[7] T. Yoshioka, "Speech Enhancement in Reverberant Environments," Ph.D. dissertation, Kyoto University, 2010.

[8] N. D. Gaubitch, M. Jeub, T. H. Falk, P. A. Naylor, P. Vary, and M. Brookes, "Performance Comparison of Algorithms for Blind Reverberation Time Estimation from Speech," in *Proc. International Workshop on Acoustic Signal Enhancement*, Aachen, Germany, 2012, pp. 1–4.

[9] N. Lopez, Y. Grenier, G. Richard, and I. Bourmeyster, "Low Variance Blind Estimation of the Reverberation Time," in *Proc. International Workshop on Acoustic Signal Enhancement*, Aachen, Germany, 2012.

[10] K. Kinoshita, T. Nakatani, and M. Miyoshi, "Spectral Subtraction Steered by Multi-Step Forward Linear Prediction for Single Channel Speech Dereverberation," in *Proc. International Conference on Acoustics, Speech and Signal Processing, 2006.*, vol. 1, Toulouse, France, 2006, pp. 817–820.

[11] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *J. R. Stat. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.

[12] Y. Ephraim and D. Malah, "Speech Enhancement Using a Minimum Mean-Square Error Log-Spectral Amplitude Estimator," *Acoustics, Speech, and Signal Processing, IEEE Transactions on*, vol. 33, no. 2, pp. 443–445, 1985.

[13] K. Furuya and A. Kataoka, "Robust Speech Dereverberation Using Multichannel Blind Deconvolution With Spectral Subtraction," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 15, no. 5, pp. 1579–1591, 2007.

[14] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least Angle Regression," *The Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[15] H. Kameoka, T. Nakatani, and T. Yoshioka, "Robust Speech Dereverberation Based on Non-Negativity and Sparse Nature of Speech Spectrograms," in *Proc. International Conference on Acoustics, Speech and Signal Processing, 2009*, Taipei, Taiwan, 2009, pp. 45–48.

[16] Y. Ephraim and D. Malah, "Speech Enhancement Using a Minimum Mean Square Error Short-Time Spectral Amplitude Estimator," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 32, no. 6, pp. 1109–1121, 1984.

[17] M. Jeub, M. Schäfer, H. Krüger, C. Nelke, C. Beaugeant, and P. Vary, "Do We Need Dereverberation for Hand-Held Telephony?" in *Proc. Int. Congress on Acoustics (ICA)*, Sydney, Australia, 2010.

[18] E. Lehmann and A. Johansson, "Diffuse Reverberation Model for Efficient Image-Source Simulation of Room Impulse Responses," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, no. 6, pp. 1429–1439, 2010.

[19] T. Falk, C. Zheng, and W. Chan, "A Non-Intrusive Quality and Intelligibility Measure of Reverberant and Dereverberated Speech," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, no. 7, pp. 1766–1774, 2010.